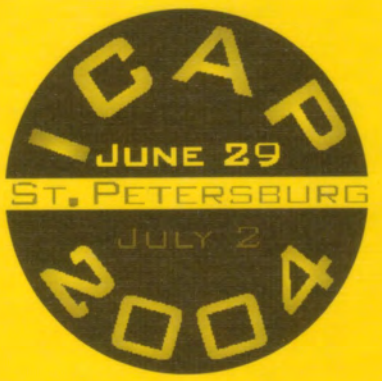

# ICAP 2004

Proceedings of the 8th International Computational Accelerator Physics Conference

St. Petersburg, Russia, June 29–July 2, 2004




*Editors*

**Dmitri Ovsyannikov**
**Martin Berz**
**Kyoko Makino**

# ICAP 2004

Proceedings of the 8th International Computational Accelerator Physics Conference

St. Petersburg, Russia, June 29 – July 2, 2004


*Editors*

Dmitri Ovsyannikov

*St. Petersburg State University, Russia*

Martin Berz
Kyoko Makino

*Michigan State University, MI, USA*

ELSEVIER

doi:10.1016/j.nima.2005.11.162

# Editorial

We are happy to present these Proceedings of the Eighth International Computational Accelerator Physics Conference (ICAP, 2004), which was held in June 29–July 2, 2004 at Saint-Petersburg State University in St. Petersburg, Russia.

The main aim of the conference was to bring together researchers working on different theoretical and practical problems of accelerating devices, including their design, construction, maintenance, and use. It is our pleasure to say that the Conference was successful in achieving this goal. Forty-six institutions from all around the world—from Germany, Russia, the United States, and many other countries—had their work presented at ICAP 2004. One hundred and thirty-nine talks and poster reports were given by 127 participants. The 4-day program of the Conference consisted of 21 oral sessions divided into subfields, and one common poster session, which provided enough room and time for discussions and communication.

The best works in each of the ten conference topics were selected to form this volume of proceedings. We hope this book will be useful for many researchers in computational accelerator physics.

Dmitri Ovsyannikov, Martin Berz, Kyoko Makino
(*Guest Editors*)

0168-9002/$ - see front matter 
doi:10.1016/j.nima.2005.11.163

# Organization

**ICAP 2004**

8th International Computational Accelerator Physics Conference

***Organized by***

Saint-Petersburg State University: V.I. Zubov, Institute of Computational Mathematics & Control Processes, and Faculty of Applied Mathematics & Control Processes

***In cooperation with***

D.V. Efremov, Institute of Electrophysical Apparatus (NIIEFA), Saint-Petersburg, Russia
Joint Institute of Nuclear Research, Dubna, Russia
Lawrence Berkeley National Laboratory, USA
Michigan State University, USA
Moscow Engineering Physics Institute (State University), Russia
Moscow Institute of Physics and Technology (State University), Russia
Technische Universität Darmstadt, Germany
US Department of Energy

***Supported by***

Russian Foundation for Basic Research
Federal Agency for Nuclear Energy (Russia)

**Chairman**

D. Ovsyannikov, St. Petersburg State University, Russia

**Organizing Committee**

M. Berz, Michigan State University, USA
B. Bogdanovich, Moscow Engineering Physics Institute (State University), Russia
R. Brinkmann, DESY, Germany
S. Chattopadhyay, JLab, USA
A. Dragt, University of Maryland, USA
A. Friedman, LLNL, USA
I. Hoffmann, GSI, Germany
K. Ko, SLAC, USA
N. Kudryavtsev, Moscow Institute of Physics and Technology (State University), Russia
K. Makino, Michigan State University, USA
D. Ovsyannikov, St. Petersburg State University, Russia
S. Russenschuck, CERN, Switzerland
R. Ryne, LBNL, USA
G. Shirkov, Joint Institute of Nuclear Research, Russia
A. Skrinsky, Budker Institute of Nuclear Physics, Russia
M. Vorogushin, D.V. Efremov Institute of Electrophysical Apparatus, Russia
Th. Weiland, Technische Universität Darmstadt, Germany
K. Yokoya, KEK, Japan

**Program Committee**

S. Andrianov, St. Petersburg State University, Russia
M. Berz, Michigan State University, USA
A. Dragt, University of Maryland, USA
A. Friedman, LLNL, USA
I. Hoffmann, GSI, Germany
K. Ko, SLAC, USA
K. Makino, Michigan State University, USA
D. Ovsyannikov, St. Petersburg State University, Russia
R. Ryne, LBNL, USA
A. Skrinsky, Budker Institute of Nuclear Physics, Russia
Th. Weiland, Technische Universität Darmstadt, Germany
K. Yokoya, KEK, Japan
I. Yudin, Joint Institute of Nuclear Research, Russia

doi:10.1016/j.nima.2005.11.164

**International Advisory Committee**

N. Angert, GSI, Germany
M. Berz, Michigan State University, USA
S. Chattopadhyay, JLab, USA
Y. Cho, ANL, USA
P. Colestock, FNAL, USA
A. Dragt, University of Maryland, USA
A. Friedman, LLNL, USA
U. Gensch, DESY, Germany
B. Herrmannsfeldt, SLAC, USA
S. Holmes, FNAL, USA
C. Johnstone, FNAL, USA
S. Kawata, Utsunomiya University, Japan
K. Ko, SLAC, USA
A. Luccio, BNL, USA
S. Machida, KEK, Japan
K. Makino, Michigan State University, USA
A. Mondelli, SAIC, USA
A. Mosnier, SACLAY, France
S. Myers, CERN, Switzerland
G. Peters, US DOE
A. Richter, Technische Universität Darmstadt, Germany
U. van Rienen, University of Rostock, Germany
R. Ryne, LBNL, USA
A. Skrinsky, Budker Institute of Nuclear Physics, Russia
G. Vignola, INFN, Italy
A. Wrulich, PSI, Switzerland
K. Yokoya, KEK, Japan

# List of participants

Andrianov, Serge; St. Petersburg State University, St. Petersburg, Russia; sandrianov@yandex.ru
Ackermann, Wolfgang; Technische Universität Darmstadt, Institut für Theorie Elektromagnetischer Felder, Darmstadt, Germany; ackermann@temf.de
Alexandrov, Vladimir; Joint Institute of Nuclear Research, Dubna, Russia; Vladimir.Alexandrov@sunse.jinr.ru
Amundson, James; Fermilab, Batavia, IL, USA; amundson@fnal.gov
Antonov, Andrey; St. Petersburg State University, St. Petersburg, Russia; antonov_emsm@hotmail.com
Antropov, Igor; St. Petersburg State University, St. Petersberg, Russia; igor_antropov@mail.ru
Artamonov, Stanislav; Petersburg Nuclear Physics Institute, Gatchina, Russia; start@mail.pnpi.spb.ru
Balk, Monika; TU Darmstadt, Darmstadt, Germany; balk@temf.tu-darmstadt.de
Becker, Reinard; Institut für Angewandte Physik der Johann Wolfgang Goethe Universität, Frankfurt, Germany; rbecker@physik.uni-frankfurt.de
Berg, J. Scott; Brookhaven National Laboratory, Upton, NY, USA; jsberg@bnl.gov
Berz, Martin; Michigan SU, East Lansing, MI, USA; berz@msu.edu
Bobyleva, Liudmila; Joint Institute of Nuclear Research, Dubna, Russia; perel@nu.jinr.ru
Boine-Frankenheim, Oliver; GSI, Darmstadt, Germany; o.boine-frankenheim@gsi.de
Boscolo, Ilario; Milan University, Milano, Italy; ilario.boscolo@mi.infn.it
Candel, Arno; ETH Honggerberg, Zurich, Switzerland; candel@itp.phys.ethz.ch
Craievich, Paolo; Sincrotrone Trieste—ELETTRA, Trieste, Italy; paolo.craievich@elettra.trieste.it
Dehler, Micha; Paul Scherrer Institut, Villigen, Switzerland; Micha.dehler@psi.ch
Drivotin, Oleg; St. Petersburg State University, St. Petersburg, Russia; drivotin@yandex.ru
Durkin, Aleksandr; Moscow Radiotechnical Institute, Moscow, Russia; apdurkin@mtu-net.ru
Dzergach, Adolf; Moscow Radiotechnical Institute, Moscow, Russia; adzergatch@mtu-net.ru
Dyatlov, Andrey; St. Petersburg State University, St. Petersburg, Russia; nick@apmath.spbu.ru
Edamenko, Nikolay; St. Petersburg State University, St. Petersburg, Russia; nick@apmath.spbu.ru
Emanov, Fedor; Budker Institute of Nuclear Physics, Novosibirsk, Russia; fedor@gs.nsc.ru
Ermoshina, Marina; St. Petersburg State University, St. Petersburg, Russia; m-ermoshina@yandex.ru
Galea, Raphael; Columbia University/Max-Planck-Institut für Physik, Munich, Germany; galea@nevis.columbia.edu
Gjonaj, Erion; Technische Universität Darmstadt, TEMF, Darmstadt, Germany; gjonaj@temf.de
Golub, Yurii; Moscow Radiotechnical Institute, Moscow, Russia; yurigolub@mtu-net.ru
Gorbachev, Valery; Saratov State University, Saratov, Russia; gorbachevvp@info.sgu.ru
Gorbatchev, Eugene; Joint Institute of Nuclear Research, Dubna, Russia; gorbe@sunse.jinr.ru
Grote, Johannes; Michigan SU, East Lansing, MI, USA grotejoh@msu.edu
Gutnic, Michael; Institut de Recherche Mathematique Avancée, Université Louis Pasteur, Strasbourg Cedex, France; gutnic@math.u-strasbg.fr
Ivanov, Andrey; Budker Institute of Nuclear Physics, Novosibirsk, Russia; A.V.Ivanov@inp.nsk.su
Ivanov, Valentin; SLAC, Menlo Park, CA, USA; ivanov@slac.stanford.edu
Johnstone, Carol; Fermilab, Batavia, IL, USA; cjj@fnal.gov
Kabel, Andreas; SLAC, Stanford, CA, USA; akabel@slac.stanford.edu

doi:10.1016/j.nima.2005.11.165

Kawata, Shigeo; Utsunomiya University, Utsunomiya, Japan; kwt@cc.utsunomiya-u.ac.jp
Kazarinov, Nikolay; Joint Institute of Nuclear Research, Dubna, Russia; nyk@lnr.jinr.ru
Kikuchi, Takashi; Utsunomiya University, Utsunomiya, Japan; kwt@cc.utsunomiya-u.ac.jp
Kotina, Elena; St. Petersburg State University, St. Petersburg, Russia; ekotina@compmath.spbu.ru
Kozynchenko, Sergey; St. Petersburg State University, St. Petersburg, Russia; Sergey_Kozyntchenko@hotmail.com
Kozynchenko, Vladimir; St. Petersburg State University, St. Petersburg, Russia;
Krasilnikov, Mikhail; DESY, Zeuthen, Germany; mikhail.krasilnikov@desy.de
Krasnykh, Anatoly; SLAC, Stanford, CA, USA; krasnykh@SLAC.Stanford.EDU
Kurakin, Vyacheslav; Lebedev Physical Institute, Moscow, Russia; kurakin@x4u.lebedev.ru
Kvardakov, Vasily; Budker Institute of Nuclear Physics, Novosibirsk, Russia; kvardakoff@ngs.ru
Lau, Thomas; TU Darmstadt; Institut für Theorie Elektromagnetischer Felder, Darmstadt, Germany; lau@temf.de
Lebedev, Nikolay; Joint Institute of Nuclear Research, Dubna, Russia;
Lebedev, Valeri; Fermilab, Batavia, IL, USA; val@fnal.gov
Lebrun, Paul; Fermilab, Batavia, IL, USA; lebrun@fnal.gov
Li, Zenghai; SLAC, Menlo Park, CA, USA; lizh@slac.stanford.edu
Limberg, Torsten; DESY, Hamburg, Germany; torsten.limberg@desy.de
Lund, Steven; Lawrence Livermore National Laboratory, Berkeley, CA, USA; SMLund@lbl.gov
Makarov, Alexandr; Joint Institute of Nuclear Research, Dubna, Russia; makarov@nusun.jinr.ru
Makino, Kyoko; Michigan SU, East Lansing, MI, USA; makino@msu.edu
Mamaev, Yuri; State Polytechnic University, St. Petersburg, Russia; mamaev@spes.stu.neva.ru
Manikonda, Shashikant; Michigan State University, East Lansing, MI, USA; manikond@msu.edu
Masunov, Eduard; Moscow Engineering Physics Institute (State University), Moscow, Russia; masunov@dinus.mephi.ru
Matyshev, Alexander; State Polytechnic University, St. Petersburg, Russia; matyshev@mail.spbstu.ru
Mehrenberger, Michel; Université Louis Pasteur IRMA, Strasbourg, France; mehrenbe@math.u-strasbg.fr
Merkuriev, Sergey; St. Petersburg State University, St. Petersburg, Russia;
Miginsky, Sergey, Budker Institute of Nuclear Physics, Novosibirsk, Russia; S.V.Miginsky@inp.nsk.su
Mityanina, Natalia; Budker Institute of Nuclear Physics, Novosibirsk, Russia; mityanina@inp.nsk.su
Miyazaki, Shuji; Utsunomiya University, Utsunomiya, Japan; dt030106@cc.utsunomiya-u.ac.jp
Mytrochenko, Viktor; National Science Center ''Kharkov Institute of Physics & Technology'', Kharkov, Ukraine; mitvic@kipt.kharkov.ua
Nikiforov, Konstantin; St. Petersburg State University, St. Petersburg, Russia; nikiforov_k@mail.ru
Nikulin, Eugene; Institute for Nuclear Research RAS, Moscow, Russia; nikulin@inr.ru
Ostroumov, Petr; Argonne National Laboratory, Argonne, IL, USA; ostroumov@phy.anl.gov
Ovsyannikov, Dmitri; St. Petersburg State University, St. Petersburg, Russia; dovs@compmath.spbu.ru
Ovsyannikov, Aleksander; St. Petersburg State University, St. Petersburg, Russia; ovs@compmath.spbu.ru
Perelstein, Elkuno; Joint Institute of Nuclear Research, Dubna, Russia; perel@nu.jinr.ru
Perelstein, Oleg; Moscow SU, Moscow, Russia; oleg.perelstein@mail.ru
Petrenko, Alexey; Budker Institute of Nuclear Physics, Novosibirsk, Russia; petrenko@ngs.ru
Petrov, Valery; Efremov Institute of Electrophysical Apparatus, St. Petersburg, Russia; petrov_vi@rbcmail.ru
Petrunin, Vladimir; Efremov Institute of Electrophysical Apparatus, St. Petersburg, Russia; petrunin@lek.ru
Pickulin, Eugene; Saint-Petersburg Electrotechnical University ''LETI'', St. Petersburg, Russia; eugene_pickulin@mail.ru
Poklonskiy, Alexey; St. Petersburg State University/Michigan SU, St. Petersburg, Russia; poklonsk@fnal.gov

Polyakova, Rimma; Joint Institute of Nuclear Research, Dubna, Russia;
pevgeny@jinr.ru
Qiang, Ji; Lawrence Berkeley National Laboratory, Accelerator and Fusion Research Division, Berkeley, CA, USA;
jqiang@lbl.gov
Rabtsoun, Sergey; Joint Institute of Nuclear Research, Dubna, Russia;
Riabov, German; Petersburg Nuclear Physics Institute, Gatchina, Russia;
riabov@mail.pnpi.spb.ru
Riabov, V.; Petersburg Nuclear Physics Institute, Gatchina, Russia;
riabov@mail.pnpi.spb.ru
Rienen, Ursula van; Universität Rostock, Rostock, Germany;
ursula.van-rienen@etechnik.uni-rostock.de
Rubtsova, Irina; St. Petersburg State University, St. Petersburg, Russia;
Ruhl, Hartmut; University of Nevada, Reno, Reno, NV, USA;
ruhl@physics.unr.edu
Russenschuck, Stephan; CERN, Geneva, Switzerland;
stephan.russenschuck@cern.ch
Rybitskaya, Tatyana; Budker Institute of Nuclear Physics,
Novosibirsk, Russia;
tanij-r@mail.ru
Sagan, David; Cornell University, Ithaca, NY, USA;
dcs16@cornell.edu
Schmidt, Frank; CERN, Geneva, Switzerland;
Frank.Schmidt@cern.ch
Senichev, Yurij; FZ Jülich, Jülich, Germany;
Y.Senichev@fz-juelich.de
Shavrina, Tatiana; Joint Institute of Nuclear Research, Dubna, Russia;
Sheshin, Evgeniy; Moscow Institute of Physics and Technology (State University), Dolgoprydny, Moscow region, Russia;
sheshin@lafeet.mipt.ru
Shevchenko, Sergey; Institute for Analytical Instrumentation RAS,
St. Petersburg, Russia;
shev@iai.rssi.ru
Shevtsov, Vladimir; Joint Institute of Nuclear Research, Dubna, Russia;
Vladimir.Shevtsov@sunse.jinr.ru
Shirkov, Grigori; Joint Institute of Nuclear Research, Dubna, Russia;
shirkov@sunse.jinr.ru
Shishanin, Oleg; Moscow State Industrial University, Moscow, Russia;
shisha-n@msiu.ru
Sideris, Ioannis; Northern Illinois University, DeKalb, IL, USA;
sideris@nicadd.niu.edu
Sidorin, Anatoly; Joint Institute of Nuclear Research, Dubna, Russia;
sidorin@jinr.ru
Smirnov, Alexander; Joint Institute of Nuclear Research, Dubna, Russia;
smirnov@jinr.ru
Snopok, Pavel; St. Petersburg State University/Fermilab, St. Petersburg, Russia;
snopok@fnal.gov
Sonnendrücker, Eric; IRMA, Strasbourg, France;
sonnen@math.u-strasbg.fr
Spachmann, Holger; CST GmbH, Darmstadt, Germany;
Holger.Spachmann@cst.com
Stepanchuk, Viktor; Saratov State University, Saratov, Russia;
stepanchukvp@info.sgu.ru
Svistunov, Yury; Efremov Institute of Electrophysical Apparatus, St. Petersburg, Russia;
svistunov@luts.niiefa.spb.su
Tiunov, Mikhail; Budker Institute of Nuclear Physics, Novosibirsk, Russia;
M.A.Tiunov@inp.nsk.su
Toprek, Dragan; VINCA institute of Nuclear Sciences, Laboratory 011, Belgrade, Serbia & Montenegro;
toprek@eunet.yu
Trubnikov, Grigory; Joint Institute of Nuclear Research, Dubna, Russia;
trubnikov@jinr.ru
Tsyganov, Yuri; Joint Institute of Nuclear Research, Dubna, Russia;
tyura@sungns.jinr.ru
Tuzikov, Alexey; Joint Institute of Nuclear Research, Dubna, Russia;
atuzikov@mail.ru
Tverskoy, Mikhail; Petersburg Nuclear Physics Institute, Gatchina, Russia;
Tyukhtin, Andrey; St. Petersburg State University, St. Petersburg, Russia;
tyukhtin@pobox.ru
Varajun, Marina; St. Petersburg State University, St. Petersburg, Russia;
MIVarajun@newmail.ru
Vasyukhin, Nikita; FZ Jülich, Jülich, Germany;
N.Vasyukhin@fz-juelich.de
Vinogradov, Sergey; Moscow Radiotechnical Institute, Moscow, Russia;
svv@chat.ru
Vladimirova, Lyudmila; St. Petersburg State University, St. Petersburg, Russia;
Vyuga, Elena; St.-Petersburg State University of Engineering and Economics,
St. Petersburg, Russia;
sl_vyuga@mail.ru
Walstrom, Peter; Los Alamos National Laboratory, Los Alamos, USA;
walstrom@lanl.gov

Warnock, Robert; Stanford Linear Accelerator Center (SLAC), Stanford University, Stanford, CA, USA; warnock@slac.stanford.edu
Weiland, Thomas; TU Darmstadt, Darmstadt, Germany; thomas.weiland@temf.de
Wipf, Susan; DESY, Hamburg, Germany; mpyszw@mail.desy.de
Wolfheimer, Felix; TU Darmstadt; Institut für Theorie Elektromagnetischer Felder, Darmstadt, Germany; wolfheimer@temf.de
Yan, Yiton; SLAC, Menlo Park, CA, USA; yan@slac.stanford.edu
Yaramyshev, Stepan; GSI (Germany)/ITEP (Russia), Darmstadt, Germany; s.yaramishev@gsi.de
Yaskina, Tatiana; Budker Institute of Nuclear Physics, Novosibirsk, Russia; yask@gorodok.net
Yudin, Ivan; Joint Institute of Nuclear Research, Dubna, Russia; yudin@jinr.ru
Zadorozhny, Vladimir; Glushkov Institute of Cybernetics, National Academy of Sciences of Ukraine, Kiev, Ukraine; zvf@compuserv.com.ua
Zagorodnov, Igor; TU Darmstadt, Darmstadt, Germany; zagor@temf.tu-darmstadt.de
Zhukov, Vladimir; St. Petersburg State University, St.Petersburg, Russia; dvzhuk@nevsky.net
Zhukov, Denis; St. Petersburg State University, St. Petersburg, Russia; denis.v.zhukov@hotmail.ru
Zuev, Yurii; Efremov Institute of Electrophysical Apparatus, St. Petersburg, Russia; Yuri_Zuev@mail.ru

# Conference photos

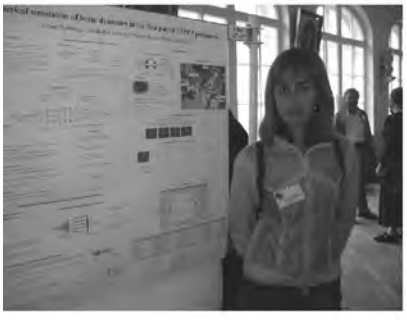

doi:10.1016/j.nima.2005.12.047









Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) xvii–xxi
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Contents

ICAP 2004
Proceedings of the 8th International Computational Accelerator Physics Conference
St. Petersburg, Russia, June 29–July 2, 2004

Editors: Dmitri Ovsyannikov, Martin Berz, Kyoko Makino



doi:10.1016/S0168-9002(06)00133-1

**Section 3. Field Computation and Analysis**



**Section 4. Precision Modeling of Dynamics and Fields**



**Section 5. High Current and High Emittance Beams**

**Section 6. High Performance Computing, Computer Modeling and Visualization**



**Section 7. Design of Accelerators and Components**

**Section 8. Charged Particle Beam Generation**



**Section 9. Beam Dynamics and Optimization**

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 1–10
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Long-term stability of the Tevatron by verified global optimization ☆

Martin Berz*, Kyoko Makino, Youn-Kyung Kim

*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*

Available online 28 November 2005

**Abstract**

The tools used to compute high-order transfer maps based on differential algebraic (DA) methods have recently been augmented by methods that also allow a rigorous computation of an interval bound for the remainder. In this paper we will show how such methods can also be used to determine rigorous bounds for the global extrema of functions in an efficient way. The method is used for the bounding of normal form defect functions, which allows rigorous stability estimates for repetitive particle accelerator. However, the method is also applicable to general lattice design problems and can enhance the commonly used local optimization with heuristic successive starting point modification. The global optimization approach studied rests on the ability of the method to suppress the so-called dependency problem common to validated computations, as well as effective polynomial bounding techniques. We review the linear dominated bounder (LDB) and the quadratic fast bounder (QFB) and study their performance for various example problems in global optimization. We observe that the method is superior to other global optimization approaches and can prove stability times similar to what is desired, without any need for expensive long-term tracking and in a fully rigorous way.


*PACS:* 29.20.Dh; 29.27.Bd; 02.60.Pn; 06.20.Dk

*Keywords:* Maps with remainder; Taylor model; Global optimization; Linear dominated bounder LDB; Quadratic fast bounder QFB; COSY INFINITY; COSY-GO; Symbolic-numeric; Interval; Taylor model

## 1. Introduction

In this paper we describe a practical method to rigorously assess the long-term stability of storage rings and other repetitive systems. The approach is based on methods of rigorous global optimization and will be applied to the study of the dynamics of the Tevatron at Fermilab. The theoretical foundation of the stability estimate is the method of normal forms [1,2] that allows to determine a family of three approximate invariants $I$ of the motion. The defect of these invariants $I$, i.e. the quantity

$$d = \max(I(\mathscr{M}) - I)$$

where $\mathscr{M}$ is the transfer map describing one revolution, can be used to provide bounds for minimum stability. In fact, if the distance of the outermost approximate invariant to the disallowed region is $a$, then the number of iterations necessary to reach this region is apparently at least $a/d$. Thus, bounding $d$ from above allows to assert stability, and the sharpness of the bound directly determines the quality of the stability estimate.

High-order map methods relating final and initial conditions of phase space variables for one revolution of the accelerator by the transfer map $\mathscr{M}$ have grown into a very widely used tool since their inception [3]; many details can be found in [2]. Recently, methods have been developed that allow not only the computation of the maps themselves, but also rigorous bounds for their remainders [4–6]. As we shall illustrate, these methods can also be used very beneficially for the problem of rigorously finding the global maximum or minimum of a function, and thus for the rigorous estimate of stability time of the Tevatron.

☆This work was supported by the US Department of Energy, the National Science Foundation, the German National Merit Foundation, and an Alfred P. Sloan Fellowship. We thank Ray Moore for many fruitful discussions.

*Corresponding author.
*E-mail addresses:* berz@msu.edu (M. Berz), makino@msu.edu (K. Makino), kimyounk@msu.edu (Y.-K. Kim).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.035

In the following sections we will briefly outline general ideas about rigorous global optimization and how they can be enhanced using the tools to compute remainder bounds for maps. We will illustrate the performance of these methods with various examples, and then apply them to the study of the Tevatron normal form defect problem. We will arrive at a rigorous stability estimate for the Tevatron that is practically relevant.

## 2. Rigorous global optimizers

Rigorous, validated, or verified global optimization characterizes the attempt of finding bounds of the global maximum or minimum of a sufficiently smooth function over a given domain. This task is distinct from the widely known and applied method of local optimization in that not only a nearby local extremum is sought. Furthermore, enclosures for the extrema and the points where they are assumed are obtained in a fully mathematically rigorous way, including the accounting of all effects of numerical inaccuracy.

### 2.1. Interval-based methods

The simplest method for validated global minimization is the interval branch-and-bound approach that successively studies sub-boxes of the original domain and attempts to prove that they cannot contain the minimum, upon which they are discarded. Usually these algorithms are based on a stack of boxes to be studied, which is initialized with the original full domain, as well as a rigorous upper bound for the minimum, the so-called cutoff, which is initialized with the result of a function evaluation in interval arithmetic at a suitably chosen point.

There are many variations used in practice, but the most elementary approach is based on picking the oldest box from the stack, evaluating the function on the center of the box to possibly obtain an update for the cutoff, and determining a lower bound of the function over the box by interval evaluation. If this lower bound exceeds the current cutoff, the box can be discarded; if it does not, then the box will be split along its longest direction and the two resulting boxes will be added to the stack for further study. By executing all operations in outward rounded interval arithmetic, full rigor of the argument is retained. This is not the place to do justice to the wide field of interval global optimization, but we rather refer to some of the common references on the topic [7–10].

Different from local optimization, the complexity of which often scales polynomially in dimension, that of global optimization can frequently grow exponentially with dimension. This is exemplified by the function $f_n$ on $[-1,1]^n$ in $n$ variables $x_i$ given by $f(x_1,\ldots,x_n)=\sum_{i=1}^{n}\cos(x_i/2\pi)$; apparently, each term in the sum assumes a minimum for $x_i=\pm\frac{1}{2}$, so that there are $2^n$ local minima. While local optimizers will settle into one of these depending on the starting point chosen, global optimizers at the very least have to probe all of them.

Performing global optimization using interval tools can introduce further complications. The two most important ones are that

(1) Interval methods are known for the fact that the resulting range bounds contain potentially significant overestimation, depending on the complexity of the function; this is frequently referred to as the *dependency problem*.
(2) It is usually observed that as the global optimization algorithm progresses, the number of boxes in the stack that lie in the vicinity of a local minimum remains nearly constant for a long time and thus all these neighboring boxes have to be split in each step, slowing down the process of elimination [11]; this is frequently referred to as the *cluster effect*.

### 2.2. Control of the dependency problem

It was recognized in Ref. [12] that the computation of a map with remainder bound of a functional dependency can significantly reduce the dependency problem. The reason for this simple yet very powerful observation is that the dominating part of the description of the function is caught by the Taylor coefficients. Thus, any operations on representations of functions merely require floating point arithmetic for the treatment of coefficients, and thus no overestimation occurs (with the exception of small effects due to the limited mantissa length and the resulting floating point errors, which are rigorously accounted for [13]). Overestimation does occur, as with all interval arithmetic, within the arithmetic of the remainder bound; but the influence of this remainder bound is many orders of magnitude smaller than that of the coefficients and thus its harm is reduced accordingly.

To illustrate this phenomenon, let us consider the following simple one-dimensional function:

$$f(x)=\sum_{i=0}^{30}(-1)^i\frac{x^{2i}}{(2i)!}$$

which is recognized to be the power series representation of $\cos(x)$ up to order 60; over the domain $[0,4\pi]$, it represents $\cos(x)$ to an accuracy of better than $10^{-15}$. In Fig. 1 we show the overestimation of the true range of the function (which is assumed to be that of $\cos(x)$) for the naive interval method, as well as the more advanced approaches of mean value form and centered form (see for example Ref. [9]) for domains of width $2^{-j}$ for the values $j=1,\ldots,8$ around $x=\pi/4$. The results are compared with what is obtained by first determining the map representation of the function, and then evaluating the polynomial part of this map representation in interval arithmetic. Apparently all methods become sharper for larger $j$, but the sharpness of

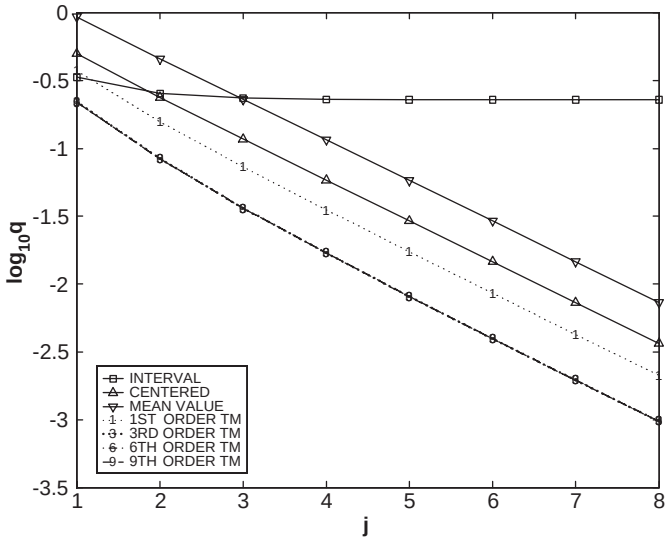



Fig. 1. Relative overestimation for the range bounding by the interval-, mean value-, and centered form methods, as well as with maps with remainders of orders 1, 2, 6, and 9 at $x = \pi/4$.

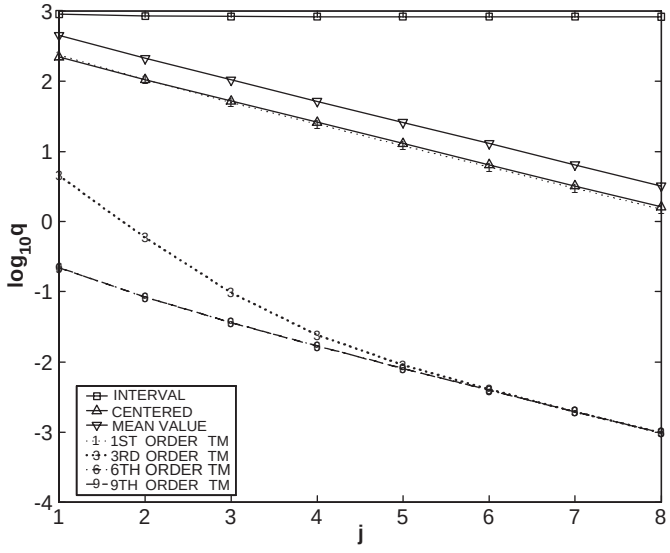



Fig. 2. Relative overestimation for the range bounding by the interval-, mean value-, and centered form methods, as well as with maps with remainders of orders 1, 2, 6, and 9 at $x = 2\pi + \pi/4$.

the map approach is superior because the map representation has less dependency than the original problem.

Performing the same comparison again for $x = 2\pi + \pi/4$, where the function has the same form except that the dependency problem is more significant, we see that the map approach provides a significant improvement, as seen in Fig. 2. Indeed, the higher order map methods achieve the same amount of overestimation as in the previous case, while the interval-, centered form- and mean value form methods suffer from overestimation due to the dependency problem.

A substantially more detailed study of the behavior of the method for many other cases can be found in Ref. [14].

### 2.3. The linear dominated bounder, LDB

As we saw, the map representation with remainder bound can significantly reduce the dependency problem. However, it is even possible to obtain sharper bounds yet by replacing the interval evaluation of the polynomial part with more sophisticated methods. In particular, we will in the following study the so-called linear dominated bounder (LDB) first introduced in Ref. [15]. It is based on using the map representation to determine a linear function that is a lower bound to the original function, and use this linearization and simple linear constraint methods to successively reduce the domain that can contain the extremum.

Within the framework of validated global optimization, after reducing the dependency problem with map methods as shown above, this will now lead to a tool to effectively reduce or eliminate boxes. While there is nothing that can help the inherent exponential complexity of high-dimensional problems, it will help improve practical performance significantly. Given a domain $D$, we consider the representation of the function $f$ by a Taylor polynomial $P$ and a rigorous bound for the remainder $I$, which can be obtained by the above-mentioned methods [4–6] such that

$$f(x) \in P(x) + I \quad \text{for all } x \in D.$$

In case we are away from a stationary point, the linear part of $P$ will dominate the behavior of the representation. The linear dominated bounder utilizes the linear part as a guideline for iterative domain reduction to bound $P$. Specifically, the algorithm is as follows.

**LDB Algorithm**

(1) Re-expand $P$ at the mid-point $c$ of $D$, call the resulting polynomial $P_m$ and the centered domain $D_1$.
(2) Make the linear coefficients $L_i$ of $P_m$ all positive by flipping coordinate directions as necessary; call the resulting polynomial $P_+$.
(3) In step $n$, compute interval bounds of the linear ($I_1$) and nonlinear ($I_h$) parts of $P_+$ in $D_n$. The minimum is then bounded by $[M, M_{\text{in}}] := \underline{I}_1 + I_h$. If applicable, lower $M_{\text{in}}$ by the current cutoff, the actual function value at the lower endpoints, and that at the midpoint.
(4) Let $d = \text{width}([M, M_{\text{in}}])$. If $d$ lies below a termination value, stop. Otherwise, if $L_i \neq 0$, the domain containing the minimum can be restricted to $\overline{D}_{n+1,i} := \min(\underline{D}_{n,i} + d/L_i, \bar{D}_{n,i})$. Re-expand $P_+$ at the mid-point $c$ of $D_{n+1}$. Prepare the new coefficients $L_i$, and continue with step 3.

Any errors associated with re-expansion and estimating point values are included in the remainder error bound interval. If $f$ is monotonic, the exact bound is often enclosed with high accuracy. If only a threshold cutoff test is needed, the resulting domain reduction or elimination is often very effective.

We now consider the performance of the LDB approach for the problem discussed in the previous sections. Figs. 3 and 4 contain the behavior of Figs. 1 and 2 on the left, as well as the results of using the LDB bounder on the right. Observe that even in a high-dependency case, the LDB bounder significantly outperforms the other methods, achieving accuracies that exceed those of conventional interval-based tools by 10 orders of magnitude.

### 2.4. The quadratic fast bounder, QFB

The natural next idea of the bounding of the polynomial is to explicitly bound the quadratic part of $P$. This will help in cases where the linear part alone is not dominating, for example, in the proximity of a local minimizer. Exact range bounding of a general quadratic polynomial has a complexity that scales exponentially with dimension and can thus be expensive. A preliminary scheme of a quadratic dominated bounder (QDB) is discussed in Ref. [15], and a more advanced method as well as the bounder QFB that will be discussed in the following is introduced in Ref. [16].

However, obtaining a lower bound of the quadratic polynomial near a local minimizer, which is the most important problem in global optimization, turns out to be much simpler. Indeed, in sufficient proximity of an isolated interior minimizer of the function $f$, the Hessian of $f$ is positive definite, and so the purely quadratic part of a representation $P+I$ that locally encloses $f$ also has a positive definite Hessian matrix $H$. The actual definiteness can be tested in a validated way using the common LDL or extended Cholesky decomposition [16]. The quadratic fast bounder (QFB) provides a lower bound of $P+I$ cheaply when the purely quadratic part is positive definite. It is based on the following observation.

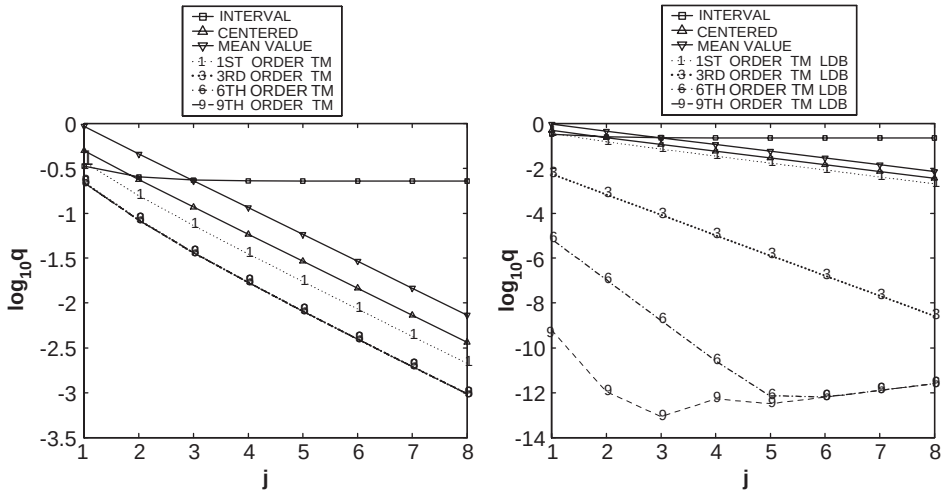



Fig. 3. Overestimation for the range bounding by the interval-, mean value-, centered form methods, as well as with maps with remainders of orders 1, 2, 6, and 9 at $x = \pi/4$ for naive interval evaluation of $P$ (left) and using the LDB bounder (right).

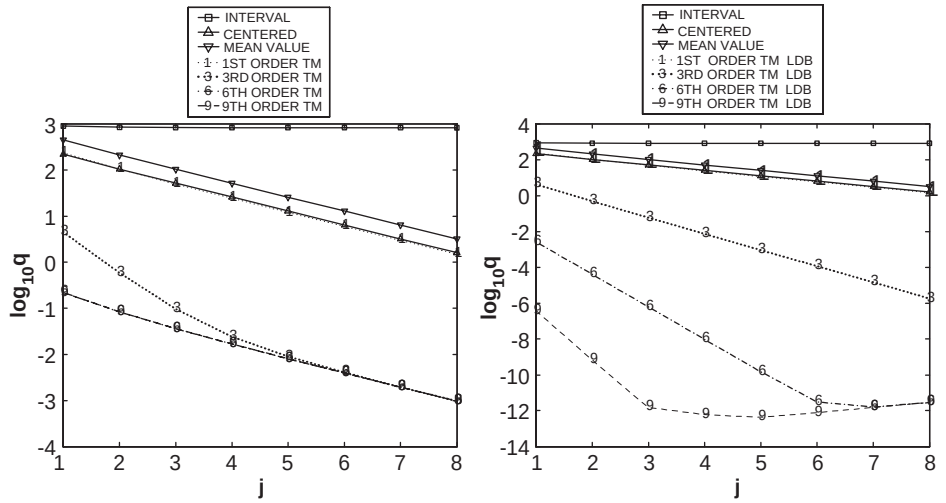



Fig. 4. Overestimation for the range bounding by the interval-, mean value-, centered form methods, as well as with maps with remainders of orders 1, 2, 6, and 9 at $x = 2\pi + \pi/4$ for naive interval evaluation of $P$ (left) and using the LDB bounder (right).

Given the quadratic form

$$Q(x) = \tfrac{1}{2} x^t \cdot H \cdot x + a \cdot x + b$$

where $H$ is symmetric, we determine the so-called Ordered LDL (OLDL) Decomposition (L: lower diagonal with unit diagonal, D: diagonal) of $H$ as follows:

(1) Pre-sort rows and columns by the size of their diagonal elements.
(2) Successively execute conventional $L^t DL$ decomposition step in interval arithmetic, beginning by representing every element of $H$ by a thin interval; in step $i$:
   (a) If $l(D(i,i)) > 0$ proceed to the next row and column.
   (b) If $u(D(i,i)) < 0$, exchange row and column $i$ with row and column $i+1$, $i+2, \ldots$ If a positive element is found, proceed as in step (a). If none is found, stop.

Similar to what is done in conventional Cholesky or LDL decomposition, it is useful to give careful consideration to the case in which $0 \in D(i,i)$, or $D(i,i)$ is too small. In this case, it is useful to apply a small correction $C$ to $H$, i.e. to study $H + C$ instead of $H$, such that all elements of $D$ are clearly positive or negative. As is typical also in non-validated LDL decomposition, $C$ is usually chosen to be diagonal, and $|C|$ is lumped into the remainder bound of the original problem. The resulting LDL decomposition has the important property that sufficiently near a local minimizer, $D$ will contain only positive elements that are sufficiently away from zero.

In the wider vicinity of a local minimum, the method may still possibly determine that all diagonal terms of $D$ can be proven positive, in which case positive definiteness is asserted. If this is not the case, at least the upper part where $D(i,i) > 0$ will describe a large positive semi-definite subspace.

In the following we study the case in the near proximity of a minimizer where the OLDL decomposition succeeds to assert that $H$ is positive definite. We will now use that knowledge to determine a sharp lower bound for $f$. Let $f \in P + I$ over $D$. We decompose the polynomial $P$ into two parts and write

$$P + I = (P - Q) + I + Q.$$

Then a lower bound for $P + I$ is obtained as

$$l(P + I) = l(P - Q) + l(Q) + l(I). \tag{1}$$

For the purpose of the QFB algorithm, we choose

$$Q = Q_{x_0} = \tfrac{1}{2}(x - x_0)^t H (x - x_0) \tag{2}$$

with some suitable $x_0 \in D$.

Since $H$ is positive definite, $l(Q_{x_0}) = 0$, and the value 0 is attained (at $x = x_0$). The remaining $P - Q_{x_0}$ does not contain pure quadratic terms anymore, but consists of linear as well as third and higher order terms $P_{>2}$. If $x_0$ is chosen to be the minimizer of the quadratic part $P_2$ of $P$ in $D$, then $x_0$ is also a minimizer of the remaining linear part (a consequence of the Kuhn–Tucker conditions), and so the lower bound estimate (1) is optimally sharp. Thus, by choosing $x_0$ sufficiently close to the minimizer in $D$ of $P_2$, the contribution of $P_2 - Q_{x_0}$ to the lower bound can become arbitrarily small.

A simple and efficient way to determine a sequence $x_0^{(n)}$ of candidates for $x_0$ is based on determining the "feasible descent direction"

$$g_i^{(n)} = \begin{cases} -\dfrac{\partial Q}{\partial x_i} & \text{if } x_i^{(n)} \text{ inside} \\ \min\left(-\dfrac{\partial Q}{\partial x_i}, 0\right) & \text{if } x_i^{(n)} \text{ on right} \\ \max\left(-\dfrac{\partial Q}{\partial x_i}, 0\right) & \text{if } x_i^{(n)} \text{ on left} \end{cases}$$

and to move in the direction of $g^{(n)}$ until we hit the bounding box or a one-dimensional quadratic minimum along the line. The method is very fast, can cover large ground per step, and in the terminology of constrained optimization, can change the set of active constraints very quickly. As a result, we obtain an inexpensive third order cutoff test that requires very few, if any, iterations to determine a useful $x_0$.

### 2.5. The validated global optimizer, COSY-GO

For the example problems of validated global optimization in the next sections, we apply three branch-and-bound methods available in the code COSY-GO [17]. The first one is the optimizer utilizing the LDB and QFB algorithms. We compare the performance with two other optimizers; one based on mere interval bounding (IN) and one based on bounding with centered form (CF). The subdomain box list management is performed in the same way for all three optimizers. At each step in which a subdomain is being studied for possible elimination, the following tasks are performed:

(1) A lower bound $l$ of the function is obtained using various bounding schemes in a hierarchical manner. If the lower bound is above the cutoff value, the box is eliminated; if not, the box is bisected.
   (a) Interval bounding of the polynomial $P$ is utilized for all optimizers; if it fails to eliminate the box, then additional tests are performed.
   (b) For the CF optimizer, centered form bounding is performed. For the COSY-GO optimizer, as a first test the polynomial part is evaluated in interval arithmetic. When it fails to eliminate the box, the LDB bounding and possible domain reduction is executed. If it also fails to eliminate the box, and if the quadratic part of the polynomial $P$ is positive definite, QFB bounding is performed.
(2) The cutoff value is updated using various schemes.
   (a) The conventional midpoint test is performed for all optimizers.

(b) For the COSY-GO optimizer, the linear and quadratic parts of $P$ are utilized to obtain a potential cutoff update. In particular, if the quadratic part of the polynomial is positive definite, the minimizer of the quadratic polynomial is tested. If the quadratic part is not positive definite, the minimizer of the quadratic part in the direction of the negative gradient is tested.

## 3. Validated global optimization—illustrative examples

In the following, we study the performance of the various versions of validated optimizers discussed above for some illustrative examples. We also compare with the performance of non-validated local optimization.

### 3.1. The Rosenbrock function

As a first example, we consider a relatively benign-looking function of two variables that contains a single local minimum; but the minimum occurs along a long and very shallow parabolic valley. This function originally proposed by Rosenbrock has the form

$$f_R(x,y) = 100(y-x^2)^2 + (1-x)^2$$

and it apparently assumes its minimum of 0 at $(x,y) = (1,1)$.

We study the function over the domain $[-1.5, 1.5]^2$, where its values range from the minimum 0 to more than 1400 near the point $(-1.5, -1.5)$. Fig. 5 shows a three-dimensional rendering of the function and the seemingly innocent parabolic valley generated by the first term, as well as a set of logarithmic contour lines for function values $0.1 \times 10^{i/2}$ for $i = 1, \ldots, 13$, revealing the substructure of the parabolic valley generated by the second term.

This function causes difficulties even for powerful conventional local minimizers because as soon as the optimizer is probing not exactly inside the valley, the direction of steepest descent is nearly perpendicular to the valley and hence in the almost completely wrong direction. Table 1 shows the performance of various tried-and-true optimizers in COSY [18,19], the LMDIF algorithm based on steepest descent with various enhancements, the Simplex algorithm which is particularly powerful for non-smooth problems, as well as the Anneal approach based on stochastic search of the global minimum by simulated annealing. The starting point was chosen at $(-1.2, 1.0)$, which is near the valley but on the opposite side of the true minimum. For LMDIF and Simplex, an accuracy tolerance of $10^{-12}$ and a maximum number of steps of 100,000 was specified, and Anneal was given a total number of 100,000 annealing steps. Table 1 shows the performance of the three non-validated optimizers. It can be seen that the usually quite powerful LMDIF algorithm performs rather poorly and is in this example significantly outperformed by the Simplex algorithm.

Now we utilize various validated global optimization tools to study the problem. In particular, we use the common Moore–Skelboe algorithm and determine ranges of the function by the mere interval evaluation method, and also the usually more accurate centered form method. From the perspective of interval methods, the function has a rather benign form with little dependency, and the squares that appear in the two terms can even be handled exactly. We compare these methods with the COSY-GO optimizer. Table 2 shows the performance of these three methods. It is apparent that for COSY-GO, the number of processing steps, i.e. the number of boxes considered, compares very favorably even with the performance of the best non-validated optimizer, aside from the fact that it of course determines a rigorous global optimum. On the other hand, the Interval and CF methods each require approximately one order of magnitude more steps, but still compare rather favorably to LMDIF and Anneal. It is also interesting to note that in a significant number of cases, LDB could reduce the size of a box under consideration without actually fully eliminating it.

Table 1
Performance of various local optimizers for finding the minimum of the Rosenbrock function in $[-1.5, 1.5]^2$ from the starting point $(-1.2, 1.0)$

| | LMDIF | Simplex | Anneal |
|---|---|---|---|
| Number of steps | 100,000 | 225 | 100,000 |
| Error in $f(x,y)$ | $1 \times 10^{-10}$ | $2 \times 10^{-13}$ | $3 \times 10^{-4}$ |
| Error in $(x,y)$ | $2 \times 10^{-5}$ | $4 \times 10^{-7}$ | $6 \times 10^{-3}$ |

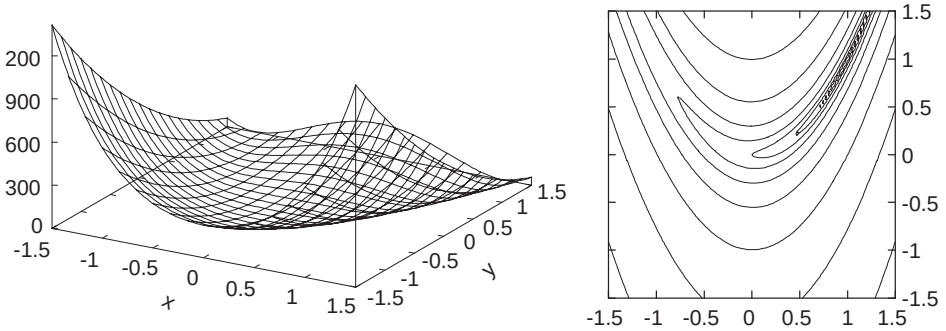



Fig. 5. The Rosenbrock function; three-dimensional rendering (left) and contour lines $0.1 \times 10^{i/2}$ for $i = 1, \ldots, 9$.

Table 2
Performance of various validated global optimizers for finding the minimum of the Rosenbrock function in $[-1.5, 1.5]^2$

| | IN | CF | COSY-GO |
|---|---|---|---|
| Total box processing steps | 1325 | 1325 | 143 |
| Max number of active boxes | 47 | 47 | 9 |
| Retained small boxes ($<10^{-6}$) | 15 | 15 | 1 |
| LDB domain reduction steps | – | – | 43 |

We now study in more detail the performance of the validated global optimizers. To this end we show all boxes that were studied in the process within the original domain box. Fig. 6 shows these for both the interval- and COSY-GO methods. Apparently, both methods successfully eliminate relatively large boxes away from the minimizer, while the boxes tend to get smaller and smaller as the minimizer is approached; however, the size of the boxes rejected by COSY-GO is significantly larger.

On the right of Fig. 6 we show the number of currently active boxes, as well as the value of the current upper bound of the minimum, the so-called cutoff value. For a

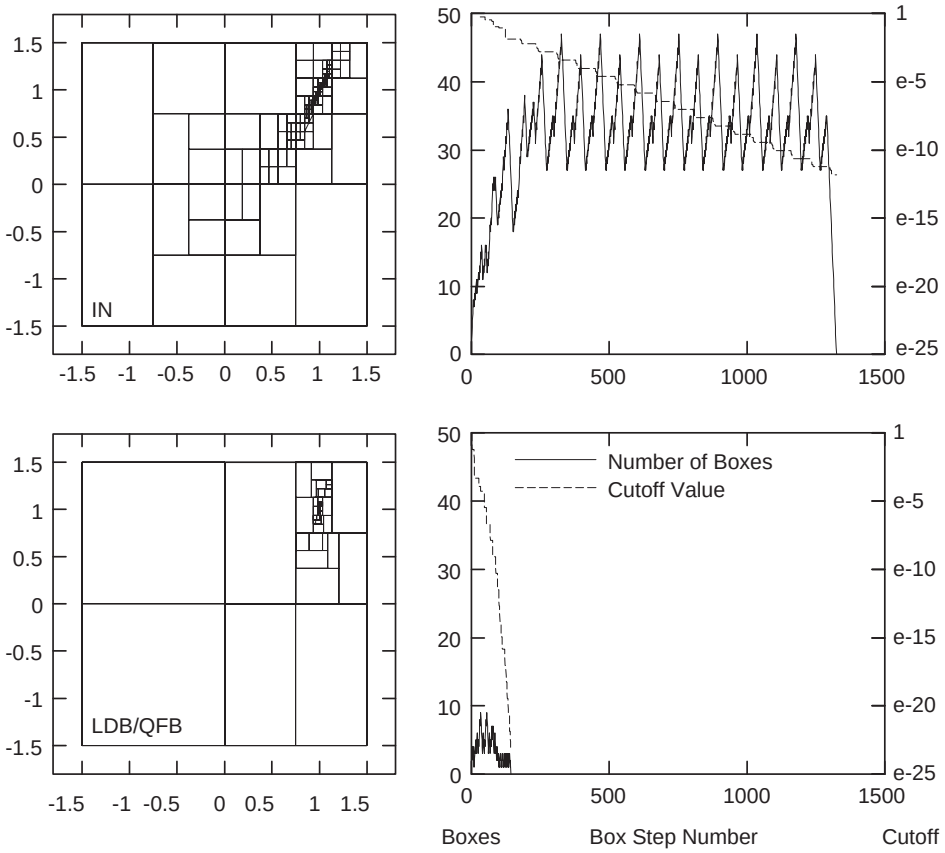



Fig. 6. The minimum search for the Rosenbrock function in $[-1.5, 1.5]^2$ by the interval and the LDB/QFB optimizers. Left: subdomain boxes studied. Right: number of active boxes and cutoff value as function of step number.

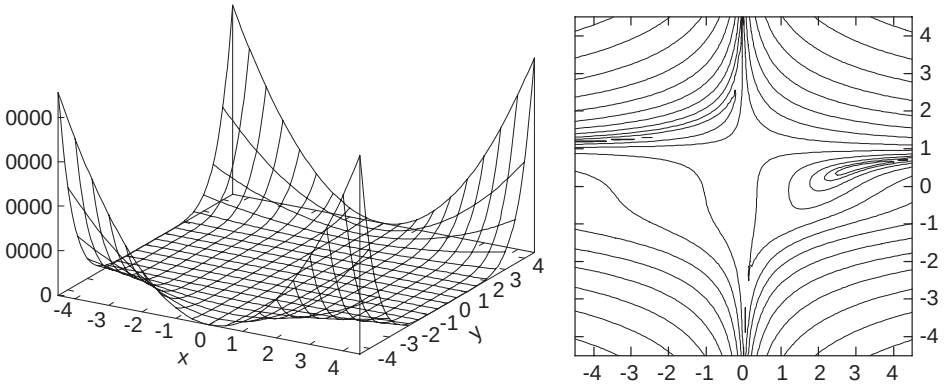



Fig. 7. The Beale function; three-dimensional rendering (left) and contour lines $0.1 \times 10^{i/2}$ for $i = 1, \ldots, 13$.

long time, the interval method maintains a nearly constant number of around 35 active boxes, while COSY-GO never require more than 10 active boxes, and soon keeps only one or two.

### 3.2. The Beale function

The next example is the Beale function [20]

$$f_B(x_1,x_2) = (1.5 - x_1(1-x_2))^2 + (2.25 - x_1(1-x_2^2))^2 + (2.625 - x_1(1-x_2^3))^2.$$

The problem is to find the minimum in the initial domain $[-4.5,4.5]\times[-4.5,4.5]$; one easily verifies that the function assumes the minimum 0 at $(3,0.5)$. Fig. 7 shows the behavior of the function, exhibiting a nearly flat extended valley containing the minimum. However, logarithmic contour lines of levels $0.1\times 10^{i/2}$ for $i=1,\ldots 13$ show a subtle substructure of the valley, having a trench-like shape near the minimizer $(3,0.5)$, but another trench of small function values in the upper left quadrant.

In particular this additional "trench" of small function values makes optimization difficult. We begin the study by using the three default optimizers in COSY, with an accuracy tolerance of $10^{-12}$ and a maximum of 100,000 iterations. Starting at $(x,y)=(4,-4)$ yields reasonable convergence in under 1000 steps for both LMDIF and Simplex, as shown in Table 3. On the other hand, the situation changes drastically when using as a starting value the point $(x,y)=(-4,4)$. In this case, the local optimizers get caught by the wrong trench of the objective function, and fail to get near the global minimum. The simulated annealing tool still succeeds in finding a reasonable approximation of the global minimum; the details of the performance are summarized in Table 4.

On the other hand, all validated global optimizers have no difficulty finding the global minimum accurately.

Table 3
Performance of various local optimizers for finding the minimum of the Rosenbrock function from the starting point $(4,-4)$ in $[-4.5,4.5]^2$

| | LMDIF | Simplex | Anneal |
|---|---|---|---|
| Number of steps | 825 | 123 | 100,000 |
| Error in $f(x,y)$ | $7\times10^{-13}$ | $3\times10^{-13}$ | $3\times10^{-2}$ |
| Error in $(x,y)$ | $2\times10^{-6}$ | $4\times10^{-7}$ | $8\times10^{-2}$ |

Table 4
Performance of various local optimizers for finding the minimum of the Beale function from the starting point $(-4,4)$ in $[-4.5,4.5]^2$

| | LMDIF | Simplex | Anneal |
|---|---|---|---|
| Number of steps | 100,000 | 100,000 | 100,000 |
| Error in $f(x,y)$ | $6\times10^{-1}$ | $4\times10^{0}$ | $1\times10^{-4}$ |
| Error in $(x,y)$ | $2\times10^{1}$ | $5\times10^{0}$ | $3\times10^{-2}$ |

Table 5
Performance of various validated global optimizers for finding the minimum of the Beale function in $[-4.5,4.5]^2$

| | IN | CF | COSY-GO |
|---|---|---|---|
| Total steps | 3407 | 3285 | 353 |
| Max boxes | 236 | 234 | 52 |
| Retained boxes | 25 | 25 | 3 |
| LDB reductions | – | – | 108 |

Table 6
Performance of the validated global optimizer COSY-GO for a generic normal form defect function

| Dimension | CPU time (s) | Max list | Total boxes |
|---|---|---|---|
| 2 | 5.747 | 11 | 31 |
| 3 | 38.49 | 44 | 172 |
| 4 | 346.9 | 357 | 989 |
| 5 | 3970 | 2248 | 6641 |
| 6 | 57,842 | 17,241 | 49,821 |

Significant differences exist, however, in the speed with which this is achieved; the results are summarized in Fig. 8 and Table 5. Square expressions in $f_B$ are executed as multiplications.

We observe no significant advantage in the CF optimizer compared to the interval optimizer, both of which maintain a list of about 60 active boxes for an extended time. On the other hand, the COSY-GO optimizer significantly outperformed both others because of more efficient box rejection and LDB domain size reduction, and requires a number of boxes comparable to the number of steps needed for the local optimizers in the case of a favorable initial condition.

## 4. Long-term stability of the Tevatron

We now return to the study of long-term stability of the Tevatron storage ring at collision. We utilize the four-dimensional map of lattice with parameters as described and optimized by Snopok et al. [21]. To illustrate the behavior of the normal form defect function, Fig. 9 shows a two-dimensional projection in which the two normal form radii are frozen at $5\times10^{-4}$, and the two normal form angles moving from 0 to $2\pi$. Already in these low-dimensional projections it becomes apparent that the functions have a large number of local minima and maxima, which makes finding their global extrema difficult.

In order to assess the performance of the COSY-GO global optimizer, we attempt a comparison with the validated global optimizer GLOBSOL [7]. As it turned out, the very high demands on the sharpness of the upper bound of the maximum of the normal form defect function did not allow the use of GLOBSOL for the specific problem at hand. So for the purpose of comparison, we

chose a different, less demanding normal form defect bounding problem based on the polynomials available at bt.pa.msu.edu. In Tables 6 and 7 we show some parameters describing the performance of COSY-GO and GLOBSOL for subspaces of different dimensionality. GLOBSOL succeeds to complete only the two-dimensional case in a reasonable time, while COSY-GO succeeds to complete even the six-dimensional case in similar time. For the two-dimensional case, COSY requires much less than 1% of the CPU time of GLOBSOL. The maximum list length and total number of boxes studied are rather manageable (Table 8).

We next apply COSY-GO to the realistic Tevatron problem. The one sigma emittance of the beam translates

Table 7
Performance of the validated global optimizer GLOBSOL for a generic normal form defect function

| Dimension | CPU time (s) | Max list | Total Boxes |
|---|---|---|---|
| 2 | 18,810 | | 4723 |
| 3 | $>562,896$ | | – |

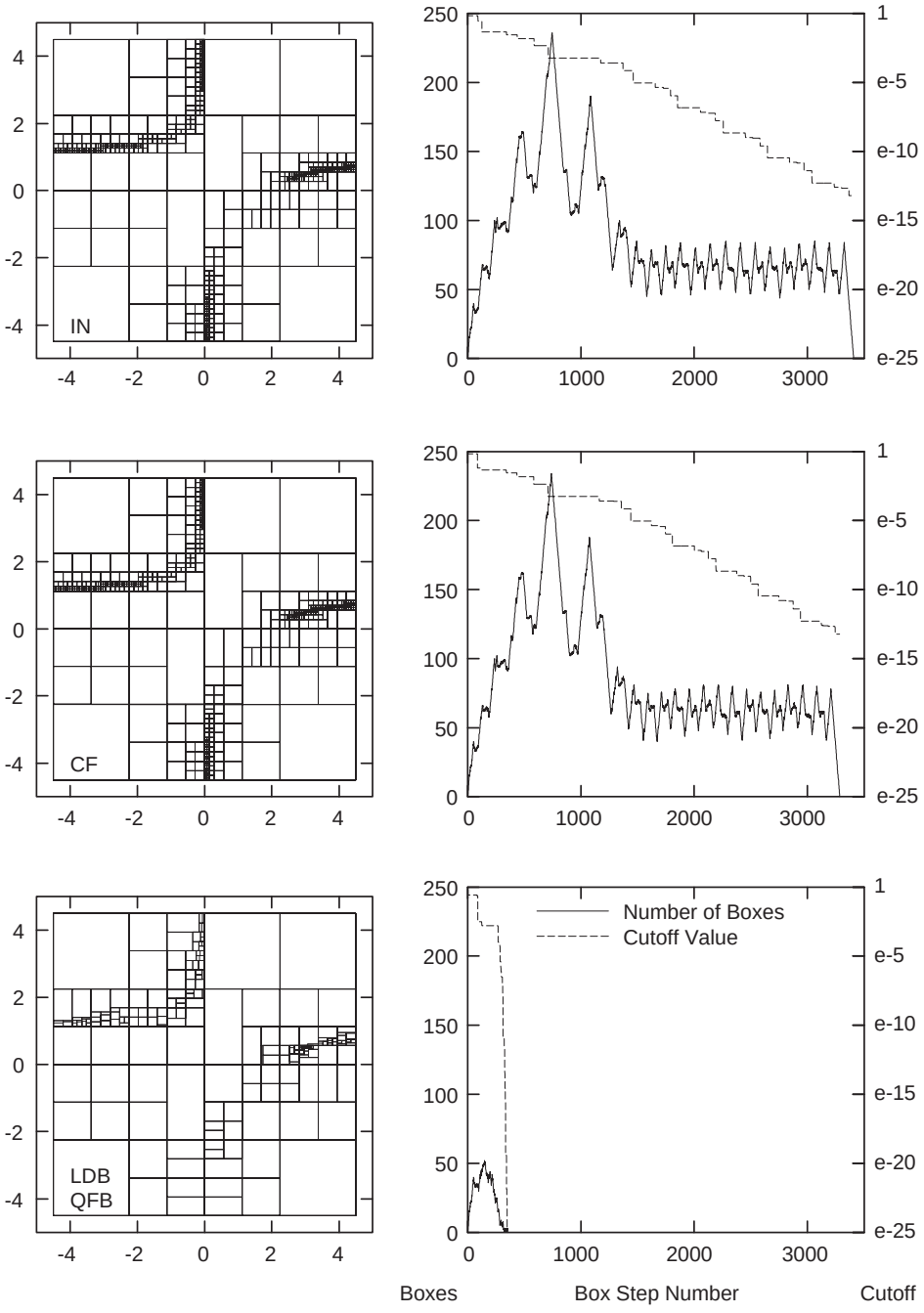



Fig. 8. Minimum search for the Beale function in $[-4.5, 4.5]^2$ by the interval, the centered form, and the LDB/QFB optimizers. Left: subdomain boxes studied. Right: number of active boxes and cutoff value as a function of step number.

Table 8
Global bounds obtained for six regions in normal form space for the Tevatron. Also computed are the guaranteed minimum transversal iterations

| Region | Boxes studied | CPU time (s) | Bound | Transversal iterations |
|---|---|---|---|---|
| $[0.2, 0.4] \times 10^{-4}$ | 82,930 | 30,603 | $0.859 \times 10^{-13}$ | $2.3283 \times 10^{8}$ |
| $[0.4, 0.6] \times 10^{-4}$ | 82,626 | 30,603 | $0.587 \times 10^{-12}$ | $3.4072 \times 10^{7}$ |
| $[0.6, 0.9] \times 10^{-4}$ | 64,131 | 14,441 | $0.616 \times 10^{-11}$ | $4.8701 \times 10^{6}$ |
| $[0.9, 1.2] \times 10^{-4}$ | 73,701 | 13,501 | $0.372 \times 10^{-10}$ | $8.0645 \times 10^{5}$ |
| $[1.2, 1.5] \times 10^{-4}$ | 106,929 | 24,304 | $0.144 \times 10^{-9}$ | $2.0833 \times 10^{5}$ |
| $[1.5, 1.8] \times 10^{-4}$ | 111,391 | 26,103 | $0.314 \times 10^{-9}$ | $0.95541 \times 10^{5}$ |

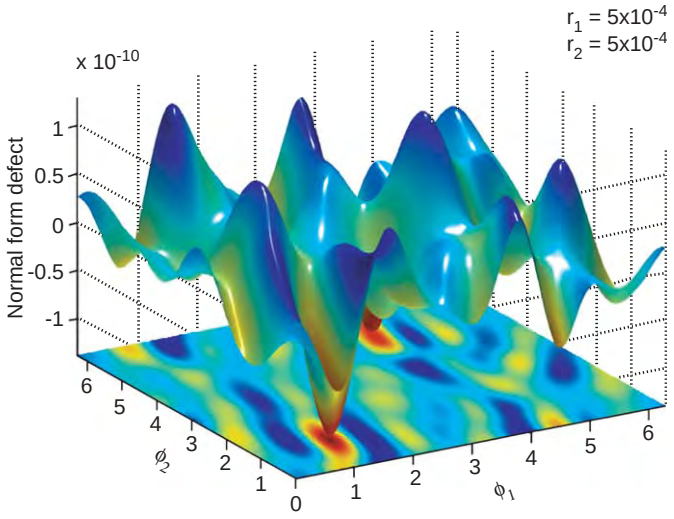



Fig. 9. Projection of the normal form defect function. Dependence on two angle variables for the fixed radii $r_1 = r_2 = 5 \times 10^{-4}$.

into a normal form radius of approximately $0.12 \times 10^{-4}$. We study the magnitude of the normal form defect bound for various different radius bands, and determine the ratio of width of band to normal form defect bound, which gives the minimum transversal iterations for the band. Fig. 9 shows the result of the calculations; it becomes apparent that it requires more than $2.7 \times 10^{8}$ revolutions for a particle starting with a normal form radius below $0.2 \times 10^{-4}$ to be lost.

## References

[1] M. Berz, Differential algebraic formulation of normal form theory, in: M. Berz, S. Martin, K. Ziegler (Eds.), Proceedings of Nonlinear Effects in Accelerators, IOP Publishing, London, 1992, p. 77.
[2] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999 Also available at ⟨http://bt.pa.msu.edu/pub⟩.
[3] M. Berz, Particle Accelerators 24 (1989) 109.
[4] M. Berz, K. Makino, New approaches for the validation of transfer maps using remainder-enhanced differential algebra, Nucl. Instr. and Meth. A 519 (2004) 53.
[5] M. Berz, K. Makino, Reliable Comput. 4 (4) (1998) 361.
[6] K. Makino, M. Berz, Suppression of the wrapping effect by Taylor model based validated integrators, submitted for publication, also MSUHEP-040910, available at ⟨http://bt.pa.msu.edu/pub⟩.
[7] R.B. Kearfott, Rigorous Global Search: Continuous Problems, Kluwer, Dordrecht, 1996.
[8] E. Hansen, Global Optimization using Interval Analysis, Marcel Dekker, New York, 1992.
[9] R.E. Moore, Interval Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1966.
[10] R.E. Moore, Methods and Applications of Interval Analysis, SIAM, Philadelphia, 1979.
[11] R.B. Kearfott, K. Du, J. Global Optim. 5 (1994) 253.
[12] K. Makino, M. Berz, Reliable Comput. 5 (1) (1999) 3.
[13] N. Revol, K. Makino, M. Berz, J Logic and Algebraic Programming 64 (1) (2004) 135 University of Lyon LIP Report RR 2003-11, MSU Department of Physics Report MSUHEP-030212, ⟨http://bt.pa.msu.edu/pub⟩.
[14] K. Makino, M. Berz, Int. J. Pure and Appl. Math. 6 (3) (2003) 239 available at ⟨http://bt.pa.msu.edu/pub⟩.
[15] K. Makino, Rigorous Analysis of Nonlinear Motion in Particle Accelerators. Ph.D. Thesis, Michigan State University, East Lansing, Michigan, USA, 1998, also MSUCL-1093.
[16] K. Makino, M. Berz, The LDB, QDB, and QFB bounders, Technical Report MSUHEP-40617, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, June 2004.
[17] K. Makino, M. Berz, Nucl. Instr. Meth., in print.
[18] M. Berz, J. Hoefkens, K. Makino, COSY INFINITY Version 8.1—programming manual, Technical Report MSUHEP-20703, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2002, see also ⟨http://cosy.pa.msu.edu⟩.
[19] M. Berz, K. Makino, COSY INFINITY Version 8.1—user's guide and reference manual, Technical Report MSUHEP-20704, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2002, see also ⟨http://cosy.pa.msu.edu⟩.
[20] More, Garbow, Hillstrom, ACM Trans. Math. Software 7 (1981) 17.
[21] P.V. Snopok, C.J. Johnstone, M. Berz, D.A. Ovsyannikov, A.D. Ovsyannikov, Nucl. Instr. and Meth., in print.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 11–19
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Beam dynamics optimization: Models, methods and applications

D.A. Ovsyannikov[a,*], A.D. Ovsyannikov[a], M.F. Vorogushin[b], Yu.A. Svistunov[b], A.P. Durkin[c]

[a]*SPbSU, St-Petersburg, Russian Federation*
[b]*NIIEFA, St-Petersburg, Russian Federation*
[c]*MRTI, Moscow, Russian Federation*

Available online 28 November 2005

**Abstract**

A special class of problems attracting attention of numerous researchers is represented by the tasks associated with the beam dynamics optimization in accelerators of charged particles. In this paper the problem of beam dynamics optimization is considered as a control theory problem. Problem statement is considered on the pattern of RFQ channel.


*PACS:* 29.27.Bd; 02.60.Pn; 02.30.Yy

*Keywords:* Beam Dynamics; RFQ structure; Numerical optimization; Mathematical modeling; Control theory

## 1. Introduction

Experts in different accelerator laboratories accumulated wide experience in RFQ designing [1–7]. Nevertheless new tasks on high-power, high-energy CW linac designing (for example, for ADT applications) appeared recently. Such accelerators must be practically free of beam losses. This main condition placed more stringent requirements upon beam quality at the input of each linac part. Because RFQ is used as beginning part in the above accelerators, beam high-quality requirements is related to RFQ primarily. It means that existing experience in RFQ designing must be enriched by new approaches and new methodology.

This paper suggests new mathematical models, which allow joint optimization of program motion and an ensemble of perturbed motions.

The suggested approach allows to construct optimization methods for selecting program motion with simultaneous optimization of various dynamical characteristics of ensemble of trajectories, which satisfy corresponding equations in deviations from program motion [8–12].

Considering the problems of beam physics we can understand program motion as the given motion. In particular, we can consider the motion of synchronous particle as the program motion. The variation of amplitude of accelerating wave and the variation of synchronous phase along the accelerator can be considered as control functions.

New approach to solving optimization problem for charged particles dynamics in accelerators includes: construction of mathematical model of controlled dynamical process; choice of control functions or parameters of optimization; construction of quality functionals, which allow efficient evaluation of various characteristics of examined controlled motions; analytical representation of the functionals variations, which allow to construct various methods of optimization for quality functionals; construction of methods and algorithms of optimization.

The suggested approach is applied to the optimization of charged particles dynamics in RFQ channel. Corresponding codes are developed.

Application of developed methods shows their effectiveness.

*Corresponding author. Tel.: +7 812 4287179; fax: +7 812 4287189.
*E-mail address:* dovs@compmath.spbu.ru (D.A. Ovsyannikov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.002

## 2. Problem statement

Let us consider following mathematical model of control described by a system of integro-differential equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(t,x,u), \tag{1}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = F_1(t,x,y,u) + \int_{M_{t,u}} F_2(t,x,y,z_t)\rho(t,z_t)\,\mathrm{d}z_t = F(t,x,y,u), \tag{2}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho(t,y)\,\mathrm{div}_y F(t,x,y,u) \tag{3}$$

with initial conditions

$$x(0) = x_0, \tag{4}$$

$$y(0) = y_0 \in M_0, \tag{5}$$

$$\rho(0,y(0)) = \rho_0(y_0), \quad y_0 \in M_0. \tag{6}$$

Here $t \in T_0 = [0,T]$ is independent variable, $x \in R^n$ and $y \in R^m$ are vectors of phase variables, $u = u(t)$ is an $r$-dimensional function, $T$ is fixed number, $\rho = \rho(t,y)$ is scalar function, $\rho_0(y_0)$ is some non-negative continuous function. Set

$$M_{t,u} = \{y_t \mid y_t = y(t,x(t),y_0,u), y_0 \in M_0, x(0) = x_0\}$$

is a cross-section at moment $t$ of beam of trajectories $y(t,x(t),y_0,u)$, $y_0 \in M_0$, under fixed control $u = u(t)$ and according to the program motion $x(t)$. In other words $M_{t,u}$ is a shift of set $M_0$ along trajectories of system (2). The set $M_0 \subset R^m$ is compact, non-zero measure. Vector function $f(t,x,u)$ of dimension $n$ is supposed to be continuous with its partial derivatives of the first order. Vector functions $F_1(t,x,y,u)$ and $F_2(t,x,y,z)$ of dimension $m$ are supposed to be continuous together with partial derivatives up to second order. We assume that allowed controls $u = u(t)$, $t \in T_0$, form some class $D$ of piecewise continuous function accepting values from a compact set $U \in R^r$. $R^n$, $R^m$, $R^r$ are $n$, $m$ and $r$-dimensional Euclidean spaces.

Let us agree that subsystem (1) describes dynamics of program (marked) motion, and subsystem (2) describes dynamics of motions disturbed along initial condition that is trajectories' beam. In particular, subsystem (2) can be considered as equations of deviations of program motions. Eq. (3) is equation of change of density distribution of particles $\rho = \rho(t) = \rho(t,y(t))$ along trajectories (2) with given initial function $\rho_0(y_0)$ of density distribution of particles in the set $M_0$.

Solution of system (1)–(3) under initial equations (4)–(6) and fixed control function $u = u(t)$ will be a trajectory $x(t) = x(t,x_0,u)$ and beam of trajectories $y(t,y_0) = y(t,x(t),y_0,u)$, $y_0 \in M_0$, together with density distributions of particles $\rho(t,y(t,y_0))$ on according trajectories $y(t,y_0)$, turning Eqs. (1)–(3) into identity. Let us note that solution of subsystem (1) can be considered independent of subsystem (2) and Eq. (3).

Introduced mathematical model of control accounts interaction of particles in beam. Here vector function $F_1$ determines the influence of external fields on particles, and vector function $F_2$, interaction of particles. Let us note that only function $F_1$ depends on control $u$.

It is important to note that mathematical model of charged particles dynamics described above represents a system of Vlasov equations. But function $F_2$, describing particles interaction here, is smooth enough and describes some smoothed interaction between a particle and an ensemble of all other particles. Thus we get smoothed models under various numerical solutions of Vlasov equation, for example by a way of macro particles. It is possible to show that solution of system (1)–(3) under given assumptions exists and it is unique. Moreover, vector function $y(t,y_0)$ is continuous and continuously differentiable along $y_0$ [13].

Let us introduce following functionals:

$$I_1(u) = \int_0^T \varphi_1(t,x(t),u(t))\,\mathrm{d}t + g_1(x(T)). \tag{7}$$

$$I_2(u) = \int_0^T \Phi(w_1(t))\,\mathrm{d}t + G(w_2), \tag{8}$$

where

$$w_1(t) = \int_{M_{t,u}} \varphi_2(t,x(t),y_t,\rho(t,y_t),u(t))\,\mathrm{d}y_t, \tag{9}$$

$$w_2 = \int_{M_{T,u}} g_2(y_T,\rho(T,y_T))\,\mathrm{d}y_T. \tag{10}$$

Here $\Phi$, $G$, $\varphi_1$, $\varphi_2$, $g_1$, $g_2$ are non-negative, continuously differentiable functions.

Functional (7) characterizes program motion dynamics, and functional (8) estimates behavior of beam trajectories.

Let us introduce following functional:

$$I(u) = I_1(u) + I_2(u), \tag{11}$$

simultaneously assessing dynamics of program motion and particle beam dynamics in account with their density distribution and their interactions.

## 3. Optimality conditions

Let us consider admissible controls $u(t)$ and $\tilde{u}(t)$. Define $\Delta u(t) = \tilde{u}(t) - u(t)$. Then variation of functional (11) can be represented in following way:

$$\delta I(u,\Delta u) = -\int_0^T (\chi^* \Delta_u f - \Delta_u \varphi_1)\,\mathrm{d}t - \int_0^T \int_{M_{t,u}} \begin{pmatrix} \mu^* \Delta_u F_1 + v \cdot \Delta_u\,\mathrm{div}_y F_1 \\ -\Phi'(w_1)\Delta_u \varphi_2 \end{pmatrix} \mathrm{d}y_t\,\mathrm{d}t. \tag{12}$$

Here symbol $\Delta_u$ means increment of function under change of only the variable $u$, for example, $\Delta_u f = f(t,x,u+\Delta u) - f(t,x,u)$). Auxiliary functions $\chi(t)$, $\mu(t,y_t)$, $v(t,y_t)$ satisfy along trajectories (1)–(3) the following linear integro-differential equations:

$$\begin{aligned}\frac{\mathrm{d}\chi}{\mathrm{d}t} = & -\frac{\partial f^*}{\partial x}\chi + \frac{\partial \varphi_1^*}{\partial x} \\ & - \int_{M_{t,u}} \left(\frac{\partial F^*}{\partial x}\mu + v\frac{\partial \operatorname{div}_y F}{\partial x}\right)^* \mathrm{d}y_t \\ & + \Phi'(w_1)\int_{M_{t,u}} \frac{\partial \varphi_2}{\partial x}\,\mathrm{d}y_t, \end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\mathrm{d}\mu}{\mathrm{d}t} = & -\left(\frac{\partial F}{\partial y} + E\cdot \operatorname{div}_y F\right)^*\mu \\ & - v\frac{\partial \operatorname{div}_y F^*}{\partial y} + \Phi'(w_1)\frac{\partial \varphi_2^*}{\partial y} \\ & - \rho(t,y_t)\int_{M_{t,u}} \begin{pmatrix} \dfrac{\partial F_2(t,x,z_t,y_t)^*}{\partial z}\mu(t,z_t) \\ +v(t,z_t)\dfrac{\partial \operatorname{div}_y F_2(t,x,z_t,y_t)^*}{\partial z}\end{pmatrix}\mathrm{d}z_t, \end{aligned} \tag{14}$$

$$\frac{\mathrm{d}v}{\mathrm{d}t} = -v\cdot \operatorname{div}_y F + \Phi'(w_1)\cdot\left(\varphi_2 - \rho\frac{\partial \varphi_2}{\partial \rho}\right), \tag{15}$$

and terminal conditions

$$\chi(T) = -\frac{\partial g_1(x(T))^*}{\partial x}, \tag{16}$$

$$\mu(T) = -G'(w_2)\cdot\frac{\partial g_2(y_T,\rho_T)^*}{\partial y}, \tag{17}$$

$$v(T) = -G'(w_2)\cdot\begin{pmatrix} g_2(y_T,\rho_T) \\ -\rho_T\dfrac{\partial g_2(y_T,\rho_T)}{\partial \rho}\end{pmatrix}. \tag{18}$$

Let us note that while getting variation (12) auxiliary functions $\chi$, $\mu$, $v$ also known in control theory as conjugate functions has been used. They play a role analogous of Lagrangian multipliers in calculus of variations.

Let us introduce functions

$$H_1(t,x,\chi,u) = \chi^* f(t,x,u) - \varphi_1(t,x,u), \tag{19}$$

and

$$\begin{aligned} & H_2(t,x,y,\rho,w_1,\mu,v,u) \\ & \quad = \mu^* F_1(t,x,y,u) + v\cdot \operatorname{div}_y F_1(t,x,y,u) \\ & \qquad - \Phi'(w_1)\cdot\varphi_2(t,x,y,\rho,u). \end{aligned} \tag{20}$$

Then variation of functional (12) can be rewritten in the following way:

$$\delta I(u,\Delta u) = -\int_0^T \Delta_u H_1\,\mathrm{d}t - \int_0^T\int_{M_{t,u}} \Delta_u H_2\,\mathrm{d}y_t\,\mathrm{d}t, \tag{21}$$

where $x$, $y$, $\rho$ are solutions of system (1)–(3) according to control $u = u(t)$; $w_1(t)$ is determined by formula (9); auxiliary functions $\chi$, $\mu$, $v$ satisfy Eqs. (13)–(15) accordingly with terminal conditions (16)–(18).

Variations (12), (21) are non-classical variations of functional (11) because in general case they are non-linear on $\Delta u$. Such variations are used in mathematical control theory when getting optimality conditions in a form in Pontryagin maximum principle. In the given case optimality conditions can be represented as

**Theorem 1.** *Let $u^{\mathrm{o}} = u^{\mathrm{o}}(t)$ is optimal control, i.e. it delivers a minimum to functional* (11). *Then under $t \in T_0 = [0,T]$ we have*

$$\max_{u\in U} H^{\mathrm{o}}(t,u) = H^{\mathrm{o}}(t,u^{\mathrm{o}}(t)). \tag{22}$$

Here

$$\begin{aligned} H^{\mathrm{o}}(t,u) = & H_1(t,x_t^{\mathrm{o}},\chi_t^{\mathrm{o}},u) \\ & + \int_{M_{t,u^{\mathrm{o}}}} H_2(t,x_t^{\mathrm{o}},y_t^{\mathrm{o}},\rho_t^{\mathrm{o}},w_1^{\mathrm{o}},\mu_t^{\mathrm{o}},v_t^{\mathrm{o}},u)\,\mathrm{d}y_t^{\mathrm{o}}, \end{aligned}$$

where functions $\chi_t^{\mathrm{o}} = \chi^{\mathrm{o}}(t)$, $\mu_t^{\mathrm{o}} = \mu^{\mathrm{o}}(t,y_t^{\mathrm{o}})$, $v_t^{\mathrm{o}} = v^{\mathrm{o}}(t,y_t^{\mathrm{o}})$, $w_1^{\mathrm{o}} = w_1^{\mathrm{o}}(t)$, $x_1^{\mathrm{o}} = x^{\mathrm{o}}(t)$, $y_1^{\mathrm{o}} = y^{\mathrm{o}}(t)$, $\rho_t^{\mathrm{o}} = \rho^{\mathrm{o}}(t,y_t^{\mathrm{o}})$ correspond optimal control $u^{\mathrm{o}} = u^{\mathrm{o}}(t)$.

From Eq. (21) comes representation of classical variation of functional (11)

$$\delta_{\mathrm{cl}} I(u,\Delta u) = -\int_0^T \left\{\frac{\partial H_1}{\partial u} + \int_{M_{t,u}} \frac{\partial H_2}{\partial u}\,\mathrm{d}y_t\right\}\Delta u(t)\,\mathrm{d}t. \tag{23}$$

**Definition.** We will call function $q(t)$, $t \in T_0$, *admissible direction* at point $u \in D$ along set $D$, if there exists such $\varepsilon_0 > 0$, that $(u(t) + \varepsilon\cdot q(t)) \in D$, when $\varepsilon \in [0,\varepsilon_0]$.

Let us consider any admissible direction $q(t)$ at point $u^{\mathrm{o}} = u^{\mathrm{o}}(t)$. Let $\Delta u_\varepsilon = \varepsilon\cdot q(t)$, where $\varepsilon \in [0,\varepsilon_0]$. Then for optimal control $u^{\mathrm{o}} = u^{\mathrm{o}}(t)$ we have

$$I(u^{\mathrm{o}} + \Delta u_\varepsilon) - I(u^{\mathrm{o}}) = \Delta I(u^{\mathrm{o}},\Delta u_\varepsilon) \geqslant 0.$$

With the use of classical variation (23) of functional (11) it is possible to formulate following optimality condition.

**Theorem 2.** *Let $u^{\mathrm{o}}(t)$ is optimal control. Then we have*

$$\delta_{\mathrm{cl}} I(u^{\mathrm{o}},q) \geqslant 0. \tag{24}$$

*under all admissible directions $q(t)$ at point $u^{\mathrm{o}} \in D$.*

Further we present optimality conditions that are based on the representation of the minimized functional via the solution of special partial differential equations.

Let us consider system (1)–(3) and suppose that $F_2 = 0$, i.e. we do not take into account interaction between charged particles in this case.

Consider functional (11), where $I_1(u)$ is defined by (7), and

$$I_2(u) = \int_0^T \int_{M_{t,u}} \varphi_2(t, x(t), y_t, \rho(t, y_t), u(t))\, \mathrm{d}y_t\, \mathrm{d}t + \int_{M_{T,u}} g_2(x(T), y_T, \rho(T, y_T))\, \mathrm{d}y_T.$$

Consider $v_1 = v_1(t, x)$, $v_2 = v_2(t, x, y, \rho)$, which satisfy special partial differential equations

$$\frac{\partial v_1}{\partial t} + \frac{\partial v_1}{\partial x} f(t, x, u(t)) + \varphi_1(t, x, u(t)) = 0, \tag{25}$$

$$\frac{\partial v_2}{\partial t} + \frac{\partial v_2}{\partial x} f(t, x, u(t)) + \frac{\partial v_2}{\partial y} F(t, x, y, u(t)) + \left(v_2 - \rho \frac{\partial v_2}{\partial \rho}\right) \mathrm{div}_y\, F(t, x, y, u(t)) + \varphi_2(t, x, y, \rho(t, y), u(t)) = 0 \tag{26}$$

with terminal conditions

$$v_1(T, x) = g_1(x), \tag{27}$$

$$v_2(T, x, y, \rho) = g_2(x, y, \rho). \tag{28}$$

Introduce the following functions:

$$w_1(t, x, u) = \frac{\partial v_1(t, x)}{\partial t} + \frac{\partial v_1(t, x)}{\partial x} f(t, x, u) + \varphi_1(t, x, u), \tag{29}$$

$$w_2(t, x, y, \rho, u) = \frac{\partial v_2(t, x, y, \rho)}{\partial t} + \frac{\partial v_2(t, x, y, \rho)}{\partial x} f(t, x, u) + \frac{\partial v_2(t, x, y, \rho)}{\partial y} F(t, x, y, u) + \left(v_2(t, x, y, \rho) - \rho \frac{\partial v_2(t, x, y, \rho)}{\partial \rho}\right) \mathrm{div}_y\, F(t, x, y, u) + \varphi_2(t, x, y, \rho, u) = 0. \tag{30}$$

Evidently

$$w_1(t, x, u(t)) \equiv 0, \quad w_2(t, x, y, \rho, u(t)) \equiv 0. \tag{31}$$

Hence we obtain new representation for functionals

$$I_1(u) = v_1(0, x_0), \tag{32}$$

$$I_2(u) = \int_{M_0} v_2(0, x_0, y_0, \rho_0(y_0))\, \mathrm{d}y_0. \tag{33}$$

Let

$$\tilde{u} = u + \Delta u, \quad \tilde{x}(t) = x(t, x_0, \tilde{u}), \quad \tilde{y}(t) = y(t, y_0, \tilde{u}).$$

Then we obtain

$$\Delta I(\Delta, u) = I(\tilde{u}) - I(u) = \int_0^T w_1(t, \tilde{x}(t), \tilde{u}(t))\, \mathrm{d}t + \int_0^T \int_{M_{t,\tilde{u}}} w_2(t, \tilde{x}(t), \tilde{y}(t), \tilde{\rho}(t, \tilde{y}(t)), \tilde{u}(t))\, \mathrm{d}\tilde{y}_t\, \mathrm{d}t. \tag{34}$$

Using the results of the works [13–16] the variation of functional (11) in this case can be presented as

$$\Delta I = \delta I + \mathrm{o}(\mu), \quad \delta I = \delta I_1 + \delta I_2. \tag{35}$$

where

$$\delta I_1 = \int_0^T w_1(t, x(t), \tilde{u}(t))\, \mathrm{d}t, \tag{36}$$

$$\delta I_2 = \int_0^T \int_{M_{t,u}} w_2(t, x(t), y_t, \rho(t, y_t), \tilde{u}(t))\, \mathrm{d}y_t\, \mathrm{d}t. \tag{37}$$

From expressions (36), (37), using needle-shaped variation, we obtain necessary conditions for optimality as follows.

**Theorem 3.** *Suppose that $u^{\mathrm{o}} = u^{\mathrm{o}}(t)$ is an optimal control, and the functions $v_1^{\mathrm{o}}(t, x), v_2^{\mathrm{o}}(t, x, y, \rho(t, y))$ are solutions of the Eqs.* (25), (26) *with terminal conditions* (27), (28) *under optimal control $u^{\mathrm{o}}$. Then there exists the relationship*

$$\min_{u \in U} \left\{ w_1^{\mathrm{o}}(t, x^{\mathrm{o}}(t), u) + \int_{M_{t,u^{\mathrm{o}}}} w_2^{\mathrm{o}}(t, x^{\mathrm{o}}(t), y_t, \rho^{\mathrm{o}}(t, y_t), u)\, \mathrm{d}y_t \right\} = w_1^{\mathrm{o}}(t, x^{\mathrm{o}}(t), u^{\mathrm{o}}(t)) + \int_{M_{t,u^{\mathrm{o}}}} w_2^{\mathrm{o}}(t, x^{\mathrm{o}}(t), y_t, \rho^{\mathrm{o}}(t, y_t), u^{\mathrm{o}}(t))\, \mathrm{d}y_t = 0. \tag{38}$$

Classical variation of functional (11) based on the above approach we can get from formula (35).

## 4. Modeling of longitudinal motion

Let us consider beam longitudinal motion in RFQ channel. It declares by following Eq. [3]

$$\frac{\mathrm{d}^2 z}{\mathrm{d}\tau^2} = \frac{eUk\sigma}{2W_0} I_0(kr) \cos(kz) \cos(\omega\tau + \varphi_s), \tag{39}$$

where $\tau = ct$, $U$ is intervane voltage, $k = 2\pi/L$, $L$ is modulation period, $e$ and $W_0$ are particle charge and rest energy, respectively $\omega = 2\pi c/\lambda$, $\lambda$ is RF wave length, $r$ is distance between particle and channel axis, $\sigma = m^2 - 1/(m^2 I_0(ka) + I_0(mka))$. The non-relativistic approximation $\gamma \cong 1$ is used.

Let us resolve RF field into two traveling harmonics and average over harmonics $\omega\tau + kz$. Then we obtain

$$\frac{d^2z}{d\tau^2} = \frac{eUk\sigma}{4W_0}\cos(\omega\tau - kz + \varphi_s).$$

For synchronous particle $\omega\tau + kz$. Let us introduce variable $\psi = \omega\tau - kz = k(z - z_s)$. Taking $L = \beta_s\lambda$ and using equations for $\beta_s$ and $z - z_s$, we obtain

$$\psi''_\tau + \frac{2L'_\tau}{L}\psi'_\tau + \frac{L''_\tau}{L}\psi - \Omega^2(\cos\varphi_s - \cos(\varphi_s + \psi)) - \frac{ek^2}{W_0}\frac{\partial U_c}{\partial\psi} = 0.$$

where $\Omega^2 = 4\pi eUT/W_0L^2$, $T = (\pi/4)\sigma$, $U_c$ is Coulomb field potential. Using relations

$$\frac{\dot{L}}{L} = \frac{\Omega^2}{\omega}\cos\varphi_s,$$

$$\frac{L''_\tau}{L} = \frac{\Omega^2}{\omega}\frac{(UT)'_\tau}{(UT)}\cos\varphi_s + \frac{\Omega^2}{\omega}(\cos\varphi_s)'_\tau - \left(\frac{\Omega^2}{\omega}\cos\varphi_s\right)^2$$

and taking $\Omega^2 = \Omega_0^2\eta(W_n/W)$, where $W_n$ and $W$ are input and current energies of particles, $\Omega_0^2 = 2\pi e(UT)_{max}/W_n\lambda^2$, $\eta = UT/(UT)_{max}$. If we convert to independent variable $\zeta = \Omega_0\tau$, we finally obtain

$$\psi'' + 2\frac{\Omega_0}{\omega}\eta\frac{W_n}{W}\cos\varphi_s\psi' + \left(\frac{\Omega_0}{\omega}\frac{W_n}{W}\cos\varphi_s\eta' + \frac{\Omega_0}{\omega}\eta\frac{W_n}{W}(\cos\varphi_s)' - \left(\frac{\Omega_0}{\omega}\eta\frac{W_n}{W}\cos\varphi_s\right)^2\right)\psi - \eta\frac{W_n}{W}(\cos\varphi_s - \cos(\psi + \varphi_s)) - \frac{ek^2}{W_0\Omega_0^2}\frac{\partial U_c}{\partial\psi} = 0. \quad (40)$$

Eq. (40) together with synchronous particle equation

$$\left(\frac{W}{W_n}\right)' = 2\frac{\Omega_0}{\omega}\eta\cos\varphi_s \quad (41)$$

determines longitudinal motion in RFQ channel in equivalent traveling wave approximation.

In traveling wave approximation we see that longitudinal motion is determined only by one given parameter $\kappa = \Omega_0/\omega$ and by laws of $\eta$ and $\varphi_s$ variations.

We propose to calculate longitudinal component of Coulomb field as ones for cylinder with constant radius and with uniform distribution in each transverse cross-section. In longitudinal direction periodical modulation take place along axis $Z$. In this case each macroparticle can be presented as uniform charge disc. One hundred discs are enough for reliable estimates.

If $\tilde{U}$ is Coulomb field potential with charge unit in beam, then Coulomb term in Eq. (40) has a form

$$\alpha\left(\frac{W_n}{W}\right)^{3/2}\frac{\partial\tilde{U}}{\partial\psi},$$

where

$$\alpha = \frac{I}{\varepsilon_0 c\beta_n(UT)_{max}}$$

## 5. Optimization criteria

Further we will consider Eqs. (40), (41).

Let

$$0 \leqslant \zeta \leqslant \bar{T}, \quad x = W/W_n, \quad y = (y_1, y_2)^* = (\psi, \psi')^*,$$

$$M_0 = \{y \mid -\pi \leqslant y_1 \leqslant \pi, \theta_1 \leqslant y_2 \leqslant \theta_2\}.$$

As the aim of RFQ structure optimization we consider following: insurance of maximal capture of particles under the acceleration conditions; obtaining required or maximal possible output energy.

If optimization is made only for these criteria, then final results can be obtained with high defocusing parameters and very small transverse acceptance. It means that restriction on defocusing parameter must be added

$$A_{def} = \frac{2k^2\eta|\sin\varphi_s|}{(L/L_0)^2} = \frac{2k^2u_1|\sin u_2|}{x} \leqslant A, \quad (42)$$

Parameter A usually lies in the range (A = )0.01–0.015.

If high-current beam is accelerated, then this beam must not be pinched in longitudinal direction in optimal mode. In this case new limitation demands monotonic variation of rms beam width$\langle\Delta\varphi\rangle^2$. In ideal this limitation looks as

$$\frac{d\langle\Delta\varphi\rangle^2}{d\zeta} \leqslant 0. \quad (43)$$

For the beam dynamics analyzing we introduce functionals

$$I_1(u) = c_1\int_0^{\bar{T}}\varphi_1(A_{def})\,d\zeta + c_2(x(\bar{T}) - \bar{x})^2, \quad (44)$$

$$I_2(u) = c_3\int_0^{\bar{T}}\Phi(w_1(\zeta))\,d\zeta, \quad (45)$$

$$w_1(\zeta) = \frac{d}{d\zeta}\int_{M_{\zeta,u}}(\psi^2\cdot\rho)\,dy_\zeta = \int_{M_{\zeta,u}}2\psi\psi'\rho(\zeta,\psi,\psi')\,dy_\zeta = \int_{M_{\zeta,u}}2y_1y_2\rho\,dy_\zeta,$$

$$I(u) = I_1(u) + I_2(u). \quad (46)$$

Here $c_1$, $c_2$, $c_3$—non-negative weight constants; $\varphi_1$ and $\Phi$ are penal functions, which can be introduced in

following way:

$$\varphi_1(A_{\text{def}}) = \begin{cases} 0 & A_{\text{def}} \leqslant \bar{A}, \\ (A_{\text{def}} - \bar{A})^{2p} & A_{\text{def}} > \bar{A}, \end{cases}$$

$$\Phi(w_1) = \begin{cases} 0 & w_1 \leqslant 0, \\ w_1^{2q} & w_1 > 0, \end{cases}$$

where $p$, $q$ are certain positive integer constants.

$A_{\text{def}}$ is defocusing factor; $\bar{x}$ is fixed value, which determines energy of synchronous particle at the output of accelerator; $\rho = \rho(\zeta, \psi, \psi')$ is density distribution of particles.

Functional (44) integrally estimates the deviation of defocusing factor from given value and accounts the deviation of synchronous particle velocity from given one at the output of accelerator. Functional (45) characterizes velocity of variation of mean square phase distribution. In this connection, condition $w_1 \leqslant 0$ ensures the monotonicity of grouping and simultaneously ensures the capture of particles into acceleration mode.

In order to finish definition of optimization task for longitudinal motion it is needed to parameterize functions $\eta(\zeta)$ and $\varphi_s(\zeta)$. As far as there are no oscillations with frequency $\omega$ in external force it is enough to define several plot points and intermediate point values determined by interpolation (linear in the simplest cases). In practice, 20 points is enough.

So longitudinal motion optimization can be reduced to search values of functions $\eta(\zeta)$ and $\varphi_s(\zeta)$ in number of extraction points.

On the basis of variation (12) for the functional (46) numerical methods of optimization can be constructed, which allow joint minimization of functionals (44), (45). The realization of these methods shows their effectiveness.

## 6. Modeling and optimization of transverse motion

Let us consider transverse motion optimization. Further we may use different approaches for solving the problems of transverse motion.

Let us suppose that longitudinal motion optimization by above scheme was successful. As a result we obtain functions $\eta(\zeta)$ and $\varphi_s(\zeta)$. Using relations between variable $\zeta$ and cell number $n = 2\zeta/\kappa$, we obtain for each cell value $U(n)T(n)$, $\varphi_s(n)$, $A_{\text{def}}(n)$ and factor of density increasing by beam bunching $\theta(n)$ connected with rms values of beam width. Now we are needed to find optimal values of intervane voltage $U(n)$ and aperture $a(n)$ or mean radius $R_0(n)$.

It is known that focusing period in RFQ is defined by three parameters: focusing parameter $B$ (proportional to $U/R_0^2$), defocusing parameter $A_{\text{def}}$ and Coulomb parameter proportional to $\theta I/\varepsilon$. Here $R_0$ is value near the mean radius, $I$ is beam current, $\varepsilon$ is beam emittance. For ideal form of vane shape—$R_0^2 = a^2/(1 - \sigma I_0(ka))$. Each parameter triplet (if it is placed inside stability zone) corresponds to matched radius $r$ maximal over period. Relation $E = (R_0^2/r^2)\varepsilon$ defines transverse channel acceptance. We know values for defocusing and Coulomb factors obtained by longitudinal motion calculation. Because RF field intensity is limited by some value then value $U/R_0$ is also limited by $E_b$. Using maximal value we obtain that focusing factor is proportional to $E_b/R_0$. Both values $r$ and $R_0$ are decreased with focusing parameter growth. Value $\tilde{R}_0(n)$ corresponds to maximal acceptance and is the required optimal value.

By next step we find $U(n) = E_b\tilde{R}_0(n)$, $\tilde{T}(n) = (UT)(n)/U(n)$, and find $m(n)$ and $a(n)$ on the basic of $T$ and $R_0$.

Transverse motion calculations were not needed if we use proposed method. Instead, it is enough to find periodical solution for channel envelope in each cell.

Another approach is to use different equations describing transverse motion and developed corresponding optimization methods for finding optimal parameters. These methods are given an account, for example, in the works [2,10–18]. The case is considered when particles were uniformly distributed along the beam cross-section , this allow us to use self-consistent distributions of Kapchinsky–Vladimirov and to design according mathematical models of optimization. Minimal radius $a$ and intervane voltage $U_L$ are used as controls while already found longitudinal motion controls are preserved.

Conducted calculations have shown the effectiveness of step-by-step optimization on various stages of the process of the optimization. Thus the minimization of the defocusing factor on the stage of the optimization of longitudinal motion allows focusing of the beam with the conservation of chosen dynamics of longitudinal motion. Gradual consideration of following tasks seemed reasonable:

(1) maximization of the acceptance of the transverse motion under fixed motion of synchronous particle,
(2) maximization of the acceptance of the channel under various phases of particles,
(3) optimization of beams matching or its other characteristics by the introduction of special functionals.

## 7. Numerical simulation

The code tools described above were used for RFQ designing. Main results are presented below to demonstrate power new code BDO-RFQ.

The following parameters were preset:

| | |
|---|---|
| Accelerating particles | Protons |
| Input energy | 0.095 MeV |
| Output energy | 5 MeV |
| Beam current | 100 mA |
| Beam emittance | $0.15\pi$ cm mrad |
| Operating frequency | 352 MHz |

There were some restrictions and requirements:

| | |
|---|---|
| Length | no more than 8 m |
| Beam transmission | no less then 98% |
| Defocusing factor | no more than 0.01 |

At Figs. 1–6 plots of main parameters for four RFQ structures [4,5,9] are shown. The set of structures includes I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (versions of structures found with Russian BDO-RFQ code).

The aim of optimization in variants III and IV was to get accelerator structure with output energy 5 MeV with minimal length. With this in variant III more strict limits on change of defocusing factor had been taken into account.

Further, both variants III and IV had been calculated to output energy 6.7 MeV. During analyzing and optimization of transverse motion it was supposed that intervane voltage is equal to 100 kV.

## 8. Conclusions

The mathematical methods of many-parametric optimization (as applied to charge particle beam dynamics) were developed [2,13–15,17–19]. In principle, they enable one to solve the problem of RFQ optimization.

The possible criteria of optimization for RFQ channel designing can be: minimal length of RFQ channels and beam transmission no less than the given one. Vane modulation, intervane voltage, synchronous phase and bore aperture in each cell can be used as control parameters. But many difficulties arise if we make an attempt to realize optimization procedure in the form described above. Time needed to solve the procedure will be large because control parameter number are more than thousand. It must be added that time for one RFQ version simulation is measured by minutes, and often modern computer techniques have no power to make such optimization.

It means that new mathematical methods of optimization and simplified models for RFQ beam acceleration and focusing are needed. These models must, from one side, hardly reduce control parameter number and, from another side, give equivalent description of beam motion. Above we propose a model that allows to separate longitudinal motion optimization from transverse motion one. Moreover, we suggest new mathematical model of optimization for this case.

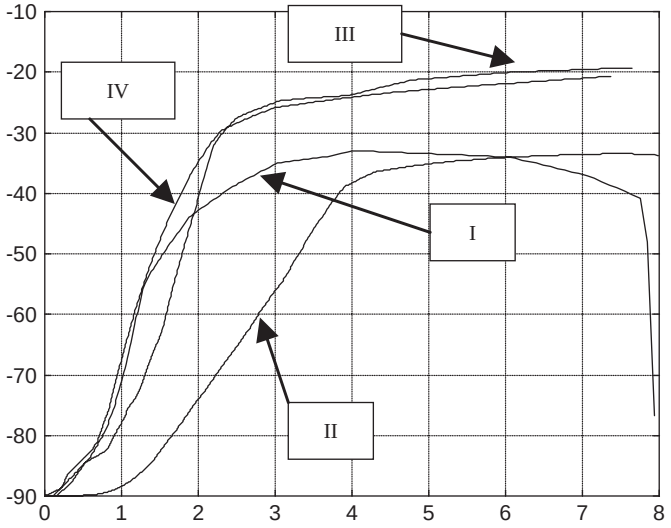



Fig. 2. Variations of synchronous phase. I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (Russia). These variations are given along accelerator in meters (axis-Z).

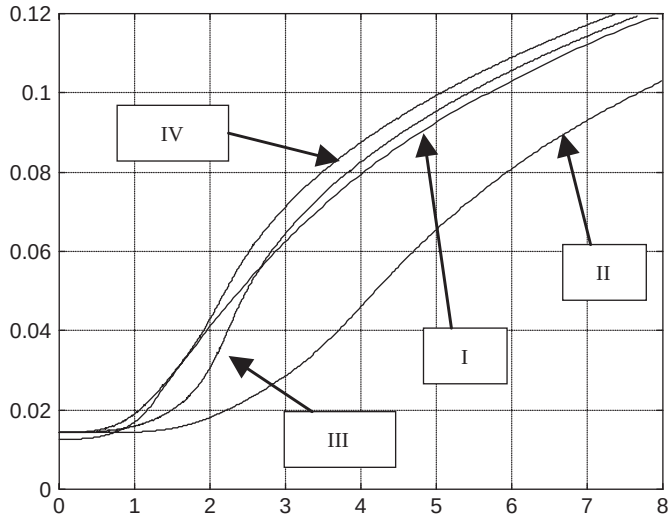



Fig. 1. Variation of reduced velocity of the synchronous particle. I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (Russia). These variations are given along accelerator in meters (axis-Z).

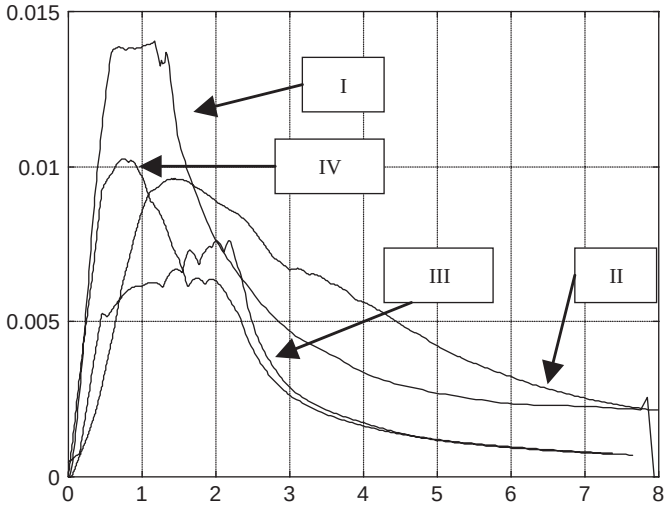



Fig. 3. Variations of defocusing factor. I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (Russia). These variations are given along accelerator in meters (axis-Z).

While designing such complicated structures as accelerators, a standard approach is usually applied when at the first stage a programmed motion is investigated, and at the second one perturbed motions are calculated. Unfortunately, it is not always possible to obtain desirable results using this approach. In this connection there is a requirement to consider new mathematical models and develop optimization methods, which allows us to realize a simultaneous optimization of the programmed motion and ensemble of perturbed motion [8–12,16].

So proposed method gives a possibility to use mathematical methods of multi-parametric optimization. Optimization of RFQ channel was based on simple model but main components of external forces and linear component of Coulomb forces are taken into account. Of course it cannot be advocated that optimal solution will be kept when other factors will be taken into account. Nevertheless it can be assumed that it will be near to optimal. Then obtained optimal solution can be used as initial approximation for procedure of further optimization or optimization in interactive mode.

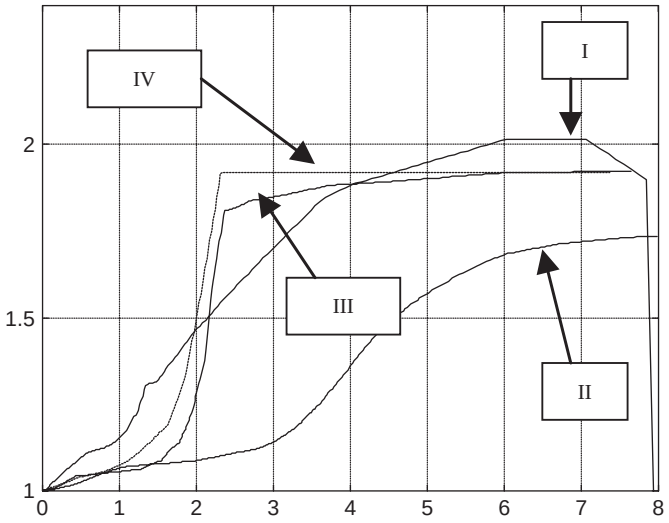



Fig. 4. Variations of vane modulation. I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (Russia). These variations are given along accelerator in meters (axis-Z).

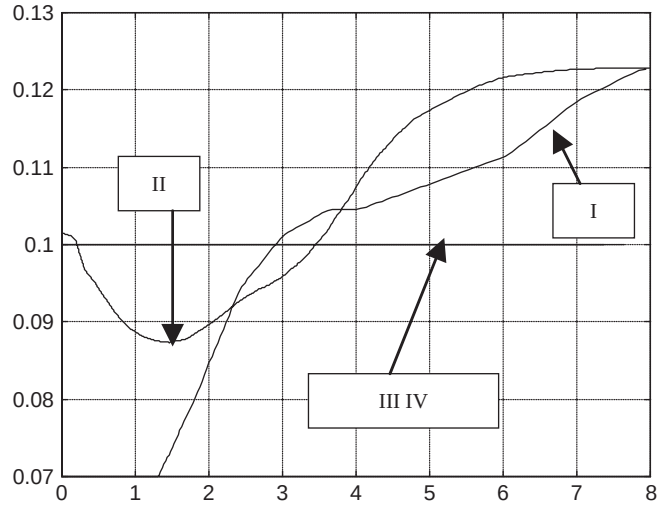



Fig. 5. Variations of intervane voltage. I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (Russia). These variations are given along accelerator in meters (axis-Z). Variations of all characteristics are given along accelerator in meters (axis-Z).

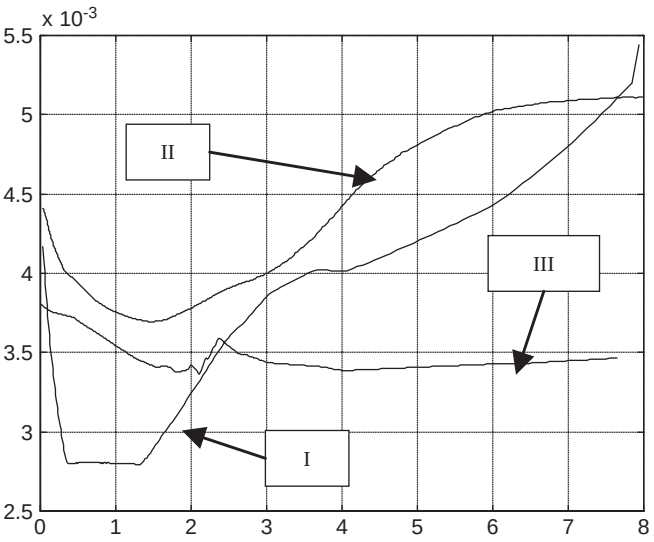



Fig. 6. Variations of average radius. I—LEDA-RFQ (USA), II—IPHI-RFQ (France), III—BDO-RFQ-1 and IV—BDO-RFQ-2 (Russia). These variations are given along accelerator in meters (axis-Z).

The above considered optimization problem can be used in optimization of charged particle dynamics in other types of accelerators, for example in waveguide or with drift tubes linear accelerators.

Representations (21), (23) and (36), (37) of variation of functional (11) and optimality conditions allow to construct various directed methods of minimization of considered functional.

## Acknowledgements

The work is supported in part by the Russian Foundation for Basic Research, grant No 03-01-00726.

## References

[1] K.R. Crandall, et al., RF quadrupole beam dynamics design studies, in: Proceedings of the Linear Accelerator Conference, Montauk, NY, 1979, BNL 51134, Upton NY, 1980, p. 205.
[2] O.I. Drivotin, A.E. Loukianova, D.A. Ovsyannikov, Yu.N. Gavrish, Yu.A. Svistunov, M.F. Vorogushin, The choice of acceleration structure for PET-System, in: Proceedings of European Particle Accelerator Conference, Barcelona, Spain, vol. 1, 1996, pp. 783–785.
[3] I.M. Kapchinsky, V.A. Teplyakov, Prib. Tekh. Eksp. N2(1970).
[4] W.P. Lysenko, et al., High energy beam transport beam line for LEDA, Proceedings of the LINAC98, Chicago, 1998, pp. 496–498.
[5] R. Ferdinand, J.-M. Lagniel, P. Mattei, S. Nath, Optimization of RFQ design, in: Proceeding of EPAC-98 Conference, June 22–26 Stockholm, Sweden, pp. 1106–1108.

[6] D. Schrage, et al., CW RFQ fabrication and engineering, in: Proceedings of the LINAC98, Chicago, 1998, pp. 679–683.
[7] R.H. Stokes, et al., The radio-frequency quadrupole: general properties and specific applications, in: Proceedings of the 11th International Conference on High-Energy Accelerators, Geneva, 1980, Birkhauser, Basel, 1980, p. 399.
[8] B.I. Bondarev, A.P. Durkin, A.D. Ovsyannikov, New mathematical optimization models for RFQ structures, in: Proceedings of the 1999 Particle Accelerator Conference, New York, 1999, pp. 2808–2810.
[9] B.I. Bondarev, A.P. Durkin, S. Vinogradov, D.A. Ovsyannikov, A.D. Ovsyannikov, RFQ optimization methods and codes, in: Proceedings of the Sixth International Computational Accelerator Physics Conference, 11–14 September 2000, Darmstadt, Germany.
[10] A.D. Ovsyannikov, Mathematical modeling and optimization dynamics of charged particles and plasma, Ph.D. dissertation, Saint-Petersburg State University, Saint-Petersburg, 1999, p. 118.
[11] A.D. Ovsyannikov, New approach to beam dynamics optimization problem, in: Proceedings of the Sixth International Computational Accelerator Physics Conference, 11–14 September 2000, Darmstadt, Germany.
[12] A.D. Ovsyannikov, Simultaneous Optimization of Programmed and Disturbed Motions, Saint-Petersburg State University, Saint-Petersburg, 2002, p. 53.
[13] D.A. Ovsyannikov, Modeling and Optimization of Charged Particle Beam Dynamics, Leningrad State University, Leningrad, 1990, p. 312.
[14] D.A. Ovsyannikov, Mathematical Methods of Beam Control, Leningrad State University, Leningrad, 1980, p. 228.
[15] D.A. Ovsyannikov, Modeling and optimization problems of charged particle beams dynamics, in: Proceedings of the Fourth European Control Conference, Brussels, 1997, pp. 390–394.
[16] D.A. Ovsyannikov, Mathematical modeling and optimization of beam dynamics, in: Proceedings of the IFAC Nonlinear Control Systems, St.-Petersburg, Russia, 2001, vol. 1, pp. 13–22.
[17] D.A. Ovsyannikov, Mathematical methods of optimization of charged particle beams dynamics, in: Proceedings of European Particle Accelerator Conference, Barcelona, Spain, vol. 2, 1996, pp. 1382–1384.
[18] D.A. Ovsyannikov, N.V. Egorov, Mathematical Modeling of Electron and Ion Beams Forming Systems, Saint-Petersburg State University, Saint-Petersburg, 1998, p. 276.
[19] C.K. Allen, Image effects and optimal controls of space-charge dominated particle beams, Ph.D. dissertation, University of Maryland, 1996.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 20–24
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Computational needs for muon accelerators ☆

J. Scott Berg

*Brookhaven National Laboratory, Building 901A, P.O. Box 5000, Upton, NY 11973-5000, USA*

Available online 29 November 2005

**Abstract**

Muon accelerators contain beam lines and components which are unlike any found in existing accelerators. Production of the muons requires targets for beams with powers which are at or beyond what has currently been achieved. Many subsystems use solenoid focusing systems where at any given point, several magnets have a significant influence. The beams that are transported can have energy spreads of ±30% or more. The required emittances necessitate accurate tracking of particles with angles of tenth of a radian and which are positioned almost at the edge of the beam pipe. Tracking must be done not only in vacuum, but also in materials; therefore, statistical fluctuations must also be included.

Design and simulation of muon accelerators requires software which can: accurately simulate the dynamics of solid and liquid targets under proton bombardment; predict the production of particles from these targets; accurately compute magnetic fields based on either a real magnet design or a model which includes end fields; and accurately design and simulate a beam line where the transported beam satisfies the above specifications and the beam line contains non-standard, overlapping elements. The requirements for computational tools will be discussed, the capabilities of existing tools will be described and compared to what is required.


*PACS:* 29.27.Bd; 29.25.−t

*Keywords:* Muon accelerators; Beam dynamics; Simulation; Target

## 1. Introduction

In recent years, there has been an increased interest in machines which transport and accelerate muons instead of the more traditional electrons or protons [1], especially in light of the recent experimental results in neutrino physics [2–8].

Unlike the more commonly used electrons and protons, muons decay in a relatively short amount of time, and this places strong requirements on any system that is transporting or accelerating them. Furthermore, the method of muon production, bombarding a target with protons and capturing the muons that result from the decay of the produced pions, leads to the muons being produced in a very large phase space volume. As a result, the muon beams will almost completely fill the beam pipe, while at the same time particles in the beam will make angles with respect to the reference orbit of as much as 0.3 rad (especially in the ionization cooling section). There may be an energy range of a factor of two within a given beam. Furthermore, for acceleration, there has recently been interest in using FFAGs [9]. These are circular machines with a beamline that accepts a factor of two or more in energy. These are conditions which are not often encountered in traditional particle accelerators, and therefore must be handled very carefully.

The large number of muons that must be produced requires that a large number of protons hit the target. For several reasons, this generally should be done in a small number of high-energy pulses, rather than a large number of low-energy pulses or even continuously. The RF power systems for acceleration and cooling require too much power and/or cooling to be run continuously, so they are generally run in a pulsed mode; the average power used in these systems will be proportional to their repetition rate. For collider applications, fewer pulses will lead to a higher

☆ Work supported by US Department of Energy contract DE-AC02-98CH1088.

*E-mail address:* jsberg@bnl.gov.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.054

luminosity. For neutrino applications where the detector is not buried deep underground, background elimination will be more effective with a lower pulse rate. The high average power and pulse energy on the target will often damage or destroy the target; modeling and predicting this is important to muon accelerator design.

Finally, the requirements of rapid acceleration and cooling as well as large longitudinal phase space acceptance require high RF gradients. The gradient which can be achieved in cavities is limited by RF breakdown. To complicate matters, these cavities are in a high magnetic field, which has a strong effect on the breakdown [10]. This phenomenon is not well-understood at this point, making predictive computations difficult. Hence, this topic will not be addressed here in detail. Understanding and predicting that process is important for the design of cavities for muon cooling and for the prediction of achievable RF gradients.

## 2. Beam dynamics

A muon beam transport line must deal with beams having an especially large phase space area. In addition, many beamlines must accommodate a beam whose central energy and position will vary over time as the beam is accelerated (specifically, FFAGs). Many accelerator design and tracking codes, often without stating so, implicitly assume that at least some dimensions of the beam's phase space are small. Any code which is used in analyzing muon accelerators must be cautious to include all dynamical effects.

### 2.1. Magnetic fields

To transport a large-emittance beam, it is necessary to have a highly compact lattice. Fig. 1 shows an example of the lattice for a cooling cell. There are solenoids around the cavities and at the ends of the cavities. The longitudinal field as a function of position in the cell is shown in Fig. 2.

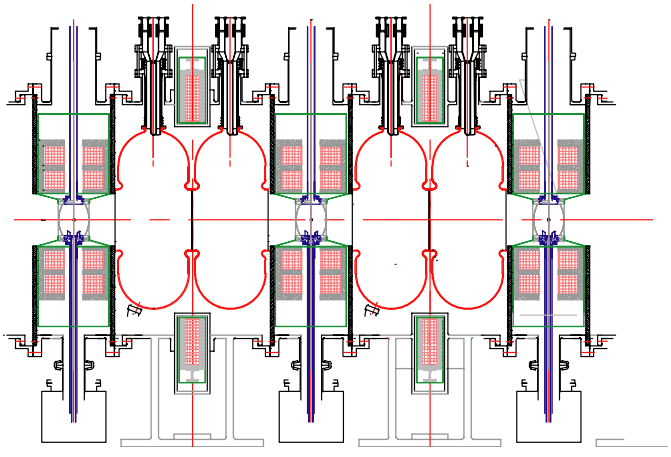


Fig. 1. Example of two cells of a cooling lattice.

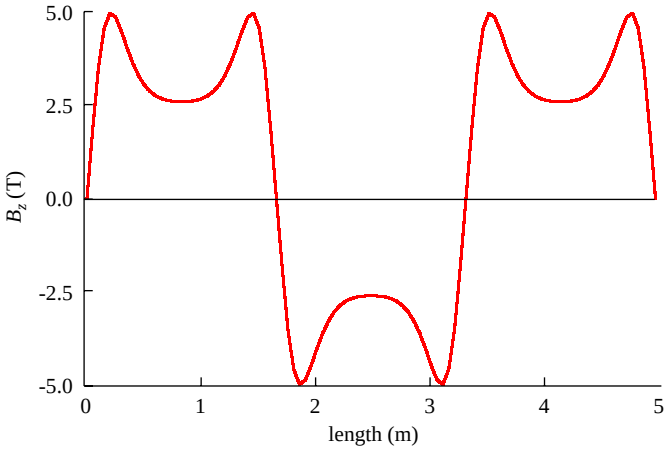



Fig. 2. Longitudinal field as a function of position in a cooling cell.

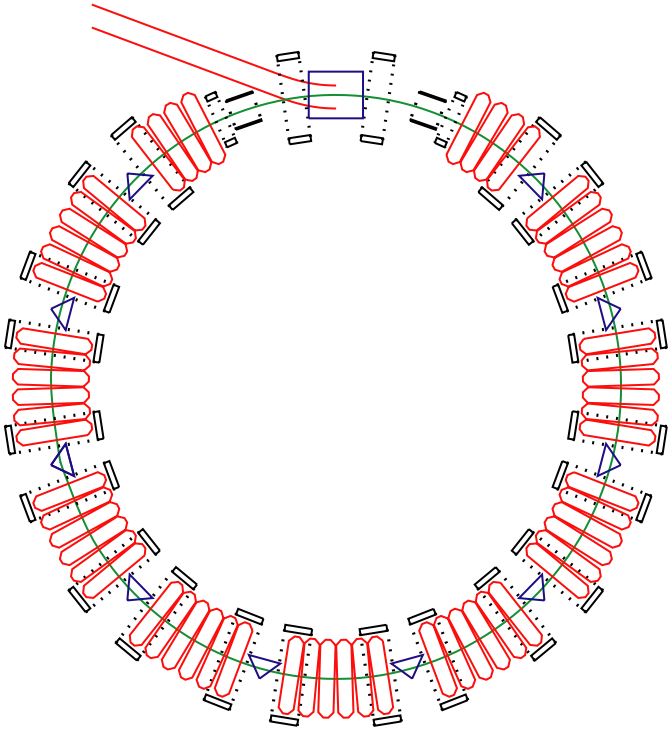


Fig. 3. Layout of a cooling ring.

It can be seen from this field that the solenoids are not well-approximated by constant-field separated magnets. In fact, the variation of the field with position leads to significant nonlinearities, the correct modeling of which is critical for understanding the performance of the lattice.

For another example, consider the RFOFO cooling ring shown in Fig. 3. This ring consists exclusively of solenoids, but the solenoids are tilted so as to produce a bending field. The horizontal and vertical fields on the circle shown in Fig. 3 are shown in Fig. 4. As discussed before, this field cannot be computed by looking at one magnet at a time,

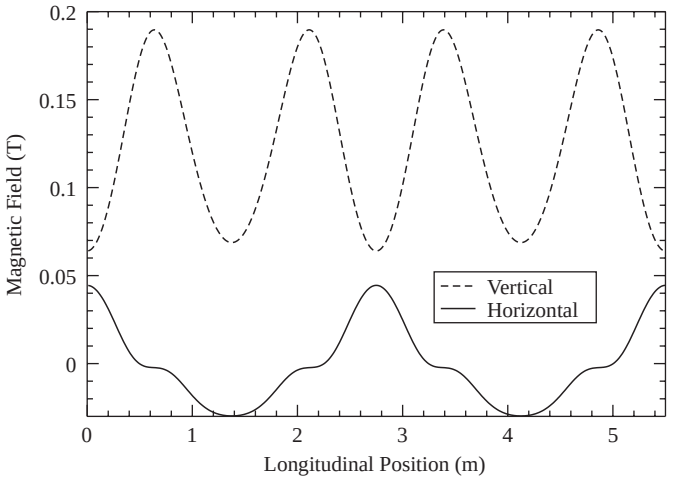



Fig. 4. Horizontal and vertical fields in a cooling ring.

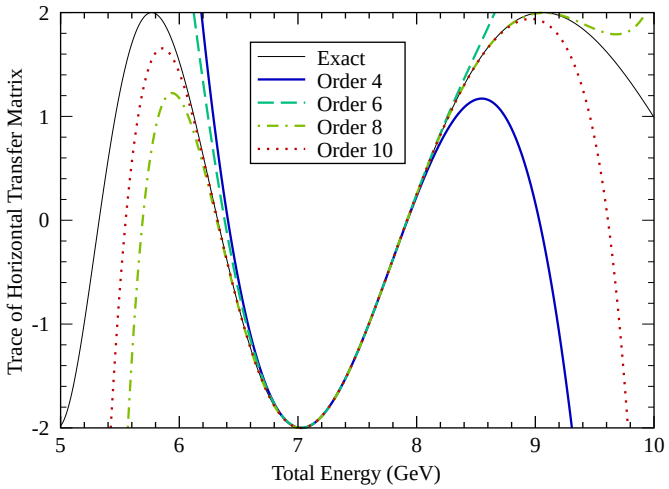



Fig. 5. Trace of the horizontal transfer matrix for 10 cells of an FFAG lattice computed using a truncated power series [11].

and the longitudinal variation of the field will produce significant nonlinearities which must be accurately computed. But this further illustrates a problem of coordinate system representation. The optimal way to represent the particle positions for this ring is most likely by deviations from the circle shown in Fig. 3. However, many codes would instead use a coordinate system based on a path that curves as a particle at a "reference energy" moving in the magnetic would. As can be seen from Fig. 4, this would lead to a trajectory that has non-uniform curvature, and bends in the vertical plane as well. Ensuring that the resulting reference curve bends by the appropriate amount per cell and does not have a vertical displacement from beginning to end would be very challenging, and would in fact require one to start out working in some coordinate system that is independent of the magnetic field in the first place. Thus, a code for complex systems such as this must be capable of working in a coordinate system that is not determined by the magnetic fields. Since in this case, no particle actually follows the reference curve (but should stay near it), one must be sure to handle RF synchronization properly, presumably by making a preliminary pass to find a closed orbit.

Many nonlinear effects result not from intentionally introduced nonlinearities, but from nonlinearities that are the result (due to Maxwell's equations) of the longitudinal variation of linear fields (Figs. 2 and 4). Any tracking or analysis code must be able to describe the fields that a particle sees in a way that is consistent (at least to some level of accuracy) with Maxwell's equations. While this can often be done with a field map, a field map for a complex lattice can be extremely large, and will often require an external magnet design code to compute. Thus, the capability of finding a field from more compact information (such as the longitudinal variation of multipole components) is an extremely useful capability of any tracking or analysis code for muon accelerators.

### 2.2. Truncated power series

Many accelerator design codes use a truncated power series representation as part of their analysis. When a beam has a large phase space area, there is a concern with the rate of convergence for the power series at large amplitudes.

In most cases, a truncated power series will give an adequate description of the transfer map through a short section of the lattice. However, the transfer map for a longer section of lattice may not give an accurate representation of the dynamics. Fig. 5 shows the trace of the transfer matrix (which should be between $-2$ and 2 for stability) for 10 cells (each cells contains a single triplet) of an FFAG lattice. Note that for even a 10th order power series, the trace of the transfer matrix is so inaccurate as to give the incorrect results for the range of stability. It is in fact possible to construct a sufficient anomalous triplet lattice where the power series fails to converge within the operating range of the lattice.

The basic difficulty is that in composing two truncated power series and then re-truncating, information is lost in the final truncation.

This should not lead to ruling out truncated power series as a useful technique for analyzing these lattices, but they should be used with extreme caution. Maps made for a long section of lattice are more likely to have problems than those made from shorter sections. Many algorithms based on truncated power series require a map for the entire ring [12]. Some analysis can be performed by making a power series at a fixed energy for several different energies. In fact, most FFAG analysis occurs by finding a closed orbits at different energies and computing the linear maps about those closed orbits [13].

### 2.3. Analysis techniques

In the design of muon accelerators, many of the questions asked are different than those that are asked in traditional accelerators. This is true especially for cooling: one may want to compute the equilibrium emittance, the rate of cooling, and various related merit factors.

Muon cooling systems have generally been designed using tracking. However, using tracking has many difficulties. Often one is trying to compare some merit factor for two different designs, and the statistical fluctuations due to using a finite number of particles make differences in the merit factor difficult to ascertain. Using more particles would be helpful, but due to the complexity of the computation of the fields as described above, it becomes prohibitive to run large numbers of particles. Thus, improving the efficiency of field computations would allow more effective machine design. Making codes such as ICOOL [14] run on parallel computers would also be useful, and will probably be necessary for any final machine design.

Ideally, we would like to have techniques available for cooling lattice design that do not involve simply tracking particles. Such an analysis code should be able to compute the usual quantities (linear maps, closed orbits), but should also include effects such as cooling and multiple scattering (in some averaged way). Some theoretical work on this subject has been done, but under restricted conditions [15]. This theory has yet to be implemented in any analysis codes.

One may prefer a library which implements accelerator computations instead of a single code; this would allow one to use a higher-level language to compute complex computations on the results of simpler lattice computations.

## 3. Target

To achieve acceptable physics results, muon accelerators require the production of large quantities of muons. The weak interactions of neutrinos with matter means that to produce a reasonable number of events, one must start with large quantities of muons. For a muon collider, the diffuse phase space in which the muons are produced requires that luminosity be achieved with large numbers of muons instead of very small beams.

In designing the muon production system, the first thing one must know is the number of pions and muons produced. For the design of the subsequent capture systems, one must also know their energy and angular spectrum. One must be able to compute particle production to compute energy deposition in the target area materials so as to design the cooling system. And one must understand how materials will behave under this irradiation, both for the purposes of radiation protection as well as material lifetime and degradation of material properties.

There are several codes which compute particle interactions with matter [16–20]. However, there are uncertainties in the results of those codes by as much as 30% in some regimes [21,22]. These uncertainties may arise both from the algorithms in the codes as well as the unknown nature of the physical processes. If we want to improve prediction accuracy for target design, these uncertainties must be reduced.

Target damage and destruction is a significant problem to be avoided for any high-power target. It is believed that if the energy deposition in and properties of the material are known, existing codes can predict the point at which the target will be damaged [23]. However, the material properties, such as the coefficient of thermal expansion and the yield strength, will change under irradiation [24]. One must be able to predict this change computationally, which at this point requires experimental results to give parameters to the computation.

To avoid the problem of breaking a solid target, a liquid mercury jet has been proposed [25]. The evolution of this jet under proton bombardment and in a magnetic field must be simulated to ascertain whether the jet will be sufficiently stable to be hit by a proton beam, whether earlier portions of the jet will interfere with later portions of the jet or with pion production, and other kinds of effects (see Fig. 6). Parts of the liquid mercury will become vapor as a result of the proton energy deposition and the subsequent pressure waves in the mercury. This "cavitation" process is essential to correctly model the mercury jet. Code has been written which models the fluid using a two-phase equation of state [27]. This has been only done in two dimensions; it can in principle be done in three dimensions, but is computationally expensive. Another approach is to treat the fluid as a homogeneous "bubbly fluid" using an equation of state [28]; this has been done in three dimensions (Fig. 7). While simulations without cavitation (either model) have been performed without magnetic fields, simulations with cavitation or the bubbly fluid model and magnetic fields have not. Furthermore, as can be deduced from Fig. 6, the evolution of the jet in the nozzle (including magnetic fields) is important for producing a good target, but has not been studied as yet.

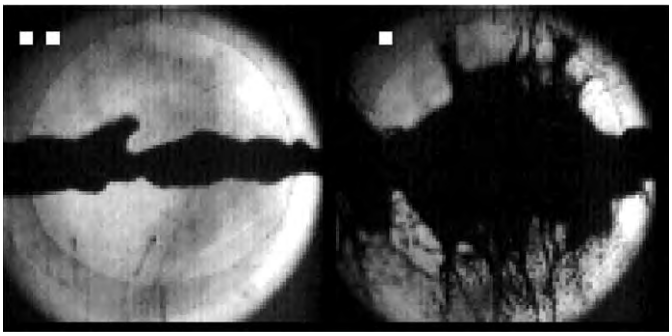


Fig. 6. Experimental results showing the mercury jet before (left) and after (right) being hit with a proton beam [26].

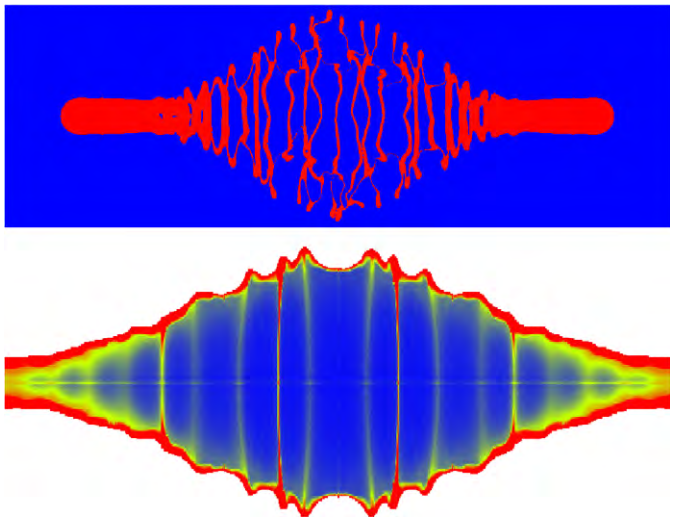


Fig. 7. Simulation of interaction between the beam and the liquid mercury jet with cavitation: top is with the two-phase model, bottom is with the homogeneous equation of state.

## 4. Conclusion

To design and predict the performance of muon accelerators, one must carefully simulate all aspects of the machine. Single-particle dynamics are well understood, but must be handled with particular care due to the characteristics of the muon beam and the beamline guiding it. Targets must be carefully studied to ensure that they will perform as needed and survive. Understanding of RF breakdown, especially in a magnetic field, must be obtained to attempt to maximize the gradients in room-temperature RF cavities.

## References

[1] M.M. Alsharo'a, et al., Phys. Rev. ST Accel. Beams 6 (2003) 081001.
[2] Y. Fukuda, et al., Phys. Rev. Lett. 81 (1998) 1562.
[3] S. Fukuda, et al., Phys. Rev. Lett. 86 (2001) 5651.
[4] Q.R. Ahmad, et al., Phys. Rev. Lett. 89 (2002) 011301.
[5] Q.R. Ahmad, et al., Phys. Rev. Lett. 89 (2002) 011302.
[6] K. Eguchi, et al., Phys. Rev. Lett. 90 (2003) 021802.
[7] M.H. Ahn, et al., Phys. Rev. Lett. 90 (2003) 041801.
[8] M. Apollonio, et al., Eur. Phys. J. C 27 (2003) 331.
[9] J. Scott Berg, et al., in: J. Chew, P. Lucas, S. Webber (Eds.), Proceedings of the 2003 Particle Accelerator Conference, IEEE, Piscataway, NJ, 2003, p. 3413.
[10] J. Norem, et al., Phys. Rev. ST Accel. Beams 6 (2003) 072001.
[11] K. Makino, M. Berz, Nucl. Instr. and Meth. A 427 (1999) 338.
[12] E. Forest, M. Berz, J. Irwin, Part. Accel. 24 (1989) 91.
[13] J.S. Berg, C. Johnstone, in: Proceedings of the 2003 Particle Accelerator Conference, J. Chew, P. Lucas, S. Webber (Eds.), IEEE, Piscataway, NJ, 2003, p. 2216.
[14] R. Fernow, in: Proceedings of the 1999 Particle Accelerator Conference, New York, 1999, A. Luccio, W. MacKay (Eds.), IEEE, Piscataway, NJ, 1999, p. 3020.
[15] C.-xi Wang, K.-Je Kim, Phys. Rev. Lett. 88 (2002) 184801.
[16] N.V. Mokhov, et al., Radiat. Prot. Dosimet, Also Fermilab report FERMILAB-Conf-04/053-AD, 2004, to appear.
[17] Nikolai V. Mokhov, Shielding Aspects of Accelerators, Targets, and Irradiation Facilities (OECD), to appear. Also Fermilab report FERMILAB-Conf-03/053, 2003.
[18] A. Fasso, A. Ferrari, P.R. Sala, in: A. King, et al. (Eds.), Advanced Monte Carlo for Radiation Physics, Transport Simulation and Applications, Springer, Berlin, 2001, pp. 159–164.
[19] A. Fasso, A. Ferrari, J. Ranft, P.R. Sala, in: A. King, et al. (Eds.), Advanced Monte Carlo for Radiation Physics, Transport Simulation and Applications, Springer, Berlin, 2001, pp. 955–960.
[20] W.R. Nelson, H. Hirayama, D.W.O. Rogers, SLAC report SLAC-265, 1985.
[21] N.V. Mokhov, et al., Fermilab report FERMILAB-FN-732, 2003.
[22] S. Chiavassa, et al., Radiat. Prot. Dosimet, to appear.
[23] N. Simos, et al., in: Proceedings of EPAC 2002, Paris, France, p. 2709.
[24] H. Kirk, et al., in: J. Chew, P. Lucas, S. Webber (Eds.), Proceedings of the 2003 Particle Accelerator Conference, IEEE, Piscataway, NJ, 2003, p. 1628.
[25] C.M. Ankenbrandt, et al., Phys. Rev. ST Accel. Beams 2 (1999) 081001.
[26] H. Kirk, et al., in: Proceedings of the 2001 Particle Accelerator Conference, Chicago, P. Lucas, S. Webber (Eds.), IEEE, Piscataway, NJ, 2001, p. 1535.
[27] R. Samulyak, Y. Prykarpatskyy, Math. Comput. Simul. 65 (2004) 431.
[28] R. Samulyak, T. Lu, Y. Prykarpatskyy, in: Lecture Notes in Computer Science, vol. 3039 Springer, Berlin, Heidelberg, 2004, pp. 653–660.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 25–31
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Computational needs for the RIA accelerator systems ☆

P.N. Ostroumov*, J.A. Nolen, B. Mustapha

*Physics Division, Argonne National Laboratory, 9700 S. Cass Ave., Argonne, IL 60439, USA*

Available online 28 November 2005

**Abstract**

This paper discusses the computational needs for the full design and simulation of the RIA accelerator systems. Beam dynamics simulations are essential to first define and optimize the architectural design for both the driver linac and the post-accelerator. They are also important to study different design options and various off-normal modes in order to decide on the most-performing and cost-effective design. Due to the high-intensity primary beams, the beam-stripper interaction is a source of both radioactivation and beam contamination and should be carefully investigated and simulated for proper beam collimation and shielding. The targets and fragment separators area needs also very special attention in order to reduce any radiological hazards by careful shielding design. For all these simulations parallel computing is an absolute necessity.


*PACS:* 29.17.+w; 29.27.−a; 41.75.−i

*Keywords:* RIA simulations; Beam dynamics; Stripper modeling; Target simulations; Fragment separators

## 1. Introduction

The Rare Isotope Accelerator (RIA) is a next generation facility for basic research with intense beams of radioactive and rare isotopes [1]. To produce such beams RIA will use primary beams of any ion from protons (up to 1 GeV) to uranium (up to 400 MeV/$u$) with beam power up to 400 kW. RIA is based on two CW superconducting (SC) linacs, a 1.4-GV driver designed to simultaneously accelerate multiple-charge-state heavy-ion beams and a ~140-MV post-accelerator designed for the efficient acceleration starting from singly charged secondary beams with masses up to $A = 240$ from ion source energies. To meet the facility requirements RIA will also include state-of-the-art electron cyclotron resonance (ECR) ion sources for the production of high-intensity heavy-ion beams, two stripping stations for beams of heaviest ions, high-power ISOL and fragmentation targets, high-resolution fragment and isobar separators and transport systems for beams with large momentum and charge spread. RIA will also have a new production scheme combining both the fragmentation and the ISOL methods by thermalizing fast radioactive ions in a gas catcher [2] to produce low-energy good-quality secondary beams. Fig. 1 shows a schematic layout of the RIA facility including four experimental areas with different secondary beam energies serving different experimental programs from Ion Traps to Astrophysics to Nuclear Structure to high-energy Nuclear Reactions.

RIA has recently been ranked as the third highest priority for future Scientific Facilities in the 20 year plan of the US Department of Energy (DOE) [3]. Following this high ranking, the RIA project received a preliminary approval (CD-0: critical decision 0) from the DOE. The research and development (R&D) for RIA involves many national laboratories and universities. An impressive progress has been made over the last few years in the different areas of RIA R&D from the ion source [4] to the driver linac [5] to the different targets [6] and fragment separators [7] to the post-accelerator [8]. In this paper, we discuss the computational needs for the full design and simulation of the RIA accelerator systems. In the next section, we review both the existing and newly developed

☆ This work is supported by the U.S. Department of Energy under contract W-31-109-Eng-38.
*Corresponding author. Tel.: +1 630 252 4897; fax: +1 630 252 9647.
*E-mail address:* ostroumov@phy.anl.gov (P.N. Ostroumov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.034

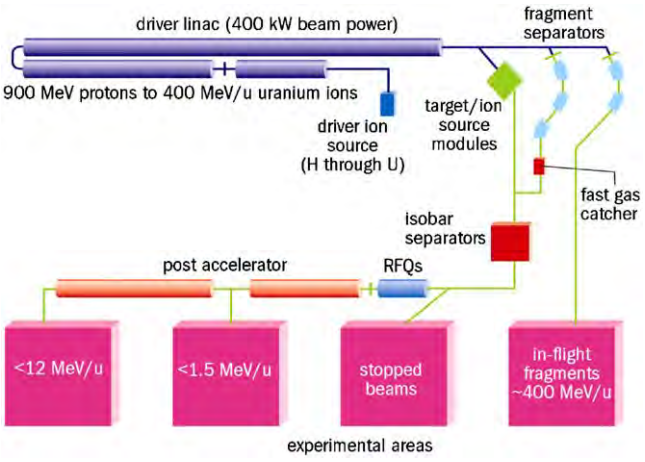



Fig. 1. Schematic layout of the RIA facility showing the three sections of the driver linac starting from the ion source, the three radioactive isotope production schemes and the four experimental areas.

tools for both accelerator design and beam dynamic simulations emphasizing the need to fully describe the new features of the RIA driver linac. In Section 3, we consider the beam-stripper interaction including elastic and inelastic interactions and the subsequent need for beam collimation and shielding. In Section 4, we discuss the production and separation of radioactive isotopes emphasizing the aspect of radioactive ion release from ISOL targets and the radiological and shielding aspects for both the ISOL and fragmentation targets.

## 2. Beam dynamics

### 2.1. Goal and scope

The ultimate goal of the beam dynamics simulations is to define the overall architectural designs for both the driver linac and the post-accelerator that satisfy the facility requirements. End-to-end simulations are necessary to study the performance and limitations of different design options in order to decide on the most-performing and cost-effective design. In order to fulfill these goals the end-to-end simulations should include:

- Up-to-date heavy-ion beam physics.
- Multiple charge state acceleration necessary to reach the high intensity goal.
- Stripper effects on the dynamics of heaviest ion beams.
- Automatic corrective steering required for simulations with errors.
- Capability of determining both the fractions and locations of any beam loss in order to optimize the design and define the tolerance to different errors.
- Capability of studying possible failure modes and ways to restore the beam by automatic retuning of the accelerator excluding the failing elements.

### 2.2. Tools for element and section design and optimization

For accelerator element design including RF cavities and the calculation of the corresponding 3D electromagnetic fields, codes such as Microwave Studio, Electromagnetic Studio and MAFIA from CST [9], ANSYS [10], Poisson/Superfish [11] and HFSS (High-Frequency Structure Simulator) [12] may be used. Recently, we have used the advanced electromagnetic code Omega-3P developed by the SLAC group [13] for more precise resonator design.

For the design of the different sections of the SC Linac including optimization and beam matching, matrix-based codes such as TRANSPORT [14], TRACE3D [15], COSY [16] and GIOS [17] are often used. For the RT front-end structure, codes like DYNAMION [18] and PARMTEQ [19] may be used.

### 2.3. Codes for full design and end-to-end simulations

A few years ago, none of the existing beam dynamics codes was able to fully describe the beam dynamics in all the elements of the RIA accelerator systems and more importantly the new features of the driver linac like multiple-charge-state acceleration and stripper simulation. A great effort has been made during the last few years toward developing new codes that allow detailed studies and end-to-end simulations of the RIA driver linac. Two new ray-tracing codes have been specially developed: the LANA code originally developed at the Institute for Nuclear Physics INR-Moscow [20] and currently supported at Michigan State University and the TRACK code [21] developed at Argonne National Laboratory. Both codes have been used to perform detailed simulations of the SC linac sections and produced similar overall results except in some of the details. For an independent validation of the two codes, an effort is currently under way to modify the existing codes PARMTEQ and IMPACT [22] for the simulation of heavy-ion beams in RIA-type accelerators. Initial comparisons of energy gain and beam second moments in the high-energy section of the RIA driver linac show very good agreement [23] with the code TRACK.

### 2.4. Simulations of two linac options using TRACK

Since the first simulations [24] the code TRACK has undergone many updates and further development [25] to either include new features or refine some of the existing ones. Among these updates we cite:

- Realistic space charge effects of multi-component ion beams.
- Realistic stripper effects including thickness fluctuations on the beam properties.
- End-to-end simulation of the driver linac starting from the ion source.
- Randomly generated misalignment and RF errors.

- Automatic and realistic beam steering based on beam position monitors.
- Capability of determining the fractions and locations of eventual beam losses.
- Parallel computing on multi-processor machines.

Fig. 2 shows the evolution of the beam envelopes along the driver linac starting from the ECR ion source to the target (bottom) as well as the particle coordinates of the 5 charge state uranium beam at the target location (top). Using this powerful simulation tool we have performed extensive simulations of two linac design options including different sources of errors on the parallel computer cluster Jazz [26] at Argonne. The goal of these simulations was first to identify the most critical errors, second to establish an error budget based on beam loss analysis and finally to compare the performances and limitations of the two linac options. The first linac option is the original Baseline design described in [5] and the second is the Triple-spoke design where the elliptical-cell cavities in the high-energy section of the linac are replaced by the triple-spoke cavities under development now at ANL [27].

Table 1 lists the errors used in these simulations and their typical amplitudes. Simulating 50 random sets (10 million particles) for each individual error and comparing to the case with no errors, we were able to identify the RF field and phase errors and the fluctuations in the strippers thicknesses as the most critical errors. For further investigation and beam loss analysis of both designs we simulated different combinations of these errors keeping other errors at their values of Table 1. Table 2 lists the different combinations of RF errors and stripper thickness fluctuations. For each combination, 200 random sets of errors were simulated with $2 \times 10^5$ particles each (a total of 40 million particles). Increasing the error amplitudes from combination 1 to 6, we noticed an increase in the longitudinal emittances for both the Baseline and the Triple-spoke designs. The Baseline design showed more sensitivity with an increase in the transverse beam emittances not observed for the Triple-spoke design. This increase in the transverse emittances reflects a possible coupling between the transverse and longitudinal motion which resulted into an increasing beam loss in the Baseline design [25]. Whereas no beam losses were observed for the Triple-spoke design even in the case of the highest errors (combination 6). Fig. 3 shows beam losses in W/m along the driver linac for both the Baseline and the Triple-spoke designs. The first two peaks on each plot correspond to the losses at the two strippers which are controlled losses to be stopped at the collimators following the strippers. For the Baseline design, uncontrolled losses are observed in the high-energy section. They are negligible for combinations 1&2, approaching the 1 W/m limit for hands-on maintenance for combinations 3&4 and about 10 W/m for combinations 5&6. Whereas no uncontrolled losses were observed for the Triple-spoke design. From these studies we conclude that the Baseline design has more limitations concerning beam losses and that the Triple-spoke design is more tolerant of errors.

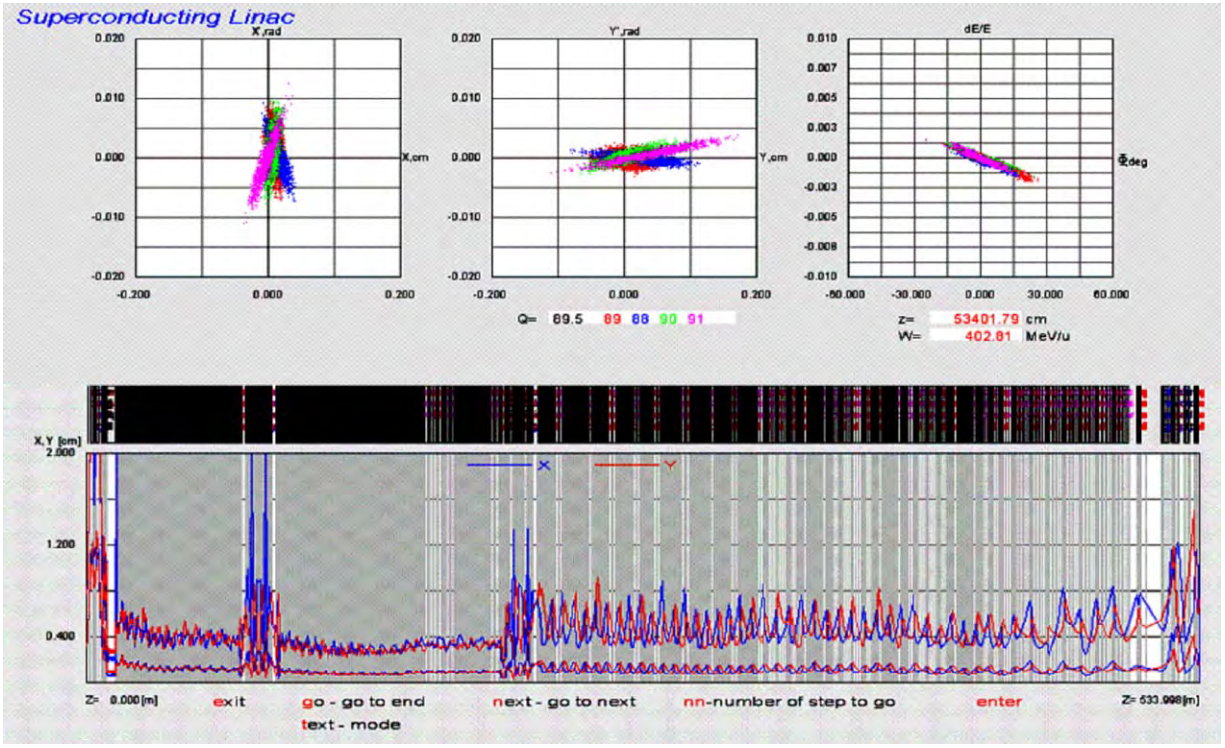



Fig. 2. End-to-end simulation of the baseline driver linac showing beam envelopes throughout the linac (bottom) and the phase–space plots of the multi-charge states beam at the linac exit (top).

Table 1
Different sources of error and their typical values

| Error | Description | Value | Distribution |
|---|---|---|---|
| 1 | Cavity end displacements | 0.05 cm (max.) | Uniform |
| 2 | Solenoid end displacements | 0.015–0.05 cm (max.) | Uniform |
| 3 | Quadrupole end displacements | 0.01 cm (max.) | Uniform |
| 4 | Quadrupole rotation | 2 mrad (max.) | Uniform |
| 5 | Cavity field error | 0.5% (r.m.s.) | Gaussian |
| 6 | Cavity phase error | 0.5° (r.m.s.) | Gaussian |
| 7 | Stripper thickness fluctuation | 5–10% (FWHM) | Gaussian |

Table 2
Combinations of RF errors and stripper thickness fluctuations used to study beam dynamics and beam losses

| Combination | RF errors | Thickness fluctuation |
|---|---|---|
| 1 | Field: 0.3%, Phase: 0.3° | 5% FWHM |
| 2 | Field: 0.3%, Phase: 0.3° | 10% FWHM |
| 3 | Field: 0.5%, Phase: 0.5° | 5% FWHM |
| 4 | Field: 0.5%, Phase: 0.5° | 10% FWHM |
| 5 | Field: 0.7%, Phase: 0.7° | 5% FWHM |
| 6 | Field: 0.7%, Phase: 0.7° | 10% FWHM |

### 2.5. Future code development

The goal of future code development is to develop a complete accelerator set-up tool based on an accurate and realistic computer model. Experience at many operating machines showed that such models are essential and save both machine time and manpower. The main future development is to develop a longitudinal phase setting procedure for multi-charge state beams in order to:

- Account for cavity-dependent field levels which is inherent to a linac comprising a large number of SC resonators.
- Minimize the effective beam emittance at the stripper location.
- Produce tunes for ions with different $q/A$ ratios to maximize the output energy.
- Develop off-normal tunes to compensate for missing resonators after an eventual failure.

Along with the longitudinal tuning procedure, a procedure for the optimization of the beam transverse envelopes is also necessary in order to:

- Determine optimum focusing fields to minimize the beam envelopes.
- Produce transverse tunes for ions with different $q/A$ ratios.
- Retune the accelerator after a failure excluding the failing resonators and or focusing elements.

## 3. Strippers

Stripping is necessary to reach intermediate and high energies in heavy-ion accelerators. By increasing the ions $q/A$ ratio, the acceleration to higher energies is more efficient and more importantly more cost-effective. However, stripping will affect both the beam intensity and quality. In the beam-stripper interactions we distinguish between elastic and inelastic processes. Elastic interactions include atomic and elastic nuclear interactions which could change the charge, the energy and the angle of the incident ion but not the ion specie. Inelastic interactions are nuclear reactions which could produce radioactive products.

Charge exchange or electron stripping is the process responsible of producing the charge state distribution of the beam after a stripper. Semi-empirical formulas and Monte-Carlo codes are available to estimate the required stripper thickness and calculate the charge state distribution. From these we cite the set of formulas originally developed by Baron et al. [28] and later updated by Leon et al. [29], the code ETACHA [30] and the code GLOBAL [31] which is more valid for energies above 100 MeV/$u$. In the case of RIA, a comparative study [32] of these formulas and codes for uranium beams led to the necessity of experimental measurements at the exact stripping energies. These measurements have already been performed with 11 MeV/$u$ uranium at Texas A&M University, 85 MeV/$u$ uranium at GSI-Darmstat and 80 MeV/$u$ bismuth at Michigan State University. Analysis of the data is under way.

Multiple Coulomb scattering with both the electrons and atoms of the stripping medium is the process responsible for the energy (energy loss and straggling) and angle (angular straggling) distributions of the beam after a stripper. For particle tracking codes like TRACK, it is important to have the correlated energy and angle distributions including the respective tails. A Monte-Carlo code such SRIM [33] may be used, however, when compared with data the calculated energy-loss straggling is 5–7 times smaller. We believe that this is due to the fact that SRIM does not include any fluctuations in the stripper thickness or the ion charge state which could be responsible for the extra-broadening of the energy peak. In order to include these effects and for fast generation of the correlated energy and angle distributions ion by ion we opt for the parameterization of SRIM results for a uranium beam and both strippers [25]. An event generator based on these parameterization is now used by the code TRACK. We are currently developing a more general stripper model for any beam-stripper combination.

The simultaneous acceleration of multiple charge states in the RIA driver linac will not only reduce the losses at the strippers but also reduce the need for shielding. However, due to the high intensity, radioactivation and possible

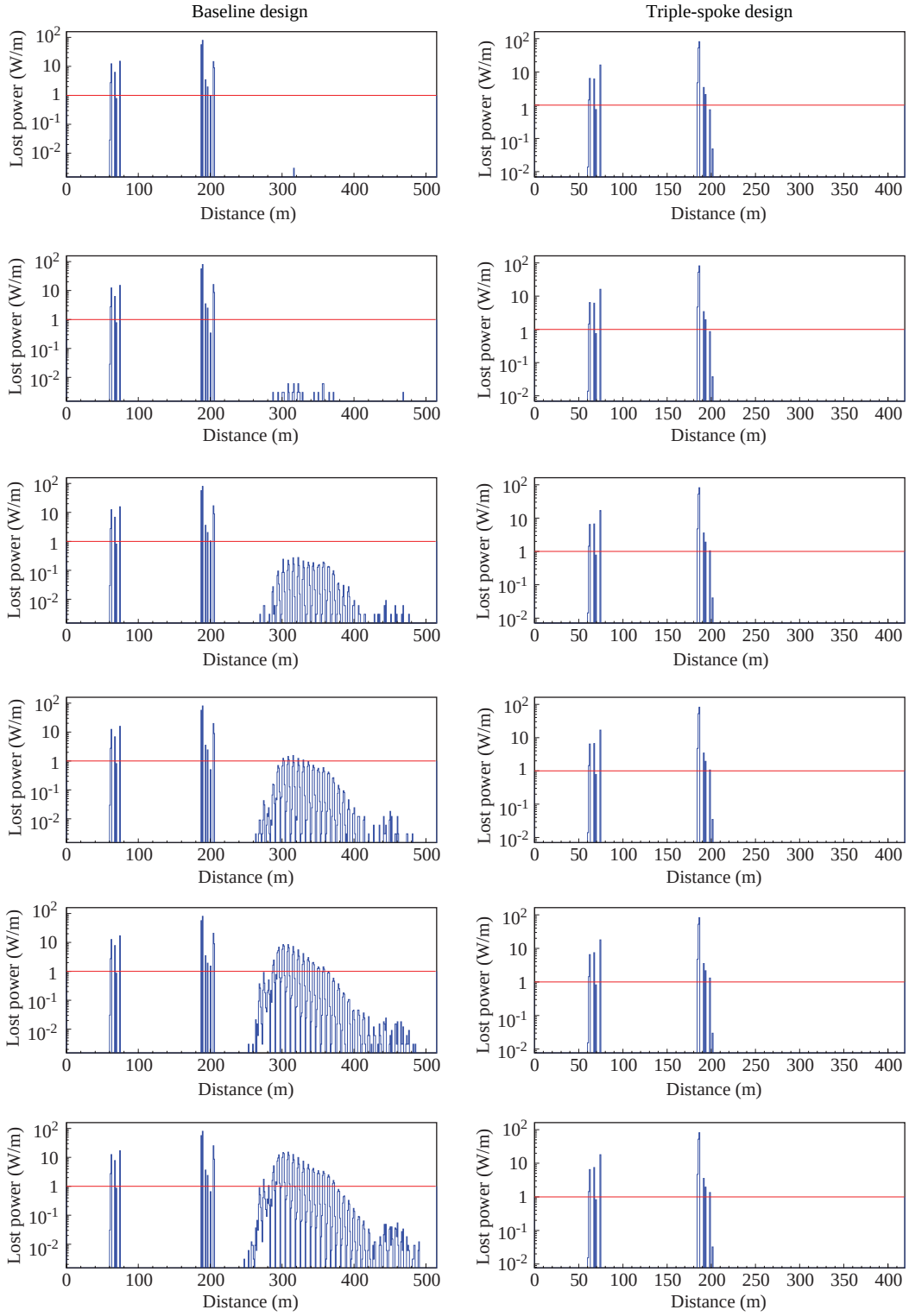



Fig. 3. Beam power lost along the accelerator in W/m for both the baseline and Triple-spoke design and the error combinations of Table 2. The plots are for combination 1 (top) to combination 6 (bottom), respectively. The horizontal line shows the 1 W/m limit to not exceed for hands-on maintenance. The first two peaks on each figure correspond to the losses at the two strippers which are controlled losses.

beam contamination by nuclear reactions products become an issue. Knowing the fractions and distributions of these radioactive products is very important for the appropriate design of beam collimation and shielding. Depending on the stripper material and the beam energy and mass, different mechanisms of inelastic interactions may take place. They become more important at energies $>50\,\mathrm{MeV}/u$. The dominant processes are the fragmentation for most heavy

ions and fission for the heaviest ones. In order to determine the interacting beam fraction and the nuclei produced by fragmentation, we may use the semi-empirical formula EPAX [34] for production yields. EPAX is energy-independent and most valid at higher energies. Monte-Carlo codes based on the two-step model consisting of an Intra-Nuclear Cascade step (ISABEL [35], INCL [36],...) followed by a Fission-Evaporation step (ABLA [37], PACE [38],...) are available and may be used. A Recently developed heavy-ion-reactions event generator LAQGSM [39] seems to be promising. In addition to the production yields these codes could also produce full energy and momentum distributions of the fragments. This information could be used for further tracking using a beam dynamics code in order to study related beam losses and possible beam contamination. We are currently investigating these issues. For the design of beam dumps and shielding, codes such MCNPX [40] and MARS [41] may be used. For both codes, development is under way to include heavy-ion capabilities, by incorporating the LAQGSM event generator. A newly developed particle and heavy-ion transport code PHITS [42] is also used. Monte-Carlo tracking of heavy-ion reactions products with such general purpose codes will be very computer-intensive and require parallel processing.

## 4. Targets and fragment separators

Targets have a lot in common with strippers except that they will receive a much higher beam power due to the higher energy and more importantly that most of the beam will be stopped. Only a fraction of the secondary beam (the selected isotope) is to be transmitted for immediate use or re-acceleration.

In the case of ISOL targets where a light-ion beam is incident on a thick target of a heavy element, the beam will be stopped in the target. To handle such high beam powers (up to 400 kW) two target design options are been developed: the two-step target [43] and the tilted-foil target [44]. In both cases, the reaction products have to first diffuse through the target material then effuse through the target enclosure to be ionized in an ion source and extracted. This process known as the release process is chemical dependent and require the use of an isobar separator to select a given isotope. We have recently developed a Monte-Carlo code package to simulate the release process [45] where the diffusion is based on a theoretical model [46]. The effusion is simulated by tracking the produced ions inside the target-ion source system [47] using the geometry and tracking capabilities of Geant-4 [48]. These targets will be very hot and radioactive, appropriate cooling and shielding are absolutely required. Prototyping as well as code development for full simulations of these two aspects are under-way.

In the case of fragmentation targets where a heavy-ion beam is incident on a thin target of a light element, up to a third of the beam power will be deposited in the target. A windowless flowing liquid-lithium target has already been built and tested for this purpose [49]. In this case, the non-interacting primary beam and the secondary beam (of all reaction products) will leave the target into a small angle cone with energies close to the primary beam energy. A fragment separator will be used to first deflect the non-interacting beam into a beam dump and second to separate and select the isotope of interest from the rest of the beam which should also be stopped in specially designed beam stops. Although the demonstration of the liquid-lithium target solved the problem of power dissipation in the target, the problems of power dissipation and induced radioactivity in the fragment separator and its beam dumps are yet to be investigated. For beam dumps and shielding, the same codes as for the stripper areas (MNCPX and MARS after developing heavy-ion capabilities as well as the code PHITS) could be used.

## 5. Summary

The basic design and simulation tools for the RIA accelerator systems exist. Parallel computing is vital for extensive simulations of the full system. Further code development is under way in all areas: accelerators, strippers, targets and fragment separators.

## References

[1] J.A. Nolen, Nucl. Phys. A 734 (2004) 661.
[2] G. Savard, et al., Nucl. Instr. and Meth. B 204 (2003) 582.
[3] DOE Office of Science web site @ ⟨http://www.sc.doe.gov⟩.
[4] D. Leitner, ECR Ion sources for RIA, RIA Facility Workshop, March 9–13, 2004, Organized jointly by the NSCL and the ANL Physics Division @ ⟨http://meetings.nscl.msu.edu/ria2004/talks_dl.php⟩.
[5] P.N. Ostroumov, Phys. Rev. ST. Accel. Beams 5 (2002) 030101.
[6] J.A. Nolen, et al., Nucl. Phys. A 701 (2002) 312c.
[7] B.M. Sherrill, Fragment Separator at RIA, RIA Facility Workshop, March 9–13, 2004, Organized jointly by the NSCL and the ANL Physics Division @ ⟨http://meetings.nscl.msu.edu/ria2004/talks_sbd.htm⟩.
[8] P.N. Ostroumov, et al., in: Proceedings of the RIA R&D Workshop, Washington, DC, August 26–28, 2003 @ ⟨http://www.c-ad.bnl.gov/RIA⟩, paper 2.4.3.
[9] CST, Darmstadt, Germany web site @ ⟨http://www.cst.de/⟩.
[10] ANSYS web page @ ⟨http://www.ansys.com⟩.
[11] Poisson/Superfish, LAACG' web page @ ⟨http://laacg1.lanl.gov⟩.
[12] HFSS, ANSOFT web page @ ⟨http://www.ansoft.com⟩.
[13] Y. Sun, Ph.D. Thesis, Stanford University, 2002.
[14] K.L. Brown, The ion optical program TRANSPORT, Technical Report 91, SLAC, 1979.
[15] K.R. Crandall, TRACE 3-D documentation, Report LA-11054-MS, Los Alamos, 1987.
[16] M. Berz, COSY INFINITY, Version 8 User's Guide and Reference Manual, MSU, 1999.
[17] H. Wollnik, J. Brezina, M. Berz, Nucl. Instr. and Meth. A 258 (1987) 408.
[18] A.A. Kolomiets, et al., in: S. Myers, et al. (Eds.), Proceedings of the Sixth European Particle Accelerator Conference, Stockholm, Sweden, June 22–26, 1998, p. 1201.

[19] T.P. Wangler, et al., in: Proceedings of the RIA R&D Workshop, Washington, D.C., August 26–28, 2003 @ ⟨http://www.c-ad.bnl.gov/RIA⟩, paper 2.0.1.
[20] D.V. Gorelov, P.N. Ostroumov, in: R. Pascual (Ed.), Proceedings of the Fifth European Particle Accelerator Conference, Universitat Autonoma de Barcelona, Sitges, Spain, 1996, p. 1271.
[21] P.N. Ostroumov, V.N. Aseev, TRACK—a code for beam dynamics simulation in SC linac with 3D electric and magnetic fields, ANL Technical report, 2003.
[22] R.D. Ryne, et al., MaryLie/IMPACT: A Parallel Beam Optics Code with Space Charge, ICAP 2002 Abstracts and Posters @ ⟨http://icap02.pa.msu.edu/papers/⟩.
[23] R.D. Ryne, et al., in: Proceedings of the RIA R&D Workshop, Washington, D.C., August 26–28, 2003 @ ⟨http://www.c-ad.bnl.gov/RIA⟩, paper 2.0.2.
[24] P.N. Ostroumov, Phys. Rev. ST. Accel. Beams 3 (2000) 030101.
[25] P.N. Ostroumov, V.N. Aseev, B. Mustapha, Phys. Rev. ST. Accel. Beams 7 (2004) 090101.
[26] Jazz, LCRC web page @ ⟨http://www.lcrc.anl.gov⟩.
[27] K.W. Shepard, et al., Phys. Rev. ST Accel. Beams 6 (2003) 080101.
[28] E. Baron, et al., Nucl. Instr. and Meth. A 328 (1993) 177.
[29] A. Leon, et al., At. Data Nucl. Data Tables 69 (1998) 217.
[30] J.P. Rozet, et al., Nucl. Instr. and Meth. B 107 (1996) 67.
[31] C. Scheidenberger, et al., Nucl. Instr. and Meth. B 142 (1998) 441.
[32] M. Portillo, Ph.D. Thesis, Michigan State University, 2002.
[33] J.F. Ziegler, J.P. Biersack, U. Littmark, The Stopping and Range of Ions and Solids, Pergamon Press, NY, 1996. The code SRIM @ ⟨www.srim.org⟩.
[34] K. Summerer, B. Blank, Phys. Rev. C 61 (2000) 034607.
[35] Y. Yariv, Z. Fraenkel, Phys. Rev. C 20 (1979) 2227.
[36] J. Cugnon, Nucl. Phys. A 462 (1987) 751.
[37] J.J. Gaimard, K.-H. Schmidt, Nucl. Phys. A 531 (1991) 709.
[38] A. Gavron, Phys. Rev. C 21 (1980) 230.
[39] K.K. Gudima, S.G. Mashnik, A.J. Sierk, User Manual for the code LAQGSM, LANL Report LA-UR-01-6804 @ ⟨http://lib-www.lanl.gov/la-pubs/00818645.pdf⟩.
[40] L. Waters, MCNPX web page @ ⟨http://www.mcnpx.lanl.gov⟩
[41] N.V. Mokhov, MARS web page @ ⟨http://www-ap.fnal.gov/MARS⟩
[42] H. Iwase, et al., J. Nucl. Sci. Tech. 39 (2002) 1142.
[43] M. Portillo, et al., Nucl. Instr. and Meth. B 194 (2002) 193.
[44] B. Mustapha, J.A. Nolen, I.C. Gomes, Nucl. Phys. A 746 (2004) 437c.
[45] B. Mustapha, J.A. Nolen, Nucl. Instr. and Meth. B 204 (2003) 286.
[46] J. Crank, The Mathematics of Diffusion, Clarendon Press, Oxford, 1975.
[47] B. Mustapha, J.A. Nolen, Nucl. Instr. and Meth. A 521 (2004) 59.
[48] Geant-4 web page @ ⟨http://geant4.web.cern.ch/geant4/⟩.
[49] C.B. Reed, J.A. Nolen, J.R. Specht, V.J. Novick, P. Plotkin, Nucl. Phys. A 746 (2004) 161c.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 32–35
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Emittance growth as mesh artefact

R. Becker*, R.A. Jameson

*Institut für Angewandte Physik der, Universität Frankfurt, Robert-Mayer-Strase 2-4, 60054 Frankfurt, Germany*

Available online 21 November 2005

**Abstract**

The space charge beam spreading calculations of the laser ion source group of CERN [K. Hanke, et al., Rev. Sci. Instrum. 73 (2002) 783] has been recomputed with IGUN. Instead of only looking to the total value of beam spreading, we also investigated the behaviour of the rms-emittance and found the following unexpected results:

1) Even zero emittance beams show considerable emittance growth.
2) The emittance growth of the central part depends on both the number of meshes used and the number of trajectories used.
3) The most prominent emittance growth occurs at the beam boundary, which is independent of the number of trajectories used and only depends on the number of meshes inside the beam. The more meshes used, the less this emittance growth will be.

Generally, emittance growth is associated with the non-linear electric field components, integrated up along the particles trajectories. By using a fractional emittance analysis, we found two different contributions to emittance growth, which might explain the observed results:

(a) the deposition of space charge and the interpolation of fields inside the meshes produces errors and provides "noise" to the particle ray tracing algorithm.
(b) at the beam boundary, the filling of the space charge map depends very much on the relative location of the beam boundary to the discrete mesh structure.

This mesh artefact is inherent to all programs that need meshes for the evaluation of the influence of space charge on the potential distribution. When calculating a "real" beam with given emittance, it is good advice to compare this calculation with the emittance growth, obtained for a beam of zero emittance. Furthermore, a variation of the number of trajectories and the number of meshes will allow to extrapolate to the unaccessible use of an infinite number of trajectories and meshes.


*PACS:* 02.70.Bf; 07.05.Tp; 07.77.Ka; 29.25.−t; 29.27.Ac

*Keywords:* Emittance growth; Mesh artefact; Particle beams; Space charge

## 1. Introduction

Beam spreading by space charge is well understood and solved for non-thermal "laminar" beams of charged particles by numerical integration of the beam boundary equation. This requires that the density inside the beam is uniform. Therefore, the numerical simulation of this ideal case with space charge optics programs, like CPO3 [2], KOBRA [3], or IGUN [4–6], by the CERN laser ion source group [1] served as a test for the comparison of these different programs. CPO3 is a 3D boundary element program, KOBRA uses finite differences on a cartesian

*Corresponding author.
*E-mail addresses:* rbecker@physik.uni-frankfurt.de (R. Becker), jameson@postman.riken.jp (R.A. Jameson).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.009

mesh, while IGUN is based on finite differences on a rotational symmetric 2D square mesh. For the beam spreading of 60 and 120 mA $Ta^{20+}$ ions at 60 kV extraction voltage, all programs agreed reasonably well with results from the numerically integrated beam boundary equation. Only for the emittance, small differences have been seen between the programs as well as between simulations and experiments [1].

When recomputing the IGUN results, we started with a laminar beam, having zero emittance, in order to correctly compare with the numerical integration of the beam boundary equation. All trajectories should be self-similar to the beam boundary and no emittance should be seen after beam spreading. However, a considerable emittance has been observed, depending on the parameters used for the simulation. Similar effects have been reported earlier [7,8], but no explanation has been given. Emittance growth with zero initial emittance is a very unpleasant situation, because it makes all predictions on emittance growth questionable, which are obtained for "real" beams with such programs. We have used the tool of fractional emittance analysis, provided by IGUN, to acquire more insight about the origin of emittance growth of laminar beams.

## 2. Fractional emittance analysis

IGUN, as a simulation program with axisymmetric option, can use radial evenly distributed starting points for ray tracing. This "saves" on the number of rays and speeds up the computation accordingly. By this, trajectories are started with currents, which increase with radius. Therefore, the calculation of the rms-emittance needs weighting of the averages of radii $r_i$ and angles $\alpha_i$ by the currents $I_i$, which are attributed to trajectory $i$ :

$$E_{\mathrm{rms}}(j) = \sqrt{\frac{\sum_{i=1}^{j} r_i^2 I_i \times \sum_{i=1}^{j} \alpha_i^2 I_i - \sum_{i=1}^{j} (r_i \alpha_i)^2 I_i}{\sum_{i=1}^{j} I_i}}. \quad (1)$$

Furthermore, the emittance calculation is done after finding the virtual crossover of the final beam radii and angles and sorting all trajectories according to their radius. Then, the rms-emittance can be calculated for all inner beam parts, stopping summation at the according index $j$. This gives fractional rms-emittances, which will be considered in the following.

## 3. Simulations with IGUN©

The example of the CERN beam spreading comparison is shown in Fig. 1. Here, 20 trajectories started equidistant from the axis up to a radius of 15 mm ( = meshes) entering a tube of twice the diameter from the left to simulate a beam of 60 mA $Ta^{20+}$ ions at 60 keV total energy.

The representative trajectories spread out by about 17%, which is in good agreement with the numerical integration of the beam boundary equation. The equipotential lines show the potential increase towards the axis of about 1.5 kV, which is higher at the left, where the beam is narrower. The unexpected results are seen in Fig. 2, which shows the numerically determined radial current density profiles as well as the rms-emittances in selected positions of the second half of the above example: the initially laminar beam with zero emittance exhibits a considerable rms-emittance, although the profiles look quite reasonable.

Theoretically, emittance growth is associated with radially non-linear fields, through which the representative particles are passed. A uniform beam with constant and radius-independent current density should not experience non-linear fields. However, the space charge allocation method to mesh nodes is not uniform, as seen in the profiles of Fig. 2. In order to understand more about the emittance growth, we have undertaken a systematic variation of the parameters, which can be chosen to calculate the above example: these are the number of representative particles and the number of meshes used inside the beam. The results are shown in diagrams of fractional emittances versus trajectories used, in Fig. 3 for 100 trajectories, in Fig. 4 for 1000 trajectories. In both diagrams, three curves are plotted for the cases, where 7.5, 15, and 30 meshes lie inside the starting beam, e.g. a mesh resolution of 2, 1, and 0.5 mm. In each diagram and for all curves, we see 2 characteristic features. The inside beam

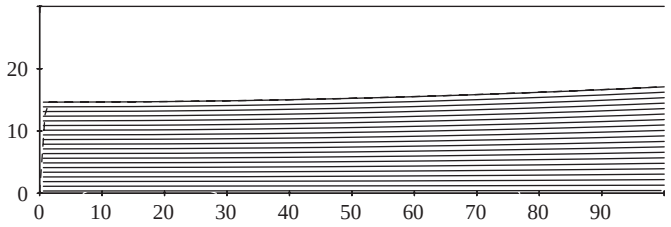



Fig. 1. Beam spreading of a $Ta^{20+}$ beam of 15 mm diameter over a length of 100 mm. Beam current 60 mA, total beam energy 60 keV.

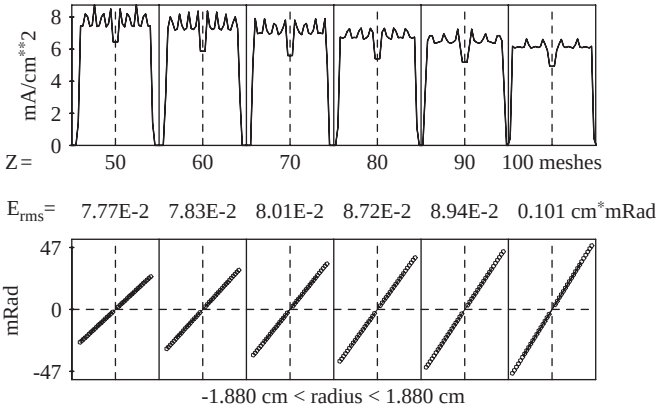



Fig. 2. Current density beam profiles (upper panels), rms-emittances (central line), and emittance plots for the beam spreading example of Fig. 1. Clearly, the non-uniform distribution of space charge can be seen, as well as the variation of this in different axial profile positions. The initial zero emittance has increased at 50 mm to 0.08 and at 100 mm to 0.1 cm mRad.

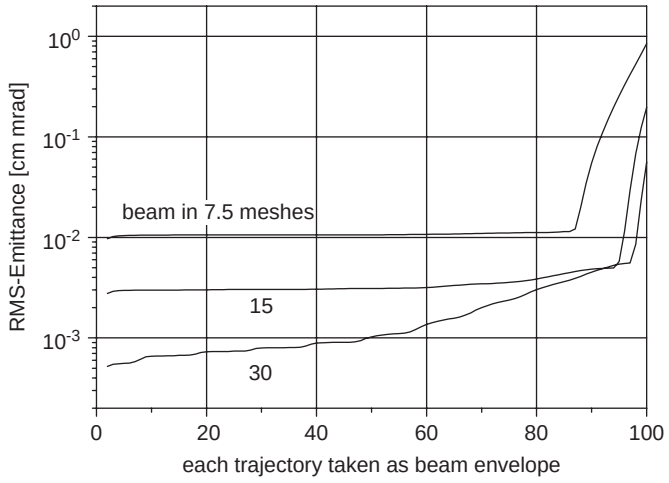



Fig. 3. Fractional emittances for a 3 cm diameter 60 KeV $Ta^{20+}$ beam, started with zero emittance and spreading over 10 cm. The beam is split up into 100 representative particles and 7.5, 15, and 30 radial meshes have been used for the 15 mm radius of the initial beam.

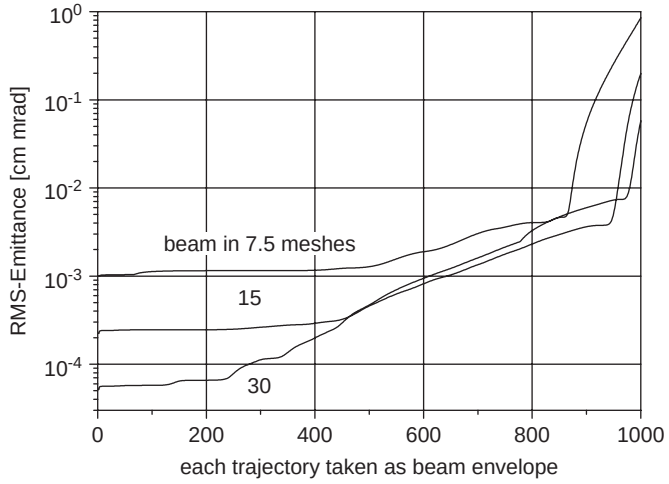



Fig. 4. Fractional emittances for a 3 cm diameter 60 KeV $Ta^{20+}$ beam, started with zero emittance and spreading over 10 cm. The beam is spilt up into 1000 representative particles and 7.5, 15, and 30 radial meshes have been used for the 15 mm radius of the initial beam.

has a lower and almost constant emittance, which is lower, if the meshes are more. For the trajectories in the last mesh at the beam boundary, the emittance suddenly rises by one to two orders of magnitude; the more, the lower the number of meshes are used, almost independent of the number of trajectories. Obviously, the emittance growth in the beam centre has a different origin than the one at its edge. In the centre, the growth is less if more trajectories are taken and also when more meshes are provided inside the beam; at the edge, the number of trajectories is almost unimportant, and the growth is less if more meshes are used. The emittance growth at the edge always seems to be much larger than in the beam centre and the fractional emmitances increase in radial direction within one mesh.

The observed behaviour of emittance growth for a beam, starting with zero emittance, is well explained by looking at the origin of non-linear fields: in the centre, the non-uniform space charge distribution, coming from the allocation algorithm, causes spikes and grooves in the profiles, as seen in Fig. 2. These become smoother with increasing number of trajectories as well as with an increasing number of meshes. At the beam boundary, the non-linear field is caused by the non-uniform filling of the space charge on the mesh nodes just inside and just outside the beam and this does not depend on the number of trajectories.

## 4. How to recognize the emittance growth artefact

Finally, the question may arise, how to distinguish between "real" and artificial emittance growth in a simulation, which starts with a beam of non-zero emittance. According to the experiences and explanations found for the above test case, a clear discrimination of the central emittance growth requires the comparison of a fractional emittance analysis for different numbers of trajectories as well as for different numbers of meshes, used inside the beam. Then the latter also will show much larger emittance growth at the beam boundary. Excellent experience with an increase of accuracy by applying extrapolation of numerical solutions with different kind of discretisation [9,10] suggests to use this procedure also here. The radial function of the increase of the fractional emittance should be extrapolated to arbitrary fine discretisation, using results obtained with different sets of discretisation parameters. An alternative approach is the use of a distorted radial mesh with highest mesh density at the beam boundary [11], which reduces the emittance growth artefact, but does not allow to discriminate it.

## 5. Conclusions

Emittance growth in beam spreading calculations of high-intensity ion beams, starting with zero emittance, has detected two kinds of mesh artefacts for finite difference simulation programs, like IGUN©: the numerical procedure to allocate space charge to mesh nodes causes small but distinct differences of the space charge of neighbouring nodes. The solution of the Poisson solver will then produce non-linear radial electric fields, even in the case of a homogeneous beam. This will create emittance growth for the inner part of the beam; the smaller it becomes, the more trajectories and meshes will be used. The most prominent mesh artefact occurs at the beam edge: the filling of the space charge map here causes very non-uniform space charges and hence electrical fields. This effect is independent of the number of trajectories used, but only depends on the number of meshes.

In case of real simulations, these mesh artefacts can be discriminated even for simulated ion beams with non-zero emittance, by varying the mesh resolution and the number of trajectories. Results obtained for the radial fractional emittance dependence then may be extrapolated to

infinitely fine discretisation, to reveal the correct emittance growth by non-linear electric fields, without the hiding effect of mesh artefacts.

## References

[1] K. Hanke, et al., Rev. Sci. Instrum. 73 (2002) 783.
[2] F. H. Read, F. Bowring, cpo@electronoptics.com
[3] P. Spädtke, KOBRA-3D, INP, Junckernstraße 99, D-65205 Wiesbaden, Germany
[4] R. Becker, W.B. Herrmannsfeldt, Rev. Sci. Instrum. 63 (1992) 2756.
[5] R. Becker, http://accelconf.web.cern.ch/AccelConf/e98/PAPERS/THP44G.PDF.
[6] http://www.egun-igun.com.
[7] J. Staats et al., http://accelconf.web.cern.ch/accelconf/p99/AI/AIFRAME.HTM.
[8] G. Pöplau, et al., Proceedings of EPAC-2000, Sittges, http://accelconf.web.cern.ch/accelconf/e00/PAPERS/TUP5B08.pdf.
[9] G. Martinez, R. Becker, LNCS 18 (2001) 153.
[10] R. Becker, et al, Nucl. Instr. and Meth. A 519 (2004) 49.
[11] G. Pöplau, et al., Proceedings of EPAC-2002, Paris, http://accelconf.web.cern.ch/accelconf/e02/PAPERS/WEPRI059.pdf.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 36–39
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Canonical harmonic tracking of charged particles in circular accelerators

V. Kvardakov*, E. Levichev

*The Budker Institute of Nuclear Physics, Novosibirsk, Russia*

Available online 6 December 2005

**Abstract**

Harmonic tracking is a method used to study non-linear particle dynamics in a circular accelerator. The tracking algorithm is based on numerical solution of the Hamilton equations of motion. An essential feature of the method is the approximation of Hamiltonian perturbation terms by a finite number of azimuthal harmonics, which provides an effective tool for optimization of non-linear particle motion. The equations of motion are solved canonically, with the first-order prediction made using the explicit Lie transformation. The major features of harmonic tracking are presented and examples of its application are discussed.


*PACS:* 41.85.Ct; 29.27.Bd

*Keywords:* Accelerators; Non-linear dynamics; Hamiltonian; Resonance; Tracking; Dynamic aperture

## 1. Introduction

Harmonic tracking (HT) is a method for research into the non-linear motion of a particle in a circular accelerator [1]. The method is based on numerical calculation of the Hamiltonian function of the perturbation potential expanded in a harmonic series. Integration is carried out over the independent azimuthal coordinate. In HT all non-linear excitation is replaced with the excitation potential, including all non-linear elements and represented by an infinite harmonic series.

In the present work we used expansion of the sextupole potential, since the sextupole lens is the basic non-linear element intended to compensate for natural chromaticity.

It is interesting that HT allows easy (in fact, manual) operation of the perturbation harmonics describing the non-linear potential, and therefore optimization of the parameters of the non-linear system. Thus, in a circular accelerator, it is possible to ''plausibly'' describe non-linear motion of a particle using a finite number of harmonics of sextupole perturbation.

The harmonics depend on the values of forces and the arrangement of sextupole magnets, thus, it is possible to choose sextupole magnets in such a way that their non-linear influence does not lead to reduction of the area of stability.

HT is realized with using the method of canonical integration of any order [2]. This method of integration of the Hamilton equations of motion is based on replacement of a continuous Hamilton system with an implicit canonical mapping. The implicit mapping is realized by the ''predictor–corrector'' scheme, in which explicit Lee mapping is used as the predictor.

## 2. Theory

The sextupole perturbation Hamiltonian in the ''action-phase'' variables takes the following form:

$$H = \nu_x J_x + \nu_z J_z + (2J_x)^{3/2} \sum_n [3A_{1n}\ \cos(\phi_x - n\theta) + A_{3n}\ \cos(3\phi_x - n\theta)] - 3(2J_x)^{1/2}(2J_z) \times \sum_n [2B_{1n}\ \cos(\phi_x - n\theta) + B_{+n}\ \cos(\ldots) + \cdots] \quad (1)$$

*Corresponding author.
*E-mail address:* kvardakoff@ngs.ru (V. Kvardakov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.074

where the amplitudes of the corresponding harmonics are determined as follows:

$$A_{jn} = \frac{1}{48\pi}\sum_i \beta_{xi}^{3/2}(\theta_i)k_s(\theta_i)\cos(j\psi_{xi} - \nu\theta_i + n\theta_i), \quad j = 1,3$$

$$B_{\pm n} = \frac{1}{48\pi}\sum_i \beta_{xi}^{1/2}(\theta_i)\beta_{zi}(\theta_i)k_s(\theta_i) \times \cos(\psi_{\pm i} - \nu_\pm\theta_i + n\theta_i), \quad \psi_\pm = \psi_x \pm 2\psi_z$$

where $\beta(\theta_i)$ and $\psi(\theta_i)$ are the amplitude and phase betatron functions and $k_s(\theta_i) = (1/B\rho)(\partial^2 B_z/\partial x^2)$ is the strength of the sextupole lens. The integrals can be replaced with the sum in approximation of "thin" lenses, when the lens length is much less than that of betatron oscillations [$k_s(\theta_i) = k_s l$ is the integrated strength of the sextupole magnets].

To solve Hamiltonian Eq. (1), it would be better to rewrite it in the variables $(x, z, P_x, P_z)$ using the following transformation of coordinates:

$$u = \sqrt{2J_u}\,\cos\varphi_u, \quad P_u = -\sqrt{2J_u}\,\sin\varphi_u, \quad u = x, z.$$

Substitution of the new variable in the initial Hamiltonian Eq. (1) yields a new Hamiltonian, which now depends only on the coordinates, harmonic values and azimuthal coordinate. Therefore, it is possible to write the Hamiltonian in a more convenient form to mark out the factors dependent on the harmonics and azimuth for the coordinates $x, z, P_x, P_z$. In the new form, the Hamiltonian is

$$H = \frac{\nu_x}{2}(P_x^2 + x^2) + \frac{\nu_z}{2}(P_z^2 + z^2) + V(x, z, P_x, P_z, \theta)$$

where

$$V(x,z,P_x,P_z,\theta) = a(\theta)x^3 + b(\theta)P_x^3 + c(\theta)x^2P_x + d(\theta)xP_x^2 + e(\theta)xz^2 + f(\theta)xP_z^2 + g(\theta)z^2P_x + h(\theta)P_xP_z^2 + i(\theta)zP_xP_z + k(\theta)xzP_z. \quad (2)$$

The 10 factors in the Hamiltonian are written in the following form:

$$\begin{aligned} a(\theta) &= \sum_m (3A_{1m} + A_{3m})\cos m\theta \\ b(\theta) &= \sum_m (A_{3m} - 3A_{1m})\sin m\theta \\ c(\theta) &= -3\sum_m (A_{1m} + A_{3m})\sin m\theta \\ d(\theta) &= 3\sum_m (A_{1m} - A_{3m})\cos m\theta \\ e(\theta) &= -3\sum_m (2B_{1m} + B_{+m} + B_{-m})\cos m\theta, \quad \text{etc.} \end{aligned} \quad (3)$$

Such a form of the Hamiltonian is rather convenient, since it depends on a finite number of factors (10). The factors can be calculated subject to a large number of harmonics of the sextupole perturbation, or subject only to the basic resonant harmonics.

The equations of motion derived from Hamiltonian Eq. (2) are as follows:

$$\begin{aligned} \frac{dx}{d\theta} &= \nu_x P_x + 3b(\theta)P_x^2 + 2c(\theta)P_x x + d(\theta)x^2 + \cdots \\ \frac{dP_x}{d\theta} &= -\nu_x x - 3a(\theta)x^2 - c(\theta)P_x^2 - 2d(\theta)xP_x + \cdots \\ \frac{dz}{d\theta} &= \nu_z P_z + 2f(\theta)xP_z + 2h(\theta)P_xP_z + \cdots \\ \frac{dP_z}{d\theta} &= -\nu_z z - 2e(\theta)xz - 2g(\theta)P_x z + \cdots . \end{aligned} \quad (4)$$

## 3. Method of canonical integration of any order

The equations of motion [Eq. (4)] were solved numerically by the method of canonical integration of any order [3]. For realization of this method it is necessary to deal with implicit equations, solution of which requires one of the methods for minimization of a finite number of elements. Newton's method can be used here. However, such an approach leads to difficulties related to the fast convergence of consecutive approximations.

Application of the Lee transformation creates no problems related to unambiguity and convergence. Moreover, it is known that within one step of integration the values $(\tilde{x}, \tilde{u})$ obtained can coincide well with the actual values $(\bar{x}, \bar{u})$.

For this reason, the following scheme was used for effective integration of the equations of motion: the Lee prediction was taken as the initial approach (predictor) and further refinement was carried out with the help of the simple iteration $\vec{u}^{(k)} = F(x, u, \vec{u}^{(k-1)})$, $k = 1, 2\ldots$ (corrector). It is evident from practice that more than two refinements are seldom required.

We can use the Lee prediction of a lower order (1,2) for the initial approach and move in small steps, or use the Lee prediction of a higher order and move in larger steps.

In our program we used the Lee prediction of the first order as the initial approach, and the required accuracy was achieved via control of the current error.

## 4. Method application for an increase in dynamic aperture

Modeling using HT allows the easy determination of which of the (five) basic resonant harmonics plays a basic role, thereby optimizing the non-linear system.

An increase in the stable motion area via setting the "required" basic resonant harmonics to zero works well for structures of the double bend achromatic (DBA) type.

Fig. 1a and b shows the horizontal phase planes $(X-P_x)$ designed for the model under research, with application of the matrix formalism in approximation of thin lenses and HT subject to 100 harmonics.

From Fig. 1, it is obvious that the results of tracking for both methods coincide with very good accuracy. The next step is to understand which of the five basic harmonics

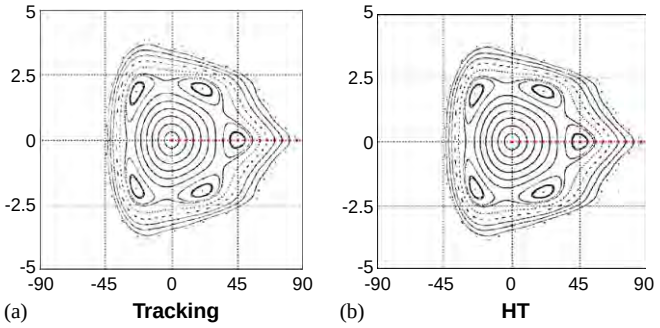



Fig. 1. Phase space trajectory for DBA lattice: (a) results obtained by matrix tracking and (b) results obtained by HT.

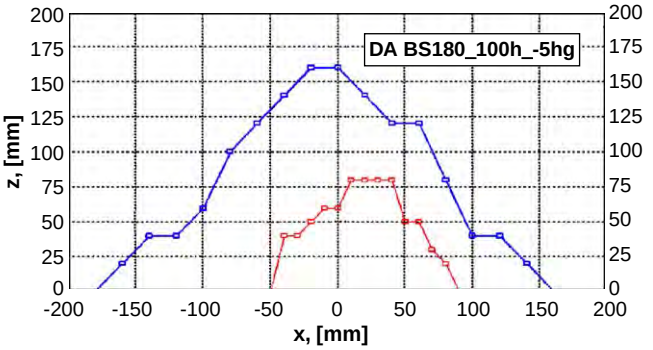



Fig. 2. Dynamic aperture for a DBA lattice. Tracking (smaller) and harmonic tracking (larger) with 200 harmonics, in which two resonant harmonics are "vanished".

represents the main contribution to the perturbation and how plausible the description of the non-linear system under research will be if only these harmonics are used.

Fig. 2 shows how the dynamic aperture varies according to the theory if only two basic resonant harmonics of perturbation are set to zero.

From consideration of the dynamic aperture presented in Fig. 2, it becomes obvious that setting the basic resonant harmonics to zero leads to an increase in the dynamic aperture size.

In addition, this example shows the utility of the HT method. The technique of setting the basic resonant harmonics to zero is applicable for circular accelerators with strong symmetry in which the DBA magnetic structure is represented by several identical cells (superperiods). In modern circular accelerators, the installations that satisfy such a structure are specialized sources of synchrotron radiation, in which the area of steady movement can be increased essentially by setting the basic resonant harmonics to zero.

## 5. Comparison with experiment

Fig. 3 shows experimental results for two-dimensional coupled sextupole resonances of $\nu_x + 2\nu_z = 24$ and $\nu_x - 2\nu_z = -7$. The experiment was carried out on the VEPP-4 M collider [3,4].

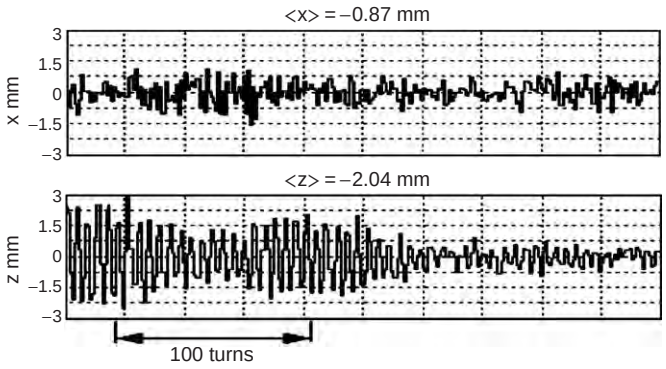



Fig. 3. Experimental results: difference resonance $\nu_x - 2\nu_z = -7$ ($\nu_x = 8.5514$, $\nu_z = 7.7773$).

The measurements were based on fast pulse (one turn) excitation of coherent oscillations of the electron beam [5]. The measurements were carried out at energy $E = 1.85$ GeV.

The typical beam current in the measurements was 1–3 mA. In this case, the rms coordinate resolution of the BPM is approximately 70 μm. Phase trajectories were restored by the position of the center of the beam from one pickup [5].

Figs. 3 and 4 show comparative results obtained experimentally and via modeling for the case of movement close to the difference resonance.

The results of measurements on the VEPP-4M collider (Fig. 3) are presented using the coordinates beam position–revolution number. Fig. 3 shows the results of the measurements (dependence of indications of the beam position pick-up on the revolution number) close to the difference resonance. In addition, Fig. 4 shows results of the modeling (including the phase trajectories), which allows estimation of the value of the basic harmonic as $B_{-,-7} \approx 0.04\,\mathrm{mm}^{-1/2}$, which coincides well with estimations made by other methods.

In Fig. 3 the fast decay of coherent betatron oscillations is evident, which for the VEPP-4 M storage ring has a large rate for the vertical coordinate. However, this is not important here, since we are interested in the first 200 turns, giving us the frequency and amplitude of the beat. Thus, omitting the decay of coherent betatron oscillations, the effect of which was not taken into account in the modeling, the concurrence of the experimental (Fig. 3) and modeling (Fig. 4) results becomes obvious.

## 6. Conclusions

A new method for research of non-linear motion in a circular accelerator was developed and realized. The method allows easy operation ("manual") of the harmonics of perturbation, and thus optimization of the parameters of the non-linear system. The utility of this method has been checked using the example of increasing the area of steady motion for some modern circular accelerators. Results of modeling by the HT method agree

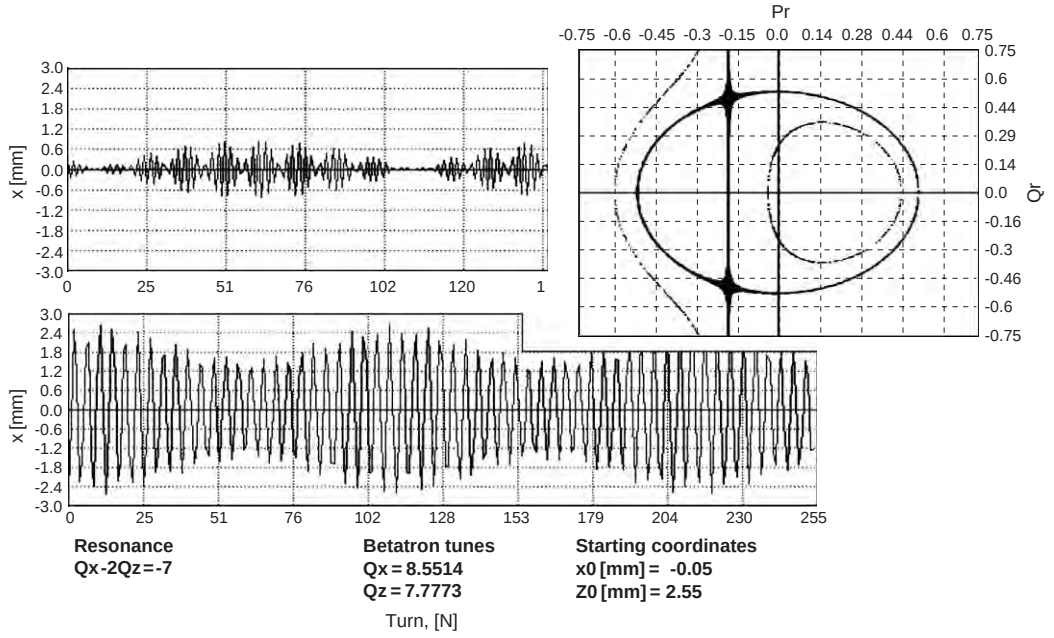



Fig. 4. Results of Ht using one resonance harmonic (difference resonance $v_x - 2v - 7$).

with the results obtained by traditional methods of tracking, as well as with experimental results.

## References

[1] R. Nagaoka, K. Yoshida, M. Hara, Nucl. Instr. and Meth. A 302 (1991) 9.

[2] V. Vecheslavov. The method of canonical integration of any order, Internal preprint 89-35, BINP, Novosibirsk, 1989.

[3] V. Kvardakov, E. Levichev, A. Naumenkov, P. Piminov, The experimental study of sextupole coupled resonances at VEPP-4M collider, in: Proceedings of the Third APAC, Korea, 2004.

[4] E. Levichev, V. Kiselev, A. Naumenkov, A dynamic aperture of VEPP-4M, in: Proceedings of the 2001 Particle Accelerator Conference (PAC2001), Chicago, IL, June 18–22, 2001, TPPH004, p. 1659.

[5] A. Kalinin, V. Smaluk, Turn-by-turn phase space diagram construction for nonlinear betatron oscillations, in: Proceedings of the Fourth European Workshop on Diagnostics for Particle Accelerators, Chester, 1999.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 40–43
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# A method to derive corpuscular–optics identities as a consequence of the static character of fields

A.A. Matyshev*

*St. Petersburg State Polytechnical University, 29 Politechnicheskaya St., St. Petersburg, 195251, Russia*

Available online 1 December 2005

**Abstract**

A method to derive identities in static corpuscular optics is described. The essence of the method involves consideration of the particle start time as a parameter. As an example, 12 identities have been derived for a single electrostatic lens in the asymptotic case.


*PACS:* 41.85

*Keywords:* Particle optics; Aberrations; Identities

## 1. Introduction

One of the central problems of modern corpuscular optics involves estimation of the accuracy of numerical calculations. One common procedure consists of calculating the electromagnetic field distribution, the charged particle trajectory, the cardinal elements and the aberration.

Unfortunately, there are not enough cases for which particle trajectories may be obtained in analytical form for an electromagnetic field with a known electrode profile. As a consequence, there is no unified test to determine the accuracy for all existing tracking methods for numerical calculations in corpuscular optics.

It is likely that for imaging systems such a test may be evolved from identities including the cardinal elements and aberrations. A group of identities exists as a consequence of the simplectic condition [1]. Such identities exist for any Hamiltonian system, including time-dependent Hamiltonian systems.

This paper describes a simple method for deriving identities for imaging corpuscular–optics systems. The method under consideration yields identities as a consequence of the static field character. Such identities exist for any static system, for which it must be noted that the system may be non-Hamiltonian.

The main point of the method proposed for deriving new identities in a wide range of cases is to regard the particle start time as a parameter. As in any static field, the particle trajectory is independent of such a parameter, and usually this parameter is not considered.

## 2. Particle start time as a parameter in static corpuscular optics

If we take into account the particle start time as a parameter and calculate any quantity in two ways, then it is possible to determine some identities.

This method is illustrated for the case of a single electrostatic lens with a plane of symmetry when an object and its image are outside the lens (i.e. an asymptotic case for both relativistic and non-relativistic particles).

In Fig. 1, a single two-dimensional or axisymmetrical lens is represented with a more general quasi-stationary electric field. A field potential in the symmetry $x$–$z$ plane may be written as $U(x,z,t) = f(t)u(x,z)$, where $u(x,z)$ is an even function of $x$. In the quasi-stationary lens, particle trajectories actually depend on the start time. If the lens is static [i.e.$f(t) = \text{const}$], then trajectories are independent of the start time.

*Tel.: +7 812 552 75 16; fax: +7 812 552 95 16.
*E-mail address:* matyshev@mail.spbstu.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.010

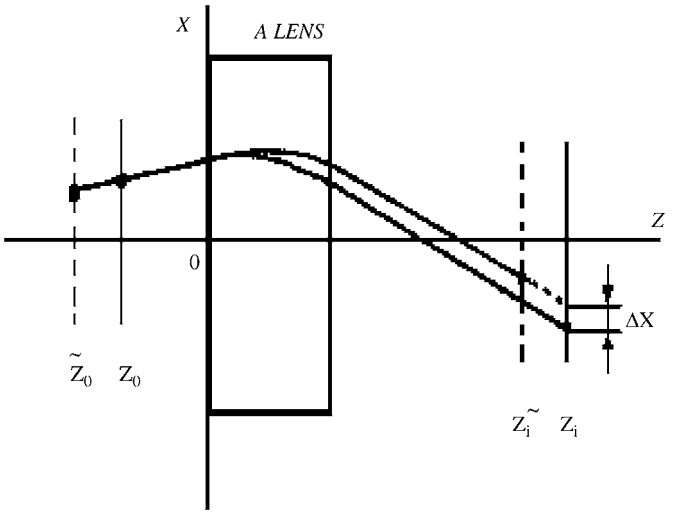



Fig. 1. Appearance of space-time aberration (STA) in a field-varying lens: $\Delta x$ is blurring of the image in a Gaussian plane due to the pulse duration $\Delta t$.

If we choose a coordinate $x$ of the point of intersection of the particle trajectory and the image plane and calculate it twice for two different particle start times, then this allows the determination of a countable set of identities. This is the essence of the proposed method. In the example under consideration, this is equivalent to calculation of the STA.

## 3. Calculation of the STA up to second-order pulse duration

Let us calculate the STA $\Delta x$ in a Gaussian plane for a more general quasi-stationary lens and equate it to zero for static lenses.

We briefly mention detailed studies of STA in Refs. [2–4]. Particles starting from an object point $(z_0, x_0)$ not at moment $t = 0$ but later at a moment $t = \Delta t > 0$ may be absolutely correctly considered as particles that start at the moment $t = 0$ from a moved point $(\tilde{z}_0, \tilde{x}_0)$. Thus, these coordinates may be written in the form:

$$\tilde{z}_0 = z_0 - \Delta d \tag{1}$$

$$\tilde{x}_0 = x_0 - \Delta d\, x'_0, z \tag{2}$$

where the quantity $\Delta d = v \cos\alpha\, \Delta t$ has the dimension of length, $v$ is the initial particle velocity, $x'_0 = \tan\alpha$, and $\alpha$ is the angle between the initial velocity vector and the $z$-axis.

Now it is only necessary to carry out some technical calculations. In fact, the object plane $z = z_0$ has a conjugated plane $z = z_i$ in the image space. Then the object plane $z = \tilde{z}_0$ should have a conjugated plane $z = \tilde{z}_i$ in the image space, as shown in Fig. 1. The point of intersection of the particle trajectory and the image plane up to the third order may be written in the form:

$$x = Mx_0 + Ax_0'^3 + Bx_0x_0'^2 + Cx_0^2x'_0 + Dx_0^3 \tag{3}$$

where the coefficients correspond to lens lateral magnification ($M$), spherical aberration ($A$), coma ($B$), distortion ($D$), astigmatism and field curvature ($C$).

All the coefficients in the last formula are functions of $z_0$ and, in the general case, of particle start time $t_0$ (and also particle energy and lens parameters) so as for delayed particles, starting from the point $(\tilde{z}_0, \tilde{x}_0)$, an analogous expression must be written in the same form:

$$\tilde{x} = \tilde{M}\tilde{x}_0 + \tilde{A}x_0'^3 + \tilde{B}\tilde{x}_0x_0'^2 + \tilde{C}\tilde{x}_0^2x'_0 + \tilde{D}\tilde{x}_0^3. \tag{4}$$

Each coefficient taken separately from Eq. (4) may be expanded into a Taylor series correct up to the second order. For example, after taking into account Eq. (1), the lateral magnification becomes

$$\tilde{M}(z_0 - \Delta d) = M - M_z\,\Delta d + 0.5\,M_{zz}(\Delta d)^2 + O[(\Delta d)^3] \tag{5}$$

where the index $z$ denotes the derivative with respect to $z_0$: $M_z = \partial M/\partial z_0$, $M_{zz} = \partial^2 M/\partial z_0^2$, and so on. The values of all functions on the right-hand side of Eq. (5) should be calculated for the definite point $(z_0, 0)$.

In addition, the spherical aberration coefficient becomes

$$\tilde{A}(z_0 - \Delta d) = A - A_z\,\Delta d + 0.5\,A_{zz}(\Delta d)^2 + O[(\Delta d)^3]. \tag{6}$$

Exactly the same equation is valid for the remaining aberration coefficients.

As the image planes are in the field-free region, the particle trajectories in the image space are straight lines $X(Z)$. Consequently, the expression for trajectory of the delayed particle crossing the corresponding image plane at point $(\tilde{z}_i, \tilde{x})$, is as follows:

$$X(Z) = \tilde{x} + \tilde{x}'(Z - \tilde{z}_i). \tag{7}$$

This straight line crosses the plane $z = z_i$ at the point $X(z_i)$, so STA may be written in the form:

$$\Delta x = X(z_i) - x = \tilde{x} + \tilde{x}'(z_i - \tilde{z}_i) - x. \tag{8}$$

To complete the STA calculation, it is necessary to determine the displacement of the image plane when the object plane moves. As the $z_i$ coordinate of the image plane is a function of the $z_0$ coordinate of the object plane, it is possible to expand the quantity $\tilde{z}_i$ into a Taylor series correct up to the second order, analogous to the Taylor series in Eqs. (5) and (6):

$$\tilde{z}_i(z_0 - \Delta d) = z_i - \frac{\partial z_i}{\partial z_0}\,\Delta d + 0.5\,\frac{\partial^2 z_i}{\partial z_0^2}\,(\Delta d)^2 + O[(\Delta d)^3]. \tag{9}$$

If we take into account the definition of longitudinal lens magnification $M_l$,

$$M_l = \frac{\partial z_i}{\partial z_0} \tag{10}$$

then Eq. (9) is easily transformed to the form:

$$z_i - \tilde{z}_i = M_1 \Delta d - 0.5 \frac{\partial M_1}{\partial z_0} (\Delta d)^2 + O[(\Delta d)^3] = M_1 \Delta d - 0.5 M_{1z} (\Delta d)^2 + O[(\Delta d)^3] \quad (11)$$

where $M_{1z} = \partial M_1 / \partial z_0$.

Substitution of the last equation into Eq. (8) yields the following expression for STA up to the second order in $\Delta d$:

$$\Delta x = X(z_i) - x = \tilde{x} - x + \tilde{x}'(z_i - \tilde{z}_i) = \tilde{x} - x + \tilde{x}'[M_1 \Delta d - 0.5 M_{1z} (\Delta d)^2]. \quad (12)$$

The STA coefficients for an axial point were calculated in Ref. [3], but approximation of the trajectory slope $\tilde{x}'$ was only up to the first order. To take into account this quantity in an approximation up to the third order, then it must be written in the form:

$$x' = \gamma x_0 + \Gamma x_0' + E x_0'^3 + F x_0 x_0'^2 + G x_0^2 x_0' + H x_0^3 \quad (13)$$

where $\Gamma$ is the angular magnification, $\gamma = g'(z_i)$ and $g(z)$ is one of the independent linear solutions of the paraxial lens equation for which $g(z_0) = 1$, $g'(z_0) = 0$.

It should be noted that the second independent linear solution of the paraxial lens equation is usually determined by the initial data $h(z_0) = 0$, $h'(z_0) = 1$. It is well known that in a conjugated plane $h(z_i) = 0$, $g(z_i) = M$ and $h'(z_i) = \Gamma$. However, the value of the quantity $g'(z_i)$, marked in this paper by the special symbol $\gamma$, is almost unknown as a function of the other lens parameters and is usually not examined in textbooks on particle optics. Furthermore, it is clear how it is possible to obtain this quantity as a function of the other lens parameters.

The coefficients $E, F, G, H$ are also not considered in textbooks; although knowledge of image blurring is of interest, aberrations of the trajectory slope are usually not of interest. For the coefficients $E, F, G$ and $H$ in the static case, aberration integrals analogous to the standard aberration integrals for the coefficients $A, B, C$ and $D$ may be derived by any known method.

Now it is possible to write an equation for the quantity $\tilde{x}'$, as has been done for the quantity $\tilde{x}$:

$$\tilde{x}' = \tilde{\gamma} \tilde{x}_0 + \tilde{\Gamma} x_0' + \tilde{E} x_0'^3 + \tilde{F} \tilde{x}_0 x_0'^2 + \tilde{G} \tilde{x}_0^2 x_0' + \tilde{H} \tilde{x}_0^3 + O[(\Delta d)^3]. \quad (14)$$

All the coefficients with tilde in the last formula may be expanded in turn correct up to the second order into Taylor series that are absolutely analogous to the Taylor series in Eqs. (5) and (6). For example, for coefficient $\tilde{\gamma}$ we have

$$\tilde{\gamma}(z_0 - \Delta d) = \gamma - \gamma_z \Delta d + 0.5 \gamma_{zz} (\Delta d)^2 + O[(\Delta d)^3]. \quad (15)$$

Finally, inserting into Eq. (12) for STA all the necessary items and expanding it into a series up to the second order in $\Delta d$, we obtain

$$\begin{aligned}
\Delta x = \Delta d [ & x_0'(-M + \Gamma M_1) + x_0(-M_z + \gamma M_1) \\
& + x_0'^3(-A_z - B + E M_1) \\
& + x_0 x_0'^2(-B_z - 2C + F M_1) \\
& + x_0^2 x_0'(-C_z - 3D + G M_1) \\
& + x_0^3(-D_z + H M_1)] \\
& + (\Delta d)^2 \left[ x_0' \left( M_z - \frac{M_{1z}}{2} \Gamma - M_1 \gamma - M_1 \Gamma_z \right) \right. \\
& + x_0 \left( \frac{M_{zz}}{2} - \frac{M_{1z}}{2} \gamma - M_1 \gamma_z \right) \\
& + x_0'^3 \left( \frac{A_{zz}}{2} + B_z + C - \frac{M_{1z}}{2} E \right. \\
& \left. -3 M_1 E^2 E_z - M_1 F \right) \\
& + x_0 x_0'^2 \left( \frac{B_{zz}}{2} + 2C_z + 3D - \frac{M_{1z}}{2} F \right. \\
& \left. -2 M_1 F F_z - 2 M_1 G \right) \\
& + x_0^2 x_0' \left( \frac{C_{zz}}{2} + 3D_z - \frac{M_{1z}}{2} G - M_1 G_z - 3 M_1 H \right) \\
& \left. + x_0^3 \left( \frac{D_{zz}}{2} - \frac{M_{1z}}{2} H - M_1 H_z \right) \right]. \quad (16)
\end{aligned}$$

It is obvious that in the general case the series for the STA should contain an infinite sequence of terms. This series is a power series in powers of $\Delta d$, and the coefficients of the series are in their turn a series of all odd combinations of $x_0^m (x_0')^n$, which is typical for geometrical aberrations in the static case.

For static lenses, the STA is identically equal to zero because each of the terms of the series in Eq. (16) is identically equal to zero as well. For expansion up to the second order in $\Delta d$ and up to the third order in $x_0, x_0'$, this gives 12 identities.

## 4. Analysis of the identities

Let us consider the first two identities. First we have

$$M = \Gamma M_1. \quad (17)$$

Recall that the Helmholtz–Lagrange invariant $M\Gamma = p_0/p_i$ exists, where $p_0$ and $p_i$ are the initial particle momentum in the object plane and final particle momentum in the image plane, respectively. Then the angular magnification may be excluded from identity (17). Such a procedure yields a connection between the lateral and longitudinal lens magnifications:

$$M_1 = \frac{M}{\Gamma} = \frac{p_i}{p_0} M^2. \quad (18)$$

The last has long been well known in both classical light optics, where it is called the Maxwell formula [5], and particle optics [6].

In contrast to the first identity, the second identity is much less well known. It allows expression of the quantity $\gamma$ through angular magnification of a lens. In fact, we have

$$\gamma = g(z_i) = \frac{M_z}{M_1} = \frac{p_0}{p_i}\frac{M_z}{M^2} = -\frac{\partial \Gamma}{\partial z_0} = -\Gamma_z. \tag{19}$$

It must be emphasised that both identities connect the first-order lens parameters and may be useful for checking existing numerical results in the paraxial approximation.

When the asymptotic case is considered (i.e. an object and its image are in drift space) it is obvious that the value of $g(z_i)$ is independent of object position $z_0$, so that $\gamma = \text{const}(z_0)$. This gives the possibility of treating identity (19) as a differential equation for the quantity $\Gamma(z_0)$ with the trivially obtained solution

$$\Gamma = -\gamma z_0 + \text{const} = a z_0 + b \tag{20}$$

where the coefficients $a$ and $b$ are any real numbers.

Furthermore, according to the Helmholtz–Lagrange invariant the lateral magnification can be written in the form

$$M = \frac{p_0}{p_i}\frac{1}{(a z_0 + b)}. \tag{21}$$

If we recall now that the asymptotic aberration coefficients are polynomials with respect to inverse lateral magnification [7], then it is possible to state that the asymptotic aberration coefficients are polynomials of the same degree with respect to object position $z_0$.

Further analysis shows that among the remaining 10 identities, two link the magnifications of a lens as follows:

$$M_z - \frac{M_{1z}}{2}\Gamma - M_1\gamma - M_1\Gamma_z = 0 \tag{22}$$

and

$$\frac{M_{zz}}{2} - \frac{M_{1z}}{2}\gamma - M_1\gamma_z = 0. \tag{23}$$

It is easy to show that identity (22) is a consequence of identity (19) and identity (23) yields the equation $\Gamma_{zz} = 0$, which has function (20) as the solution, i.e. yields $\Gamma$ as a linear function of $z_0$ without additional argument.

Finally, the remaining eight identities link the third-order aberration coefficients. Apparently, it was not previously known that these identities are a consequence of the static character of fields. These identities consist of lens magnifications, position and slope aberration coefficients. They are

$$-A_z - B + E M_1 = 0 \tag{24}$$

$$-B_z - 2C + F M_1 = 0 \tag{25}$$

$$-C_z - 3D + G M_1 = 0 \tag{26}$$

$$-D_z + H M_1 = 0 \tag{27}$$

$$\frac{A_{zz}}{2} + B_z + C - \frac{M_{1z}}{2}E - 3M_1E^2E_z - M_1F = 0 \tag{28}$$

$$\frac{B_{zz}}{2} + 2C_z + 3D - \frac{M_{1z}}{2}F - 2M_1FF_z - 2M_1G = 0 \tag{29}$$

$$\frac{C_{zz}}{2} + 3D_z - \frac{M_{1z}}{2}G - M_1G_z - 3M_1H = 0 \tag{30}$$

$$\frac{D_{zz}}{2} - \frac{M_{1z}}{2}H - M_1H_z = 0. \tag{31}$$

It may be shown that some slope aberration coefficients may be expressed via position aberration coefficients, but this is outside the scope of the present paper. For example, it may be shown [8] that identity (24) in real cases allows expression of the third-order lens coma via the third-order spherical aberration, its first derivative with respect to $z_0$ and a linear lens magnification.

## 5. Summary

A new method for deriving identities in static corpuscular optics was proposed and illustrated.

A total of 12 identities were derived for imaging systems, some of which have seemingly been derived for the first time by the method proposed. Apart from clarifying a difficult problem on links between different aberration coefficients, the identities may form the basis for the creation of universal tests for checking numerical calculations in corpuscular optics, which have been lacking until now.

Analogous identities may be obtained for real aberration coefficients for static magnetic or combined lenses. If we take into consideration the fifth-order aberration coefficients, then new identities should be obtained.

## References

[1] H. Wollnik, Optics of Charged Particles, Academic Press, Orlando, FL, 1987.
[2] A.A. Matyshev, Isotrajectory Corpuscular Optics, Nauka, St. Petersburg, 2000 (in Russian).
[3] A.A. Matyshev, Proc. SPIE 4187 (2000) 42.
[4] A.A. Matyshev, On a method to derive identities in static corpuscular optics, Abstracts of the CPO-6 Conference, Maryland, 2002.
[5] M. Born, E. Wolf, Principles of Optics, Pergamon Press, oxford, 1964.
[6] P.W. Hawkes, E. Casper, Principles of Electron Optics, vol. 1, Academic Press, London, 1989.
[7] M. Szilagyi, Electron and Ion Optics, Plenum Press, New York, 1988.
[8] A.A. Matyshev, On differential relationship between aberration coefficients of static lenses, Abstracts of the CPO-6 Conference, Maryland, 2002.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 44–46
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Application of Monte-Carlo method for design and optimization of beam lines

G.A. Riabov*, V.G. Riabov, M.G. Tverskoy

*Petersburg Nuclear Physics Institute, Orlova Roscha, Gatchina, Leningrad district, 188350, Russia*

Available online 21 November 2005

**Abstract**

Monte-Carlo algorithms developed in PNPI RAS for calculation and optimization of particle beam are presented. We discuss the MEZON code for simulation of the particle behavior in the beam line and the OPTIMUM code for optimization of the beam parameters. Combination of Monte-Carlo GEANT3 calculations with MEZON code provides the very effective tool for designing the beam lines with absorbers.


*PACS:* 29.20.Mr; 29.27.Eg; 29.27.Eh

*Keywords:* Monte-Carlo method; Beam calculation and optimization

## 1. Introduction

Many computer codes were developed for particle beam design. However, several problems appeared during the stage of R&D of beam lines for 1 GeV synchrocyclotron and required further development of existing methods of calculations and introduction of new approaches. This article is concerned with application of Monte-Carlo methods for the design and optimization of the beam lines.

## 2. Simulation of the primary and secondary beams

In most of existing codes, the envelopes of the primary beam are calculated with the use of transfer matrix and phase ellipse in "hard edge" approximation. For many applications, this approach is inadequate. Calculation of secondary beams is performed with the use of beam line acceptance under the assumption of uniform distribution of particles in phase space and it is not physically justified. The deficiency of existing codes is especially evident in calculation of beams of the third generation, for example, muon beams produced as a result of decay on the flight of π-mesons. In this case, local source of muons does not exist and they can be produced at any point along the beam line. To reduce the number of simplifying assumptions and to get more detailed information on beam behavior, it was decided to use Monte-Carlo method widely used for simulation of passage of neutrons through the matter, simulation of particle detectors at modern colliers, etc. In beam optics, Monte-Carlo method consists of tracing and accumulation of large number of particle trajectories passing through the channel. Initial conditions of the particles in the target or in the source are random variables. On every step of trajectory integration, deviation of particle from the channel axis is compared to limiting transverse apertures. If particle falls out of the aperture, then it is considered to be lost. MESON code is based on such an approach [1]. Simulation method enables to get any characteristics of the beam such as absolute intensity, beam losses in particular elements of the transport line, spatial distribution and timing structure of the beam, momentum spectrum, etc. It also takes into account 3D geometry of the target, dependence of differential cross-section on energy and angle, and variation of density of primary beam. If the length of π-meson trajectory is equal

*Corresponding author. Tel.: +7 812 713 0857; fax: +7 812 713 0476.
*E-mail address:* riabov@mail.pnpi.spb.ru (G.A. Riabov).

0168-9002/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.011

to decay length, i.e. a random variable, then π-meson disappears giving rise to muon. Transformation of π-meson to muon changes momentum vector of particle, which is defined by decay angles ($\theta$ and $\varphi$) according to well-known formulas of kinematics. Decay angles were treated as random variables also.

To increase the number of simulated muon trajectories, parameters of one π-meson generated in the target are assigned to pion cluster. All particles from the same cluster decay in the same point, but parameters of the decays are simulated for every π-meson independently. Thus, the cluster makes up a set of muons with different momenta and decay angles.

Such an approach let us to simulate so-called "meson cloud" effect. It was shown that 30–50% of separated muons originate not from the beam of π-mesons captured in channel but from π-mesons decayed in the vicinity of target or on the input of the channel where pion flow is higher by factor of 1000. This effect, earlier observed in experiment, has been quantitatively described.

## 3. Optimization of beam lines

Besides simulation of beam parameters in the given beam line, there is a problem of optimization of beam parameters. This problem consists of selection of lay-out and fields of magnetic elements with the purpose of optimization of some beam parameters such as intensity, intensity density, momentum resolution, etc. The problem of optimization consists of two parts: (a) calculation of the beam parameters and (b) development of algorithm for search of optimum. Both these tasks can be solved with Monte-Carlo method. Algorithm of MESON code can be used for determination of beam parameters and, as a consequence, for calculation of any function of these parameters called quality function. Mathematically, the task of optimization can be reduced to search global maximum of the quality function depending on many arguments. Random search method with statistical prediction algorithm (so-called method of statistical gradient) was used for search of global maximum. On the basis of this method, the OPTIMUM code was developed [2].

Optimization algorithm is as follows [3]. Parameters of the beam line under optimization ($x_i$, where $i = 1 \ldots n$) are random variables which are normally distributed with mean $x_i^0$ and dispersion $\sigma_i$. Full set of the parameters constitute a random vector:

$$\vec{X} = (x_1, x_2, ..., x_i, ..., x_n),$$

where $i = 1,2,\ldots n$; $n$—number of parameters.

Series of such vectors $\vec{X}^k$ where $k = 1, 2, \ldots, N$ ($N \leqslant n$) is generated and corresponding quality functions $\Phi(\vec{X}^k)$ are calculated. Maximum value of the function $\Phi_m = \max \Phi(\vec{X}^k)$ achieved at $\vec{X}^m$ gives the first approximation of the global maximum. To improve the approximation, weight of every variable is calculated as

$$g_i = \frac{1}{N}\sum_{K=1}^{N}(\Delta\Phi^K \Delta X_i^K)/\sigma_i^2,$$

where $\Delta\Phi^K = \Phi^K - \Phi^0$ and $\Delta X_i^K = X_i^K - X_i^0$.

More convenient to use are the normalized weights: $y_i = g_i/\sum_{i=1}^{n} g_i^2$, where $n$ is the number of parameters. New approximation of the maximum is defined by $X_i^r = X_i^m + \varepsilon y_i$, where $\varepsilon$ is the scale factor being one of input parameters.

Then, new quality function $\Phi_r = \Phi(\vec{X}^r)$ is calculated and compared to $\Phi_m$. If $\Phi_m < \Phi_r$, then one more step is performed in the same direction. Otherwise, previous value is used as starting one in the next cycle of calculations. The variation range for random vectors, sigma and epsilon is narrowed in the process of optimization. The optimization is finished when variations of quality function and random vector become smaller than some given value.

The main features of this method are follows:

1. The random search provides a possibility to reach global maximum and the prediction algorithm improves the convergence process.
2. The number of calculations of quality function is smaller than in the regular optimization methods that may be important when number of variables is great.
3. Constrains on the search region can be easily included in the search algorithm.

Both programs MESON and OPTIMUM were successfully used for design of π- and muon beam lines at 1 GeV synchrocyclotron. Experimental measurements of beams of π-mesons and muons showed very good agreement with results of simulations.

## 4. Design of the variable energy proton beam obtained by the moderation of the primary beam in the degrader

Monte-Carlo method was also used to design the proton beam of variable energy (200–900 MeV) at 1 GeV synchrocyclotron. The energy variation of the beam is provided by the copper degrader. It is common practice to decrease energy of the proton beam by using a degrader. Application of Monte-Carlo method to this problem makes possible to find all parameters of the beam in detail. Simulation algorithm consists of two major stages. On the first stage, parameters of the beam after degrader are simulated with well-known GEANT3 code. On the second stage, the coordinates and momentum vector for every proton after degrader are used as input data for MESON and OPTIMUM codes. To get 300–500 particles on the output of the beam channel, one has to simulate $10^6$ particles after degrader that looks quite acceptable. Experimental studies showed good agreement with results of calculations of intensity, beam size, momentum spread

and magnetic field in the dipoles and quadruples of the beam line [5].

## 5. High-energy beam scattering on collimator walls

In beam optics, the collimators are used to restrict the beam size and its divergence. However, the collimator in high-energy optics is not identical to diaphragm in the light optics. Due to multiple scattering, significant part of the beam hitting the collimator walls can re-enter the collimator opening. This is a well-known effect quantitatively analyzed in Ref. [4]. Proposed approach let to study effect of multiple scattering on the collimator walls and its influence on the beam quality numerically in detail. In this approach, passage of protons through the system consisting of degrader and collimator with an opening of $4 \times 6\,\mathrm{cm}^2$ is simulated with GEANT3 code. Calculations showed that beam intensity after real collimator is 25% higher than that in the case of ideal collimator which absorbs all the particles falling out of aperture. Effective opening of the collimator can be considered to be larger than its real value by 0.5 cm. Due to multiple scattering in the collimator walls, the beam after collimator is enriched by particles with wide angular and energy distribution. Two major conclusions can be derived from these observations: collimator should not be used in front of registering apparatus; collimators of width less than 0.5–1 cm should not be used because of large contribution of particles scattered in the walls. Density of particles with large divergence angle is inversely proportional to the distance from collimator. So it is reasonable to collimate high-energy beam on the long distance from detector. Moreover, deflecting magnets installed in the beam line reduce density of particles with lower energies due to particle separation by momentum. Therefore, it is desirable to provide the magnetic analysis after collimator. In our case of the beam line having a length of $\sim$30 m and containing a deflecting magnet, the contribution of particles scattered in collimator does not exceed a few percent. Thus, the combination of GEANT3 code and beam line simulation codes MESON and OPTIMUM provides reliable tool for the designing beam lines with degrader.

## 6. Conclusions

Methods developed in PNPI to design various beam lines are presented in this article. Codes developed for tracing of particles through the magnet channels (MESON) and for optimization of beam parameters (OPTIMUM) are described. It is shown that GEANT3 code in combination with Monte-Carlo simulation of beam behavior in the beam line represents reliable tool for detailed calculation of beam lines with degraders. The developed methods have been tested at 1 GeV synchrocyclotron and the experimental studies showed a good agreement with results of calculation. The developed codes are multi-purpose ones and can be used for beam line design in different accelerators.

## References

[1] N.K. Abrossimov, V.A. Volchenkov, G.A. Riabov, Preprint LNPI-205, 1978 and Computer program for the calculation the primary and secondary beam by using Monte-Carlo method (MESON), in: Proceedings of Fourth National Particle Accelerator Conference, vol. 1, Nauka, Moskow, 1974, p. 258 (in Russian).
[2] N.K. Abrossimov, V.A. Volchenkov, G.A. Riabov, Proceedings of Sixth National Particle Accelerator Conference, vol. 2, Nauka, Moskva, 1978, p. 175 (in Russian).
[3] D.M. Kuzmenkov, V.I. Chernetskiy, Algorithms and Programs for the Random Search, Zinatne, Riga, 1969, p. 148 (in Russian).
[4] E.D. Courant, Rev. Sci. Instrum. 22 (12) (1951) 1033.
[5] N.K. Abrossimov, E.M. Ivanov, Yu.T. Mironov, G.A. Riabov, M.G. Tverskoy, Proton beam of variable energy—the new tool for investigation of radiation effects at PNPI synchrocyclotron, 2003 IEEE Radiation Effects Data Workshop Record, ISBN 0-7803-8127-0, 2003, pp. 145–148.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 47–49
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Mad-X a worthy successor for MAD8?

F. Schmidt*

*CERN, 1211 Geneva 23, Switzerland*

Available online 4 January 2006

**Abstract**

MAD-X is the successor at CERN to MAD8, a program for accelerator design and simulation with a long history. We had to give up on MAD8 since the code had evolved in such a way that the maintenance and upgrades had become increasingly difficult. In particular, the memory management with the Zebra banks seemed outdated. MAD-X was first released in June, 2002. It offers most of the MAD8 functionality, with some additions, corrections, and extensions. The most important of these extensions is the interface to PTC, the Polymorphic Tracking Code by E. Forest. The most relevant new features of MAD-X are: languages: C, Fortran77, and Fortran90; dynamic memory allocation: in the core program written in C; strictly modular organization, modified and extended input language; symplectic and arbitrary exact description of all elements via PTC; Taylor Maps and Normal Form techniques using PTC. It is also important to note that we have adopted a new style for program development and maintenance that relies heavily on active maintenance of modules by the users themselves. Proposals for collaboration as with KEK, Japan and GSI, Germany are therefore very welcome.


*PACS:* 07.05.T

*Keywords:* Accelerator optics design; Symplectic integrator; Normal Form analysis

## 1. Introduction

Originally it was planned to replace the MAD8 [1] code by a C++ version called MAD9 [2]. A project like this made sense at the time since MAD8 had become unmaintainable due to its complexity and its old fashioned structure all written in Fortran77. This grand new start was intended to make use of the powerful C++ object oriented tools for a transparent code structure, very general and flexible element and lattice description and making use of modern Normal Form techniques. In practice, it turned out that the MAD9 structure was so complex that bug fixes became very time consuming. Moreover there were serious timing issues that could not be resolved leading to very slow performance of the code compared to MAD8.

After many months of futile attempts to make use of MAD9 to do the much needed LHC design and simulation studies the LHC optics design team decided to look for an alternative. A crash program was started to rewrite MAD8 with a core program written in C with interfaces to independent modules written in either C or Fortran. In the summer of 2002 a first version of MAD-X [5] was released. In the meantime the code has matured to fully cover all issues concerning the LHC. The first section describes the present status of MAD-X proper.

Structurally and concerning the physics MAD-X offers nothing new with respect to MAD8. In particular Normal Form tools cannot be easily added. Moreover, we have suppressed all problematic modules like the thick lens tracking and modules needed for small machines, which had to be omitted due to lack of symplecticity. To overcome this deficiency we have linked MAD-X with the PTC [6] code described in the second section.

Lastly, the future plans for MAD-X are outlined.

## 2. Mad-X proper

The main new features of MAD-X can be summarized as follows:

- Core part in C with dynamic memory allocation.
- Truly independent modules with interfaces to the core for data access.

*Tel.: +41 022 7675245.
*E-mail address:* Frank.Schmidt@cern.ch.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.249

Table 1
Module keepers, people in bold are collaborators from outside CERN

| Module | Description | Keeper | Status |
|---|---|---|---|
| C6T | SixTrack [3] converter | F. Schmidt | OK |
| CORORBIT | Orbit correction | W. Herr | OK |
| DYNAP | Tracking postprocessing | F. Zimmermann | OK |
| EMIT | Emittance, radiation | R. Assmann | OK |
| ERROR | Error assignment | W. Herr | OK |
| IBS | Intra-beam scattering | D. Brandt | OK |
| MAKETHIN | Thinlens converter | H. Burkhardt | OK |
| MATCH | Matching procedures | O. Brüning | OK |
| PLOT | Plotting | T. d'Amico | OK |
| PTC_NORMAL | Normal Form coefficients | A. Bolshakov | Prepared |
| PTC_TRACK | Thicklens lattice tracking | A. Bolshakov | Prepared |
| SURVEY | Machine survey | A. Verdier | OK |
| SPACE-CHARGE | Space charge effect | O. Boine Frankenheim | Planned |
| SXF | Standard eXchange Format [4] | F. Pilat | OK |
| TOUSCHECK | Touscheck effect | C. Milardi | Prepared |
| THREADER | Beam threading | T. Risselada | OK |
| TWISS | Optics parameters | F. Schmidt | OK |
| THINTRACK | Thinlens lattice tracking | A. Verdier | OK |

- Making use of existing and debugged modules of MAD8 in Fortran77.
- Retaining only those features of MAD8 that are sound and concentrate on those modules that are needed for the LHC design.
- Use PTC for small machines and Normal Form calculations.
- Adding constructs to the input language like: WHILE and IF .. ELSE .. ENDIF .
- Powerful "macro" structures.
- Improved table handling.
- CVS version management.
- Linux, MAC and Windows versions available.

In the past CERN has assigned considerable manpower for writing accelerator design codes. For MAD-X it was decided to replace this dedicated person by a team of module keepers which still represents a substantial amount of manpower, but less work for the individual module keeper. In addition, there is one MAD-X custodian who is responsible for the overall functioning of the code. We expect that this team approach of module keepers, who are mostly themselves active users of their module, will lead to faster bug fixes and better, more usable modules.

In Table 1 a list is given of all relevant modules with name, purpose and module keeper (people in "bold" are collaborators external to CERN).

On the MAD-X website [7] one can find a "News" link which shows the changes between versions, the documentation based on "html" files and derived from them a "ps" and a "pdf" version, a "Keyword and Subject Index", a link to "Source and binaries" and one link to "Examples" for all modules and a facility to report bugs found in MAD-X by its users. Lastly, one can subscribe to a MAD-X newsgroup and a mailing list.

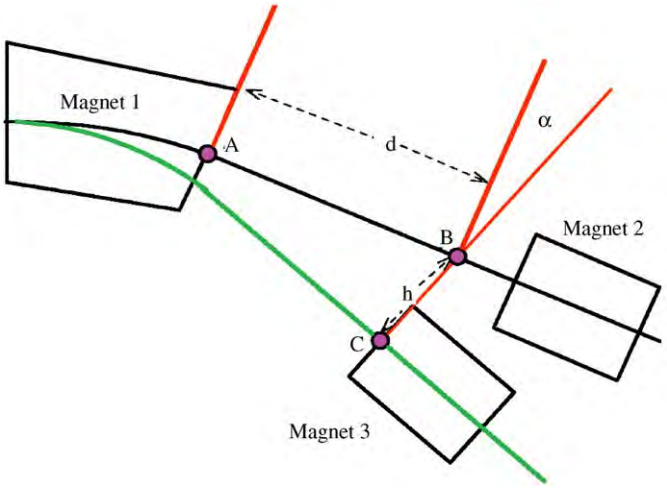



Fig. 1. Patching 2 beam lines (taken from Ref. [6]).

## 3. PTC

Forest's Polymorphic Tracking Code PTC [6] is a kick code or symplectic integrator and therefore ideally suited to describe all elements symplecticly and to arbitrary exactness. The degree of exactness is determined by the user and the speed of his computer. The code is written in an object oriented fashion using Fortran90. Therefore, it becomes much easier to describe arbitrarily complex accelerator structures. The other main advantage is that the code is inherently based on the map formalism [8] and provides MAD-X with all the associated tools.

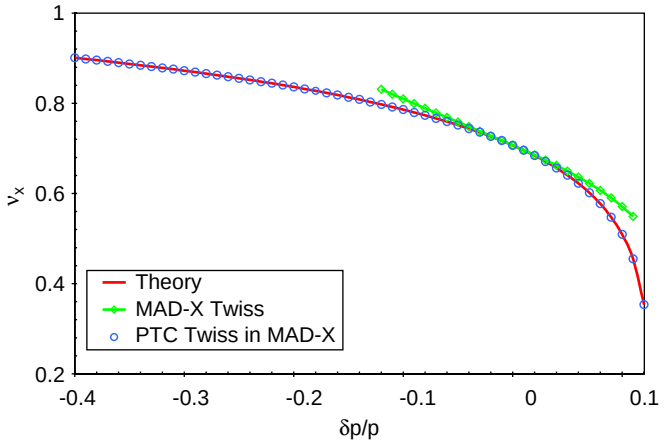



Fig. 2. Off-momentum tune of simple cyclotron.

One particular advantage is the fact that PTC allows to treat more complex beam line arrangements that can no longer be described by a simple sequence of elements. Fig. 1 shows an example in which the beam comes back to the same element but with a different energy. PTC handles all the coordinate system transformations by a built-in "patch"ing mechanism.

PTC allows to treat elements correctly even for very large momentum deviations. This becomes apparent in the simple cyclotron example that can be described analytically. The authors of Ref. [9] have demonstrated that MAD8 disagrees at large momentum deviation. In Fig. 2 one finds the same problem with MAD-X. However, using the very same MAD-X input file as an input for PTC, one can perfectly reproduce the analytical result.

## 4. Outlook

The latest MAD-X version V2.11 is quite mature for the LHC design work. There is, however, still a significant effort needed to guarantee full integrity of the code, in particular concerning memory management issues. The main upgrade path is to make full use of PTC which is now solidly linked to MAD-X. In particular, it is planned to use nonlinear coefficients calculated with PTC in the MAD-X matching.

## References

[1] H. Grote, F.C. Iselin, CERN/SL/90–13(AP) (Rev. 5).
[2] F.C. Iselin, J. Jowett, J. Pacin, A. Adelmann, MAD version 9, in: Proceedings of EPAC2000.
[3] F. Schmidt, CERN SL/94–56 (AP) (Rev.11.2001), 1994.
[4] H. Grote, et al., RHIC/AP/155.
[5] H. Grote, F. Schmidt, MAD-X—An Upgrade from MAD8, in: Proceedings of PAC2003.
[6] E. Forest, E. McIntosh, F. Schmidt, KEK Report 2002–3, CERN–SL–2002–044 (AP).
[7] H. Grote, F. Schmidt, MAD-X Webpage, ⟨http://www.cern.ch/Frank.Schmidt/Xdoc/mad-X.html⟩.
[8] M. Berz, É. Forest, J. Irwin, Part. Accel. 24 (1989).
[9] D. Trbojevic, et al., A comparison of several lattice tools for computation of orbit functions of an accelerator, in: Proceedings of PAC2003.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 50–53
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Electron gun simulation with CST PARTICLE STUDIO

Holger Spachmann, Ulrich Becker*

*CST GmbH, Bad Nauheimer Strasse 19, 64289 Darmstadt, Germany*

Available online 1 December 2005

**Abstract**

This paper introduces CST PARTICLE STUDIO™, a specialist tool for the fast and accurate design and analysis of 3D electron guns. The new software is based on the multi-purpose electromagnetic solvers of the CST STUDIO family and incorporates their powerful modelling capabilities as well as successful algorithms of the MAFIA-TS simulators. The underlying theory of the PBA® based field solvers and the used tracking technique is presented and some examples of application demonstrate the current status of the software.


*PACS:* 41.75.–i; 41.75.Fr; 41.85

*Keywords:* Gun-code; CST MICROWAVE STUDIO; CST EM STUDIO; FIT; PBA; Space-charge limited emission

## 1. Introduction

The present paper introduces CST PARTICLE STUDIO™ as an easy to use three-dimensional simulation tool for the fast and accurate design of electron guns. The paper is structured as follows: In Section 2, the theoretical background of the Finite Integration Technique (FIT) and a short description of the CST PARTICLE STUDIO™ (PS) program are presented. Section 3 describes the PS simulation procedure. Finally some typical simulation results are presented.

## 2. The CST PARTICLE STUDIO™ program

CST PARTICLE STUDIO™ (PS) is designed for gun-simulations and incorporates powerful electromagnetic field solvers for calculating the external fields, an efficient particle tracking algorithm and sophisticated emission models describing the extraction of particles from active surfaces into free space.

*The finite integration technique*: The build-in general-purpose electromagnetic solvers are based on the FIT, first proposed by T. Weiland in 1977 [1]. The FIT provides a general spatial discretization scheme usable for different electromagnetic applications of arbitrary geometry, e.g. static problems or calculations in frequency- and time-domain and is successfully applied in the well-known MAFIA simulation software [4]. Unlike most numerical methods, the FIT discretises the integral form of Maxwell's equations, rather than the differential one, on a pair of dual interlaced grids. Its formulation results in the so called Maxwell Grid Equations

$$\mathbf{C}\,\widehat{\mathbf{e}} = -\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{b}}, \quad \widetilde{\mathbf{C}}\,\widehat{\mathbf{h}} = \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{d}} + \widehat{\mathbf{j}} + \widehat{\mathbf{j}}_{\mathrm{ext}}, \tag{1a}$$

$$\widetilde{\mathbf{S}}\,\widehat{\mathbf{d}} = \mathbf{q}, \quad \mathbf{S}\,\widehat{\mathbf{b}} = \mathbf{0} \tag{1b}$$

and the constitutive material relations $\widehat{\mathbf{d}} = \mathbf{M}_\varepsilon\,\widehat{\mathbf{e}}$, $\widehat{\mathbf{b}} = \mathbf{M}_\mu\,\widehat{\mathbf{h}}$ and $\widehat{\mathbf{j}} = \mathbf{M}_\kappa\,\widehat{\mathbf{e}}$. Since particle guns are usually driven by static electric and magnetic fields, fast and efficient numerical solvers for the solution of the governing electrostatic and magnetostatic equations

$$\text{E-static: } \widetilde{\mathbf{S}}\mathbf{M}_\varepsilon\widetilde{\mathbf{S}}^{\mathrm{T}}\Phi = \mathbf{q}, \quad \text{M-static: } \mathbf{S}\mathbf{M}_\mu\mathbf{S}^{\mathrm{T}}\Phi_m = \mathbf{q}_m \tag{2}$$

with $\mathbf{q}_m = \mathbf{S}\mathbf{M}_\mu\widehat{\mathbf{h}}_i$, $\widehat{\mathbf{h}}_e = -\mathbf{S}^{\mathrm{T}}\mathbf{q}_m$, $\widetilde{\mathbf{C}}\widehat{\mathbf{h}}_i = \widehat{\mathbf{j}}$ and $\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_e + \widehat{\mathbf{h}}_i$ are implemented in CST PARTICLE STUDIO™.

*Corresponding author
*E-mail address:* becker@cst.com (U. Becker).
*URL:* http://www.cst.com.

0168-9002/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.075

*The emission model*: The emission model describes the particle generation process at the emitting surface. Up to now two types of emission models are implemented: Fixed emission and space-charge limited emission. The space-charge limited emission depends on the strength of the external field at the emitting surface: Particles are emitted as long as their reaction to the external fields decreases the extracting force to a level, where further particle extraction is suppressed. The amount of emitted particles can be determined in two ways: either by applying the Child–Lingmuir diode equation [2] or by emitting the whole surface-charge [3] of the electrode (see Fig. 1). Both schemes can optionally be applied. The Child–Lingmuir based model, depicted in Fig. 1a, evaluates the current at a certain distance $d$ of the emitting cathode with

$$J_s = \frac{4}{9}\varepsilon\sqrt{2\frac{q_e}{m_e}}\frac{(\Phi(d)-\Phi(0))^{3/2}}{d^2}. \quad (3)$$

In contrast the divergence based model (see Fig. 1b) measures the divergence of each emitting cell and assigns the cell-charge $Q = \widehat{d}_1 + \widehat{d}_2 + \widehat{d}_3 + \widehat{d}_4$ to the emitted particles.

*The particle tracking process*: The particles are pushed through the computational domain by interpolating the field-values to their location and calculating the EM-forces. The deposited space-charge in each cell and the particles' current is monitored and mapped to the grid. Hereby the timestep is adapted to the highest particle velocity ensuring a minimal number of pushing steps. Thus this adaptive timestepping ensures a fast and efficient tracking algorithm. Having finished the tracking of all particles through the computational domain, the resulting space-charge vector is used to modify the right-hand side of the electrostatic potential equation (Eq. (2)). Calculating the updated electric field yields the starting point for the next iteration step. The relative difference of the space-charge vector from one iteration step to the next serves as accuracy control mechanism.

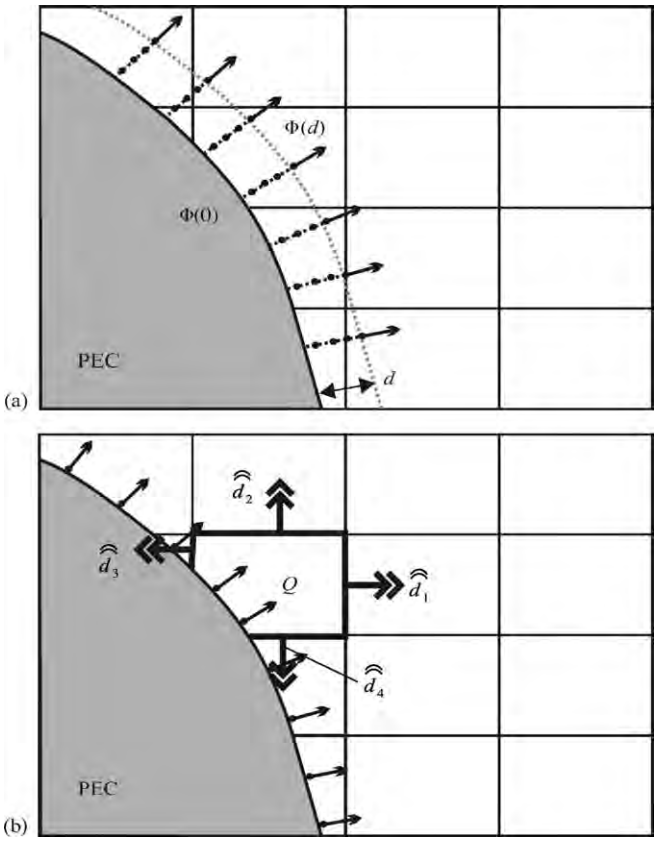



Fig. 1. Theoretical model of the Child–Lingmuir based emission model (a) and the divergence based model (b).

## 3. Simulation procedure

Once a physical model of the device under study has been established, it can be modelled in the 3D simulator in three steps: preprocessing, solving and postprocessing.

*Preprocessing*: The preprocessing of a structure (geometric description, material properties, boundary conditions, mesh generation) is made in an intuitive environment. The geometric modelling is based on the powerful ACIS kernel and allows definition of numerous basic shapes, Boolean operations (addition, subtraction, intersection, etc.) transformations (rotation, translation, mirroring) etc. However, the strength of PS's user interface goes beyond the basic ACIS functionality: it is the intuitive way of working and modelling, such as parameterization of geometrical details at any point of the construction process or the availability of comfortable pick features. The field sources can be either potentials, charges, magnetizations or currents through free-definable coils. The particle sources are defined by selecting one or more faces of a solid.

*Solving*: Two different field solvers are at hand for solving electrostatic and magnetostatic problems. At first, boundary conditions have to be defined whereby Neumann, Dirichlet and open boundary operators are supported. After the computation of the electric and magnetic fields, the tracking process can be started: either a simple tracking of particles through the pre-computed EM-fields is performable or an iterative solving process can be triggered—the gun-iteration.

*Postprocessing*: The solution of the gun-iteration are the emitted current, perveance and trajectory data as well as the electric and magnetic field distribution inside the gun. A large number of additional postprocessing facilities are available: visualization of electric and magnetic field in different representations (scalar, vector, component-wise), potential-distribution and current density. One- and two-dimensional integrals along free definable paths and over faces as well as electric and magnetic forces and torques can be evaluated and an animation of the flight path is possible showing the dynamic of the particles. Phase space diagrams of the particle data are in preparation.

*Special PS features*: A part of the PS features make it especially suited for the simulation of large configurations:

- *Geometry import*: Available import formats comprise classical 2D formats such as DXF and 3D generic

formats like SAT, STEP or STL. Also models of commercially available CAD packages as Pro/E, Catia 4&5, Autodesk's Inventor & AutoCAD and many more are importable.

- *Model export*: Available export formats are SAT, STL, IGES, STEP and MAFIA-DRC.
- *Full parameterization*: All dimensions can fully be parametrized.
- *Optimization capability*: PS includes a versatile tool for multi-dimensional-constrained optimization of the EM-simulation.
- *Result export feature*: Various results such as Energy or trajectory data can be exported in ASCII format.
- *Macro programming language*: A Visual Basic for Application (VBA) oriented macro language gives access to all PS features.
- *Integration into workflow*: PS can be both: OLE client or OLE server.

## 4. Results

The models exemplified here are the DC-part of a magnetron, a pierce-type gun and a simple multi-beam gun. For details of the pierce-gun please see Ref. [3]. The modelled geometries of the magnetron and the pierce-gun together with their particle trajectories are shown in Fig. 2. Both configurations were simulated with more than 10 000 particles and a relative accuracy of $10^{-3}$. The pierce-gun was calculated with $1.2 \times 10^6$ cells and the calculation was performed in less than an hour on a standard 2.4 GHz PC with 2 GB RAM.

PS is capable of defining several sources with different properties as well as one source consisting of several disjunct emitting surfaces. Thus multi-beams guns can easily be modelled. The triangularisation of the emitting surfaces and the trajectories of such a simple multi-beam gun are shown in Fig. 3.

## 5. Conclusions

PS offers a user-friendly environment for the design and optimization of particle guns. Due to the memory- and CPU-efficient FIT and PBA® method in conjunction with an adaptive timestepping of the tracking algorithm, fast and accurate 3D gun-simulations can be performed without the disadvantage of staircase approximations. This is especially important for such large simulations as multi-beam guns or complicated structures. The next development steps aim towards a more flexible and user-oriented

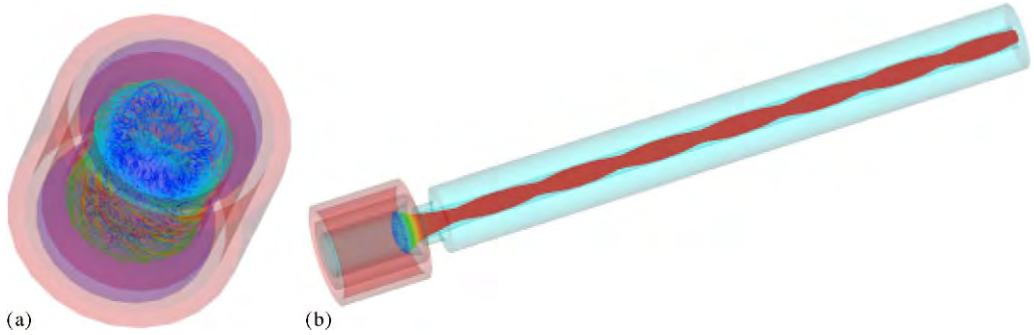



Fig. 2. The configuration of the DC part of the magnetron (a) and the pierce-type gun (b). The trajectories of the particles are displayed.

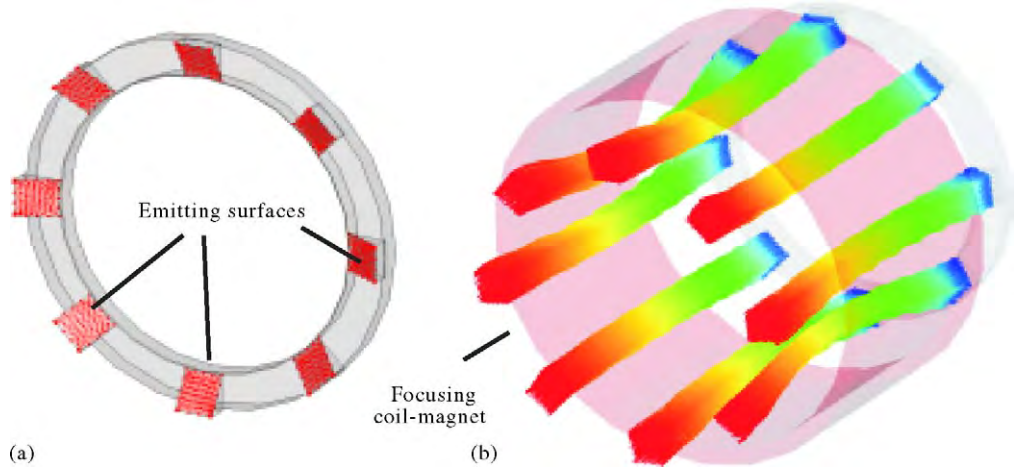



Fig. 3. (a) The discretization of the emitting surface of a simple multi-beam gun and (b) the trajectory of the particle-beam.

result handling and the implementation of tracking methods for time-dependent fields (Particle in Cell techniques).

## References

[1] T. Weiland, Electron. Commun. AE 31 (1977) 116.
[2] E. Gjonaj, T. Weiland, Conformal modeling of space-charge-limited emission from curved boundaries in particle simulations, in: Proceedings of the 2003 Particle Accelerator Conference (PAC 2003), Portland, Oregon, 2003, pp. 3563–3565.
[3] E. Gjonaj, T. Weiland, 3D-Modeling of Space-Charge-Limited Electron Emission. A Charge Conserving Algorithm, Proceedings of the 11th Biennial IEEE Conference on Electromagnetic Field Computation (CEFC, 2004), Seoul, June 2004, p. 10.
[4] CST GmbH: MAFIA EM Simulation Software, Ver. 4.x, www.cst.com, Computer Simulation Technology GmbH, Darmstadt, Germany, 2004.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 54–57
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Open boundaries for particle beams within fit-simulations

M.C. Balk*, R. Schuhmann, T. Weiland

*Technische Universität Darmstadt, Institut für Theorie Elektromagnetischer Felder, Schloßgartenstr. 8, 64289 Darmstadt, Germany*

Available online 1 December 2005

**Abstract**

A method is proposed to simulate open boundary conditions for charged particle beams with $v<c$ in time domain or frequency domain within the Finite Integration Technique (FIT).

Inside the calculation domain the moving charged particles are represented by a line current. Further, the simulated field components at the boundary of the calculation domain have to be modified for an undisturbed transmission of the space-charge field. This can be realised by a "scattered field" formulation.

The method is verified by several calculations.



*Keywords:* Open boundaries; Particle beams; Scattered field formulation; FIT

## 1. Introduction

In the Finite Integration Technique (FIT [1]) as well as in FDTD the state variables are located on a primary and a dual grid. For example the electric grid voltage $\hat{e}$ is located on edges and the magnetic grid flux $\hat{\hat{b}}$ on facets of the primary grid. In an analog manner, the magnetic voltage $\hat{h}$ and the electric flux $\hat{\hat{d}}$ are located on the dual grid. In contrast to FDTD, $\hat{e}$, $\hat{\hat{b}}$, $\hat{h}$ and $\hat{\hat{d}}$ are integrated quantities—for instance $\hat{e}=\int \vec{E}\,\mathrm{d}\vec{s}$ is integrated over normal edges and $\hat{\hat{b}}=\int \vec{B}\,\mathrm{d}\vec{A}$ over normal facets of the grid. With these variables Maxwell's equations can be discretised in integral form. This yields the following set of algebraic equations, where the analytical operators **curl** and **div** have been transformed to topological matrix operators $\mathbf{C},\mathbf{S}$ on the primary and $\widetilde{\mathbf{C}},\widetilde{\mathbf{S}}$ on the dual grid. Vectors $\hat{\mathbf{e}},\hat{\mathbf{h}},\hat{\hat{\mathbf{b}}}$ and $\hat{\hat{\mathbf{d}}}$ consist of all corresponding state variables.

$$\mathbf{C}\hat{\mathbf{e}}=-\frac{\mathrm{d}\hat{\hat{\mathbf{b}}}}{\mathrm{d}t} \tag{1}$$

$$\widetilde{\mathbf{C}}\hat{\mathbf{h}}=\frac{\mathrm{d}\hat{\hat{\mathbf{d}}}}{\mathrm{d}t}+\hat{\hat{\mathbf{j}}} \tag{2}$$

$$\mathbf{S}\hat{\hat{\mathbf{b}}}=0 \tag{3}$$

$$\widetilde{\mathbf{S}}\hat{\hat{\mathbf{d}}}=\mathbf{q} \tag{4}$$

$$\hat{\hat{\mathbf{d}}}=\mathbf{M}_{\varepsilon}\hat{\mathbf{e}} \tag{5}$$

$$\hat{\hat{\mathbf{b}}}=\mathbf{M}_{\mu}\hat{\mathbf{h}} \tag{6}$$

$$\hat{\hat{\mathbf{j}}}=\mathbf{M}_{\kappa}\hat{\mathbf{e}}. \tag{7}$$

Further on, Maxwell's equations are coupled by the material relations (5)–(7). Throughout the paper we use only Cartesian grids in combination with a conformal modelling technique [5]. For further information see [1].

## 2. Line current

For the simulation of particle beams an appropriate representation for moving charged particles is needed. In our case, where the particles move in one of the coordinate directions ($z$) and do not considerably change their electromagnetic characteristics while passing the structure to be simulated, a line current $\hat{\hat{j}}_{\mathrm{L}}$ along a grid line is a sufficient approach.

A simple idea of implementing such a line current would be imposing the grid currents by $\vec{J}=\rho\vec{v}$ or in discretised form $\hat{\hat{j}}_i=(q_i/\tilde{\Delta}z_i)v$, where $v$ is the particle velocity. However, the discrete form of the continuity

*Corresponding author.
*E-mail address:* balk@temf.tu-darmstadt.de (M.C. Balk).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.076

equation

$$\widetilde{\mathbf{S}}\hat{\tilde{\mathbf{j}}}_{\mathrm{L}} = -\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} \tag{8}$$

has to be satisfied to prevent the accumulation of unphysical charges in the grid. The simplest way to fulfill Eq. (8) in combination with $\hat{\tilde{j}}_i = (q_i/\tilde{\Delta}z_i)v$ is to match the timestep $\Delta t$ such that $mv\Delta t = \tilde{\Delta}z_i$ with $m \in \mathbf{N}$, where $\tilde{\Delta}z_i$ is the dual grid step in the direction of the beam propagation. As soon as the grid is non-equidistant, the continuity (8) is violated by $\hat{\tilde{j}}_i = (q_i/\tilde{\Delta}z_i)v$. Thus, a more general approach [2,3] is the direct implementation of Eq. (8)

$$\hat{\tilde{j}}^{(n)}_{z,i} - \hat{\tilde{j}}^{(n)}_{z,i-1} = -\frac{q_i^{(n+1/2)} - q_i^{(n-1/2)}}{\Delta t} \tag{9}$$

where $n$ is the time index according to $t = n\Delta t$.

Some special treatment at the boundaries is needed, which will not be given here.

## 3. Boundary conditions

At the boundaries of the finite computation domain the electromagnetic field has to be modified to model infinite space. For these purposes, we utilise the scattered field formulation (see Fig. 1), which is equivalent to surface currents $\hat{\tilde{j}}_{\mathrm{s}}$ and $\hat{\tilde{m}}_{\mathrm{s}}$ according to Huygen's principle.

The part of the calculation domain containing the line current indicated by $\hat{\tilde{j}}_{\mathrm{L},i}$ is regarded to be a total field region. The scattered field domain is depicted in grey. The total field can be calculated from $\hat{\mathbf{e}}^t = \hat{\mathbf{e}}^s + \hat{\mathbf{e}}^i$, where $s$ indicates scattered field and $t, i$ total field and incident field, respectively. The incident field is the eigenfield of the moving charges in an homogeneous environment. Therewith, one can modify the curl operations at the boundary. For instance, the local evaluation of (1) yields

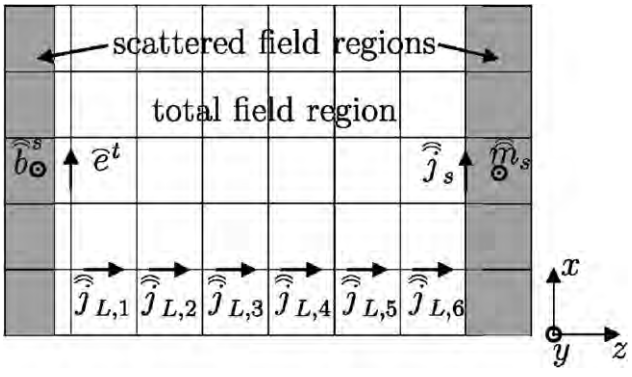



Fig. 1. Total/scattered field domain.

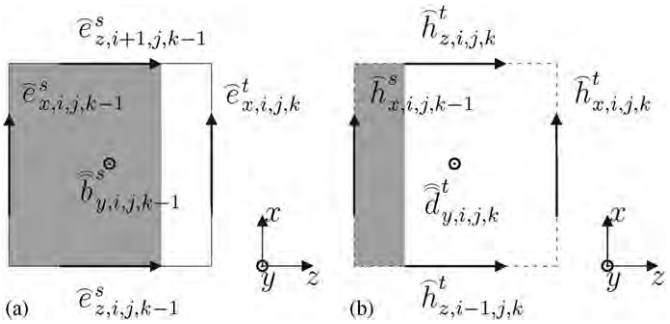



Fig. 2. Modification of electric (a) and magnetic curl (b) for scattered field formulation.

according to Fig. 2(a)

$$-\frac{\mathrm{d}\hat{\tilde{b}}^s_{y,i,j,k-1}}{\mathrm{d}t} = \underbrace{\hat{e}^t_{x,i,j,k} - \hat{e}^s_{z,i+1,j,k-1} - \hat{e}^s_{x,i,j,k-1} + \hat{e}^s_{z,i,j,k-1}}_{\text{contained in } \mathbf{c}\hat{e}} - \hat{e}^i_{x,i,j,k}. \tag{10}$$

Analogously, the result of the local evaluation of (2) based on Fig. 2(b) is

$$\frac{\mathrm{d}\hat{\tilde{d}}^t_{y,i,j,k}}{\mathrm{d}t} = \underbrace{\hat{h}^t_{x,i,j,k} - \hat{h}^t_{z,i,j,k} - \hat{h}^s_{x,i,j,k-1} + \hat{h}^t_{z,i-1,j,k}}_{\text{contained in } \widetilde{\mathbf{C}}\hat{\mathbf{h}}} - \hat{h}^i_{x,i,j,k-1}. \tag{11}$$

Depending on the realisation of the time derivative these boundary conditions can be implemented either in time domain or in frequency domain. The method of calculating the incident field needed in Eqs. (10) and (11) is described in the next section.

## 4. Numerical "incident field" solution

The incident field mentioned in the section before, is caused by moving charged particles in a homogeneous environment—thus no scattered field exist. Here and in the following, we assume particles moving in waveguides—more precisely, in beam tubes. An analytical field solution cannot be applied, since the method must fulfill some consistency laws in the discrete sense.

In time domain we use the explicit leapfrog scheme where the time derivative is replaced by central differences and electric and magnetic field are allocated on a staggered time grid.

The dynamic "incident field" calculation will be reduced to a static field calculation. Therefore, a new grid is introduced, where the transverse topology is identical to the base grid topology and the longitudinal grid step equals $\Delta z = v\Delta t$. Further, we transform the grid into the rest frame of the charge by a Lorentz transformation. In the rest frame of the charge the static field can be determined by solving a discrete Poisson's equation. The resulting electrostatic field has to be retransformed into the

laboratory frame to obtain the desired dynamic field. With this procedure, one gains the dynamic electric field $\widehat{e}^{i(n)}$ for one instance of time $n$. Since the longitudinal grid step of the calculation domain is related to the velocity of the particles by $\Delta z = v\Delta t$, the charge is moving from one cell to the next within one timestep. Thus, solving Poisson's equation with the charge in different cells yields the incident electric field at different timesteps.

For a full description of the incident field, we still need the magnetic field components. According to the leapfrog scheme, these are allocated at half timesteps $\widehat{h}^{i(n+1/2)}$. Also, electric and magnetic fields are allocated on a staggered spatial grid. This property can be utilised, when the dynamic magnetic field components have to be determined from the electrostatic fields in the rest frame. Further, we know the velocity of the moving charges. This is also the velocity of the corresponding electromagnetic field in a homogeneous environment. With this information, a direct relation between the retransformed electric and magnetic field can be found. If we reduce $\Delta t$ such that $\Delta z = (2 * m + 1)v\Delta t$ with $m \in \mathbf{N}^+$ the incident magnetic field can be evaluated by

$$\widehat{h}^{i(n+1/2)}_{x,i,j,k+1/2} = -\frac{\beta}{\mu c}\frac{\widetilde{\Delta}x}{\Delta y}\widehat{e}^{i(n-m)}_{y,i,j,k} \tag{12}$$

$$\widehat{h}^{i(n+1/2)}_{y,i,j,k+1/2} = \frac{\beta}{\mu c}\frac{\widetilde{\Delta}y}{\Delta x}\widehat{e}^{i(n-m)}_{x,i,j,k} \tag{13}$$

where again the electric grid voltages $\widehat{e}^{i(n-m)}_{y,i,j,k}$ and $\widehat{e}^{i(n-m)}_{x,i,j,k}$ have been obtained by solving Poisson's equation in the rest frame of the charges and a retransformation afterwards. The reduction of the timestep with $m = 0$ would violate the Courant criterium, so the smallest possible $m$ ensuring stability is chosen, which results in an only slightly reduced time step compared to the standard Courant timestep.

In the frequency domain the incident field is found by an additional field calculation within a homogeneous waveguide with open boundary conditions according to Ref. [4]. The calculation domain used in this reference calculation needs to consist of only a few grid cells in longitudinal direction when matching the obtained field with an appropriate phase difference.

## 5. Numerical results

The method has been implemented in a code using the same grid data as the commercial software tool CST MicrowaveStudio [5]. An example for validation of the proposed time domain code is a pickup/kicker electrode as used in the stochastic cooling system in the GSI storage ring. These electrodes serve for detecting and deflecting the beam. For further information see [6,7].

The pickup output $U_{\mathrm{PU}}$ when excited by a beam is related to the kicker input $U_{\mathrm{K}}$ when exciting the electrode structure [7]. This offers the opportunity to calculate the beam/electrode coupling in two different reciprocal ways.

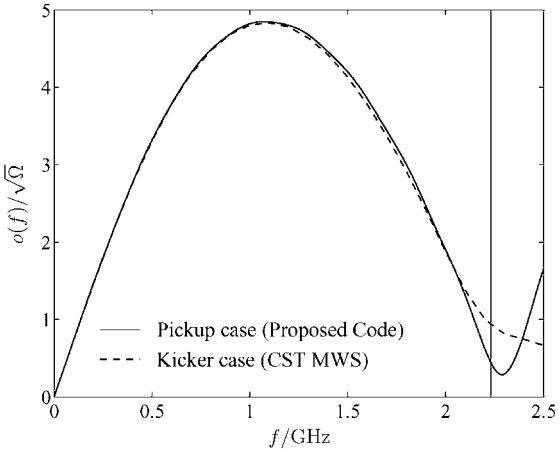



Fig. 3. Reciprocity calculation: Pickup case by proposed code, kicker case by CST MicrowaveStudio.

One way is the excitation of the electrode, which results in electromagnetic fields interacting with the beam. This interaction is evaluated using CST MicrowaveStudio [5]. The line current—introduced in Section 2—serves as excitation in the second way. The resulting output at the electrode is evaluated by means of the proposed code. The velocity of the beam is $v = 0.75c$, the number of grid points is $N_{\mathrm{P}} = 89,640$. Fig. 3 shows a good agreement between the two calculation approaches. The increase of the difference around 2.2 GHz is due to propagating and nearly propagating modes in the beam tube—cut-off frequency of the first mode is 2.23 GHz—where the reciprocity correlation is no longer valid.

## 6. Conclusion

A realisation of open boundary conditions for particle beams has been proposed. The velocity of the particles can be chosen arbitrarily. The method is based on the application of a scattered field formulation. Crucial to the method is a correct evaluation of the incident field for the given computational grid. This is implemented in time domain by a static field calculation in the charge's rest frame and a retransformation of the fields. In the frequency domain the known dependency of the electromagnetic fields is used for a separate calculation of the beam in a homogeneous waveguide. The proposed method and the corresponding code have been verified by several validation examples, showing a good accuracy compared to existing algorithms.

## References

[1] T. Weiland, Time Domain Electromagnetic Field Computation with Finite Difference Methods, Internat. J. Numer. Modelling 9 (1996) 295.
[2] M.B. Timm, Wake fields of short ultra-relativistic electron bunches, Der andere Verlag, Ph.D. Thesis, 2000.
[3] M. Dohlus, Numerical wakefield computation in time domain, unpublished, 1995.

[4] U.v. Rienen, T. Weiland, Impedance calculation above cut-off with URMEL-I, Proceedings of the First European Particle Accelerator Conference EPAC, Rome 1988, pp. 890–892.
[5] CST GmbH, Computer Simulation Technology, Bad Nauheimer Str. 19, 64289 Darmstadt, http://www.cst.com
[6] F. Nolden, et al., Stochastic Cooling at the ESR, Nucl. Instr. and Meth. Phys. Res. 441 (2000) 219.
[7] J. Bisognano, C. Leeman, Stochastic cooling, AIP Conf. Proc. 87 (1982) 584.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 58–61
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# The short-range wakefields in the BTW accelerating structure of the ELETTRA LINAC

P. Craievich[a,*], T. Weiland[b], I. Zagorodnov[b,1]

[a]*Sincrotrone Trieste-ELETTRA, S.S. 14 Km 163.5, 34012 Trieste, Italy*
[b]*Technische Universität Darmstadt-TEMF, Darmstadt, Germany*

Available online 1 December 2005

**Abstract**

Future FEL operations in the ELETTRA LINAC require a high quality beam with an ultra short bunch. The knowledge of the short-range wakefields in the backward traveling wave (BTW) accelerating structure is needed to predict the beam quality in term of the single bunch energy spread and emittance. To calculate the effect of the longitudinal and transverse wakefields we have used the time domain numerical approach with a new implicit scheme for calculation of wake potential of short bunches in long structure [A. Novokhatski, M. Timm, T. Weiland, Transition dynamics of the wake fields of ultra short bunches, Proceeding of the ICAP California, USA, 1998, I. Zagorodnov, T. Weiland, Calculation of transverse wake potential for short bunches, ICAP, 2002]. The wake potentials of the BTW structure are calculated numerically for very short bunches and analytical approximations for wake functions in short and long ranges are obtained by fitting procedures based on analytical estimations.


*PACS:* 29.17.+w; 41.20.−q; 41.60.Cr; 41.75.Ht

*Keywords:* Wakefields; Longitudinal wake function; Transverse wake function

## 1. Introduction

The FERMI@ELETTRA project [1] aims to construct a single-pass FEL user-facility in the spectral range 100–10 nm using the existing normal conducting 1.0 GeV linac. In order to raise the beam energy after pre-injector from 100 MeV to 1 GeV, the linac includes seven accelerating sections. Each section is a backward traveling wave (BTW) structure composed by 162 nose cone cavities magnetically coupled and operated in the $3/4\pi$ mode. The total length of the accelerating structure is about 6 m [2]. To avoid undesirable beam degradations, in term of energy spread and emittance, the wakefields effects have to be carefully considered. We have studied the longitudinal and transverse cases using the time domain code ECHO with a new implicit scheme for calculation of the wake potential of short bunches in long structures [3,4]. We have considered the wakefields evolution for bunches of different lengths passing through a single cell, a multi-cell and a complete accelerating structure [2]. This paper reports the longitudinal and transverse wakefields calculations for the complete BTW accelerating structure.

## 2. Longitudinal wake function of the BTW structure

The wake potentials of Gaussian bunch with length ranging from 1000 to 50 μm are calculated for a whole BTW accelerating structure. In Fig. 1 (left) the calculated longitudinal wake potentials (solid lines) are reported. To find an analytical approximation of the wake function we have chosen a combination of periodic [5] and one cell [6] dependence since the BTW structure can be treated as a periodic structure of finite length.

From the fit of the numerical wake potentials we have obtained an analytical expressions approximating the wake

*Corresponding author.
*E-mail address:* paolo.craievich@elettra.trieste.it (P. Craievich).
[1]This work was supported in part by the Deutsche Forschungsgemeinschaft under contract WE1239/22.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.069

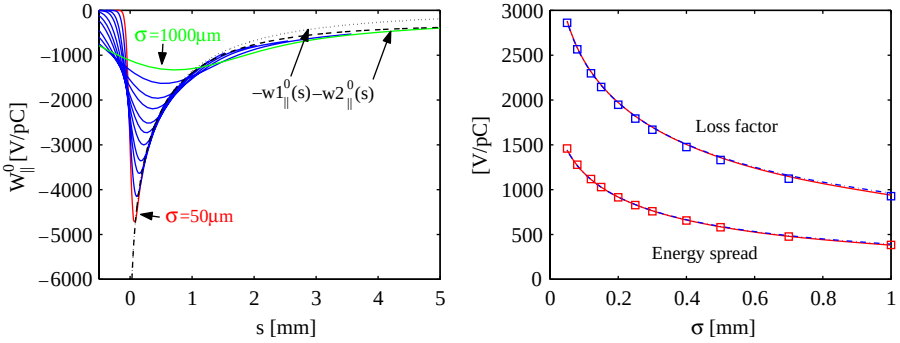



Fig. 1. Left: Longitudinal wake potentials (solid lines) and longitudinal wake functions of the BTW structure. Right: Comparison of numerical (box) and analytical (solid and dashed lines) integral parameters as function of bunch length $\sigma$.

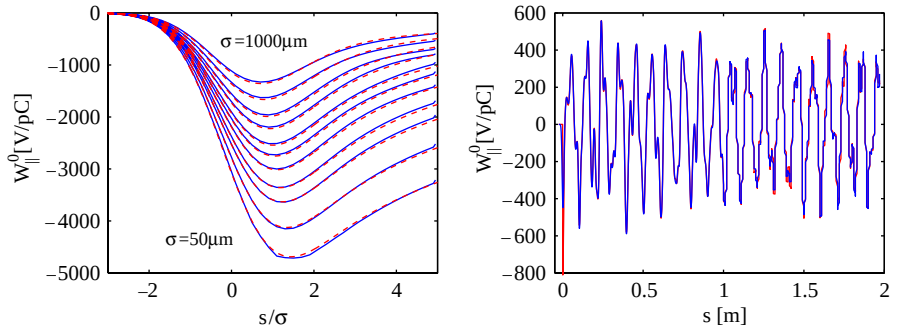



Fig. 2. Left: Longitudinal numerical (blue solid lines) and analytical (red dashed lines) wake potentials of the BTW structure. Right: The long range numerical (blue line) and analytical (red line) longitudinal wake potentials for Gaussian bunch with $\sigma = 5$ mm in the BTW structure.

function (in $[V/pC]$):

$$w1^0_{\parallel}(s) = A_{01}e^{-\sqrt{s/s_0}} + \frac{A_{02}}{\sqrt{s}}, \quad (1)$$

where $A_{01} = 7300$, $s_0 = 3.2 \times 10^{-4}$ and $A_{02} = 3.4$. Fig. 1 (left) shows the longitudinal wake function (1) (black dotted line), which tends for small $s$ to be an envelope function to the wakes and fits the numerical results up to 1.5 mm. To find an analytical approximation up to 5 mm, we have added to expression (1) an additional term with $\sqrt{s}$ dependence. Taking into account the last correction term the fit gives the following analytical expression (in $[V/pC]$):

$$w2^0_{\parallel}(s) = A_{01}e^{-\sqrt{s/s_0}} + \frac{A_{02}}{\sqrt{s}} + A_{03}\sqrt{s}, \quad (2)$$

where $A_{01} = 7450$, $s_0 = 3.1 \times 10^{-4}$, $A_{02} = 3$ and $A_{03} = 3000$. The previous relation approximates the longitudinal wake function on a wider range compared to expression (1). Fig. 1 (left) shows the longitudinal wake function (2) (black dashed line) that tends to be an envelope function to the wakes up to 5 mm. Fig. 1 (right) plots the numerical (box) and analytical (lines) loss factors and energy spread as function of bunch length $\sigma$. The coincidence of the numerical and analytical loss factors and energy spread can be seen for both the analytical expressions 1 (red solid lines) and 2 (blue dashed lines). Fig. 2 (left) presents the calculated longitudinal wake potentials (blue solid lines) together with analytical approximations (2) (red dashed lines). A more detailed analysis of Figs. 1 (left) and 2 (left) show that the analytical expression 2 approximates very well the longitudinal wake function up to 5 mm.

To estimate long range wakefields the wake potential for Gaussian bunch with $\sigma = 5$ mm is calculated for a distance up to 2 m after the bunch. The calculation is carried out with code ECHO for complete $\approx 6$ m long BTW structure. The longitudinal wake function can be approximated by the following expression:

$$w_{\parallel}(s) = -\theta(s)2\sum_{i=1}^{\infty} K_i \cos\left(\frac{2\pi}{c} f_i s\right), \quad (3)$$

where the frequencies $f_i$ and amplitudes $K_i$ have been calculated with a direct analysis of the numerical wake potentials using the Prony–Pisarenko algorithm [7]. The Prony–Pisarenko algorithm is a method to fit a set of decaying oscillation characterized by amplitudes, phases and damping constants to a given curve or data set and it is used in this example as alternative to the discrete Fourier transform. To obtain an approximation of the long range wake function we keep in (3) only a finite number ($N = 20$) of addends corresponding to the lowest frequencies. The results are shown in Table 1. Fig. 2 (right) shows the long range numerical (blue line) and analytical (3) (red line) longitudinal wake potentials for Gaussian bunch with RMS length $\sigma = 5$ mm. An excellent coincidence of the curves can be seen.

Table 1
The lowest frequencies and their amplitudes for long range longitudinal wake function of the BTW structure ($f_i$ in GHz and $K_i$ in [V/pC])

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| $f_i$ | 2.98 | 3.00 | 5.38 | 5.41 | 6.77 | 6.85 | 8.24 |
| $K_i$ | 123.18 | 57.84 | 32.24 | 15.57 | 37.14 | 14.46 | 75.69 |
| $f_i$ | 8.37 | 9.36 | 10.68 | 10.82 | 11.61 | 11.86 | 12.00 |
| $K_i$ | 35.73 | 14.48 | 9.65 | 6.44 | 23.11 | 26.09 | 8.19 |
| $f_i$ | 14.38 | 14.59 | 16.51 | 17.32 | 17.68 | 20.22 | |
| $K_i$ | 33.11 | 25.04 | 14.08 | 21.31 | 21.51 | 30.90 | |

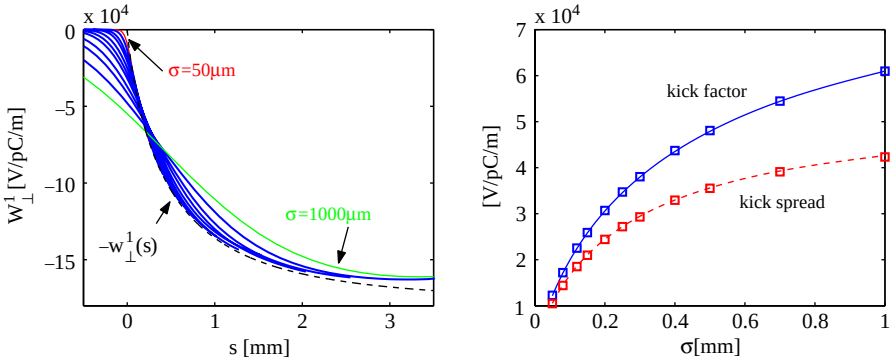



Fig. 3. Left: Transverse wake potentials (solid lines) and transverse wake functions (black dashed line) of the BTW structure. Right: Comparison of numerical (box) and analytical (line) kick factor (blue) and kick spread (red) as function of bunch length $\sigma$.

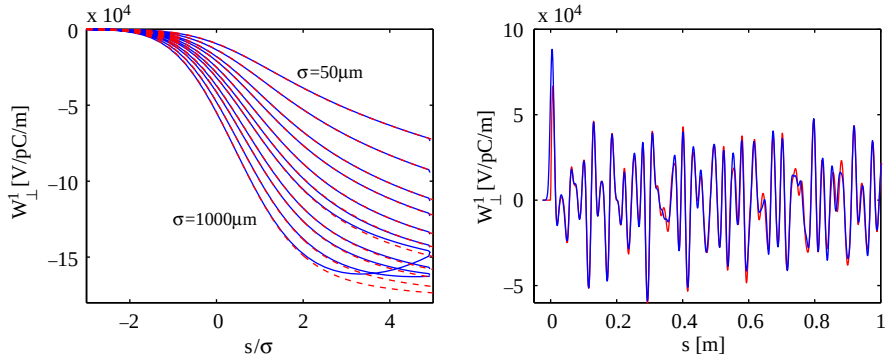



Fig. 4. Left: Transverse numerical (blue solid lines) and analytical (red dashed lines) wake potentials of the BTW structure. Right: The long range numerical (blue line) and analytical (red line) transverse wake potentials for Gaussian bunch with $\sigma = 5$ mm in the BTW structure.

## 3. Transverse wake function of the BTW structure

Fig. 3 (left) shows the calculated transverse wake potentials (solid lines) for different bunch length $\sigma$. As in the previous case, the BTW structure is treated as a periodic structure of finite length and to find an analytical approximation for the transverse wake function, a combination of periodic [5] and one cell [6] dependence was chosen. The expression for the wake function is obtained with a fit of the numerical wake potentials (in $[V/pC/m]$):

$$w_{\perp}^{1}(s) = A_{11}\left[1 - \left(1 + \sqrt{\frac{s}{s_1}}\right)e^{-\sqrt{s/s_1}}\right] + A_{12}\sqrt{s}, \qquad (4)$$

where $A_{11} = 1.7 \times 10^5$, $s_1 = 1.2 \times 10^{-4}$ and $A_{12} = 8.5 \times 10^4$.

Fig. 3 (left) shows the transverse wake function (4) (black dashed line), which tends to be an envelope function to the wakes up to distance $\sigma = 2$ mm after the bunch center. Fig. 3 (right) presents the numerical (box) and analytical kick factors (blue line) and kick spreads (red dashed line) as function of bunch length $\sigma$. Fig. 4 (left) plots the calculated transverse wake potentials (blue solid lines) together with their analytical approximation (4) (red dashed lines). For transverse case no additional term is introduced and the wake function fits the results up to 2 mm.

To estimate the long range transverse wakefields the wake potential for Gaussian bunch with $\sigma = 5$ mm is

Table 2
The lowest frequencies and their amplitudes for long range transverse wake function of the BTW structure ($f_i$ in GHz and $K_i \times 10^{17}$)

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $f_i$ | 4.94 | 6.89 | 6.92 | 8.53 | 9.91 | 10.41 | 12.03 | 12.15 | 12.91 |
| $K_i$ | 8.67 | 3.41 | 5.96 | 0.72 | 39.67 | 1.46 | 15.51 | 23.09 | 5.96 |
| $f_i$ | 14.04 | 15.02 | 15.14 | 16.14 | 16.67 | 17.21 | 18.94 | 20.62 | 21.47 |
| $K_i$ | 2.81 | 18.44 | 20.13 | 11.32 | 24.10 | 5.43 | 8.23 | 19.28 | 4.88 |

calculated for a distance up to 2 m after the bunch. The transverse wake function can be approximated by the following expression:

$$w_{\perp}(s) = \theta(s)2\sum_{i=1}^{\infty} K_i \frac{c}{2\pi f_i}\sin\left(\frac{2\pi}{c} f_i s\right), \quad (5)$$

where the frequencies $f_i$ and amplitudes $K_i$, as in the longitudinal case, have been obtained with the help of the Prony–Pisarenko algorithm. To obtain an approximation of the long-range wake function we keep in (5) only a finite number ($N = 18$) of addends corresponding to the lowest frequencies. The results are shown in Table 2. Fig. 4 (right) shows the long range numerical (blue line) and analytical (5) (red line) transverse wake potentials for Gaussian bunch with $\sigma = 5$ mm up to 1 m after the bunch. We can see that the wake function (5) is a good approximation of the long range numerical wake.

## 4. Conclusion

Calculations of the wakefields for short bunches passing through a complete BTW accelerating structure have been presented. The short range longitudinal and transverse wake potentials have been calculated in time domain with the code ECHO. From the numerical results analytical approximations of the point-charge wake functions were found. For the analytical model we have chosen a combination of periodic and one cell dependences. In the longitudinal case the term that describes the finite structure (one cell behavior) is very small compared to the periodic structure term. Hence in range of $\sigma$ we have considered, the longitudinal wakes shows mainly a periodic structure behavior. In addition, for better fitting of the data up to 5 mm, we have used an additional term in the model. Furthermore, the long range longitudinal and transverse wakefields were estimated by a direct analysis of the numerical wake potential for Gaussian bunch calculated for a distance up to 2 m after the bunch.

## References

[1] Fermi@Elettra machine project update, Sincrotrone Trieste, 2003.
[2] P. Craievich, T. Weiland, I. Zagorodnov, Report ST/M-04/02, Sincrotrone Trieste, 2004.
[3] A. Novokhatski, M. Timm, T. Weiland, Transition dynamics of the wake fields of ultra short bunches, Proceeding of the ICAP, California, USA, 1998.
[4] I. Zagorodnov, T. Weiland, Calculation of transverse wake potential for short bunches, ICAP, 2002.
[5] K. Bane, Short-range dipole wakefields in accelerating structure for the NLC, SLAC-PUB-9663, LCC-0116, 2003.
[6] K. Bane, M. Sands, Wakefields of very short bunches in an accelerating cavity, SLAC-PUB-4441, 1987.
[7] T.W. Park, C.S. Burrus, Digital Filter Design, Wiley, New york, 1987.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 62–65
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Analysis of multimode wakefield generation in accelerating structures with conductive dielectric loading

A.D. Kanareykin[a], A.V. Tyukhtin[b,*]

[a]*Euclid Techlabs LLC, Solon, OH, USA*
[b]*Radiophysics Department, St. Petersburg University, Ul'yanovskaya Str. 1, Petrodvorets, St. Petersburg 198504, Russia*

Available online 21 November 2005

**Abstract**

Wakefield generation by short electron bunches traveling along the vacuum channel of an accelerating structure loaded with a conductive ceramic dielectric has been investigated. The dielectric response is defined as no dispersion of the real part of permittivity and linear increase of loss tangent versus frequency. This model is in good agreement with measured microwave properties of some materials tested for Dielectric Wakefield Acceleration experiments. It has been shown that the dielectric conductivity has an essential impact on the wakefield. The loss factor influences most strongly on short electron bunch generation of the multimode wakefields.


*PACS:* 41.75.Lx; 41.75.Ht; 29.17

*Keywords:* Wakefield acceleration; Multimode wakefield; Electron bunch; Conductive dielectric

## 1. Introduction

In Ref. [1] the first experimental results of a new wakefield acceleration scheme was demonstrated. The multibunch driven, multimode, dielectric wakefield accelerator was performed at the Argonne Wakefield Accelerator (AWA) as well [2]. In this experiment, a bunch train of 4–5 nC electron bunches, separated by 760 ps, was passed through a 60 cm long dielectric-lined cylindrical waveguide. The separation was chosen to match the net acceleration wavelength of the multimode dielectric loaded structure. By carefully measuring the energy spectrum of the four beams after they passed through the waveguide, it was shown that the wakefield is indeed enhanced by a train of periodically spaced electron bunches. This work represents the first experimental demonstration of the concept and also shows that multipulse operation of wakefield accelerators is worth further investigation.

It should be mentioned that in Ref. [2] the loss factor of the material was not taken into account which was reasonable at that time because the bunch length $\sigma$ was equal to 4.5 mm and the high-order modes did not affect the accelerating gradients significantly. At the same time, in order to get the high accelerating gradient and increased acceleration efficiency, the short 1–2 mm electron bunches will be used [3] that are available at AWA now.

In this paper, we present numerical simulations of Cherenkov wakefield generation by short electron bunches traveling along the vacuum channel of an accelerating structure loaded with a conductive ceramic dielectric. We consider ultrarelativistic beams corresponding to the beam parameters available at the AWA.

## 2. Properties of materials

It should be mentioned that the dispersive properties of the dielectric can affect significantly the electromagnetic field generation caused by the fast moving electron bunches. The method of the Cherenkov radiation analysis in dielectric media with the various dispersion models has been developed, in particular, in Ref. [4]. The same approach has been used in Ref. [5] for the Cherenkov radiation generated by the particles moving in the

*Corresponding author. Fax: +7 812 428 72 40.
*E-mail addresses:* alexkan@euclidconcepts.com (A.D. Kanareykin), tyukhtin@pobox.ru (A.V. Tyukhtin).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.015

waveguide filled material with the resonant dispersion. It was shown that this kind of dispersion defines the radiation spectrum by reducing the high mode amplitudes significantly.

The dielectric constant of the low-loss ceramics and some another dielectric is almost flat in the frequency range from few GHz and up to few hundreds GHz, at the same time imaginary part of the dielectric response linearly increases with frequency [6,7]. The dispersion of these materials is a relaxation. It is described approximately by means of Debye model where the real and imaginary part of the dielectric constant are defined as

$$\varepsilon'(\omega) = \mathrm{Re}\,\varepsilon(\omega) = \frac{\varepsilon_0 + \varepsilon_\infty \omega^2\tau^2}{1+\omega^2\tau^2},$$
$$\varepsilon''(\omega) = \mathrm{Im}\,\varepsilon(\omega) = \frac{(\varepsilon_0 - \varepsilon_\infty)\omega\tau}{1+\omega^2\tau^2}, \quad (1)$$

where $\varepsilon_0$ and $\varepsilon_\infty$ are the dielectric constant values at the upper and low limit frequency band, respectively (it should be noticed that $\varepsilon_0 > \varepsilon_\infty$) and $\tau$ is a relaxation time. The frequencies of the harmonics contribute to the wakefields of the bunch under the condition of $\varepsilon_0\omega\tau \ll 1$. One can see from Eq. (1) that dispersion of the real part of the dielectric permittivity is of second infinitesimal order while the imaginary part is of first order one with the linear dependence of a frequency:

$$\varepsilon' \approx \varepsilon_0, \quad \varepsilon'' \approx \omega\tau_0, \quad (2)$$

where $\tau_0 = (\varepsilon_0 - \varepsilon_\infty)\tau$.

At the frequencies $>100$ GHz (or even less for some specific materials) some dielectrics show the resonance dispersion instead of the relaxation one. Taking into account just one resonant frequency one can write the definition of the dielectric constant as $\varepsilon = 1 + \omega_L^2(\omega_0^2 - \omega^2 - i\nu\omega)^{-1}$, where $\omega_0$ and $\omega_L$ are a resonance and plasma frequencies, respectively, $\nu$ is the attenuation parameter. It is worth nothing that under the condition of $\omega \ll \omega_0$, $\nu \ll \omega_0$ the real part of the dielectric constant is weakly dependent on the frequency while the imaginary part shows a small value increasing linearly: $\varepsilon' \approx 1 + \omega_L^2/\omega_0^2$, $\varepsilon'' \approx (\omega_L^2\nu/\omega_0^4)\omega$. Thus, with the above assumption the resonance dispersion approximation gives the same formulas (2) where one has to define the constants of $\varepsilon_0$ and $\tau_0$.

The proposed model validation at the wide frequency range has been tested and proofed with the numerous experimental data of the ceramic loading testing at the wakefield acceleration in recent decade [1–2,8].

## 3. Analytical results

The basic structure is very simple—a cylindrical, dielectric loaded waveguide with an axial vacuum channel is inserted into a conductive sleeve. A high charge (typically 20–40 nC), short, (1–4 mm) electron drive beam generates $TM_{01}$ mode Cherenkov radiation (wakefields) while propagating down the vacuum channel. Following at a delay adjusted to catch the accelerating phase of the wakefield is a second electron witness beam. The witness beam is accelerated to high energy by the wakefield produced by the drive beam. A series of proof of principle experiments have been successfully performed at Argonne's Advanced Accelerator Test [1] and AWA facilities [2]. The initial energy of the AWA beam is 15 MeV that corresponds to $\beta = 0.9994$. We can consider in all following calculations $\beta \to 1$ at the same estimation accuracy.

It is usually supposed that the charge bunch distribution is Gaussian: $p(\zeta) = 1/\sqrt{2\pi}\sigma \exp(-\zeta^2/2\sigma^2)$, where $\zeta = z - Vt$, $z$ is the axis of the structure, $\beta = cV$ is the velocity of the bunch. It should be mentioned that $\sigma$ corresponding to the longitudinal bunch size is much less than the first $TM_{01}$ mode wavelength of the Cherenkov radiation generated by the bunch.

Electromagnetic field simulations for the dielectric waveguide have been developed previously in Ref. [4]. At the same time there were no wakefield simulations presented for the dielectric loading with the dispersion of the dielectric constant.

We used the iterations method for the dispersion equation solution under condition of the small loss tangent values: $\tan\delta = \omega\tau_0/\varepsilon_0 \ll 1$. The modes frequencies of the dielectric loaded waveguide with the dielectric constant of $\varepsilon_0$ and no dispersion have been taken as zero approximation of the equation solution. It yields to the mode frequency corrections that are imaginary. Omitting all intermediate simulation one can write approximate expression for the longitudinal electrical field behind the bunch as

$$E_z = \sum_{m=1}^{\infty} E_{mz} \exp\left(-\delta_m\left|\frac{z}{c} - t\right|\right)\cos\left(\omega_m\left(\frac{z}{c} - t\right)\right), \quad (3)$$

where

$$E_{mz} = -\frac{4q}{cb}\frac{\sqrt{\varepsilon_0 - 1}\psi_1(s)}{(\mathrm{d}/\mathrm{d}\omega)((sb/2)\psi_1(s) - \varepsilon\psi_0(s))} \times \exp\left(-\frac{\omega^2\sigma^2}{2c^2}\right)\Big|_{\substack{\omega=\omega_m \\ s=s_m}}, \quad (4)$$

$$\psi_0(s) = J_1(sb)N_0(sa) - J_0(sa)N_1(sb),$$
$$\psi_1(s) = J_0(sb)N_0(sa) - J_0(sa)N_0(sb), \quad (5)$$

and $s = (\omega/c)\sqrt{\varepsilon_0 - 1}$. Here, $J_n(x)$ and $N_n(x)$ are Bessel and Neyman functions accordingly, $a$ is the waveguide radius, $b$ is the channel radius. Proper value and corresponding proper frequency $\omega_m = cs_m/\sqrt{\varepsilon_0 - 1}$ are given by the dispersion equation

$$sb\psi_1(s) - 2\varepsilon\psi_0(s) = 0. \quad (6)$$

At the same time values $\delta_m$ defining the modes attenuation are connected with the frequencies with the following expression:

$$\delta_m = \frac{\omega\tau_0}{\sqrt{\varepsilon_0 - 1}}\left[\frac{\omega}{2\sqrt{\varepsilon_0 - 1}} - \frac{2c\psi_0(s)}{b\psi_1(s) + bs\mathrm{d}\psi_1(s)/\mathrm{d}s - 2\varepsilon_0\mathrm{d}\psi_0(s)/\mathrm{d}s}\right]\Big|_{\substack{\omega=\omega_m \\ s=s_m}}. \quad (7)$$

The right part of expression (7) utilizes the $\omega_m$ and $s_m$ values calculated in the case of nonconductive dielectric with the dielectric constant of $\varepsilon = \varepsilon_0$ as a solution of Eq. (6). It should be noticed that all of $\delta_m$ values are positive ones.

## 4. Numerical simulations

We present here some examples of the Dielectric Wakefield Accelerator parameters for the high-mode attenuation study of the acceleration field. The accelerating structure parameters are $\varepsilon_0 = 38.1$, $a = 1.44\,\text{cm}$, $b = 0.5\,\text{cm}$ that corresponds to the wakefield accelerator parameters of the experiment presented in Ref. [2] (the fundamental mode frequency is 1.582 GHz). This structure is a typical one for the multimode dielectric wakefield accelerators. We defined loss tangent value as $\tan\delta = 10^{-4}$ at 1 GHz that gives at the arbitrary frequency $\tan\delta = 10^{-13}\nu$.

Fig. 1a and b corresponds to the short electron bunch of $\sigma = 1\,\text{mm}$. Fig. 1a gives harmonics amplitudes of component $E_{mz}$ in the case of loss less dielectric and in the case of conductive dielectric for distance $\zeta = 100\,\text{cm}$. As we see the wakefield spectrum of the second case is characterized by essential decreasing amplitudes of high harmonics. Therefore, we obtain essential decreasing maximums of wakefield in Fig. 1b. The second negative peak of the accelerating field is reduced 1.5 times in comparison with the case of no dispersion, and the fourth one is decreased 2.5 times. The bunch with $\sigma = 2\,\text{mm}$ does not show such strong dependence of the accelerating field amplitude on the imaginary part of the dielectric constant, the fourth peak reduction does not exceed 1.5. For the bunches with $\sigma \geqslant 4\,\text{mm}$ after the four cycles of the field there was no significant magnitude reduction, but it will become apparent at the long enough distance behind the bunch.

Fig. 2 shows the negative extremums of the field depending on $\tan\delta$ (it relates to the frequency 1 GHz) in the cases of $\sigma = 1$ and 4.5 mm. One can see the monotonous magnitude decreasing. For the bunch with $\sigma = 1\,\text{mm}$, the significant magnitude reduction began at a loss factor value of $\tan\delta = 10^{-5}$, and for the bunch with $\sigma = 4.5\,\text{mm}$ we found the same phenomenon for the lossy material with $\tan\delta = 10^{-4}$. It should be noticed that the same bunch with $\sigma = 4.5\,\text{mm}$ was used in Ref. [2] for the multibunch acceleration demonstration and our present study results confirmed that for the parameters published in Ref. [2] and for the loss factor of the material in the range of $10^{-5}-10^{-4}$ at 1 GHz the conductivity impact on the accelerating field magnitude is negligible. But for the more short bunches to be used in upcoming experiments, the loss factor of the material becomes a critical issue.

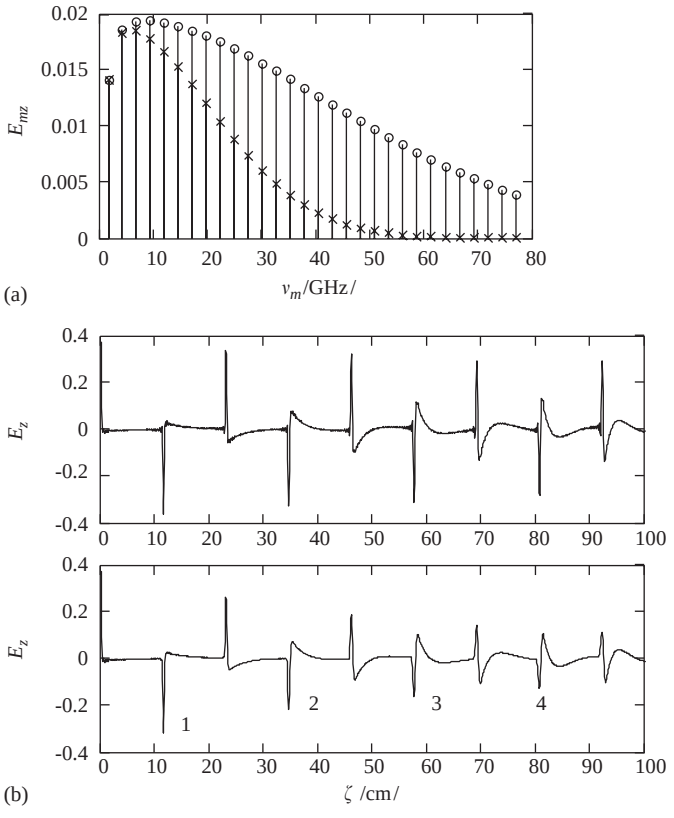



Fig. 1. (a) Amplitudes of harmonics of longitudinal electrical field (MV/m): the case of lossless dielectric (circles); the case of conductive dielectric for the distance $\zeta = 100\,\text{cm}$ (crosses). Parameters of the structure: $\varepsilon_0 = 38.1$, $a = 1.44\,\text{cm}$, $b = 0.5\,\text{cm}$, $\tan\delta = 10^{-4}$ at 1 GHz, $q = -1\,\text{nC}$, $\sigma = 1\,\text{mm}$. (b) Dependence of $E_z$(MV/m) on the distance to the beam $\zeta = ct - z$(cm) in the case of lossless dielectric (top) and in the case of conductive dielectric (bottom). Parameters are identical as in (a).

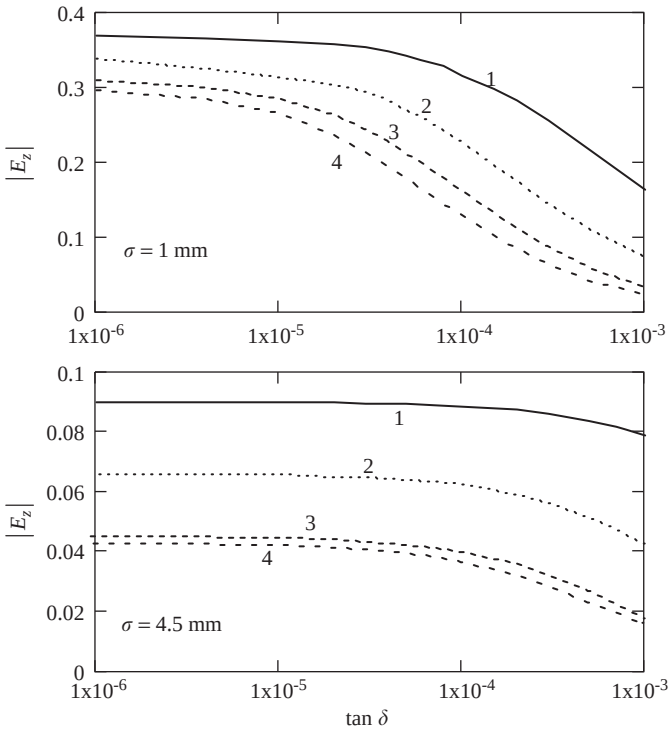



Fig. 2. The negative extremums marked in Fig. 1b depending on the loss tangent at 1 GHz.

Summarizing we would like to point out that in Dielectric Wakefield Accelerator development the loss factor of the dielectric loading has to be taken into account and the impact of the loss tangent on the accelerator parameters has to be studied, especially for the multibunch multimode acceleration where the high modes contribution into the accelerating gradient becomes significant. It should be noticed that the loss factor of the accelerator loading of the typical microwave ceramics material increases linearly versus the frequency in the range of 1–100 GHz and therefore critically affects the high mode magnitudes. One can overcome the loss factor influence by utilizing only low-loss materials like alumina and tantalum compositions for the high dielectric constant loading as well as for the interbunch distance adjustments inside the bunch train.

## References

[1] W. Gai, P. Schoessow, B. Cole, R. Konecny, J. Norm, J. Rosenzweig, J. Simpson, Phys. Rev. Lett. 61 (1988) 2756.
[2] J.G. Power, M.E. Conde, W. Gai, A. Kanareykin, R. Konecny, P. Schoessow, Phys. Rev. ST-AB 3 (2000) 101302-1.
[3] M.E. Conde, W. Gai, C. Jing, R. Konecny, W. Liu, J.G. Power, H. Wang, Z. Yusof, Proceedings of the 2003 Particle Accelerator Conference, 2003, p. 2032.
[4] B.M. Bolotovskiy, Sov. Phys. Usp. 4 (1962) 781.
[5] A.V. Tyukhtin, Technical Phys. Lett. 30 (2004) 605.
[6] B.A. Rotenberg, Ceramic Capacitor Dielectrics, RFFI, St. Petersburg, 2000.
[7] J. Petzelt, T. Ostapchuk, A. Pashkin, I. Rychetsky, J. European Ceramic Soc. 30 (2003) 2627.
[8] A.D. Kanareykin, I.L. Sheinman, E.A. Nenasheva, N.F. Kartenko, W. Gai, J. Simpson, Physics at the Turn of the 21st Century, St. Petersburg, Russia, 1998, p. 57.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 66–68
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# A differential algebraic integration algorithm for symplectic mappings in systems with three-dimensional magnetic field

P. Chang[a], S.Y. Lee[b], Y.T. Yan[c,*]

[a]National Synchrotron Radiation Research Center, Hsinchu 30077, Taiwan
[b]Department of Physics, Indiana University, Bloomington, IN 47405, USA
[c]Stanford Linear Accelerator Center, Stanford, CA 94309, USA

Available online 21 November 2005

**Abstract**

A differential algebraic integration algorithm is developed for symplectic mapping through a three-dimensional (3-D) magnetic field. The self-consistent reference orbit in phase space is obtained by making a canonical transformation to eliminate the linear part of the Hamiltonian. Transfer maps from the entrance to the exit of any 3-D magnetic field are then obtained through slice-by-slice symplectic integration. The particle phase-space coordinates are advanced by using the integrable polynomial procedure. This algorithm is a powerful tool to attain nonlinear maps for insertion devices in synchrotron light source or complicated magnetic field in the interaction region in high energy colliders.


*PACS:* 29.27. – a; 41.85. – p; 29.20. – c; 02.70.Wz

*Keywords:* Differential algebras; Symplectic integration; Particle tracking; Transfer maps

## 1. Introduction

In most storage rings, magnetic field can usually be well approximated by dominant transverse components. Thus, a single longitudinal component of the magnetic vector potential $A_s(x, y, s)$ is sufficient to describe the system. Because the Hamiltonian can be separated into a term that depends only on coordinates and another on momenta, the phase-space coordinates can be advanced with the conventional symplectic integrators [1].

However, a general 3-D magnetic field cannot be represented by a vector potential $A_s$ alone. The vector potential needs at least two transverse components, e.g. $A_x$ and $A_y$. The Hamiltonian can no longer be separated into the coordinate-dependent-only and the momenta-dependent-only terms. This paper presents a new algorithm to obtain self-consistent symplectic maps for general 3-D magnetic field. In Section 2, a differential algebraic Drift-kick-kick-drift (DKKD) procedure is presented to obtain the self-consistent reference orbit and the Hamiltonian expanded around the reference orbit. In Section 3, we discuss symplectic mapping and particle tracking. The conclusion is given in Section 4.

## 2. The Hamiltonian and the determination of reference orbit

After being normalized to the mechanical momentum magnitude of a reference synchronous particle, the dimensionless Hamiltonian for a charged particle in a three-dimensional magnetic field can be given, in Frenet–Serret coordinate system, by

$$H = -A_s - \left(1 + \frac{x}{\rho}\right)[(1+\delta)^2 - (p_x - A_x(x,y,s))^2 - (p_y - A_y(x,y,s))^2]^{1/2} \quad (1)$$

where $\rho$ is the radius of curvature, $\delta$ is the fractional momentum deviation, $p_x$ and $p_y$ are the canonical conjugate momenta, $A_x$, $A_y$, and $A_s$ are the magnetic vector potential components that are functions of the coordinates, $x, y, s$, where $s$ is the time-like coordinate

*Corresponding author.
*E-mail address:* yan@slac.stanford.edu (Y.T. Yan).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.016

along the longitudinal direction while $x, y$ are the transverse coordinates. Note that eliminating the $x/\rho$ term would simply yield the Hamiltonian in the local Cartesian coordinate system. The mixing of conjugate momenta $(p_x, p_z)$ with the vector potential $A_x$ and $A_y$ that are functions of the coordinates $(x, y)$ poses the challenge of obtaining symplectic Taylor maps for particle tracking. This problem is tackled by using a differential algebraic drift-kick-kick-drift (DKKD) algorithm to derive the self-consistent reference orbit and obtain Taylor map with respect to the reference orbit.

We consider a region of 3D magnetic field, longitudinally located between $s_i$ and $s_f$ with $L = s_f - s_i$. The system is divided into $N$-slices such that the length of each slice is $\Delta s = L/N$. The entry, mid-point, and exit of the $k$th slice are $s_k = s_i + (k-1)\Delta s$, $\bar{s}_k = s_i + (k-1/2)\Delta s$, and $s_{k+1} = s_i + k\Delta s$. The differential algebraic DKKD algorithm is used to obtain the particle's reference trajectory: $\vec{z}_0(\delta, s) \equiv (x_0(\delta, s), p_{x0}(\delta, s), y_0(\delta, s), p_{y0}(\delta, s))$. The Hamiltonian is expanded around this reference orbit.

(1) We evaluate the Hamiltonian $H(x, p_x, y, p_y, \delta; s_k)$, $\mathrm{d}x/\mathrm{d}s = \partial H/\partial p_x$ and $\mathrm{d}y/\mathrm{d}s = \partial H/\partial p_y$ at the slice-interface $s_k$. The coordinates of reference orbit are advanced to the mid-point $\bar{s}_k$ by

$$x_0(\delta, \bar{s}_k) = x_0(\delta, s_k) + (\Delta s/2) \cdot (\partial H/\partial p_x)|_{\vec{z}_0(\delta, s_k)} \quad (2)$$

$$y_0(\delta, \bar{s}_k) = y_0(\delta, s_k) + (\Delta s/2) \cdot (\partial H/\partial p_y)|_{\vec{z}_0(\delta, s_k)}. \quad (3)$$

(2) We now evaluate $H$ with the updated new reference coordinates, $\mathrm{d}p_x/\mathrm{d}s = -\partial H/\partial x$ and $\mathrm{d}p_y/\mathrm{d}s = -\partial H/\partial y$ at the mid-point of the slice. The conjugate momenta of reference orbit at the mid-point are advanced by

$$p_{x_0}(\delta, \bar{s}_k) = p_{x_0}(\delta, s_k) - (\Delta s/2) \cdot (\partial H/\partial x)|_{\vec{z}_0(\delta, \bar{s}_k)} \quad (4)$$

$$p_{y_0}(\delta, \bar{s}_k) = p_{y_0}(\delta, s_k) - (\Delta s/2) \cdot (\partial H/\partial y)|_{\vec{z}_0(\delta, \bar{s}_k)}. \quad (5)$$

(3) The Hamiltonian $H(x, p_x, y, p_y, \delta; \bar{s}_k)$, $\mathrm{d}p_x/\mathrm{d}s = -\partial H/\partial x$, and $\mathrm{d}p_y/\mathrm{d}s = -\partial H/\partial y$ at the mid-point can be evaluated. Changing the conjugate phase-space coordinates $(X = x - x_0, p_X = p_x - p_{x_0}, Y = y - y_0, p_Y = p_y - p_{y_0})$ at the mid-point $\bar{s}_k$, one can evaluate and save the Hamiltonian, $\tilde{H}(\delta, \bar{s}_k) = h_k + f_k$. Which describes the dynamics and is Taylor-expanded around the parameterized reference orbit. It is integrated with the previous concatenated map by a second-order symplectic integrator for obtaining the section-map.

(4) With the $\mathrm{d}p_x/\mathrm{d}s = -\partial H/\partial x$ and $\mathrm{d}p_y/\mathrm{d}s = -\partial H/\partial y$ obtained at this mid-point, the conjugate momenta of reference orbit are then advanced to the end-point $s_{k+1}$ of the $k$th slice by

$$p_{x_0}(\delta, s_{k+1}) = p_{x_0}(\delta, \bar{s}_k) - (\Delta s/2) \cdot (\partial H/\partial x)|_{\vec{z}_0(\delta, \bar{s}_k)} \quad (6)$$

$$p_{y_0}(\delta, s_{k+1}) = p_{y_0}(\delta, \bar{s}_k) - (\Delta s/2) \cdot (\partial H/\partial y)|_{\vec{z}_0(\delta, \bar{s}_k)}. \quad (7)$$

(5) The $H$ at the mid-point is re-evaluated with the updated new reference momenta. The $\mathrm{d}p_x/\mathrm{d}s = -\partial H/\partial x$ and $\mathrm{d}p_y/\mathrm{d}s = -\partial H/\partial y$ are also obtained. The coordinates of reference orbit are advanced to the end-point $s_{k+1}$ of the $k$th slice by

$$x_0(\delta, s_{k+1}) = x_0(\delta, \bar{s}_k) + (\Delta s/2) \cdot (\partial H/\partial p_x)|_{\vec{z}_0(\delta, \bar{s}_k)} \quad (8)$$

$$y_0(\delta, s_{k+1}) = y_0(\delta, \bar{s}_k) + (\Delta s/2) \cdot (\partial H/\partial p_y)|_{\vec{z}_0(\delta, \bar{s}_k)}. \quad (9)$$

Using conjugate phase-space coordinates at the end-point $s_{k+1}$, one evaluates the Hamiltonian $H(x, p_x, y, p_y, \delta; s_{k+1})$, $\mathrm{d}p_x/\mathrm{d}s = -\partial H/\partial x$ and $\mathrm{d}p_y/\mathrm{d}s = -\partial H/\partial y$. The integration procedure repeats for the next slice.

In this integration process, an important constraint is the continuity of the vector potential at the interface of each slice. Since this process uses differential algebras, the dependence of the reference orbit on $\delta$ is included up to a desired order. The transfer map is given by $\exp\{-\Delta s : \tilde{H}(x, p_x, y, p_y, \delta, \bar{s}_k) :\}$, where the canonical phase-space coordinates are with respect to the $\delta$-dependent reference orbit. Note that the reference orbit may also be obtained by solving the tedious differential equations that turn out to obey the Lorentz force law after making a canonical transformation with the generating function

$$F_2 = (x - x_0)(p_X + p_{x_0}) + (y - y_0)(p_Y + p_{y_0})$$

such that the linear part of the transformed Hamiltonian vanishes [2].

The DKKD integration algorithm does not require symplecticity. It is a natural method to derive the self-consistent reference orbit and obtain the Hamiltonian for symplectic mapping using Lie operator method. One can also use the Runge-Kutta integration methods to solve the reference orbit and derive the Hamiltonian. However, the continuity condition must be carefully implemented in order to obtain a proper Hamiltonian for transfer map.

## 3. Map concatenation

For convenience, let the state vector $\vec{z}$ represent the transverse particle canonical phase-space coordinates with respect to the reference orbit, i.e. $\vec{z} \equiv (x, p_x, y, p_y)^{\dagger}$. The Hamiltonian $H(\vec{z}, \delta, \bar{s})$ is a polynomial with a minimum order of 2 and the transfer map that advances the particle phase-space coordinates (with respect to the reference orbit) from the entrance to the exit for the $k$th slice is given by $\exp\{-\Delta s : H(z, \delta, \bar{s}_k) :\}$.

One can choose to track particles slice by slice. But this would be quite CPU time consuming. Instead, one can choose to concatenate all the slice transfer maps into one for fast tracking. To do so, we first reform each of the slice transfer maps into two major terms, one for the linear motion and the other for the nonlinear motion. For

example, the $k$th-slice transfer map would be reformed as

$$\exp\left\{-\Delta s : \sum_{j=2} H_j(\bar{s}_k) :\right\} = \exp\{: h_k + f_k :\}$$

where the linear part $h_k \equiv -\Delta s \cdot H_2(\bar{s}_k)$ and the nonlinear part $f_k \equiv -\Delta s \cdot \sum_{j=3} H_j(\bar{s}_k)$ can be separated by a symplectic integrator. Using the second order symplectic integrator, one obtains

$$\exp\{: h_k + f_k :\} = \exp\{: \tfrac{1}{2} h_k :\} \exp\{: f_k :\} \exp\{: \tfrac{1}{2} h_k :\}, \quad (10)$$

and the whole map as

$$M = \prod_{k=1}^{N} \{e^{:1/2h_k:} e^{:f_k:} e^{:1/2h_k:}\} = e^{:h:} e^{:f:}, \quad (11)$$

where $e^{:h:}$ is the concatenated linear map. The nonlinear map $e^{:f:}$ can be evaluated by the method of integrable polynomials [3]. Therefore, the initial phase-space coordinates $\vec{Z}(s_i)$ at the entrance of a subsystem can be directly mapped to the final coordinates $\vec{Z}(s_f)$ at the exit of the subsystem. One must keep in mind the continuity condition of mechanical momenta required at the entrance and the exit. If the transverse vector potentials are zero at both the entry $s_i$ and the exit $s_f$, then the transverse canonical momenta is equal to their corresponding mechanical momenta. Otherwise, one should make additional transformation between mechanical momenta and canonical momenta at the entrance and the exit.

We use a quadrupole to illustrate and check the validity of our algorithm. Conventionally, the vector potential of a quadrupole field is represented by a single longitudinal component $\vec{A} \equiv (A_x, A_y, A_s) = (0, 0, B_1/2(y^2 - x^2))$, where $B_1 = \partial B_y/\partial x$. To check the slice integration algorithm, the Quadrupole vector potential is represented by two transverse components, $\vec{A} = (B_1 xs, -B_1 ys, 0)$ for a valid comparison of the slice-by-slice integrated numerical transfer matrix to the corresponding analytic one. The parameters in our calculation are $K = B_1/B_0\rho = 2.870480\,\mathrm{m}^{-2}$, the quadrupole-length $L = 0.35\,\mathrm{m}$, and $B_0\rho = -5.00346\,\mathrm{T\text{-}m}$.

We set the entry position $s_i = 0$ such that particle's mechanical momenta are the same as the canonical momenta. However, at the exit position, $s_f = L^-$, an additional transformation has to be made for the transfer map from canonical momenta to mechanical momenta that become the canonical momenta at $s = L^+$ provided that the vector potential outside the quadrupole is 0. This is because a hard-edge quadrupole prevents us from making vector potential continuous on both boundaries. If the fringe fields have been included to achieve vector potential continuity at both boundaries, the additional transformation at $s_f = L$ would not be necessary.

The linear transfer matrix of the quadrupole agrees with the analytic formula to better than $10^{-4}$, provided that $N \geqslant 20$ longitudinal slices are used. One may gain precision by taking higher order symplectic integrators and a larger number of slices. For the second order symplectic integrator, we find the error is proportional to $1/N^2$ .

## 4. Conclusion and discussion

We have developed a slice-by-slice symplectic transfer mapping in a 3D magnetic field where the vector potential cannot be described by $A_s$ alone. We divide the system longitudinally into N-slices and implement a differential algebraic drift-kick-kick-drift (DKKD) procedure to advance the parameterized (*δ-dependent*) reference orbit phase-space coordinates and simultaneously obtain the Hamiltonian with respect to the reference orbit. Note that the reference orbit is chosen such that the Taylor expanded Hamiltonian is without the first order so that all feed-downs from higher-order multipoles are automatically included in the transfer maps. For fast particle tracking, one may concatenates slice transfer maps into one with a symplectic integrators. That the reference orbit is actually Taylor expanded in momentum deviation $\delta$ offers the path length difference and the dispersion functions. Requirement of mechanical momenta continuity can be automatically satisfied if one imposes the continuity condition of the vector potential.

## Acknowledgments

We thank C.C. Kuo, K.K. Lin, K.T. Hsu, and C.T. Chen at NSRRC for support of this work. This work is supported by National Synchrotron Radiation Research Center in Taiwan, US Department of energy Grant DE-FG02-92ER40747, US National Science Foundation Grant NSF-PHY-0244793, and US Department of energy Contract DE-AC02-76SF00515.

## References

[1] R.D. Ruth, IEEE Trans. Nucl. Sci. NS-30 (1983) 2669.
[2] P. Chang, C.S. Hsue, S.Y. Lee, Y.T. Yan, SLAC-PUB-8514 (2000) and SRRC/BD/IM/2000-06.
[3] J. Shi, Y.T. Yan, Phys. Rev. E 48 (1993) 3943.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 69–73
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Space charge algorithm for the multi ensemble model

M. Krasilnikov

*DESY Zeuthen, Germany*

Available online 29 November 2005

**Abstract**

Derived from the Vlasov equation the Ensemble Model [A. Novokhatski, T. Weiland, PAC'99, New York, March 1999] has been elaborated for fast and efficient beam dynamics simulations. The Model represents a particle beam as a set of sub-beams or Ensembles, described by coordinates of the centroid and 6D phase space correlations. Whereas a space charge routine for the Single Ensemble Model (SEM) has been developed and tested [M. Krassilnkov, et al., ICAP 2000, Darmstadt, September 2000], implementation of the space charge algorithm for the Multi Ensemble Model (MEM) needs more efforts. A space charge model based on the Multi-Centered Gaussian Expansion (MCGE) [M. Krassilnikov, T. Weiland, ICAP'02, East Lansing, USA, October 2002] implies a smooth particle density distribution within an Ensemble but it requires rather large computational efforts. This paper presents another space charge algorithm, based on the analytical solution for the electromagnetic field of an ellipsoidal 3D charge distribution [M. Comunian, et al., Phys. Rev. Spec. Topics—Acc. Beams 4 (2001) 124201]. Using this algorithm one can calculate the space charge force and its gradient inside and outside the driving Ensemble. Features of the implementation and simplifying approximations are discussed in this paper.


*PACS:* 41.75.−i

*Keywords:* Space charge; Emittance; Phase space

## 1. Introduction

The Ensemble Model [1] describes a particle beam by a set of sub-beams or Ensembles, each of them is characterized not only by average position, but also by correlations (or moments of the Ensemble distribution function) in a 6D phase space. The Single Ensemble Model (SEM) considers the moments up to the second-order and describes each Ensemble by 27 parameters (6 moments of the first and 21 moments of the second-order). This implies linear external and internal forces inside the Ensemble and therefore, Ensemble emittance invariance [1]. Beam emittance can be simulated by extension of the SEM to higher-order moments but it results in significant increasing the equation number and in complication of their form. As another approach to simulate the beam emittance the Multi Ensemble Model (MEM) [2] can be used. Nonlinear Lorentz force effects can be modeled by a set of Ensembles distributed in the phase space. Whereas external force implementation in MEM is straightforward, the space charge force algorithm needs more efforts.

## 2. Main equations of the Ensemble Model

A beam distribution function could be represented by means of a superposition of Ensembles:

$$\Psi(\vec{r},\vec{p}) = \sum_n w_n \psi_n(\vec{r},\vec{p}) \tag{1}$$

where $w_n$ is related to the $n$th Ensemble charge weight function ($\sum_n w_n = 1$). Each Ensemble is described by 6 first-order moments of the distribution function $\psi_n(\vec{r},\vec{p})$:

$$\langle \xi \rangle = \int \xi \psi_n(\vec{r},\vec{p})\, \mathrm{d}\vec{r}\, \mathrm{d}\vec{p}$$

$$\langle p_v \rangle = \int p_v \psi_n(\vec{r},\vec{p})\, \mathrm{d}\vec{r}\, \mathrm{d}\vec{p} \tag{2}$$

*E-mail address:* mikhail.krasilnikov@desy.de.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.049

and by 21 moments of the second-order:

$$M_{\xi v} = \langle \Delta\xi \cdot \Delta v \rangle = \int \Delta\xi \cdot \Delta v \cdot \psi_n(\vec{r},\vec{p})\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}$$
$$M_{\xi p_v} = \langle \Delta\xi \cdot \Delta p_v \rangle = \int \Delta\xi \cdot \Delta p_v \cdot \psi_n(\vec{r},\vec{p})\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p}$$
$$M_{p_\xi p_v} = \langle \Delta p_\xi \cdot \Delta p_v \rangle = \int \Delta p_\xi \cdot \Delta p_v \cdot \psi_n(\vec{r},\vec{p})\,\mathrm{d}\vec{r}\,\mathrm{d}\vec{p} \quad (3)$$

where $\xi, v = \{x, y, z\}$ and $\Delta\xi = \xi - \langle\xi\rangle, \Delta v = v - \langle v\rangle$.

Assuming Ensemble energy spread small, one can obtain main equation for any Ensemble parameter $\mu = \{\xi, p_v, \Delta\xi \cdot \Delta v, \Delta\xi \cdot \Delta p_v, \Delta p_\xi \cdot \Delta p_v\}$ [1,2]:

$$\frac{\partial\langle\mu\rangle}{\partial\tau} = \left\langle \frac{\partial\mu}{\partial\vec{r}}\frac{\vec{p}}{\gamma} \right\rangle + \left\langle \frac{\partial\mu}{\partial\vec{p}}\frac{\vec{F}}{mc^2} \right\rangle \quad (4)$$

where $\gamma = \sqrt{1+\vec{p}\cdot\vec{p}}$ is the normalized energy, $\vec{p} = \vec{P}/(mc)$ is the normalized momentum, $\tau = ct$, and $\vec{F}$ is an applied Lorentz force.

Using Lorentz force expansion till linear terms

$$\frac{\vec{F}}{mc^2} = \overrightarrow{\mathbf{F}}(\langle\vec{r}\rangle, \langle\vec{p}\rangle) + \hat{\mathbf{F}}^X \cdot \Delta\vec{r} + \hat{\mathbf{F}}^P \cdot \Delta\vec{p} \quad (5)$$

where matrices

$$\hat{\mathbf{F}}^X_{ij} = \frac{1}{mc^2}\left.\frac{\partial\vec{F}_i}{\partial r_j}\right|_{\langle\vec{r}\rangle,\langle\vec{p}\rangle}, \quad \hat{\mathbf{F}}^P_{ij} = \frac{1}{mc^2}\left.\frac{\partial\vec{F}_i}{\partial p_j}\right|_{\langle\vec{r}\rangle,\langle\vec{p}\rangle} \quad (6)$$

one can obtain 6 time equations for the first-order moments $\langle\vec{r}\rangle, \langle\vec{p}\rangle$:

$$\frac{\mathrm{d}\langle\vec{p}\rangle}{\mathrm{d}\tau} = \overrightarrow{\mathbf{F}}(\langle\vec{r}\rangle, \langle\vec{p}\rangle),$$
$$\frac{\mathrm{d}\langle\vec{r}\rangle}{\mathrm{d}\tau} = \hat{\mathbf{W}} \cdot \langle\vec{p}\rangle, \quad (7)$$

and 21 equations for the second-order moments:

$$\frac{\mathrm{d}\hat{\mathbf{M}}_{pp}}{\mathrm{d}\tau} = \hat{\mathbf{F}}^X \cdot \hat{\mathbf{M}}_{xp} + \hat{\mathbf{F}}^P \cdot \hat{\mathbf{M}}_{pp} + (\hat{\mathbf{F}}^X \cdot \hat{\mathbf{M}}_{xp} + \hat{\mathbf{F}}^P \cdot \hat{\mathbf{M}}_{pp})^{\mathrm{T}}$$
$$\frac{\mathrm{d}\hat{\mathbf{M}}_{xp}}{\mathrm{d}\tau} = \hat{\mathbf{V}} \cdot \hat{\mathbf{M}}_{pp} + \hat{\mathbf{M}}_{xx} \cdot (\hat{\mathbf{F}}^X)^{\mathrm{T}} + \hat{\mathbf{M}}_{xp} \cdot (\hat{\mathbf{F}}^P)^{\mathrm{T}}$$
$$\frac{\mathrm{d}\hat{\mathbf{M}}_{xx}}{\mathrm{d}\tau} = \hat{\mathbf{M}}_{xp} \cdot \hat{\mathbf{V}} + \hat{\mathbf{V}} \cdot \hat{\mathbf{M}}^{\mathrm{T}}_{xp}. \quad (8)$$

Elements of the auxiliary matrices $\hat{\mathbf{W}}, \hat{\mathbf{V}}$ used in Eqs. (7) and (8) are

$$\hat{\mathbf{W}}_{ij} = \frac{1}{\gamma_m}\left(\delta_{ij} - \frac{\hat{\mathbf{M}}_{p_i p_j}}{\gamma_m^2}\right)$$
$$\hat{\mathbf{V}}_{ij} = \frac{1}{\gamma_m}\left(\delta_{ij} - \frac{\langle p_i\rangle\langle p_j\rangle}{\gamma_m^2}\right) \quad (9)$$

where $\gamma_m^2 = 1 + \sum_{n=x,y,z}(\langle p_n\rangle^2 + \hat{\mathbf{M}}_{p_n p_n})$ is squared normalized Ensemble energy.

## 3. Space charge implementation

Since the Ensemble Model implies an internal motion even in the case of the SEM, collective effects in beam dynamics can be simulated. Despite the beam emittance remains constant in the SEM (with linear Lorentz force) there is a good agreement in beam size and beam divergence simulation [2].

The space charge implementation makes an Ensemble charge distribution function an important issue. The rigorous problem reduces to the determining the stationary 3D charge distribution (which does not explicitly depend on time), which corresponds to the linear applied forces. The distribution in which the forces are linear and the phase space area remains constant is known as microcanonical distribution [5]. A homogeneous $(x, y)$ ellipsoidal beam distribution, known as K–V distribution leads to a perfect linear space charge force within the beam radius. The space charge algorithm for the SEM is based on the homogeneously charged ellipsoid. Calculation of the Lorentz force gradient at the Ensemble center is reduced to the obtaining a resulting force at small offsets from the homogeneously charged ellipsoid and integrating over the thin shell of uncompensated charges [1].

In the case of several Ensembles (MEM) it is necessary to calculate not only the space charge gradient at the center of the driving Ensemble, but also Lorentz force and its gradient at centers of other Ensembles. The most probable macroparticles configuration is a set of overlapping Ensembles. One of the algorithms, based on distribution function expansion is the Multi-Centered Gaussian Expansion (MCGE) discussed in Ref. [3]. This approach is based on the expansion of the Ensemble charge density in spatially distributed basis functions with known solutions of the field equation. The main advantage of this algorithm is a smoothness of the distribution function, but necessity of solution of linear equation system on each integration step for each Ensemble makes this approach comparatively slow.

The model of homogeneously charged 3D ellipsoid being very useful for the calculation of the space charge force gradient in the SEM can be extended to the MEM. The uniform ellipsoidal distribution is not a solution of the Poisson–Vlasov system because the corresponding stationary distribution in the phase space is singular; nevertheless, it allows one to keep Hamiltonian character of the 3D model similar to the dynamics of the full particle system (Liouville problem) [4].

## 4. Space charge field of an ellipsoidal Ensemble distribution

The distribution function of a 3-axis homogeneous ellipsoidal Ensemble is given by

$$\psi(x,y,z) = \begin{cases} \dfrac{3\gamma}{4\pi s_x s_y s_z} & \text{if } \dfrac{x^2}{s_x^2} + \dfrac{y^2}{s_y^2} + \dfrac{\gamma^2 z^2}{s_z^2} \leqslant 1 \\ 0 & \text{if } \dfrac{x^2}{s_x^2} + \dfrac{y^2}{s_y^2} + \dfrac{\gamma^2 z^2}{s_z^2} > 1 \end{cases} \quad (10)$$

where ellipsoid semi-axes and RMS sizes are defined by matrix elements (3): $s_x^2 = 5\sigma_x^2 = 5M_{xx}$, $s_y^2 = 5\sigma_y^2 = 5M_{yy}$, $s_z^2 = 5\gamma^2\sigma_z^2 = 5\gamma^2 M_{zz}$. From Newton's potential theory [6] the electric field $\vec{E}$ in the Ensemble's rest frame is given:

$$E_i = \frac{qr_i}{s_x s_y s_z}\tilde{G}_i(a_x, a_y, a_z) \tag{11}$$

where $q$ denotes the Ensemble charge, $a_i = s_i^2 + \lambda$ is a square of the equivalent confocal ellipsoid semi-axis, parameter $\lambda = 0$ for internal point of the ellipsoid, otherwise, for an external point $\lambda$ can be determined as the positive root of the equation:

$$\frac{x^2}{s_x^2+\lambda} + \frac{y^2}{s_y^2+\lambda} + \frac{\gamma^2 z^2}{s_z^2+\lambda} = 1. \tag{12}$$

This equation determines one and only one ellipsoid passes through any point $(x, y, z)$ outside the ellipsoid. A geometrical form factor $\tilde{G}_i$:

$$\tilde{G}_i = \frac{s_x s_y s_z}{\sqrt{a_x a_y a_z}} G\left(\frac{a_i}{a_j}, \frac{a_i}{a_k}\right) \quad \text{where}$$

$$G(p,q) = 3 \cdot \int_0^1 \frac{v^2\,\mathrm{d}v}{\sqrt{(p+(1-p)v^2)(q+(1-q)v^2)}} \tag{13}$$

is constant for the internal point of the ellipsoid, so as is well known, the electric field is linear inside the ellipsoid. Here $\{i, j, k\}$ defines any permutation of the indices $\{x, y, z\}$. For external point $a_i = f_i(x, y, z)$ and $\tilde{G}_i$ determines a field decay with an offset from the Ensemble center. Eq. (12) can be interpreted in the following way: at any external point the electric field generated by an ellipsoidal uniform charge distribution is equivalent to the electric field generated by a confocal uniformly charged ellipsoid passing through the point $(x, y, z)$.

The geometrical form factor as a consequence of the Gaussian theorem for the electric field satisfies the equality:

$$G\left(\frac{a_x}{a_y}, \frac{a_x}{a_z}\right) + G\left(\frac{a_y}{a_x}, \frac{a_y}{a_z}\right) + G\left(\frac{a_z}{a_x}, \frac{a_z}{a_y}\right) = 3. \tag{14}$$

Assuming that the driving Ensemble has energy $E = mc^2\gamma = mc^2/\sqrt{1-\beta^2}$ and the center coordinate $\langle \vec{r}^d \rangle = \{x^d, y^d, z^d\}$, the Lorentz force acting from the driving Ensemble on the test particle $\{x^t, y^t, z^t, \beta_x^t = v_x^t/c, \beta_y^t = v_y^t/c, \beta_z^t = v_z^t/c\}$ is given by

$$\begin{aligned}
F_x^{\mathrm{d}\to t} &= e\gamma(1-\beta\beta_z^t)\cdot E_x^d(\vec{\Delta}^{dt}) \approx eE_x^d(\vec{\Delta}^{dt})/\gamma \\
&= \frac{eq\cdot(x^t-x^d)}{5^{3/2}\gamma^2\sigma_x\sigma_y\sigma_z}\tilde{G}_x \\
F_y^{\mathrm{d}\to t} &= e\gamma(1-\beta\beta_z^t)\cdot E_y^d(\vec{\Delta}^{dt}) \approx eE_y^d(\vec{\Delta}^{dt})/\gamma \\
&= \frac{eq\cdot(y^t-y^d)}{5^{3/2}\gamma^2\sigma_x\sigma_y\sigma_z}\tilde{G}_y \\
F_z^{\mathrm{d}\to t} &= eE_z^d(\vec{\Delta}^{dt}) + e\beta\gamma(\beta_x^t E_x^d(\vec{\Delta}^{dt}) + \beta_y^t E_y^d(\vec{\Delta}^{dt})) \approx eE_z^d(\vec{\Delta}^{dt}) \\
&= \frac{eq\cdot(z^t-z^d)}{5^{3/2}\sigma_x\sigma_y\sigma_z}\tilde{G}_z
\end{aligned} \tag{15}$$

where vector $\vec{\Delta}^{dt} = \{x^d - x^t, y^d - y^t, (z^d - z^t)\cdot\gamma\}$ takes into account Lorentz transformation for the coordinates. The matrix $\hat{\mathbf{F}}^X$ is given by

$$\begin{aligned}
\hat{\mathbf{F}}^X = \frac{eq}{mc^2}\frac{1}{\gamma^2\sigma_x\sigma_y\sigma_z} \\
\times\left\{\begin{pmatrix}\tilde{G}_x & 0 & 0\\ 0 & \tilde{G}_y & 0\\ 0 & 0 & \gamma^2\tilde{G}_z\end{pmatrix} + \begin{pmatrix}(x'-x^d)\cdot\tilde{G}_x^1\\ (y'-y^d)\cdot\tilde{G}_y^1\\ z'-z^d)\gamma^2\tilde{G}_z^1\end{pmatrix}\right. \\
\left.\times\begin{pmatrix}\frac{\partial\lambda}{\partial x} & \frac{\partial\lambda}{\partial y} & \frac{\partial\lambda}{\partial z}\end{pmatrix}\right\}
\end{aligned} \tag{16}$$

where $\tilde{G}_i^1 = \sum_{j=x,y,z}\frac{\partial\tilde{G}_i}{\partial a_j}$. Geometrical form factors can be tabulated or obtained by using some analytical approximation.

## 5. Simulation of the space charge dominated beam

For the illustration of the proposed space charge algorithm a space charge dominated electron beam (1 nC, 5 MeV) in a drift space has been simulated using Ensemble Model in comparison with conventional tracking code (ASTRA) [8]. Initial beam transverse phase space of the beam is shown in Fig. 1, where equivalent phase space ellipses depict Ensemble parameters (Fig.1b).

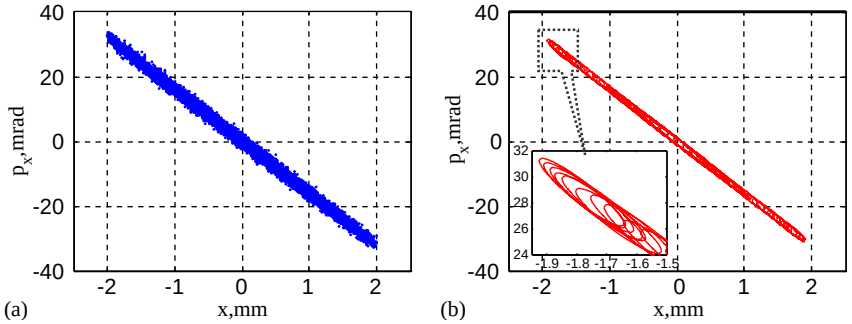



Fig. 1. Initial transverse phase space $(x, p_x)$: (a) Conventional tracking code (ASTRA), 10,000 macroparticles; (b) Ensemble Model (50 Ensembles).

Transverse RMS beam size, beam divergence and normalized beam emittance as functions of a flight time are shown in Fig. 2 for the case without space charge. The agreement between conventional tracking code (ASTRA) and Ensemble Model (even using single Ensemble) is very good, whereas computation time is much smaller for the Ensemble Model (10 variables for SEM vs. 40,000 ASTRA transverse particle coordinates).

Corresponding dependencies for the case with space charge are shown in Fig. 3. The beam size as well as divergence can be simulated with SEM with rather good agreement, but for the emittance simulations MEM should be used.

Resulting phase spaces are shown in Fig. 4. It should be noted that a discrepancy in emittance is caused mainly by a non-perfect interface between conventional macroparticles (10000 ASTRA particles) and Ensembles (50 Ensembles in MEM). To illustrate the influence of the Ensemble number

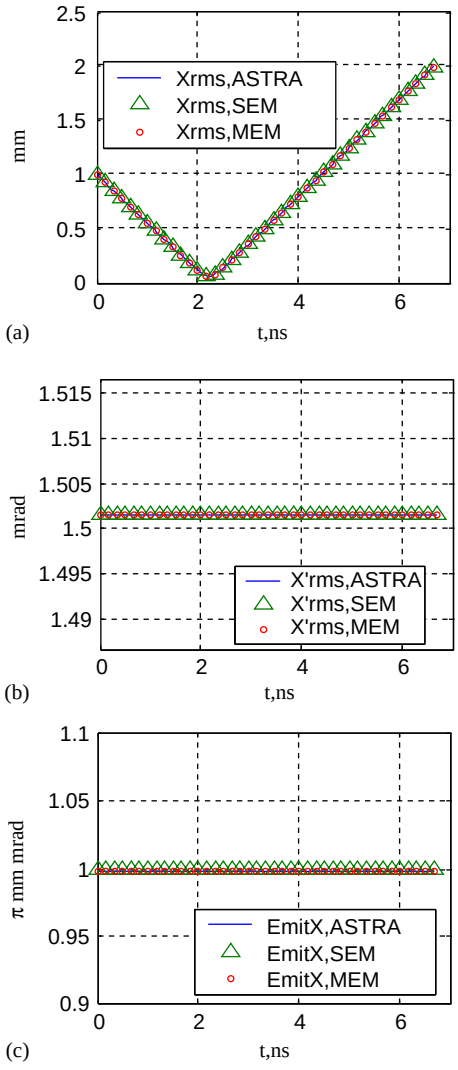



Fig. 2. Simulations of the electron beam in a drift using conventional tracking code (ASTRA), SEM and MEM (50 Ensembles) without space charge. (a) RMS beam size; (b) RMS beam divergence; (c) RMS beam emittance.

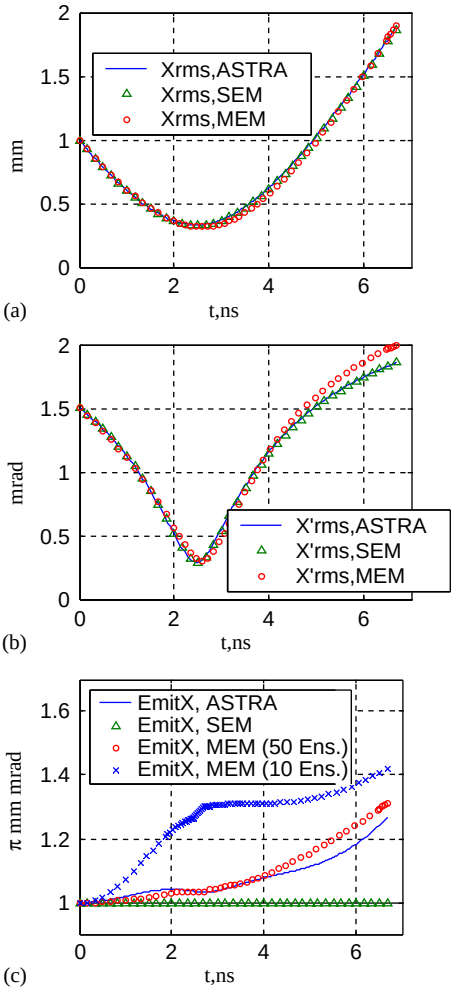



Fig. 3. Simulations of the electron beam in a drift using conventional tracking code (ASTRA), SEM and MEM with space charge. (a) RMS beam size; (b) RMS beam divergence; (c) RMS beam emittance.

the beam emittance simulated with 10 Ensembles is shown in Fig. 3c as well.

As it can be seen from Fig. 3b, the RMS beam divergence decreases till the beam waist ($t \approx 2.5$ ns), whereas for the case without space charge it is constant (Fig. 2b). It should be figured out that this takes place even for the linear space charge algorithm (SEM). From expression (25) for the space charge force one can obtain the matrix element $\hat{\mathbf{F}}^{A}_{xx} = \kappa/M_{xx}$, where $\kappa = (2/5\beta\gamma)(I/I_A)$. Differential equations for transverse phase space take a form:

$$\begin{aligned}\frac{\mathrm{d}M_{p_xp_x}}{\mathrm{d}\tau} &= 2\kappa \cdot \frac{M_{xp_x}}{M_{xx}}\\ \frac{\mathrm{d}M_{xp_x}}{\mathrm{d}\tau} &= \frac{M_{p_xp_x}}{\gamma_m} + \kappa\\ \frac{\mathrm{d}M_{xx}}{\mathrm{d}\tau} &= \frac{2}{\gamma_m} \cdot M_{xp_x}.\end{aligned} \tag{17}$$

Without solving this nonlinear system one can obtain the following integral:

$$M_{p_xp_x}(\tau) = M_{p_xp_x}(0) + \kappa\gamma_m \ln\frac{M_{xx}(\tau)}{M_{xx}(0)}. \tag{18}$$

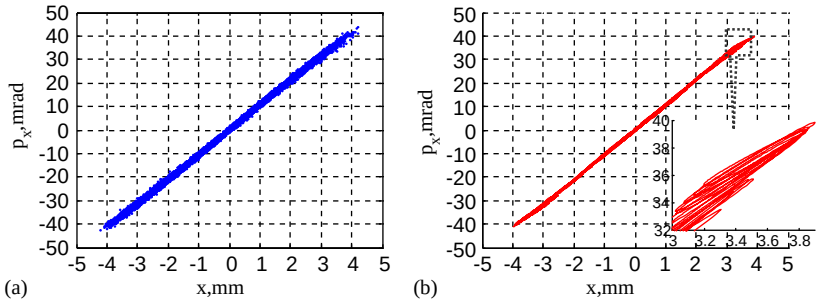



Fig. 4. Final transverse phase space: (a) Conventional tracking code (ASTRA); (b) Ensemble Model (50 Ensembles).

Hence, applying a space charge force under definite (negative) phase space correlation the logarithmic term in Eq. (18) is negative (before the focus $M_{xx}(\tau) < M_{xx}(0)$) and RMS beam divergence decreases as a result of the space charge effect.

## 6. Conclusions

Space charge algorithm for the Multi Ensemble Model (MEM) based on the 3-axis homogeneously charged ellipsoid has been developed. Analytical expressions for the Lorentz force and its gradients have been obtained. MEM simulations of the nonlinear effects in transverse phase space of the space charge dominated beam showed good agreement with conventional tracking code, demonstrating advantage in variable number and computational efficiency.

## References

[1] A. Novokhatski, T. Weiland, PAC'99, New York, March 1999.
[2] M. Krassilnkov, et al., ICAP 2000, Darmstadt, September 2000.
[3] M. Krassilnikov, T. Weiland, ICAP'02, East Lansing, USA, October 2002.
[4] M. Comunian, et al., Phys. Rev. Spec. Topics—Acc. Beams 4 (2001) 124201.
[5] I.M. Kapchinskiy, Theory of Resonance Linear Acceler., Harw. Ac. Pub., 1985.
[6] W.J. Sternberg, T.L. Smith, The Theory of Potential and Spherical Harmonics, Harmonics, University of Toronto Press, 1952.
[8] ASTRA User Manual, http://www.desy.de/~mpyflo/Astra_documentation/.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 74–76
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Features of electron dynamics at different points of a bending magnet

O.E. Shishanin*

*Moscow State Industrial University, 109280 Moscow, Avtozavodskaya St. 16, Russian Federation*

Available online 1 December 2005

**Abstract**

Peculiarities of the charged particle motion in a discrete magnetic field is described analitically. A need for considering this problem follows the difference between the synchrotron radiation properties in accelerators with straight sections on the boundaries and these properties inside a magnet.


*PACS:* 41.70.+t; 29.20.−c

*Keywords:* Dynamics; Magnet; Asymptotics; Oscillations; Amplitude; Radiation

The notion of long wavelength radiation on the boundary of straight sections and curved lines was discussed in special cases [1,2]. The ultrarelativistic electron radiation in synchrotrons is of interest to this paper. Synchrotron radiation in accelerators and storage rings was previously determined under the assumption that rotations of electrons are averaged. In this case, it was found that the emission depends significantly on the vertical betatron oscillations [3–5]. This coincides with experimental data.

Let an electron revolve in a magnetic system consisting of $N = 4$ periods. One element of the system contains one-half of a free gap of length $l$, a magnet of length $a = 2\pi R/N$ ($R$ is the radius of curvature) and again a part of a straight section.

If the origin of a cylindrical coordinate system is located symmetrically inside an accelerator, the coordinate $r$ can be represented as

$$r = \begin{cases} l_1/\cos\varphi, \varphi \in [0, \arctan(1/b)) \\ (l_2/R + R\sin\phi)/\sin\varphi, \varphi \in [\arctan(1/b), \arctan b), \\ \quad \phi \in [0, \pi/2] \\ l_1/\sin\varphi, \varphi \in [\arctan b, \pi/2] \end{cases}$$

where $\varphi$ is the azimuth angle, $b = l_1/(l/2)$, $l_1 = l/2 + R$, $l_2 = \sqrt{(l/2)^2 + l_1^2}$.

However, a Fourier-expansion of $r$ leads to double series and extra difficulties in the subsequent calculations can arise.

On the other hand, the components $H_r$ and $H_z$ of the magnetic field (or the gradient $n$ of the magnetic field) can be developed in a Fourier series on an interval $\varphi \in [0, \pi/2]$ in the form $H(\tau) = H \cdot f(\tau)$, where

$$f(\tau) = \frac{2}{\pi}p + f_1(\tau), \quad f_1(\tau) = \frac{2}{\pi}\sum_{v=1}^{\infty}\frac{(-1)^v}{v}\sin 2vp\cos v\tau,$$

$$p = \arctan((b^2-1)/2b), \quad \tau = N\varphi,$$
$$2p/\pi \approx 1/(1+k), \quad k = l/a.$$

The operator $f(\tau)$ "switches on" and "switches off" the magnetic field. Physically, of some interest is an area which is close to $\varphi \in [\arctan(1/b), \arctan b]$. Taking into account the radial betatron oscillations, we have $r = l_2 + \rho(\phi = 0)$ at the edge of magnet, $r = R + l/\sqrt{2} + \rho(\phi = \varphi = \pi/4)$ at the midpoint. In the first case the equations of small betatron vibrations in the linear approximation take the form:

$$\frac{d^2z}{d\tau^2} + \frac{(1+k)}{N^2}\frac{l_2}{R}nf(\tau)z = 0 \tag{1}$$

*Tel.: +7 95 2772620; fax: +7 95 6752256.
*E-mail address:* shisha-n@msiu.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.017

$$\frac{d^2\rho}{d\tau^2}+\frac{1}{N^2}\left[1-n\frac{l_2}{R}(1+(1+k)f_1(\tau))\right]\rho - \frac{1+k}{N^2}\int \dot{\rho} f_1(\tau)\,dt = -\frac{1+k}{N^2} l_2 f_1(\tau) \quad (2)$$

under the condition that angular velocity can be represented as follows:

$$\dot{\varphi}=\omega_0\left[\frac{1}{1+k}\left(1-\frac{\rho}{l_2}\right)+\frac{1}{l_2}\int \dot{\rho} f_1(\tau)\,dt\right] \quad (3)$$

where $\omega_0 = e_0 H_0/(mc)$ and $H_0 \sim R^{-n}$, $(n \in (0,1))$.

The solution of Eq. (1) we will seek in the form $z = \exp(i\gamma_z\tau)\cdot\varphi_z(\tau)$, where $\varphi_z(\tau)$ is a periodic function. This procedure is close to the Whittaker method and the technique of stretched parameters for the Mathieu equation in Ref. [6]. The function $\varphi_z(\tau)$ must satisfy a new differential equation

$$\frac{d^2\varphi_z}{d\tau^2}+2i\gamma_z\frac{d\varphi_z}{d\tau}-\gamma_z^2\varphi_z+\frac{n_1}{N^2}f(\tau)\varphi_z=0$$

where $n_1 = (1+k)l_2 n/R$; the solution being defined as the asymptotic series

$$\varphi_z(\tau)=\varphi_0(\tau)+\sum_{i=1}^{\infty}\varphi_i(\tau)/N^i, \quad \gamma_z=\sum_{i=1}^{\infty}\gamma_i/N^i.$$

Here, the coefficients of the same powers of the parameter $1/N$ are equal to zero. Then the chain of equations can be derived as follows:

$$\ddot{\varphi}_0=0, \quad \ddot{\varphi}_1+2i\gamma_1\dot{\varphi}_0=0,$$
$$\ddot{\varphi}_2+2i\gamma_1\dot{\varphi}_1-\gamma_1^2\varphi_0+n_1 f(\tau)\varphi_0=0$$

$$\ddot{\varphi}_3+2i\gamma_1\dot{\varphi}_2+2i\gamma_2\dot{\varphi}_1+2i\gamma_3\dot{\varphi}_0-\gamma_1^2\varphi_1-2\gamma_1\gamma_2\varphi_0 + n_1 f(\tau)\varphi_1=0 \text{ and so on.}$$

The secular terms are eliminated in the solutions of these equations.

This subsequently gives: $\varphi_0 = b, \varphi_1 = b_1$, where $b, b_1$ are constants:

$$\gamma_1=\sqrt{n\frac{l_2}{R}}, \quad \varphi_2=\frac{2n_1}{\pi}b\sum_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu^3}\sin 2\nu p\cos\nu\tau,$$
$$\gamma_2=0, \text{ etc.}$$

Eventually, the asymptotics up to $1/N^4$ can be written in the form

$$z = B\cos\tau_z\left[1+\frac{2n_1}{\pi N^2}\sum_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu^3}\sin 2\nu p\cos\nu\tau + \frac{8\gamma_1^2 n_1}{\pi N^4}\sum_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu^5}\sin 2\nu p\cos\nu\tau + \frac{2n_1^2}{\pi^2 N^4}\sum_{\mu=1}^{\infty}\frac{(-1)^\mu}{\mu}\sum_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu^3}\sin 2\mu p\sin 2\nu p \times\left(\frac{1}{(\mu-\nu)^2}\cos(\mu-\nu)\tau+\frac{1}{(\mu+\nu)^2}\cos(\mu+\nu)\tau\right) + \frac{n_1^2}{2\pi^2 N^4}\sum_{\nu=1}^{\infty}\frac{\sin^2 2\nu p}{\nu^6}\cos 2\nu\tau\right] + B\sin\tau_z\frac{4\gamma_1 n_1}{\pi N^3} \times\sum_{\nu=1}^{\infty}\frac{(-1)^\nu}{\nu^4}\sin 2\nu p\sin\nu\tau, \quad \mu\neq\nu \quad (4)$$

where $\tau_z = (v_z/N)\tau+\psi$, $B$ is the amplitude of axial oscillations, and $\psi$ is an initial phase. The frequency can be constructed in the form $v_z^2 = \gamma_1^2 + 2\gamma_1\gamma_3/N^2$ or

$$v_z^2 = n\frac{l_2}{R}+\left(\frac{l_2}{R}\right)^2\frac{\pi^2 n^2 k^2}{3(1+k)^2 N^2}.$$

By analogy, asymptotics for radial oscillations can be derived similarly. The total velocity after averaging over the period is constant.

In these formulas for the middle of magnet the parameter $l_2$ may substituted by $R + l/\sqrt{2}$.

These results will be used for an analysis of the radiation intensity. Expressions for spectral and angular distributions of the linear polarization components, obtained in Refs. [4,5], are

$$\frac{dW_\sigma(\nu)}{d\Omega}=\frac{ce^2\nu\nu'}{12\pi^4 R^2}\int_0^{2\pi}d\delta\varepsilon_1^2 K_{2/3}^2\left(\frac{\nu'}{3}\varepsilon_1^{3/2}\right),$$

$$\frac{dW_\pi(\nu)}{d\Omega}=\frac{ce^2\nu\nu'}{12\pi^4 R^2}\int_0^{2\pi}d\delta\varepsilon_1(\cos\theta-\alpha\cos\delta)^2 K_{1/3}^2 \times\left(\frac{\nu'}{3}\varepsilon_1^{3/2}\right) \quad (5)$$

where

$$\nu' = \nu(1+\hbar\omega/E), \quad \omega = \nu\omega_0, \quad \omega_0=\beta\frac{c}{R}$$

$$\varepsilon_1 = \varepsilon - 2\alpha\cos\theta\cos\delta+\alpha^2\cos^2\delta, \quad \varepsilon = 1-\beta^2\sin^2\theta.$$

The parameter $\alpha\cos\delta$ is equal to $\beta_z|_{\tau=0}$ (the slope of the electron velocity vector), where $\alpha$ has specific value for different magnetic structure.

In the more important case of storage rings we have

$$\alpha=\sqrt{\frac{A_z}{\pi}}\left[\frac{1}{\sqrt{\beta_z}}\sqrt{1+\left(\frac{1}{2}\frac{d\beta_z}{ds}\right)^2}\right]_{\varphi=0}$$

where $A_z$ is the emittance and $\beta_z$ is the beta function of vertical motion depending on orbital length $s$. Here, the azimuthal angle is recorded from the point from which the radiation is collected.

Eq. (4) for $z$, taking up to $1/N^3$, after differentiating can be represented as

$$\frac{dz}{dt} = -Bv_z\frac{\omega_0}{1+k}\sin\tau_z\left[1-\frac{2n_1}{\pi N^2}\sum_{v=1}^{\infty}\frac{(-1)^v}{v^3}\sin 2vp\cos v\tau\right] - B\frac{2n_1}{\pi N}\frac{\omega_0}{1+k}\cos\tau_z\sum_{v=1}^{\infty}\frac{(-1)^v}{v^3}\sin 2vp\sin v\tau \quad (6)$$

where accuracy is $1/N^2$.

The edge of magnet corresponds the angle $\varphi = \arctan(1/b)$ and $r$ is close to parameter $l_2$. In this case $\alpha$ approximately equals to

$$\sqrt{n}(B/R)(l_2/R)/(1+k).$$

At a centrally located points $\varphi = \pi/4$ and

$$\alpha \approx \sqrt{n}B/(R(1+k))(1+l/(2\sqrt{2}R)).$$

The parameter $B$ would be more properly selected from experimental data for a concrete position of particle. Nevertheless, these calculations revealed the correction factors for amplitude $B$.

It is necessary to note that the betatron oscillations significantly change spectral-angular and angular distributions. Quantum corrections in Eq. (5) do not play an essential role in the formation of synchrotron radiation.

Thus, the emission problem is resolved for a monoenergetic electron and the bunch effect is taken into account by averaging over the initial phases since injection of the particles is continued during some turns. In addition, $B$ is a mean square quantity because in any cross-section of the beam there is a set of variable amplitudes.

In Eq. (5) parameters $\alpha$, $\sqrt{1-\beta^2}$, and $\cos\theta \approx \psi$ ($\psi$ is the conventional term in literature) are small quantities of the same order while corrections of the order of $B^2/R^2$, $A^2/R^2$ ($A$ is the amplitude of radial oscillations) with respect to main terms were already omitted. At the $\theta \sim \pi/2$ integrals in Eq. (5) can be evaluated by passing to the Airy functions and expanding them in power series.

## References

[1] B.M. Bolotovsky, A.V. Serov, Zh. Exp. Teor. Phys. 105 (1994) 43.
[2] E.G. Bessonov, Nucl. Instr. and Meth. A 308 (1991) 135.
[3] O.E. Shishanin, in: Proceedings of the High Energy Accelerators Conference HEACC'98, Dubna, JINR, 1998, pp. 190, 193
[4] O.E. Shishanin, Zh. Tech. Fhys. 68 (1998) 133.
[5] O.E. Shishanin, J. Exp. Theor. Phys. 90 (2000) 725.
[6] A.H. Nayfen, Introduction to Perturbation Techniques, Wiley, New York, 1981.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 77–84
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Precision modeling of the internal injection and beam dynamics for a high-power RF accelerator

M.A. Tiunov*, V.L. Auslender, A.V. Ivanov, M.M. Karliner, G.I. Kuznetsov, I.G. Makarov, A.D. Panfilov, V.V. Tarnetsky

*Budker INP, Novosibirsk, Russian Federation*

Available online 1 December 2005

**Abstract**

A new high-power electron accelerator for industrial applications has been developed in Novosibirsk. The main parameters of the accelerator are: operating frequency of 178 MHz, electron energy of 5 MeV, and average beam power of up to 300 kW. The accelerator consists of a chain of accelerating cavities connected by on-axis coupling cavities with coupling slots in the walls. A triode RF gun on the base of a grid-cathode unit placed on the wall of the first accelerating cavity is used for internal injection of electrons. The paper presents the results of modeling and optimization of the accelerating structure, internal injection, and beam dynamics. A modification of the electrostatic computer code SAM designed for simulation of beam dynamics in linear RF accelerators by the long-wave approximation method is described.


*PACS:* 29.17.+w; 29.27.−a; 41.85.−p

*Keywords:* Linear accelerator; Electron beam; Injection; Dynamics

## 1. Introduction

Currently, there is interest in X-ray radiation technologies because of the high penetration ability of X-rays. This is of particular importance for the pasteurization of a wide spectrum of food products, disinfection of mail deliveries, and other applications. However, because of the low efficiency of X-ray conversion for electrons with energy below 5 MeV, the intensity of X-rays required for some industrial applications can be achieved using 5-MeV electron beams with power up to 300 kW [1].

A new RF accelerator for industrial applications, the ILU-12, is now being developed in BINP, Novosibirsk. The main parameters of the accelerator are: operating frequency of 178 MHz, electron energy of 5 MeV, and average beam power of up to 300 kW [2]. A general view of the accelerator is shown in Fig. 1.

The electrons are accelerated in a low-frequency multi-resonator standing wave structure with axis-disposed coupling resonators. This design makes it possible to decrease power losses in each resonator compared with a single-resonator accelerator (at the same average beam power) and to obtain accelerator electron efficiency of approximately 70%.

It is proposed to use internal beam injection from the grid-cathode unit, which is placed directly after the first accelerating gap. This concept permits sufficient simplification of the design, reducing the cost of the accelerator, respectively, as well as improving its reliability and reducing maintenance charges.

To successfully design the ILU-12 accelerator, the following problems have to be solved:

- Achievement of the required pulse beam current at relatively low electric field strength in the accelerating gaps compared with a single-resonator accelerator; and
- Lossless transportation of a powerful electron beam through the accelerating structure without the use of electro- and magnetostatic lenses.

*Corresponding author.
*E-mail address:* m.a.tiunov@inp.nsk.su (M.A. Tiunov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.077

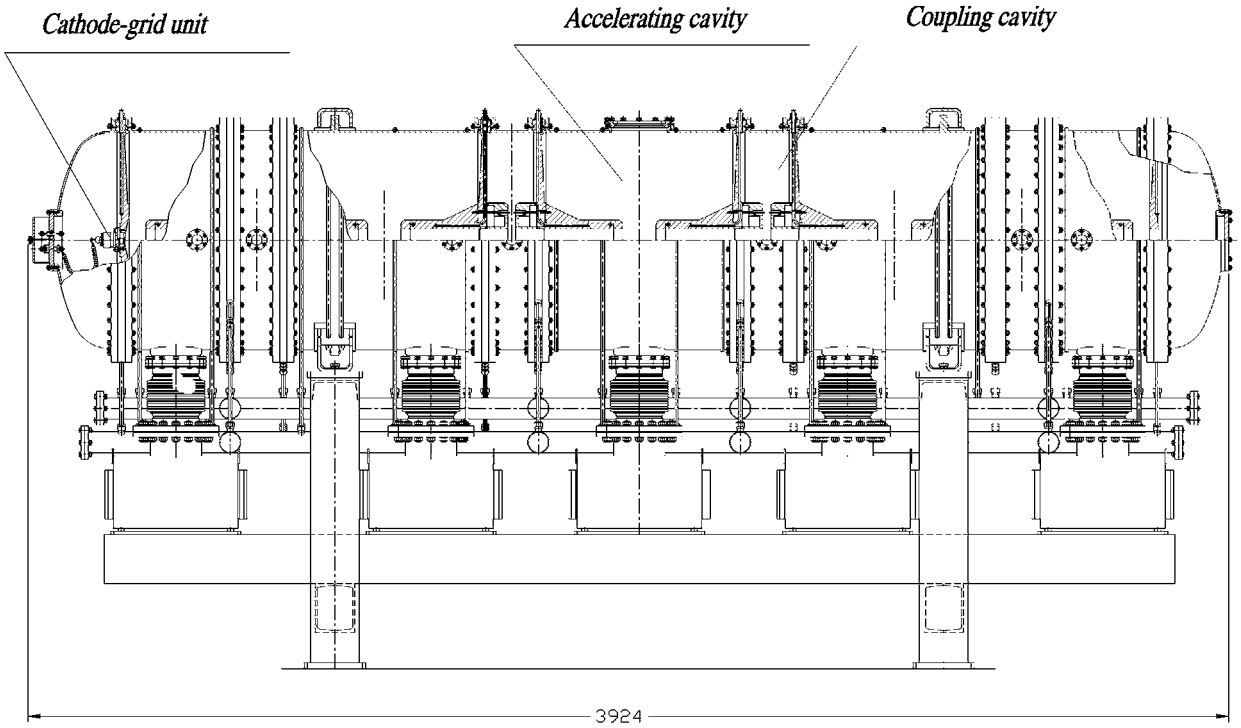



Fig. 1. General view of the accelerator.

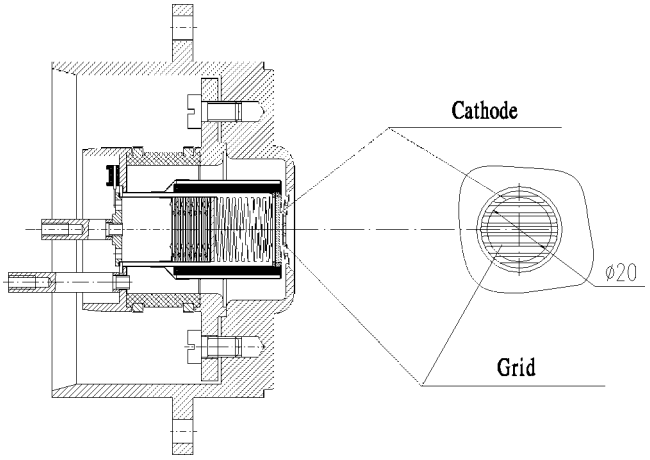



Fig. 2. General view of the grid-cathode unit.

To solve these problems, a new grid-cathode unit was designed. A general view of this important component is shown in Fig. 2. The report describes the methods and results for detailed simulation of the internal injection and beam dynamics in the ILU-12 accelerator.

Computer simulations were performed in long-wave approximation using the SAM code [3], modified for a solution of the problems mentioned above. Use of standard SUPERLANS [4] and PARMELA [5] codes is impossible because of the necessity for exact modeling of the internal beam injection.

## 2. RF field simulation using long-wave approximation

Generally, the amplitude of electric and magnetic fields in the accelerating structure is described by the appropriate Maxwell equation

$$\text{rot}\,\vec{E} = -jk \times \vec{B} \tag{1}$$

$$\text{rot}\,\vec{B} = jk \times \vec{E} \tag{2}$$

where $k = 2\pi/\lambda$.

However, since the aperture radius $a = 35\,\text{mm}$ and the maximum accelerating gap $l = 110\,\text{mm}$ in the accelerating structure are well below the wavelength $\lambda \approx 1.7\,\text{m}$, the right-hand side of the former of these equations may be omitted. This means that the static electric field, which can be easily calculated using the SAM program, may be used as the RF electric field amplitude in the accelerating gap.

From Eq. (2), the magnetic RF field in the accelerating gaps, which has only an azimuthal component and has a notable influence on the relativistic electron dynamics, may be calculated with acceptable accuracy in a long-wave approximation. Subject to the field cylindrical symmetry, Eq. (2) may be rewritten in the following integral form:

$$B_\theta(r,z) = j\left(\frac{k}{r}\right)\int_0^r E_z(r',z)r'\,\mathrm{d}r' \tag{3}$$

where $j$ is an imaginary unit that represents the magnetic field phase shift by 90° relative to the electric RF field, and $E_z$ is the longitudinal component of the static electric field.

The following estimation of the maximum relative error for the electric field long-wave approximation within the cavity accelerating gap aperture may be derived from

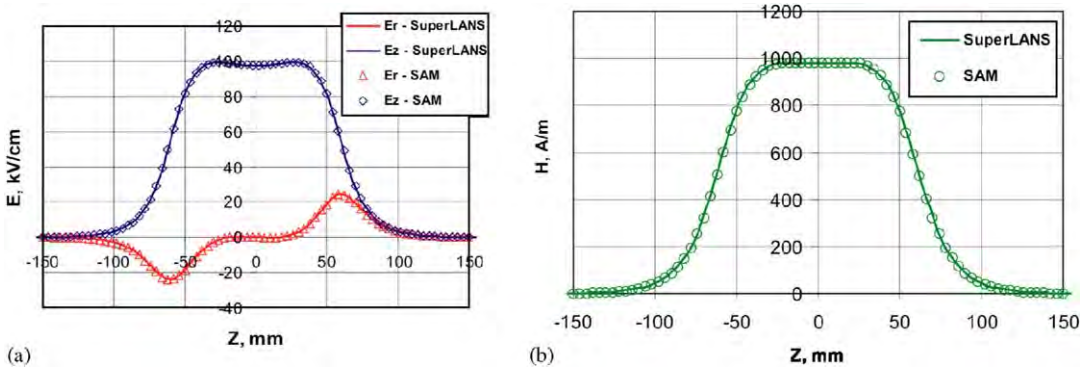



Fig. 3. Results of simulation of the longitudinal distribution of electromagnetic fields in full accelerating gap at 20-mm radius using SuperLANS and SAM codes: (a) electric field components; and (b) azimuthal magnetic field.

Eqs. (1) and (3)

$$\left(\frac{\delta E}{E}\right)_{\max} \approx \frac{(ka)^2}{4}. \tag{4}$$

For the accelerating structure chosen, this estimation is approximately 0.4%. Eqs. (3) and (4) give the following estimation for the maximum relative error for the magnetic field calculation in long-wave approximation within the cavity accelerating gap aperture:

$$\left(\frac{\delta B}{B}\right)_{\max} \approx \frac{(ka)^3}{8}. \tag{5}$$

For the accelerating structure chosen, this estimation is approximately 0.025%.

Fig. 3a shows a comparison of the longitudinal distributions in a full accelerating gap (20-mm radius) of the electric RF field components calculated by the SuperLANS program [4] designed for RF field simulations, and a static electric field calculated by the SAM program. The relative difference in electric field values does not exceed 0.2% and is in good agreement with Eq. (4) estimation.

Fig. 3b presents a comparison of longitudinal distributions in a full accelerating gap (20-mm radius) of the magnetic RF field calculated by the SUPERLANS program, magnetic field calculated by Eq. (3), and static electric field calculated by the SAM program. According to Eq. (5), the relative difference in the magnetic field value is really one order less than for the electric field and is defined by the field approximation accuracy over the mesh.

For beam injection and dynamics simulation, the whole accelerating structure aperture from the grid-cathode unit up to the accelerator output is covered by a set of uniform rectangular meshes closely attached to one another. First, the static electric field is calculated once in all the mesh nodes; this field is used as the amplitude of the electric RF field. Then the magnetic RF field amplitude is calculated according to Eq. (3). Biquadratic approximation over the node values is used for field calculation at an arbitrary point. The mesh set allows us to precisely describe RF fields near the grid-cathode unit, as well as in each of the accelerating gaps for a minimal number of nodes in all meshes.

## 3. Internal injection simulation

### 3.1. 2D model of the grid-cathode unit cell

Fig. 4 presents the grid-cathode unit model for simulation. The main geometric parameters of the unit are: cathode-grid gap $d$, grid step $h$, and wire diameter $D$. At the accelerating field value of 80 kV/cm in the grid plane, the maximum beam pulsed current of 7 A required is provided at equipotential grid and cathode and the following geometric parameters of the grid-cathode unit: $d = 1.5$ mm, $h = 3$ mm, and $D = 1$ mm. All the grid-cathode unit simulation results presented below were obtained using these geometrical parameters. The computer optimization of these parameters was carried out using the ULTRASAM program [6] in static approximation subject to the effect of emission limitation by the beam space charge.

Fig. 5 shows the simulation results for a single cell of the grid-cathode unit at equipotential grid and cathode. The flow characteristics from a single grid-cathode unit cell at 1-mm distance from the grid surface were used as starting conditions in the beam dynamics simulation. However, as may be observed from Fig. 5b, in this cross-section the current density is strongly non-uniform, and the flow phase curve is a multiple-value function. Therefore, separate descriptions of the start coordinate dependence of the current density at the cathode (Fig. 5a), output radius, and flow trajectory slope angle were used. Such an approach was previously used in the BEAM program [7]. These dependences are always single-value functions and may be easily approximated by a power series using the least-squares method (Fig. 6). For a single cell, the average grid-cathode gap transit time and electron output energy were also calculated.

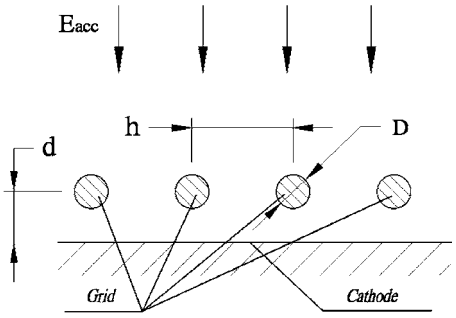



Fig. 4. Grid-cathode unit simulation model.

Simulations of the cell static characteristics listed above were carried out for a set of node values of the control cathode voltage at an accelerating field of 80 kV/cm in the grid plane. For calculation of the cell static characteristics at arbitrary values of the control cathode voltage, linear interpolation was used.

Fig. 7 presents calculation curves for the average density of the current emitted from the grid-cathode unit and the maximum electron relative transverse velocity in the grid plane versus the grid-cathode voltage.

A prototype of the proposed grid-cathode unit was designed, manufactured and tested. The results of

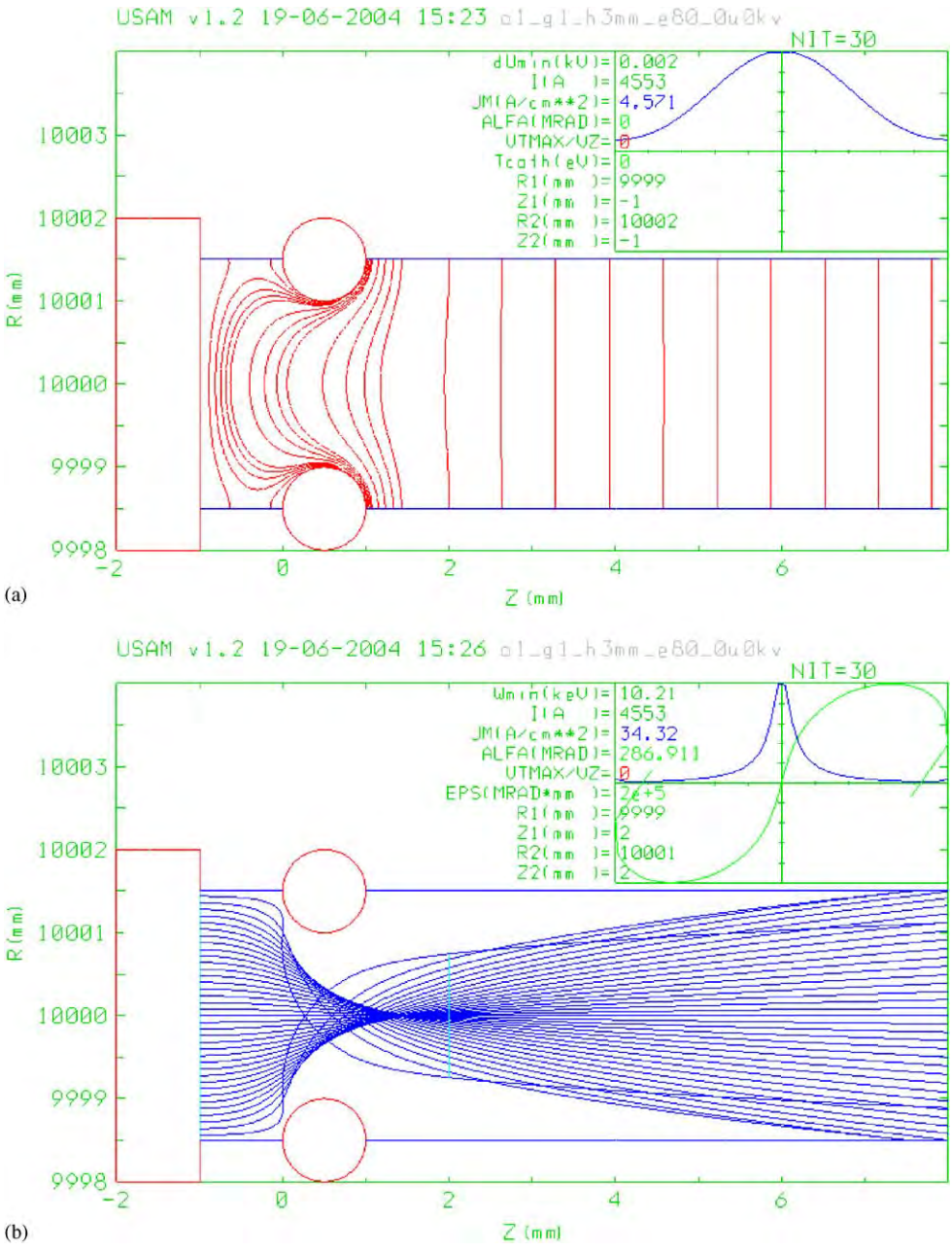



Fig. 5. Single cell of the grid-cathode unit: (a) equipotentials and current density distribution at the cathode; and (b) electron trajectories, current density distribution and flow phase curve at 1 mm from the grid surface.

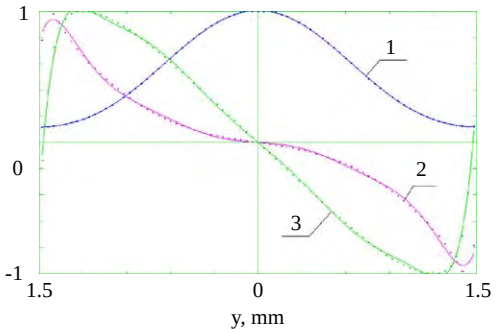



Fig. 6. Normalized cathode coordinate dependence of the main characteristics of the electron flow at zero cathode voltage (dots represent calculated values, continuous lines are approximations): 1, cathode current density; 2, output radii; and 3, output angles.

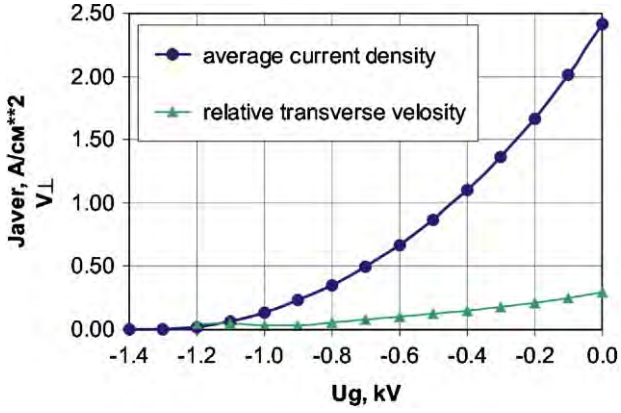



Fig. 7. Grid-cathode potential dependence of average current density and relative transverse velocity for accelerating field strength of 80 kV/cm.

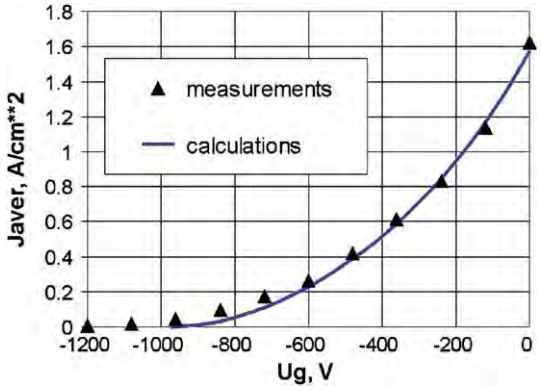



Fig. 8. Experimental and calculated grid-cathode potential dependence of average current density for the prototype grid-cathode unit.

measurements and simulation of this prototype at accelerating field strength of 60 kV/cm are shown in Fig. 8.

### 3.2. 3D model of the grid-cathode unit

A simulation 3D model of the grid-cathode unit (Fig. 9) was developed to describe beam injection into the accelerator. It considers the existence of seven various-length longitudinal slots in the grid located in front of the 20-mm cathode. In addition, the spherical form of the grid and cathode is also taken into account.

Results of a single grid cell simulation were used for calculation of emission current distributions across the slots, as well as electron coordinates and output angles. Here, the cathode start transverse coordinates for each slot was defined equidistantly, giving a total of $N_y$ points. To describe the beam parameters along the slots, they were divided into equal microcells, so that the maximum number $N_x$ was defined for the central slot. The length of each cell was determined according to the condition that the microcell central point did not leave the cathode radius.

The dynamic value of the emission current density for each microcell with transverse coordinates $(x, y)$ was calculated using the following formula:

$$J(x,y,t) = J_{\text{aver}}\left(\frac{Ug(t)}{C_{\text{E}}(x,y)\sin\omega t}\right)(C_{\text{E}}(x,y)\sin\omega t)^{3/2} \qquad (6)$$

where $J_{\text{aver}}(U)$ is the control voltage static dependence of the average emission current density (Fig. 5), $UG(t)$ is the time dependence of the grid-cathode control voltage, $C_{\text{E}}(x,y) = E_{\text{acc}}(x,y)/E_{\text{o}}$ is the accelerating RF field amplitude ratio to the field value at which the static characteristic calculations were made (80 kV/cm), and $\omega = 2\pi f$ is the accelerator operating cycle frequency.

The total emission current obtained from each microcell was then allocated along single macroparticles with various transverse coordinates in accordance with the curves presented in Fig. 6. A total of $N_y$ macroparticles with various currents started from each microcell at every given moment of time.

Similar to Eq. (6), the dynamic values of the transit time in the grid-cathode gap and the electron initial energy are

$$\Delta t(x,y,t) = \Delta t_{\text{aver}}\left(\frac{Ug(t)}{C_{\text{E}}(x,y)\sin\omega t}\right)(C_{\text{E}}(x,y)\sin\omega t)^{-1/2} \qquad (7)$$

$$W(x,y,t) = W_{\text{aver}}\left(\frac{Ug(t)}{C_{\text{E}}(x,y)\sin\omega t}\right)C_{\text{E}}(x,y)\sin\omega t \qquad (8)$$

where $\Delta t_{\text{aver}}(U)$ and $W_{\text{aver}}(U)$ are the average values of grid-cathode gap transit time and electron output energy as a function of the grid-cathode control voltage; the other parameters are the same as in Eq. (6).

## 4. Beam dynamics simulation

### 4.1. Calculation of current micropulse and beam spectrum

Eq. (6) describes the current micropulse form on the cathode. However, calculation of the macroparticle dynamics in the accelerator starts at a distance of 1 mm behind the grid-cathode unit. Here, the transit effect leads to a delay in electron departure from the grid-cathode unit and deformation of the current micropulse form.

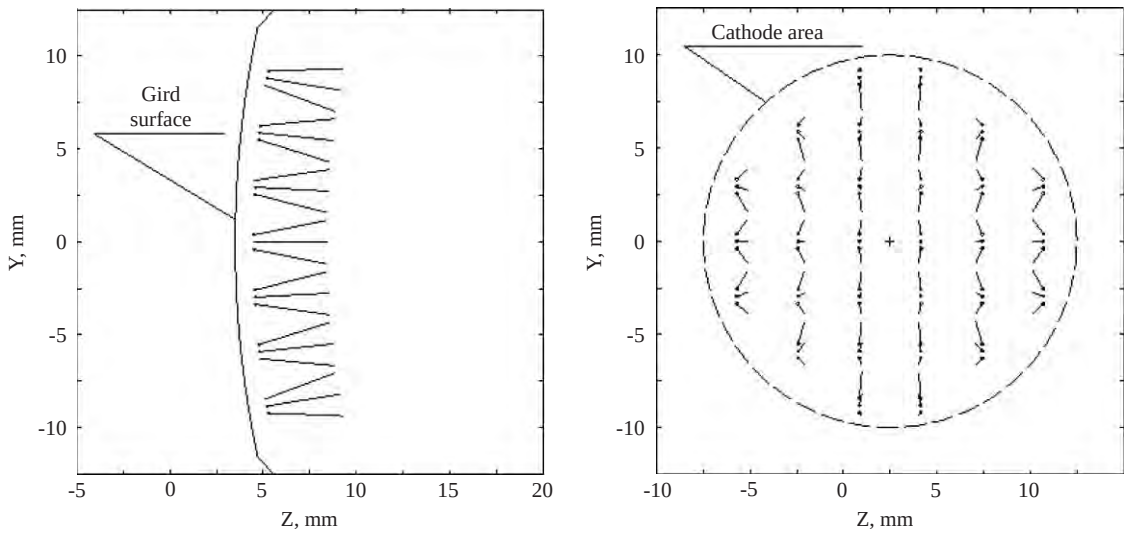



Fig. 9. 3D model of the grid-cathode unit ($N_x = 6$, $N_y = 3$): (a) $YZ$ projection; and (b) $XY$ projection.

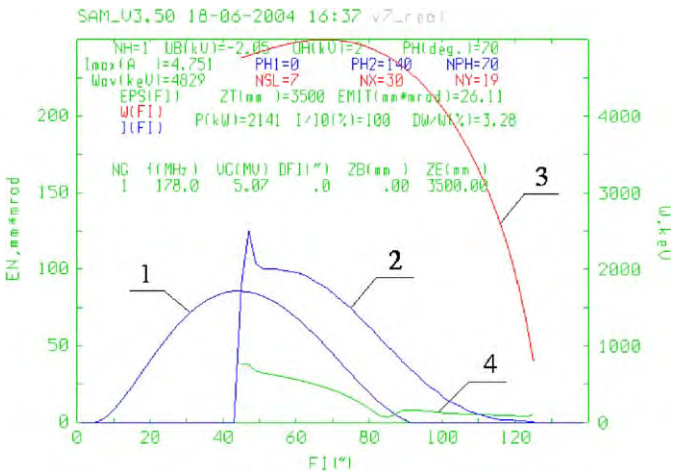



Fig. 10. Accelerating field influence on the cathode phase dependence of the current micropulse and beam output parameters: 1, 2, micropulse forms on the cathode and at a distance of 1 mm behind the grid; 3, 4, energy and normalized beam transverse emittance at the accelerator output.

To simulate the transit effect, a range of possible phases of electron injection from the cathode and their departure from the grid-cathode unit was introduced. This range was divided into $N_\varphi$ equal intervals, in the center of which the test macroparticles were launched from the cathode. The current of each test macroparticle was calculated using Eq. (6) and consideration of the cathode microcell area. Then the transit time and starting phase behind the grid were calculated from Eq. (7) for each test macroparticle.

After calculation of the parameters for all test macroparticles, calculation of the current micropulse form behind the grid was carried out. A starting phase histogram for the test particles behind the grid was constructed over the same $N_\varphi$ phase intervals, and the current of all particles was summarized in every interval.

Curves 1 and 2 in Fig. 10 show the simulation results for current micropulse form on the cathode and at a distance of 1 mm behind the grid. The control grid-cathode voltage there has the following time dependence:

$$Ug(t) = U_0 + U_1 \sin(\omega t + \varphi_1) \tag{9}$$

where $U_0 = -2.05$ kV is the constant voltage on the grid-cathode unit, and $U_1 = 2$ kV and $\varphi_1 = 70^\circ$ are the amplitude and phase shift of the additional voltage of the first harmonic. The simulations were performed for the phase range from $0^\circ$ to $140^\circ$ at $N_\varphi = 70$; the current micropulse amplitude behind the grid reached 4.75 A.

Since the main electron acceleration and transverse velocity take place at the grid-cathode unit output, the starting energy and transverse angles of macroparticles were calculated without consideration of the transit effect. Then 3D simulation of the macroparticle dynamics in the accelerator was carried out. The simulation results for the accelerating field phase dependence of the beam energy and normalized transverse emittance at the accelerator output are also presented in Fig. 10 (curves 3 and 4).

As may be observed from Fig. 10, the electron output energy depends substantially on the accelerating field phase on the cathode. This fact alone leads to the necessity of using an additional first-harmonic voltage at the grid-cathode gap to shift the current micropulse into the acceptable phase area and narrow the beam energy spectrum. As a result, average electron energy derivation from the beam average energy amounts to only 3.3%. The beneficial effect of using an additional first harmonic was experimentally proved for the present single-gap ILU-10 accelerator [2].

### 4.2. Beam transportation through the accelerator

The calculated average output power of the beam with parameters presented in Fig. 10 is 305 kW. For such a high-power beam, the problem of lossless transportation

becomes very important. Here, it is necessary to keep in mind the space charge influence on the electron transverse dynamics, especially at the first acceleration stage. Thus, for every macroparticle the effective current, which resulted in the appearance of an additional transverse electric field of the beam space charge, was calculated

$$I_{\text{eff}}(x, y, t) = J(x, y, t)\pi r_0^2 \tag{10}$$

where $J(x, y, t)$ is the dynamic value of the emission current density for a macroparticle with transverse starting coordinates $(x, y)$ and starting time $t$ calculated from Eq. (6); $r_0 = \sqrt{x^2 + y^2}$ is the macroparticle starting radius. During particle motion in the accelerator, it was assumed that the effective current did not change and the effective beam space charge transverse field had only a radial component, calculated using the following formula:

$$E_{\text{r}}^{(q)} = \frac{Z_0 I_{\text{eff}}}{2\pi r \beta \gamma^2} \tag{11}$$

where $Z_0 \approx 120\pi(\Omega)$ is the impedance of free space, $r$ is the current value of the macroparticle radius, and $\beta$ and $\gamma$ are the current values of the relative velocity and relativistic factor for the macroparticle.

Fig. 11 presents a typical view of electron trajectories in the accelerator ($N_x = 1$, $N_y = 3$, $N_\varphi = 10$). As is obvious from the figure, the beam consists of a central core formed by electrons that started from centers of slots in the grid and moved transversely to the cathode. In addition, there is a noticeable halo formed by electrons that started from the edges of slots and moved at some angle to the normal. The results presented were obtained after optimization of the sphere radius of curvature and accelerator aperture of the grid-cathode unit. Simulations proved the possibility to transport the beam through the accelerator at the expense of increasing the aperture and the RF focusing effect without using additional focusing elements.

Figs. 12 and 13 show the results of more precise simulations of beam motion through the accelerator ($N_x = 30$, $N_y = 19$, $N_\varphi = 70$). The $Z$-coordinate values in these figures correspond to the scale in Fig. 11. Fig. 12 presents the profiles calculated for average beam current density in the micropulse at the second accelerating gap input (a), in the middle between the third and fourth gaps (b), at the accelerator output (c), and at a distance of 1.5 m from the accelerator output (d). Beam transverse dimensions are shown in millimeters. It is evident that because of

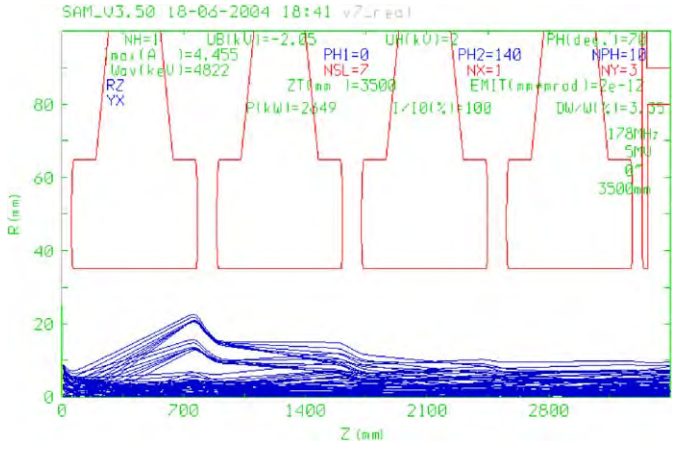



Fig. 11. Typical view of electron trajectories in the accelerator: (a) $Z = 750$ mm; (b) $Z = 2100$ mm; (c) $Z = 3500$ mm; and (d) $Z = 5000$ mm.

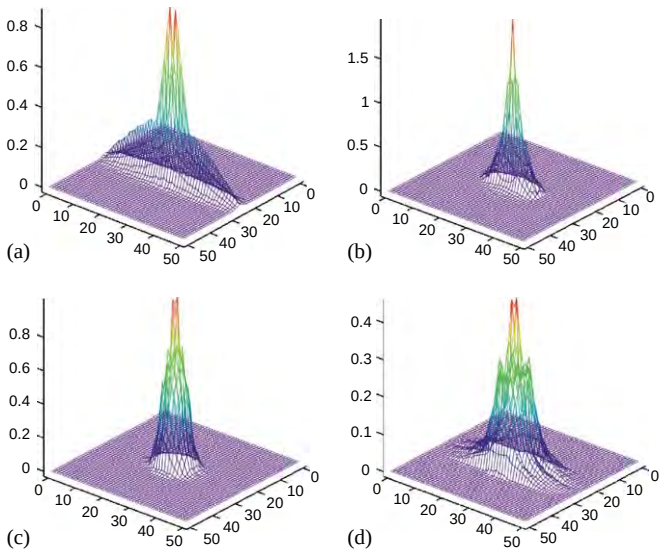



Fig. 12. Beam average current density profiles at different stages of transportation: (a) at the second accelerating gap entrance; (b) in the middle between the third and fourth gaps; (c) at the accelerator output; and (d) at a distance of 1.5 m from the accelerator output.

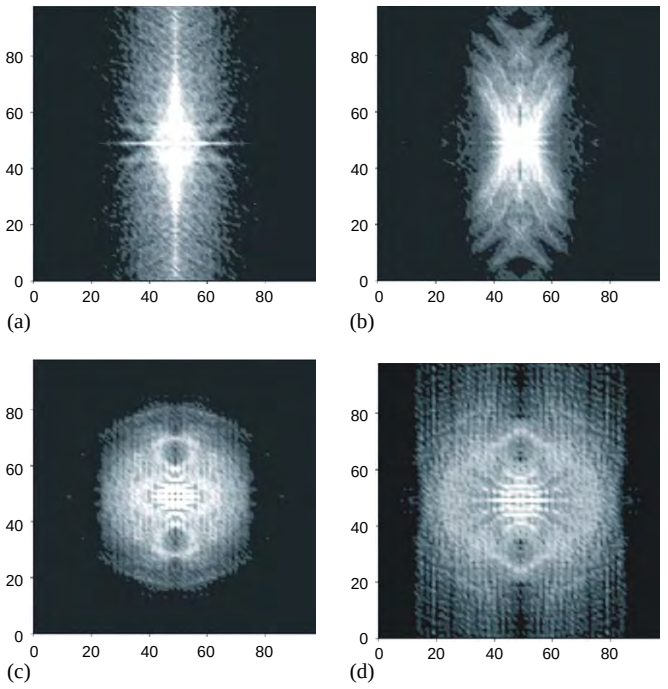



Fig. 13. Calculated lines of density level in the central part of a beam of 20 mm in diameter at various transportation stages: (a) at the second accelerating gap entrance; (b) in the middle between the third and fourth gaps; (c) at the accelerator output; and (d) at a distance of 1.5 m from the accelerator output.

injection transverse velocity spread only across the slots, the beam current density profile has an elliptical shape at the second accelerating gap entrance (Fig. 12a). It is also clear that RF focusing leads to progressive elimination of this effect at the accelerator output (Fig. 12b, c). However, in this case, overfocusing of extreme particles takes place, and the beam profile regains a slightly elliptical shape at a distance of 1.5 m from the accelerator (Fig. 12d). This effect may be eliminated by additional beam magnetic focusing in a beam-deflecting device at the target.

Fig. 13 presents results of a detailed analysis of the density profile in the central part of the beam of 20-mm diameter at the same stages as in Fig. 12. The lines of specified intensity correspond to lines of identical current density. The beam transverse dimensions are shown in the cell number of the mesh, used for density line construction. The size of one cell is $0.2 \times 0.2\,\mathrm{mm}^2$. We note that a different part of the total beam current passes through a 20-mm diaphragm at different stages of beam transportation. It is equal to 70.5% at the second accelerating gap entrance (Fig. 13a), 99.8% in the middle between the third and fourth gaps (Fig. 13b), 99.99% at the accelerator output (Fig. 13c), and 90.6% at a distance of 1.5 m from the accelerator output (Fig. 13d). Appreciable traces of the grid of the grid-cathode unit are apparent in the central part of two last pictures.

## 5. Conclusions

The results obtained prove the possibility of solving the main problems involved in creating an efficient and powerful electron accelerator for energy of 5 MeV and beam power up to 300 kW. Methods for simulation of the internal injection and beam dynamics have been developed. As a result, a new grid-cathode unit and accelerating structure have been designed. The methods described are fairly universal and may be used for the design of any accelerator satisfying the long-wave approximation.

## Acknowledgement

This work was supported by ISTC Grant #2550.

## References

[1] V. Auslender, et al., Radiat. Phys. Chem. 63 (2002) 613.
[2] M. Tiunov, et al., 5–10-MeV industrial high power electron accelerator, in: Proceedings of EPAC 2002, Paris, France, pp. 2813–2815.
[3] B. Fomel, M. Tiunov, V. Yakovlev, SAM—interactive code for evaluation of electron guns, Preprint Budker INP 96-11, 1996.
[4] D. Myakishev, V. Yakovlev, An interactive code SUPERLANS for evaluation of RF-cavities and accelerating structure, in: IEEE Particle Accelerator Conference Proceedings, vol. 5, 1991, pp. 3002–3004.
[5] L.M. Young, PARMELA, Los Alamos National Laboratory Report LA-UR-96-1835, Revised February 27, 2001.
[6] A. Ivanov, M. Tiunov, ULTRASAM—2D code for simulation of electron guns with ultra high precision, in: Proceedings of EPAC 2002, Paris, France, pp. 1634–1636.
[7] M. Tiunov, BEAM—2D code for simulation of high perveance beam dynamics in long systems, in: Proceedings of SCHEF '99, Dubna, Russia, 1999, pp. 202–208.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 85–89
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Vlasov treatment of coherent synchrotron radiation from arbitrary planar orbits ☆

R. Warnock[a], G. Bassi[b], J.A. Ellison[b]

[a]*Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309, USA*
[b]*Department of Mathematics, University of New Mexico, Albuquerque, NM 87131, USA*

Available online 1 December 2005

**Abstract**

We study the influence of coherent synchrotron radiation (CSR) on particle bunches traveling on arbitrary planar orbits between parallel conducting plates which represent the vacuum chamber. Our goal is to follow the time evolution of the phase space distribution by solving the Vlasov–Maxwell equations in the time domain. This should provide simulations with lower numerical noise than the macro-particle method, and allow one to study such issues as emittance degradation and microbunching due to CSR in bunch compressors. The fields excited by the bunch are computed in the laboratory frame from a new formula that leads to much simpler computations than usual methods. The nonlinear Vlasov equation, formulated in the interaction picture, is integrated in the beam frame by approximating the Perron–Frobenius operator. For application to a chicane bunch compressor we take steps to deal with energy chirp.


*PACS:* 41.60.−m; 41.60.Ap; 42.25.Kb; 52.65.Ff; 41.75.Fr; 41.20.−q

*Keywords:* CSR; Coherent synchrotron; Bunch compressor; Vlasov equation

## 1. Introduction

Coherent synchrotron radiation (CSR) is expected to play an important and often detrimental role in various advanced accelerator projects, for instance in linac-based coherent light sources [1] and energy recovery linacs [2]. A large concern is that CSR may cause transverse emittance growth in a bunch compressor by inducing an energy spread that is mapped into the transverse motion through dispersion. There are two principal tasks in numerical modeling of such phenomena. First, one must compute the fields produced by the particle bunch from a knowledge of past and present values of its charge/current density. Principally, the longitudinal field within the bunch is needed to find the energy change due to the field of the bunch itself. Second, one needs to find the effect of this field in subsequent evolution of the bunch form. This second problem, the question of *self-consistency*, is usually addressed by the macro-particle method.

Codes to address these issues have been developed in recent years by various authors (Li, Dohlus, Kabel, Limberg, Giannessi, Quattromini, Borland, Emma), and despite formidable complications a decent agreement of the different codes has been found in a benchmark test proposed at a meeting in Zeuthen [3]. Nevertheless, some discrepancies remain, and the problem is so complicated that there is probably room for improvement in the algorithms. We report some efforts to improve the formulation in two directions: (1) to simplify the field calculation, and (2) to study multi-particle dynamics by the Vlasov equation rather than by the macro-particle method. We hope to repeat the success of a Vlasov study of longitudinal motion in storage rings with CSR [4]. That was in a two-dimensional phase space but over time intervals very much longer than those of the four-dimensional, single-pass problem of a bunch compressor.

☆ Support from DOE grants DE-AC02-76SF00515 and DE-FG02-99ER1104 is gratefully acknowledged.

*E-mail addresses:* warnock@slac.stanford.edu (R. Warnock), gbassi@math.unm.edu (G. Bassi), ellison@math.unm.edu (J.A. Ellison).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.070

## 2. Coordinate systems and charge/current densities

We have in mind single-pass systems such as a chicane bunch compressor traversed only once by the particles [3]. With small modifications our treatment should apply as well to multi-pass systems such as storage rings. In the laboratory frame the spatial coordinates are $(Z, X, Y)$ and the independent variable is $u = ct$. The particle orbits lie between two infinite, parallel, perfectly conducting plates, which are perpendicular to the $Y$-axis and separated by a distance $h = 2g$. Points with $Y = 0$ are in the midplane, and every orbit is in some plane $Y = \text{const.} \in (-g, g)$. The $Y$-direction is "vertical". For the chicane the reference orbit (design orbit) follows the $Z$-axis initially and finally, and is in the midplane. We write the reference orbit as $\mathbf{R}_0(\beta u)$ with $\mathbf{R}_0 = (Z_0, X_0)$, where $\beta c$ is the speed of the reference particle. Generally, bold face letters refer to two-component vectors.

A point can also be specified in terms of Frenet–Serret coordinates: arc-length $s$ along the reference trajectory, and the perpendicular distance $x$ from the trajectory at $\mathbf{R}_0(s)$. Thus, $\mathbf{R} = (Z, X) = \mathbf{R}_0(s) + x\mathbf{n}(s)$ where the unit normal vector is $\mathbf{n}(s) = (-X_0'(s), Z_0'(s))$, chosen so that its $X$ component is positive. The corresponding unit tangent is $\mathbf{t}(s) = \mathbf{R}_0'(s) = (Z_0'(s), X_0'(s))$. After a change of independent variable from $u = ct$ to $s$ through standard manipulations, a convenient set of dynamical variables for motion in horizontal planes consists of the "beam frame" phase space coordinates $(z, p_z, x, p_x)$. Here $z(s) = s - \beta ct(s)$, where $t(s)$ is the time of arrival at arc-length $s$. Thus $z$ is the signed arc length from the reference particle, positive for particles in front. The conjugate variable is the relative energy deviation $p_z(s) = (E(s) - E_0)/E_0$, with $E_0 = m\gamma c^2$ the energy of the reference particle. Also $p_x(s) = v_x(s)/\beta c$ where $v_x$ is the velocity component along $\mathbf{n}$.

Clearly $Z, X$ and $u$ are determined explicitly from $z, x, s$. To obtain an explicit form for the inverse we note that we are interested in a bunch of particles with small $z$ and $x$. Expanding to lowest order in $z$ and $x$ we obtain

$$\begin{aligned}\mathbf{R} &= \mathbf{R}_0(z + \beta u) + x\mathbf{n}(z + \beta u)\\ &= \mathbf{R}_0(\beta u) + M(\beta u)\mathbf{r} + O(\kappa(z^2, xz))\end{aligned} \tag{1}$$

where $\mathbf{r} = (z, x)$ and $M = (\mathbf{t}, \mathbf{n})$ is a rotation matrix. The factor $\kappa$ in the remainder is the curvature of the reference orbit and is defined by $\mathbf{n}'(s) = \kappa(s)\mathbf{t}(s)$. Since the radius of curvature is much greater than the bunch size in our cases of interest, the remainder is negligible. Thus given $Z, X$ and $u$ we have $\mathbf{r} = M^{\mathrm{T}}(\beta u)(\mathbf{R} - \mathbf{R}_0(\beta u))$ to good approximation, where T denotes transpose.

We wish to solve the Vlasov equation for the distribution function on beam frame coordinates, $f(z, p_z, x, p_x, s)$. On the other hand, it is most convenient to solve the Maxwell equations in the laboratory frame. We must then express the lab frame charge/current density in terms of $f$. To that end suppose that $f$ has unit integral over all space and define

$$\begin{aligned}\rho(\mathbf{r}, s) &= Q\int \mathrm{d}p_z\,\mathrm{d}p_x f(z, p_z, x, p_x, s)\\ \tau(\mathbf{r}, s) &= Q\int \mathrm{d}p_z\,\mathrm{d}p_x\, p_x f(z, p_z, x, p_x, s)\end{aligned} \tag{2}$$

where $Q$ is the total charge. To an excellent approximation the lab frame charge density $\rho_{\mathrm{L}}$ is

$$\rho_{\mathrm{L}}(\mathbf{R}, Y, u) = H(Y)\rho(\mathbf{r}, \beta u), \quad \mathbf{r} = M^{\mathrm{T}}(\beta u)(\mathbf{R} - \mathbf{R}_0(\beta u)) \tag{3}$$

where $\int H(Y)\,\mathrm{d}Y = 1$ and $H(Y)$ is an arbitrary fixed vertical distribution of charge. The main approximation (beyond the use of the first order expansion (1)) is that at fixed $\mathbf{r}$ the density $\rho(\mathbf{r}, s)$ does not change appreciably when $s$ varies by an amount comparable to the bunch size. If in addition we assume that $f$ varies little when $s$ changes by a bunch size then we find a corresponding good approximation for the lab frame current density, which is defined by $\mathbf{J}_{\mathrm{L}}(\mathbf{R}, Y, u) = QH(Y)\int \mathbf{V}F(\mathbf{R}, \mathbf{V}, u)\,\mathrm{d}\mathbf{V}$, where $\mathbf{V}$ is the lab frame velocity and $F$ the lab frame phase space density with unit integral. The formula is

$$\mathbf{J}_{\mathrm{L}}(\mathbf{R}, Y, u) = \beta cH(Y)[\rho(\mathbf{r}, \beta u)\mathbf{t}(\beta u + z) + \tau(\mathbf{r}, \beta u)\mathbf{n}(\beta u + z)] \tag{4}$$

where $\mathbf{r}$ is as in Eq. (3). It is justified to expand $\mathbf{t}$ and $\mathbf{n}$ through first order in $z$. A derivation of Eqs. (3) and (4) will be reported elsewhere.

## 3. Field calculation

We calculate the electric field produced by $(\rho_{\mathrm{L}}, \mathbf{J}_{\mathrm{L}})$, but averaged over the $Y$-distribution:

$$\mathbf{E}(\mathbf{R}, u) := \langle \mathbf{E}(\mathbf{R}, \cdot, u)\rangle = \int_{-g}^{g} H(Y)\mathbf{E}(\mathbf{R}, Y, u)\,\mathrm{d}Y. \tag{5}$$

The averaged field can be computed much more quickly, and we believe that it will produce nearly the same dynamics in the $(Z, X)$ plane as the full field. To evaluate (5) we begin with the general formula for the electric field,

$$\begin{aligned}\mathbf{E}(\mathbf{R}, Y, u) = &-\frac{1}{4\pi}\int \mathrm{d}\mathbf{R}'\int \mathrm{d}Y'\xi(Y')\\ &\times\frac{\mathbf{S}(\mathbf{R}', u - [(\mathbf{R}' - \mathbf{R})^2 + (Y - Y')^2)]^{1/2}}{[(\mathbf{R}' - \mathbf{R})^2 + (Y - Y')^2)]^{1/2}}\end{aligned} \tag{6}$$

where $\mathbf{S} = \nabla\rho_{\mathrm{L}}/\varepsilon_0 + \mu_0\partial\mathbf{J}_{\mathrm{L}}/\partial t$ and $\xi(Y)$ is the effective vertical charge distribution needed to impose boundary conditions at the parallel plates by the method of images, namely $\xi(Y) = \sum_{k=-\infty}^{\infty}(-1)^k H(Y - kh)$. Here we assume that the support of $H(Y)$ is well within the interval $(-g, g)$. The term with $k = 0$ gives the potential for free space. To average the potential over $Y$ as in Eq. (5) we replace the integration variable $Y'$ by $\eta = Y' - Y$ and find the

averaged form

$$\mathbf{E}(\mathbf{R},u) = -\frac{1}{4\pi}\int \mathrm{d}\mathbf{R}' \int \mathrm{d}\eta \Phi(\eta) \times \frac{\mathbf{S}(\mathbf{R}', u-[(\mathbf{R}'-\mathbf{R})^2+\eta^2)]^{1/2}}{[(\mathbf{R}'-\mathbf{R})^2+\eta^2)]^{1/2}} \tag{7}$$

where $\Phi(\eta) = \int H(Y)\xi(Y+\eta)\,\mathrm{d}Y$. For a Gaussian $H(Y)$ with rms width $\sigma_Y$ we suppose that $\sigma_Y \ll g$ and obtain

$$\Phi(\eta) = \sum_{k=-\infty}^{\infty} \frac{(-1)^k}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{\eta-kh}{\sigma}\right)^2\right), \quad \sigma = \sqrt{2}\sigma_Y. \tag{8}$$

We assume that $\sigma$ is sufficiently small to justify replacing the Gaussians in Eq. (8) by delta functions. Thus, the averaging produces just a two-dimensional integral

$$\mathbf{E}(\mathbf{R},u) = -\frac{1}{4\pi}\sum_{k=-\infty}^{\infty}(-1)^k \int \mathrm{d}\mathbf{R}' \times \frac{\mathbf{S}(\mathbf{R}', u-[(\mathbf{R}'-\mathbf{R})^2+(kh)^2)]^{1/2}}{[(\mathbf{R}'-\mathbf{R})^2+(kh)^2)]^{1/2}}. \tag{9}$$

The integration in Eq. (9) is restricted to a very small part of the full $\mathbf{R}'$ plane, because of the small size of the bunch, but it is awkward to locate this region owing to the fact that spatial and temporal arguments of the source both depend on $\mathbf{R}'$. The bunch moves around, so to speak, during integration. The task of integration is made drastically easier if we take the temporal argument to be a new variable of integration. We first go into polar coordinates $(\zeta,\theta)$, then use the temporal argument $v$ in place of the radial coordinate $\zeta$. That is

$$\mathbf{R}'-\mathbf{R} = \zeta\mathbf{e}(\theta), \quad \mathbf{e}(\theta) = (\cos\theta, \sin\theta)$$
$$v = u - [\zeta^2 + (kh)^2]^{1/2}. \tag{10}$$

This incidentally gets rid of the small divisor in Eq. (9), giving the field simply as an integral over the source

$$\mathbf{E}(\mathbf{R},u) = -\frac{1}{2\pi}\sum_{k=0}^{\infty}(-1)^k(1-\delta_{k0}/2)\int_{-\infty}^{u-kh}\mathrm{d}v\int_{-\pi}^{\pi}\mathrm{d}\theta\,\mathbf{S}(\hat{\mathbf{R}},v) \tag{11}$$

where $\hat{\mathbf{R}} = \mathbf{R} + [(u-v)^2 - (kh)^2]^{1/2}\mathbf{e}(\theta)$. It is quite easy to determine the effective region of the $\theta$ integration in Eq. (11). Note that the source in Eq. (11) has significant values only for $\hat{\mathbf{R}}$ restricted to a bunch-sized neighborhood of $\mathbf{R}_0(\beta v)$, according to Eqs. (3) and (4). For the CSR wake field at time $u$ we are interested only in $\mathbf{R}$ in a bunch-sized neighborhood of $\mathbf{R}_0(\beta u)$. Thus the integrand is appreciable only when

$$\mathbf{R}_0(\beta u) - \mathbf{R}_0(\beta v) + [(u-v)^2 - (kh)^2]^{1/2}\mathbf{e}(\theta) = O(\Delta) \tag{12}$$

where $\Delta$ is a suitable measure of the bunch size. For $k=0$ and $u-v$ large compared to $\Delta$, this cannot be satisfied unless $\mathbf{e}(\theta)$ has nearly the same direction as $\mathbf{R}_0(\beta u) - \mathbf{R}_0(\beta v)$, which is to say that the domain of $\theta$ integration is tiny (and close to $\theta = 0$ for a chicane with small bending angle). When $u-v$ gets close to $\Delta$ the domain expands precipitously to the full $(-\pi,\pi)$. For $k \neq 0$ condition (12) cannot be met unless $u - v \gg kh$, so for image charges there are no contributions to the $v$-integral close to its upper limit.

## 4. Integration of Vlasov equation

To state the Vlasov equation in the beam frame coordinates we need the single-particle equations of motion. We assume that the only external forces are from sharp-edged bending fields. To a good approximation the equations can be linearized, except for the nonlinear CSR force. Using also the fact that $\gamma$ is large (order of $10^4$) we obtain

$$z' = -\kappa(s)x, \quad p_z' = e\mathscr{E}(\mathbf{r},p_x,s)/E_0, \quad x' = p_x, \quad p_x' = \kappa(s)p_z. \tag{13}$$

These are the equations of standard linear optics perturbed by the collective force from CSR, as in Ref. [5]. Here $\kappa(s)$ is the curvature and its sign must be taken consistent with the sign of the external magnetic field. Also, to good approximation, $\mathscr{E}(\mathbf{r},p_x,s) = (\mathbf{t}(s)+p_x\mathbf{n}(s))\cdot\mathbf{E}(\mathbf{R}_0(s)+x\mathbf{n}(s),(s-z)/\beta)$. Note that the lab frame velocity is essentially $\beta c(\mathbf{t}(s)+p_x\mathbf{n}(s))$. We have neglected the transverse electric and transverse magnetic force, since for relativistic particles one is expected to cancel the other to high accuracy (for typical small values of the curvature). The validity of this neglect will be checked at a later stage.

The unperturbed version of Eq. (13) with $\mathscr{E}=0$ can be solved explicitly [5] in terms of the lattice functions $D(s), D'(s), R_{56}(s)$. This gives the transport map $\Phi(s|0)$ from $s=0$ to arbitrary $s$, with inverse $\Phi(0|s)$. For the Vlasov formulation we use the initial conditions of the unperturbed motion as phase space variables [5]. With $\zeta = (z,p_z,x,p_x)$, the phase space coordinates will be $\zeta_0 = \Phi(0|s)\zeta$. Now $\zeta_0$ is constant in the unperturbed case, and possibly slowly varying in presence of the perturbation. The Vlasov equation for the distribution function $g(\zeta_0,s) = f(\zeta,s)$ is

$$\frac{\partial g(\zeta_0,s)}{\partial s} + \frac{\partial g(\zeta_0,s)}{\partial p_z}\frac{e}{E_0}\mathscr{E}([\Phi(s|0)\zeta_0]_1,[\Phi(s|0)\zeta_0]_3,[\Phi(s|0)\zeta_0]_4,s;g) = 0. \tag{14}$$

The Vlasov equation is to be integrated by approximating the Perron–Frobenius (PF) operator. Over a small interval $(s, s+\Delta s)$ we regard $\mathscr{E}(\mathbf{r},p_x,s;g)$ as independent of $s$, having the value it had after the previous $s$-step, and solve the initial value problem for the single-particle equations corresponding to Eq. (14) on that interval. We represent the solution of the IVP as a map $\Psi(s+\Delta s|s)(\zeta_0)$ that takes any initial $\zeta_0$ into its image under the flow. The inverse map is $\Psi(s|s+\Delta s)$. The propagation of $g$ is given by the PF operator $\mathscr{P}$ associated with $\Psi$, which is to say

$$g(\zeta_0, s+\Delta s) = \mathscr{P}g(\zeta_0,s) = g(\Psi(s|s+\Delta)(\zeta_0),s). \tag{15}$$

To apply this method we need some finite-dimensional representation of $g$ and a numerical integration of the differential equations to produce $\Psi(s|s+\Delta s)$. For the former a simple choice is local polynomial interpolation of values of $g$ on a grid in phase space, say a grid with uniform cell size $h$ in each dimension. One-dimensional interpolation is done successively in each of the four dimensions. Then in a cell labeled by grid point $\zeta^{(p)}$ the distribution is a sum over $k^4$ terms, where $k$ is the order (degree + 1) of the polynomial. Each term is a value of $g$ at a grid point near $\zeta^{(p)}$, times a product of four one-dimensional polynomials evaluated at $(\zeta_i - \zeta_i^{(p)})/h, \quad i = 1, \cdots, 4$, respectively. There are just $k$ defferent polynomials, determined entirely by the choice of $k$. This gives a fast algorithm, but it does not give a representation with overall continuity. In experience to date the lack of continuity does not seem to be a big drawback, perhaps because the grid is ususally fine enough so that jumps from one cell to the next are small. In further work, especially in attempts to use a coarser grid, some more advanced interpolation methods should be evaluated, for instance application of radial basis functions [6].

Now consider the particular case of a bunch compressor, similar to that of the LCLS design. The initial distribution has the form

$$f(\zeta_0, 0) = \frac{1}{2\pi\varepsilon_0}\exp\left[-\frac{1}{2\varepsilon_0\beta_0}(x_0^2 + (\alpha_0 x_0 + \beta_0 p_{x0})^2)\right] \times \mu(z_0)\rho_c(p_{z0}, z_0) \tag{16}$$

where $\beta_0, \alpha_0, \varepsilon_0$ are initial Twiss parameters and $x$-emittance. The factor $\rho_c$ expresses "energy chirp", a close correlation of energy with longitudinal position in the bunch:

$$\rho_c(p_{z0}, z_0) = a\,\exp\left[-\frac{1}{2\sigma_u^2}(p_{z0} - uz_0(1 + C(z_0)))^2\right]. \tag{17}$$

Here $u$ is the slope of the correlation line at $z = 0$, $\sigma_u$ is a small parameter expressing lack of the desired perfect correlation, $a$ is for normalization, and $z_0 C(z_0)$ is a small nonlinearity in the correlation, significant at the ends of the $z_0$-distribution and well represented as cubic. This nonlinearity represents effects of wake fields in the linac and curvarture of the applied r.f. field. The function $\mu(z_0)$ is approximated as parabolic over an interval, and zero elsewhere. Because of the tight correlation, the $p_{z0}$ distribution will also be nearly parabolic.

The $p_{z0}$–$z_0$ correlation will make the interpolation scheme mentioned above quite awkward, since one would need a fine mesh in both $p_{z0}$ and $z_0$, with a lot of wasted mesh points at which the distribution is negligible. It will help to make a change of variable, for example using $w_1 = (p_{z0} - uz_0)/\sigma_u$ and $w_2 = z/\sigma_{z0}$ in place of $p_{z0}$ and $z_0$, where $\sigma_{z0}$ is the initial bunch length. A convenient initial mesh for these variables is easily provided, but the mesh mush be redefined at later times. At the least it must be translated, since the radiation causes a large displacement of $w_1$.

## 5. Preliminary numerical studies

We have concentrated to date on the crucial matter of field computation, taking an assigned charge/current source rather than one determined self-consistently. The source in Eq. (11) is determined from Eq. (16) evolved in time by the backward linear map $\Phi(0|s)$. All numerical calculations reported here take $\mu(z_0)$ to be Gaussian, $C(z_0) = 0$ and use benchmark bunch compressor parameters from Ref. [3], i.e. $\gamma = 9785$, $\varepsilon_0/\gamma = 1$ mm-mrad, $\beta_0 = 40$ m, $\alpha_0 = 2.6$, $\sigma_{z0} = 200\,\mu$m, $\sigma_u = 2 \times 10^{-6}$, $u = -36\,\mathrm{m}^{-1}$, $Q = 1$ nC. The evolution of the charge density in the $Z$–$X$ plane in the absence of CSR is shown in Fig. 1. We have focused on calculations of the mean and variance of the relative energy loss. We calculate these in a first order perturbative manner where we integrate the equations of motion (13) with initial conditions at $s = 0$ chosen randomly according to Eqs. (16), (17), using the field $\mathscr{E}$ from the assigned charge/current source.

The relative energy loss, $p_{z\mathrm{L}}(s)$, for a single particle is given by $p_{z\mathrm{L}}(s) = p_z(s) - p_{z0} = (e/E_0)\int_0^s \mathscr{E}(\mathbf{r}(s'), s')\,\mathrm{d}s'$ where $\mathbf{r}(s)$ is determined from Eq. (13), and depends on the initial conditions, and $p_{z0}$ is the unperturbed relative energy deviation. Since the standard deviation of $p_{z\mathrm{L}}$ is roughly $2 \times 10^{-4}$, the statistical error with $n$ particles is

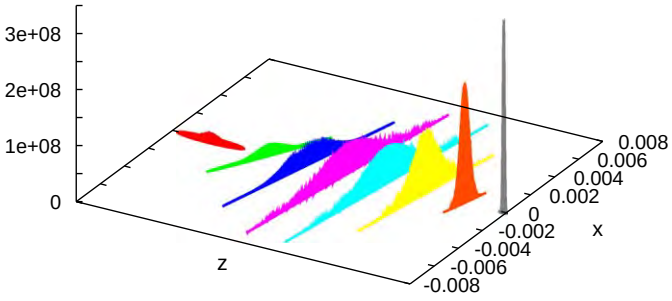



Fig. 1. Charge density in $Z$–$X$- plane at successive times, showing bunch compression.

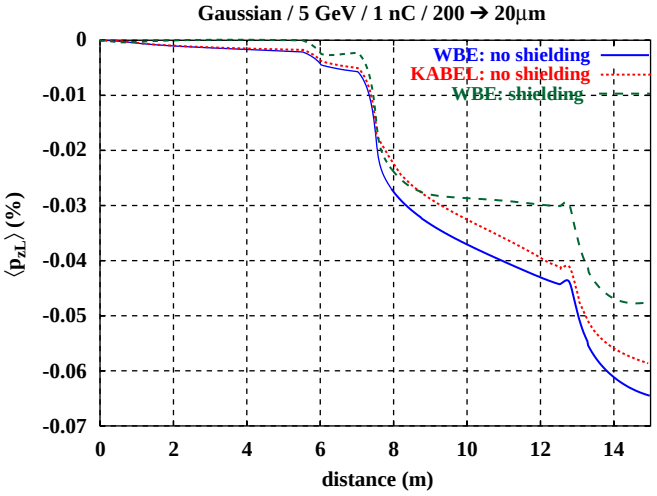



Fig. 2. Mean relative energy loss $\langle p_{z\mathrm{L}}\rangle$ vs. position in chicane.

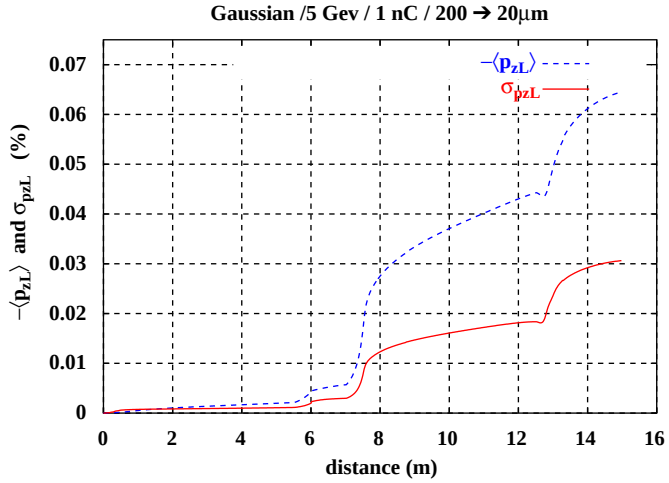



Fig. 3. Negative of $\langle p_{zL}\rangle$ (dashed blue line) and SD($p_{zL}$) (solid red line) vs. position in chicane.

$2\times10^{-4}/\sqrt{n}$. Thus we have taken 400 particles giving a statistical error of order $10^{-5}$.

Our result for no shielding is shown by the solid (blue) curve in Fig. 2 where we plot $\langle p_{zL}\rangle$ vs $s$. Note our statistical error of $10^{-5}$ is very small on the scale of the figure. The self-consistent, no shielding calculation of Kabel [3] is also shown on the figure by the dotted (red) curve and the agreement is good. Nevertheless we expect that our self-consistent calculation will make changes in the result. By the dashed (green) curve we plot our result including shielding, for interplate gap $h = 1$ cm. We added only the contribution from $k = 1$ which we believe to be dominant.

In Fig. 3 we show our results for the standard deviation $\sigma_{pzL}(S) = \langle(p_{zL}(s) - \langle p_{zL}(s)\rangle)^2\rangle^{1/2}$ of $p_{zL}$ by the dashed (blue) curve. The solid (red) curve is the negative of $\langle p_{zL}(s)\rangle$ with no shielding from Fig. 2. As in Fig. 2 $\sigma_{pzL}$ was computed using 400 particles only as fluctuations for this number of particles were small.

## Acknowledgements

Conversations with P. Emma, A. Kabel, and M. Venturini are gratefully acknowledged.

## References

[1] LCLS Design Study Report, SLAC, 1998; TESLA X-ray FEL project, see ⟨http://tesla-new.desy.de⟩.
[2] For an overview see J.B. Murphy, Proceedings of 2003 Particle Acceleration Conference, p.176, paper TOAC001.
[3] ICFA Beam Dynamics Mini-Workshop on CSR, Berlin–Zeuthen, 2002, see ⟨http://www.desy.de/csr⟩;
L. Giannessi, M. Quattromini, Phys. Rev. ST Accel. Beams 6 (2003) 120101.
[4] M. Venturini, R. Warnock, R. Ruth, J.A. Ellison, Phys. Rev. ST Accel. Beams 8 (2005) 014202.
[5] S. Heifets, G. Stupakov, S. Krinsky, Phys. Rev. ST Accel. Beams 5 (2002) 064401.
[6] M.D. Buhmann, Radial Basis Functions, Cambridge University Press, Cambridge, 2003.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 90–94
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Development of the versatile multi-particle code DYNAMION

S. Yaramyshev[a,b,*], W. Barth[a], L. Groening[a], A. Kolomiets[b], T. Tretyakova[b]

[a]*Gesellschaft fuer Schwerionenforschung, Planckstreet 1, 64291 Darmstadt, Germany*
[b]*Institute for Theoretical and Experimental Physics, B. Cheremushkinskaya street 25, 117218 Moscow, Russian Federation*

Available online 28 November 2005

**Abstract**

The multiparticle code DYNAMION (DYNAMics of ION) was written in ITEP for precise beam dynamics simulation in high-current linear accelerators, and was used world-wide for linacs design and for study of linacs in operation. In collaboration between GSI and ITEP the DYNAMION code is in use since many years for the investigation of operation of the GSI high-current heavy-ion linac UNILAC. Results of the beam dynamics simulations are described and compared with other codes as well as with experimental data. The main feature of the code is the possibility to calculate the 3D particle motion in the whole linac potentially consisting of RFQs, different types of DTL structures, transport lines and other elements. During the last years, the development of the code was focused to the increasing of accuracy and reliability of calculations. These goals were reached by an improved description of the external fields inside the code. Additionally, data from measurements or from calculations done with external codes (e.g., focusing and accelerating fields, beam emittance, etc.) are usable. In the paper, the latest features of the DYNAMION code are presented.


*PACS:* 29.17.+w; 41.75.Lx; 41.20.−q

*Keywords:* Beam dynamics simulation; Linac design; Space charge influence

## 1. Introduction

The design of high-intensity linear ion accelerators and the physics of intense beams require a tool adequate to "modern" tasks—a beam dynamics simulation code has to work with any accelerating/focusing structure taking into account all details of particle motion caused by space charge effects, coupling between oscillations in different phase planes or nonlinearities of external fields. The calculated output data have to allow deep and flexible analysis of the particle motion. To reach these goals the multi-particle code DYNAMION (DYNAMics of ION) has been written [2] and implemented to numerous projects in ITEP and several accelerator centres around the world, namely in GSI (Darmstadt, Germany), CERN (Geneva, Switzerland), ANL (Argonne, USA), LNL-INFN (Legnaro, Italy) and INR (Troitsk, Russia).

A significant part of the DYNAMION development has been done during many years of collaboration between GSI and ITEP. The code has been widely used for the investigation and optimization of the GSI heavy-ion high-current linac UNILAC (UNIversal Linear ACcelerator) [6].

## 2. GSI UNILAC

The UNILAC is designed to fill the heavy-ion synchrotron SIS up to its space charge limit for all ion species, with mass over charge ratios of up to 8.5. It consists of the 36 MHz high-current injector (HSI), a gas stripper section at an energy of 1.4 MeV/$u$ and a 108 MHz Alvarez post-stripper, accelerating ions up to of 11.4 MeV/$u$. The pre-stripper section HSI consists of the low energy beam transport (LEBT), a radio frequency quadrupole (RFQ) accelerator and two interdigital H (IH) -structure tanks. The HSI is designed for the acceleration of ions with a charge to mass ratio up to 65 and a maximum $U^{4+}$ beam current of 15 emA, resulting in 12.6 emA $U^{28+}$ after stripping in a supersonic gas jet and subsequent charge state separation.

*Corresponding author. Gesellschaft fuer Schwerionenforschung, Planckstreet 1, 64291 Darmstadt, Germany. Tel.: +49 6159 712170; fax: +49 6159 712987.
*E-mail address:* S.Yaramishev@gsi.de (S. Yaramyshev).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.018

The HSI is in routine operation since November 1999 and has achieved the design intensities for light and medium isotopes, but with a significant surplus of the primary beam current coming from the ion source. For heavy ions the beam intensities reached behind HSI are a factor of two lower than the design values. Intense beam dynamics studies of the whole UNILAC have been done by means of the DYNAMION code. Several measures were proposed and partially realized for the increase of the beam current and brilliance at the entrance of the synchrotron SYS 18 [7,9,10].

## 3. Description of the code

### 3.1. Particle motion equation

The general three-dimensional equation of charged particle motion in external and internal fields is used in the most common form and is solved numerically by time integration.

### 3.2. Calculation of the space charge effects

Space charge forces are taken into account by the calculation of the particle to particle interaction including the non-uniformity of particle distribution. To avoid artificial collisions of particles, after a discrete steps of integration a special routine is introduced. Nevertheless, due to the relatively small size of the integration steps, the probability of such collisions is low.

For the adequate calculation of the space charge influence to the dynamics of continuous beams, two neighbouring virtual bunches are created. Virtual bunches are identical to the simulated one and are used only for calculating of space charge fields.

In order to decrease the computing time a fast Poisson solver is now under development, as well as the adjustment of the code for parallel computing.

### 3.3. Input particle distribution

Several types of the input particle distributions including KV, truncated Gaussian, uniform, etc. can be generated by DYNAMION. The results of beam dynamics simulation in a linac performed by any tracking code can be transformed to the input data for the DYNAMION calculation of the following structures. The data of an emittance measurement can be used for generating of input particle distribution which includes non-uniformities of a real beam.

### 3.4. Multi-charge particle distribution

A beam coming from the heavy ion source usually is a mixture of ions with different charge states and, therefore, different energies. Several DYNAMION simulations with such beams have been carried out under space charge conditions using adequate particle distributions [3,9].

### 3.5. Beam transport lines

Different focusing elements usually installed in beam transport lines can be simulated in the DYNAMION code, including quadrupole/octupole magnetic lenses, electrostatic lenses, solenoids, bending magnets, slits, etc.

Most realistic results of beam dynamics simulations can be obtained with a detailed description of the external focusing fields. Instead of simplified approximations it can be done by means of special codes (OPERA, MAFIA, etc.) or by measurements of the electromagnetic fields.

### 3.6. RFQ accelerator

The potential distribution in an RFQ is expressed by well-known Fourier–Bessel series [1]. Analytical formulas for the coefficients of the series can be obtained in the assumption of the special hyperbolic transverse shape of electrodes, which is complicate to realize in practice. Usually, the shape of RFQ electrodes is far from the "ideal". Historically the "ideal" two-term series was used for the beam dynamics simulations and an RFQ design. Expansion of the series to eight terms, as it is done in the well-known code PARMTEQ/PARMTEQM, includes multipole effects of the ideal shape of electrodes. But, compared to the real geometry it may lead to significant errors. For the low energy heavy ion RFQ in TRIUMF, even complete loss of resonance acceleration was observed [5].

In the earlier DYNAMION version the two-term series with coefficients corrected for the real geometry of RFQ cell was implemented. It provided more reasonable results, but multipole effects were not included. In the present version of the DYNAMION code, the Laplace equation for the potential is solved on the grid by the finite element method, including all details of the electrode geometry. The expression of the eight-term series for the potential at each node of the grid is written, and the system of linear equations is solved by the least-squares method for the coefficients. The Expression for electrical field can be obtained by derivation of the formula for the potential.

The longitudinal electrical field along an RFQ cell, calculated by means of DYNAMION code, is shown in Fig. 1. Parameters of the cell correspond to the medium part of the HSI RFQ. The curves correspond to "ideal" two-term series and two-term and eight-term series for the real geometry of electrodes. The field is shown at a radius of 9/10 of the aperture. The significant difference (up to 40% in some area) between an ideal and a realistic representation of the field in an RFQ is obvious.

For the specific beam dynamics simulations in an RFQ input radial matcher, the adequate description of it is necessary. In this part, the aperture and the shape of electrodes change with a certain law to provide for the 6D matching of the beam to the RFQ. Also, the transition

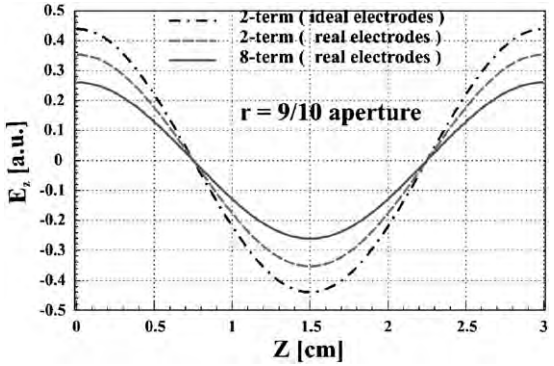



Fig. 1. Longitudinal electrical field along an RFQ cell at a radius of 9/10 of aperture calculated by means of DYNAMION code for different representation of the potential distribution.

from the resonator wall to the electrodes has to be taken into account. In the DYNAMION the Laplace equation for this area is solved on the grid by finite element method. Direct field mapping or an approximation series of 90 coefficients can be used for the calculation of the particle motion.

### 3.7. Drift tube linac (DTL)

For beam dynamics in DTLs the electrical field is calculated separately for each cell, combined by a gap and a drift tube. The real geometry of the drift tube is used including inner/outer diameters and rounding of the edges. Assuming axial geometry, the Laplace equation is solved by the finite element method. The potential $U(r,z)$ and the field components $E_z(r,z)$, $E_r(r,z)$ in a gap and inside a drift tube are approximated by series with 30 coefficients $A_n$:

$$U(r,z) = -V\left[\frac{z}{l} + \sum_{n=1}^{\infty} A_n \sin(2nkz) I_0(2nkr)\right],$$

where $r$,$z$ is the coordinates, $I_0$, $I_1$ the Bessel functions, $l$ the half of cell length, $k = \pi/l$. The voltage factors $V = \int_{-l}^{l} E_z(0,z)\,dz$ along the tank may be obtained from the design data or by measurements. The expression for electrical field can be calculated by derivation of the formula for the potential.

Several test calculations of the field carried out by the DYNAMION have shown excellent agreement with the results of the OPERA code in the working area (9/10 of the aperture). Close to the drift tube aperture the discrepancy is a few percent. Nevertheless, the accuracy of the calculations can be easily improved by increasing the number of terms in the series.

### 3.8. Stripping process and particle separation

The stripping of ions to higher charge states increases the accelerating efficiency. Simultaneously, the beam current increases several times. The process of stripping results in a wide spectrum of ions with different charge states, and a special section for the charge separation is necessary. In the DYNAMION code a special routine calculates the particle behaviour in the stripper section. Theoretical or measured charge states spectrum is introduced as input data. When each particle passes the stripper, its charge state changes randomly in accordance with the weight of each charge state in the spectrum.

## 4. Comparison of calculations and measurements

### 4.1. HSI LEBT system

Data of the beam emittance measurements in the LEBT ($X$–$X'$ phase plane) and generated macroparticle distribution are presented in Fig. 2. The level of measured intensity is shown with a grey scale (left plot). The number of particles in each bin is proportional to the measured intensity (right plot).

The GSI LEBT is intended to transport the pre-accelerated beam from the ion source to the RFQ accelerator. The matching of the beam is mainly done by means of four quadrupole lenses before the RFQ. The distribution of the magnetic field for each lens has been measured separately for different levels of the field, and is introduced into DYNAMION simulations as an input file. The magnetic field in the whole quadrupole quartet, including field overlapping between lenses, is calculated as a superposition of separate fields in accordance with the field setting in each lens (Fig. 3).

### 4.2. HSI RFQ

The HSI RFQ of the UNILAC was designed at the Frankfurt University [4] and commissioned in 1999 [6]. Its measured transmission for the high current $U^{4+}$ beam ($\approx$60%) is significantly lower than predicted by PARMTEQ (95%) and calculated by an earlier version of the DYNAMION (83%) [7].

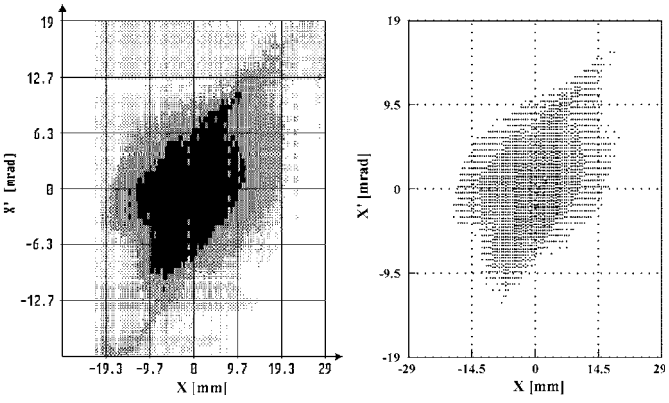



Fig. 2. Data of emittance measurement (left) and the generated particle distribution (right) in $X$–$X'$ plane for a 15 emA beam current ($U^{4+}$) at an energy 2.2 keV/$u$.

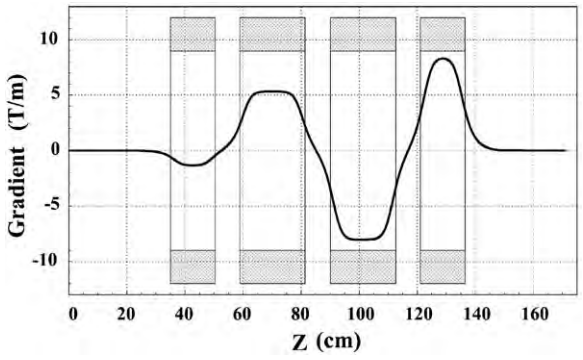



Fig. 3. Distribution of the gradient of magnetic field along the quadrupole quartet calculated as a superposition of the measured fields with recent settings for each lens.

The low transmission can be explained by a mismatch of the beam and a misalignment of electrodes. The different effects were taken into account in recent DYNAMION simulations:

- data of an emittance measurement (Fig. 2) were used for generating the particle distribution;
- the measured magnetic fields distribution in four quadrupole lenses before the RFQ, with machine settings during operation, were introduced as input data (Fig. 3);
- the electrical field in the RFQ was calculated with an eight-term series on the base of the real geometry of the resonator and electrodes;
- the measured voltage distribution along the RFQ was applied;
- the measured misalignments of the ten RFQ sections were taken into account.

The measured $U^{4+}$ beam current before the quadrupole quartet was 15 emA. Due to the limited space between the last matching quadrupole and the RFQ, it is not possible to measure the beam current exactly at the RFQ entrance. Behind the RFQ a beam current of 8.3 emA was measured.

The simulations with the DYNAMION were performed for the system LEBT RFQ, including all available information about the beam and the linac. Behind the RFQ a beam current of 8.7 emA was calculated.

### 4.3. UNILAC DTL structures

The design of the HSI IH structures was done by means of the LORASR code, based on the KONUS beam dynamics [8]. The well-known code PARMILA, which is usually used for the calculations of DTLs, is not suitable for the simulation of fields and beam dynamics of such structure, whereas the DYNAMION was successfully implemented for the task. Simulations have shown a good agreement with measurements in front of the IH section and behind it.

Numerous simulations by means of DYNAMION were carried out for the Alvarez section of the GSI UNILAC. The code was used for systematic investigations of matching a high-current beam to the periodic Alvarez DTL. The rms Twiss parameters of a realistic initial distribution were matched to the entry of the first tank. The simulated emittance growth along the whole DTL was considerably decreased by a factor of two in all three dimensions.

### 4.4. UNILAC stripper section

In the UNILAC gas stripper section the design $U^{4+}$ pulse beam current of 15 emA rises up to seven times during stripping. The $U^{28+}$ ions have to be purely separated from the neighbouring charge states under extremely high space charge influence.

The process of stripping has been calculated by DYNAMION for intense $Ar^{1+}$ and $U^{4+}$ beams. During stripping the measured spectra of ion species were used. The high resolution of the separation has been confirmed by simulations of the subsequent sections.

## 5. Conclusion

DYNAMION is a versatile multi-particle code for the simulation of acceleration, focusing and transport of space charge dominated ion beams. The general equations of particle motion are solved with precise calculation of the external fields and space charge forces. Electromagnetic external fields are represented by analytical formulas, internal Laplace solvers, field maps or field measurements. The real geometry of the accelerator elements, measured beam emittances and errors of mechanical tuning can be easily introduced into the code. Numerous features of the code give the possibility of end-to-end beam dynamics simulation even in complicate linacs, as the GSI UNILAC.

DYNAMION was successfully tested with several linacs in operation, and good agreement between calculations and measurements was demonstrated. Results of the DYNAMION calculations have been used successfully in several leading accelerator centres for the linacs design and for the optimization of operating ones. The DYNAMION code comes up as a universal, advanced tool developed for the precise solutions of various tasks in the physics of high-current linear accelerators.

## References

[1] I.M. Kapchinsky, Theory of Linear Resonance Accelerators, Energoizdat, Moscow, 1982.
[2] A. Kolomiets, et al., DYNAMION—The code for beam dynamics simulation in high current ion linac, in: Proceedings of EPAC-98, Stockholm, Sweden, 1998, pp. 1201–1203.
[3] D. Kashinsky, et al., High-current multicharge heavy ions injector linac for ITEP-TWAC Project, 14th International Symposium on Heavy Ion Inertial Fusion, HIF-2002, Moscow, Russia, 2002.

[4] A. Schempp, Design of compact RFQs, Proceedings of LINAC-96, Geneva, Switzerland, 1996, p. 53.
[5] S. Koscielniak, Correction of field errors in ISAC RFQ, Proceedings of EPAC-96, Barcelona, Spain, 1996, p. 371.
[6] W. Barth, Commissioning of the 1.4 MeV/*u* high current heavy ion linac at GSI, Proceedings of LINAC-2000, Monterey, USA, 2000, pp. 121–123.
[7] J. Klabunde, et al., Beam dynamics simulations for the GSI high current injector with the new versatile computer code DYNAMION, in: Proceedings of PAC-2001, Chicago, USA, 2001, pp. 2899–3001.
[8] U. Ratzinger, et al., Status of the 36 MHz linac cavities for the GSI high current injector, in: Proceedings of PAC-99, New York, USA, 1999, p. 3554.
[9] W. Barth, et al., Transport and injection of heavy ion beams with high brilliance for the GSI High current injector, in: Proceedings of LINAC-2002, Kuongjui, Korea, 2002.
[10] L. Groening, et al., Measurements and simulations on the beam brilliance in the universal linear accelerator UNILAC at GSI, Proceedings of LINAC-2002, Kuongjui, Korea, 2002.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 95–99
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# TE/TM alternating direction scheme for wake field calculation in 3D

Igor Zagorodnov*, Thomas Weiland

*Institut fuer Theorie Elektromagnetischer Felder (TEMF), Technische Universitaet Darmstadt, Schlossgartenstrasse 8, D-64289 Darmstadt, Germany*

Available online 28 November 2005

**Abstract**

In the future, accelerators with very short bunches will be used. It demands developing new numerical approaches for long-time calculation of electromagnetic fields in the vicinity of relativistic bunches. The conventional FDTD scheme, used in MAFIA, ABCI and other wake and PIC codes, suffers from numerical grid dispersion and staircase approximation problem. As an effective cure of the dispersion problem, a numerical scheme without dispersion in longitudinal direction can be used as it was shown by Novokhatski et al. [Transition dynamics of the wake fields of ultrashort bunches, TESLA Report 2000-03, DESY, 2000] and Zagorodnov et al. [J. Comput. Phys. 191 (2003) 525]. In this paper, a new economical conservative scheme for short-range wake field calculation in 3D is presented. As numerical examples show, the new scheme is much more accurate on long-time scale than the conventional FDTD approach.


*PACS:* 29.27.Bd; 02.60.Cb; 02.70.Bf

*Keywords:* Maxwell's equations; FDTD; ADI; Finite integration; Conformal; Wake field

## 1. Introduction

Computation of wake fields of short relativistic bunches in long structures remains a challenging problem even with the fastest computers available. It demands developing new numerical approaches for long-time calculation of electromagnetic fields. The conventional FDTD scheme [3], used in MAFIA [4] and other wake and Particle-In-Cell (PIC) codes, suffers from numerical grid dispersion and the staircase approximation problem. As an effective cure of the dispersion problem, a numerical scheme without dispersion in the longitudinal direction can be used as it was shown in Refs. [1,2].

In this paper, a new two-level economical conservative scheme for short-range wake field calculation in three dimensions is presented. The scheme does not have dispersion in the longitudinal direction and is staircase free (second-order convergent). Unlike the FDTD method [3] and the scheme developed in Ref. [2], it is based on a transversal electric–transversal magnetic (TE/TM)-like splitting of the field components in time. Additionally, it uses an enhanced alternating direction splitting of the transverse space operator that makes the scheme as computationally effective as the conventional FDTD method. Unlike the FDTD ADI method, the splitting error in our scheme is only of the fourth order. As numerical examples show, the new scheme is much more accurate on the long-time scale than the conventional FDTD approach. For axially symmetric geometries, the new scheme performs at least two times faster than the scheme suggested in Ref. [2] achieving the same level of accuracy.

## 2. Formulation of the problem

At high energies, the particle beam is rigid. To obtain the wake field, the Maxwell equations can be solved with a rigid particle distribution. The influence of the wake field on the particle distribution is neglected here; thus, the beam-surroundings system is not solved self-consistently and a mixed Cauchy problem should be considered.

The problem reads: for a bunch moving with the velocity of light $c$ and characterized by a charge distribution $\rho$, find the electromagnetic field $\vec{E}, \vec{H}$ in a domain $\Omega$ which is bounded transversally by a perfect conductor $\partial\Omega$. The

*Corresponding author. Tel.: +49 6151 164 787; fax: +49 6151 164 611.
*E-mail address:* zagor@temf.de (I. Zagorodnov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.019

bunch introduces an electric current $\vec{j}=\vec{c}\rho$ and thus we have to solve

$$\nabla\times\vec{H}=\frac{\partial}{\partial t}\vec{D}+\vec{j},\quad \nabla\times\vec{E}=-\frac{\partial}{\partial t}\vec{B},\quad \nabla\cdot\vec{D}=\rho,$$
$$\nabla\cdot\vec{B}=0, \tag{1}$$

$$\vec{H}=\mu^{-1}\vec{B},\quad \vec{D}=\varepsilon\vec{E},\quad \vec{E}(t=0)=\vec{E}_0,$$
$$\vec{H}(t=0)=\vec{H}_0,\quad x\in\Omega,$$
$$\vec{n}\times\vec{E}=0,\quad x\in\partial\Omega.$$

## 3. Implicit numerical scheme

Following the matrix notation of the finite integration technique (FIT) [5], the Cauchy problem (1) can be approximated by the *time-continuous* matrix equations:

$$\mathbf{C}\widehat{\mathbf{e}}=-\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{b}},\quad \mathbf{C}^*\widehat{\mathbf{h}}=\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{d}}+\widehat{\mathbf{j}},\quad \mathbf{S}\widehat{\mathbf{b}}=\mathbf{0},$$
$$\mathbf{S}^*\widehat{\mathbf{d}}=\mathbf{q},\quad \widehat{\mathbf{e}}=\mathbf{M}_{\varepsilon^{-1}}\widehat{\mathbf{d}},\quad \widehat{\mathbf{h}}=\mathbf{M}_{\mu^{-1}}\widehat{\mathbf{b}}. \tag{2}$$

With changing of variables, system (2) reduces to the skew-symmetric one:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{e}=\mathbf{C}_0^*\mathbf{h}+\mathbf{j},\quad \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{h}=-\mathbf{C}_0\mathbf{e},\quad \mathbf{e}=\mathbf{M}_{\varepsilon^{-1}}^{-1/2}\widehat{\mathbf{e}},$$
$$\mathbf{h}=\mathbf{M}_{\mu^{-1}}^{-1/2}\widehat{\mathbf{h}},\quad \mathbf{j}=\mathbf{M}_{\varepsilon^{-1}}^{1/2}\widehat{\mathbf{j}},\quad \tau=ct. \tag{3}$$

System (3) is a *time-continuous* and *space-discrete* approximation of problem (1). The next step is a discretization of the system in time. The field components can be split in time and the "leap-frog" scheme can be applied. Below two kinds of the splitting are considered: electric–magnetic (E/M) and TE/TM schemes.

As suggested by Yee [3], the E/M splitting of the field components yields the *explicit* FDTD scheme (E/M scheme):

$$\mathbf{e}^{n+0.5}=\mathbf{e}^{n-0.5}+\Delta\tau\,\mathbf{C}_0^*\mathbf{h}^n+\Delta\tau\,\mathbf{j}^n,$$
$$\mathbf{h}^{n+1}=\mathbf{h}^n-\Delta\tau\,\mathbf{C}_0\mathbf{e}^{n+0.5}. \tag{4}$$

Scheme (4) is a two-layer scheme:

$$\mathbf{B}\frac{\mathbf{y}^{n+1}-\mathbf{y}^n}{\Delta\tau}+\mathbf{A}\mathbf{y}^n=\mathbf{f}^n,\quad \mathbf{B}=\begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \Delta\tau\,\mathbf{C}_0 & \mathbf{I}\end{pmatrix},$$
$$\mathbf{A}=\begin{pmatrix}\mathbf{0} & -\mathbf{C}_0^*\\ \mathbf{C}_0 & \mathbf{0}\end{pmatrix},\quad \mathbf{y}^n=\begin{pmatrix}\mathbf{e}^{n-0.5}\\ \mathbf{h}^n\end{pmatrix},\quad \mathbf{f}^n=\begin{pmatrix}\mathbf{j}^n\\ \mathbf{0}\end{pmatrix}.$$

Discrete energy of electromagnetic fields can be defined as

$$E_{\mathrm{E/M}}^n=0.5\langle[\mathbf{B}-0.5\Delta\tau\,\mathbf{A}]\mathbf{y}^n,\mathbf{y}^n\rangle$$
$$=0.5(\langle\mathbf{e}^{n-0.5},\mathbf{e}^{n-0.5}\rangle+\langle\mathbf{h}^n,\mathbf{h}^{n-1}\rangle). \tag{5}$$

If the right-hand side vanishes, then the scheme is energy conserving $E_{\mathrm{E/M}}^n=E_{\mathrm{E/M}}^0$.

Scheme (4) is widely used in electromagnetic modeling. However, the FDTD algorithm causes non-physical dispersion of the simulated waves in a free-space computational lattice.

Why is zero dispersion for a special direction important? Unlike plasma problems, the charged particles in accelerators are organized and a direction of motion (the longitudinal direction) can be identified. Hence, the computational domain is very long in the longitudinal direction and relatively short in the transverse plane. Additionally, the electromagnetic field changes very fast in the direction of motion but is relatively smooth in the transverse plane.

To find the scheme, let us consider Fig. 1 and subdue an update procedure to the motion of the bunch. We suggest that a charged particle is moving in the $z$-direction with velocity of light. Additionally, let us suggest that our numerical scheme allows to take a time step $\Delta\tau$ equal to the mesh step $\Delta z$ in the $z$-direction. If at the time $\tau_0$ the particle has the position aligned with the left $z$-facet of the primary grid (see Fig. 1), then in the time $\tau_0+0.5\Delta\tau$, it will be aligned with the left $z$-facet of the dual grid and in the time $\tau_0+\Delta\tau$, it will be again aligned with the next $z$-facet of the primary grid. It suggests that we should replace the E/M time splitting of the field components in scheme (4) by a more adequate TE/TM splitting. Indeed, in the time $\tau_0$, it is reasonable to update the TM components $\mathbf{e}_x,\mathbf{e}_y,\mathbf{h}_z$ and the half time step later, namely in time $\tau_0+0.5\Delta\tau$, we have to update the TE components $\mathbf{h}_x,\mathbf{h}_y,\mathbf{e}_z$.

Following the above consideration, let us rewrite scheme (4) in the equivalent form:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{u}=\mathbf{D}_{11}\mathbf{u}+\mathbf{D}_{12}\mathbf{v}+\mathbf{j}_u,$$
$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{v}=\mathbf{D}_{22}\mathbf{v}-\mathbf{D}_{12}^*\mathbf{u}+\mathbf{j}_v \tag{6}$$

where $\mathbf{u}=(\mathbf{h}_y,\mathbf{h}_y,\mathbf{e}_z)^{\mathrm{T}}$ and $\mathbf{v}=(\mathbf{e}_x,\mathbf{e}_y,\mathbf{h}_z)^{\mathrm{T}}$.

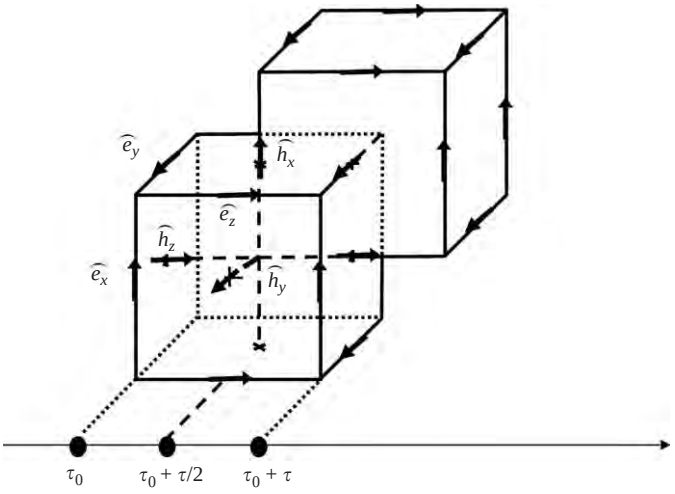



Fig. 1. Positions of the relativistic charged particle in the FIT grid in different moments of time. The time step is chosen equal to the longitudinal mesh step.

Applying to system (6) the suggested TE/TM splitting of the field in time, the following numerical scheme is obtained:

$$\frac{\mathbf{u}^{n+0.5}-\mathbf{u}^{n-0.5}}{\Delta\tau}=\mathbf{D}_{11}\frac{\mathbf{u}^{n+0.5}+\mathbf{u}^{n-0.5}}{2}+\mathbf{D}_{12}\mathbf{v}^{n}+\mathbf{j}_{u}^{n}, \tag{7a}$$

$$\frac{\mathbf{v}^{n+1}-\mathbf{v}^{n}}{\Delta\tau}=\mathbf{D}_{22}\frac{\mathbf{v}^{n+1}+\mathbf{v}^{n}}{2}+\mathbf{D}_{21}\mathbf{u}^{n+0.5}+\mathbf{j}_{v}^{n+0.5}. \tag{7b}$$

As scheme (4), scheme (7) is also two-layer one:

$$\mathbf{B}\frac{\mathbf{y}^{n+1}-\mathbf{y}^{n}}{\Delta\tau}+\mathbf{A}\mathbf{y}^{n}=\mathbf{f}^{n},\quad \mathbf{A}=-\mathbf{A}^{*},$$
$$\mathbf{Q}=\mathbf{Q}^{*},\quad \mathbf{Q}=\mathbf{B}-0.5\Delta\tau\mathbf{A}.$$

With the help of the energy inequalities method [6], we can prove

**Theorem 1.** The condition $\mathbf{Q}\equiv\mathbf{B}-0.5\Delta\tau\,\mathbf{A}\geqslant 0$ is necessary and sufficient for the stability in Hilbert space $H_Q$ (see [6]) of schemes (4) and (7) with respect to the initial data $\mathbf{y}^0$ and to the right-hand side $\mathbf{f}^n$.

Discrete energy in TE/TM scheme can be defined by the relation

$$E^{n}_{\mathrm{TE/TM}}=0.5\langle[\mathbf{B}-0.5\Delta\tau\,\mathbf{A}]\mathbf{y}^{n},\mathbf{y}^{n}\rangle = E^{n}_{\mathrm{E/M}}+O(\Delta\tau^{2}).$$

If the right-hand side vanishes, then the scheme is energy conserving $E^{n}_{\mathrm{TE/TM}}=E^{0}_{\mathrm{TE/TM}}$.

Stability condition can be rewritten in the form $\mathbf{I}-(\Delta\tau^{2}/4)\mathbf{P}_{z}^{i}\left(\mathbf{P}_{z}^{i^{*}}\right)\geqslant 0$. The last relation resembles the well-known stability condition of the explicit FDTD scheme for the one-dimensional problem. In the following, an equal mesh step $\Delta z$ in the $z$-direction will be always suggested. Then for vacuum domain with *staircase* approximation of the boundary, the stability condition reads $\Delta\tau\leqslant\Delta z$.

With the time step $\Delta\tau$ equal to the longitudinal mesh step $\Delta z$, scheme (7) does not have dispersion in the longitudinal direction. Stability condition does not contain information about the transverse mesh. Hence, the transverse mesh can be chosen independent of stability considerations.

So far, we have found a scheme which fulfills the main requirements formulated above. However, in a general case, the staircase scheme is only first-order accurate. To overcome the problem and avoid reducing of the stable time step, the uniformly stable conformal (USC) approach described in Refs. [2,7] will be used.

With the latter approach, the scheme possesses the desired features. However, it is implicit and non-economical. The economical scheme modifications, based on operator splitting, will be considered in the next section.

## 4. TE/TM-ADI2 scheme in three dimensions

In the following, we consider only modification of relation (7a). The numerical scheme using the ADI2 splitting in three dimensions has the form

$$\frac{\widehat{\mathbf{h}}_{x}^{n+0.5}-\widehat{\mathbf{h}}_{x}^{n-0.5}}{\Delta\tau}=\mathbf{M}_{\mu_{x}^{-1}}\left[\mathbf{P}_{z}\widehat{\mathbf{e}}_{y}^{n}-\mathbf{P}_{y}\frac{\widehat{\mathbf{e}}_{z}^{n+0.5}+\widehat{\mathbf{e}}_{z}^{n-0.5}}{2}\right], \tag{8a}$$

$$\frac{\widehat{\mathbf{h}}_{y}^{n+0.5}-\widehat{\mathbf{h}}_{y}^{n-0.5}}{\Delta\tau}=-\mathbf{M}_{\mu_{y}^{-1}}\left[\mathbf{P}_{z}\widehat{\mathbf{e}}_{x}^{n}-\mathbf{P}_{x}\frac{\widehat{\mathbf{e}}_{z}^{n+0.5}+\widehat{\mathbf{e}}_{z}^{n-0.5}}{2}\right], \tag{8b}$$

$$\begin{aligned}\mathbf{W}^{\mathrm{e}}_{\mathrm{ADI2}}\frac{\widehat{\mathbf{e}}_{z}^{n+0.5}-\widehat{\mathbf{e}}_{z}^{n-0.5}}{\Delta\tau}&=\mathbf{M}_{\varepsilon_{z}^{-1}}\mathbf{P}_{y}^{*}\left(\widehat{\mathbf{h}}_{x}^{n-0.5}+\frac{\Delta\tau}{2}\mathbf{M}_{\mu_{x}^{-1}}\left[\mathbf{P}_{z}\widehat{\mathbf{e}}_{y}^{n}-\mathbf{P}_{y}\widehat{\mathbf{e}}_{z}^{n-0.5}\right]\right)\\&+\mathbf{M}_{\varepsilon_{z}^{-1}}\mathbf{P}_{x}^{*}\left(-\widehat{\mathbf{h}}_{y}^{n-0.5}-\frac{\Delta\tau}{2}\mathbf{M}_{\mu_{y}^{-1}}\left[-\mathbf{P}_{z}\widehat{\mathbf{e}}_{x}^{n}+\mathbf{P}_{x}\widehat{\mathbf{e}}_{z}^{n-0.5}\right]\right)\\&+\mathbf{M}_{\varepsilon_{z}^{-1}}\widehat{\mathbf{j}}_{z}\end{aligned} \tag{8c}$$

where

$$\begin{aligned}\mathbf{W}^{\mathrm{e}}_{\mathrm{ADI2}}=&\left(\mathbf{I}+\frac{\Delta\tau^{2}}{8}\mathbf{M}_{\varepsilon_{z}^{-1}}\mathbf{P}_{y}^{*}\mathbf{M}_{\mu_{x}^{-1}}\mathbf{P}_{y}\right)\\&\times\left(\mathbf{I}+\frac{\Delta\tau^{2}}{4}\mathbf{M}_{\varepsilon_{z}^{-1}}\mathbf{P}_{x}^{*}\mathbf{M}_{\mu_{y}^{-1}}\mathbf{P}_{x}\right)\\&\times\left(\mathbf{I}+\frac{\Delta\tau^{2}}{8}\mathbf{M}_{\varepsilon_{z}^{-1}}\mathbf{P}_{y}^{*}\mathbf{M}_{\mu_{x}^{-1}}\mathbf{P}_{y}\right).\end{aligned}$$

If material matrices $\mathbf{M}_{\mu^{-1}}$, $\mathbf{M}_{\varepsilon^{-1}}$ are diagonal, then system (8c) has only product of tri-diagonal matrices on the left-hand side and can be resolved easily.

However, the conformal scheme with the diagonal material matrices reduces the stable time step. To restore stability condition and possibility to use the time step $\Delta\tau=\Delta z$, we will use modification of USC method [7]. The last approach results in modified non-diagonal but symmetric matrices $\mathbf{M}_{\mu_{x}^{-1}}$, $\mathbf{M}_{\mu_{y}^{-1}}$. Other material matrices in scheme (7)–(8) remain diagonal. It means that solution of Eq. (8c) requires additional efforts.

To overcome the problem, an iterative method can be applied [2]. Already the first iteration (which we refer as TE/TM-ADI2(1) scheme) produces accurate and stable results.

## 5. Numerical examples calculated by three-dimensional code

Finally, we discuss the results of numerical computations with fully three-dimensional realization of TE/TM-ADI2(1) schemes (7) and (8). Two test problems are considered.

To be able to check the accuracy of the three-dimensional realization of the TE/TM scheme, we have

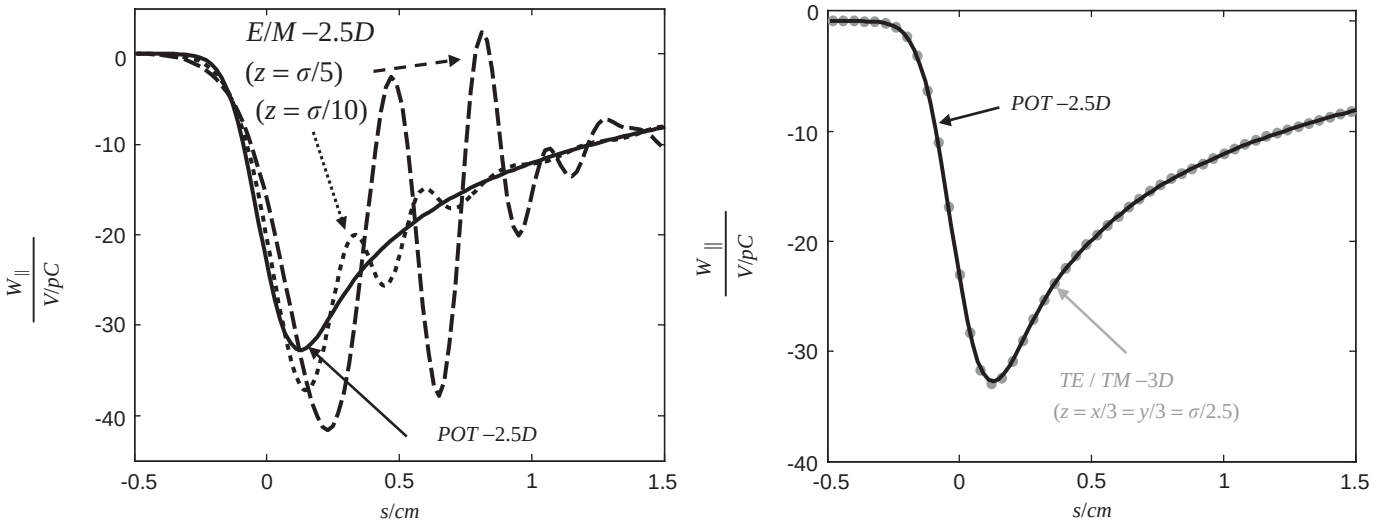



Fig. 2. Comparison of the wake potentials obtained by different methods for structure consisting of 20 TESLA cells excited by the Gaussian bunch with $\sigma = 1$ mm.

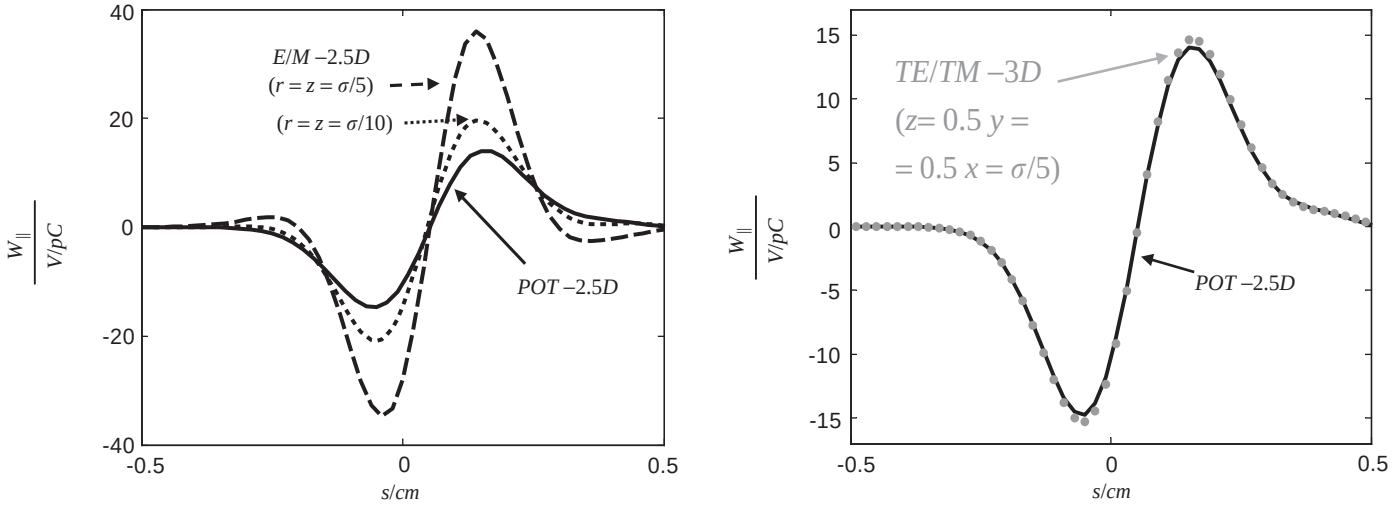



Fig. 3. Comparison of the wake potentials obtained by different methods for the round collimator.

chosen for the tests only rotationally symmetric structures. However, in the three-dimensional calculations, the symmetry of the structures was not exploited.

As the first example, we consider a structure consisting of the 20 TESLA cells supplied with infinite ingoing and outgoing pipes with diameter 35 mm.

Fig. 2 shows the longitudinal wake potential for Gaussian bunch with RMS length $\sigma = 1$ mm moving on the axis. The left figure shows results calculated by two (2.5D)-dimensional codes. The solid line (POT-2.5D) corresponds to the reference solution obtained with the vector potential method [2]. The two other lines show results obtained with different mesh resolution from MAFIA code [4] based on the classical Yee's scheme (E/M-2.5D). The observed oscillations are due to dispersion error of the Yee's scheme. The right figure shows comparison of the reference solution (solid line) with the results obtained with the help of three-dimensional schemes (7) and (8) (marked as TE/TM-3D).

It can be seen that the three-dimensional TE/TM-ADI2 scheme produces very accurate results even with the coarse meshes. Indeed, in the example, the three-dimensional code uses only 2.5 mesh points on $\sigma$ in the longitudinal direction.

As the last example, we use the round collimator example. Fig. 3 demonstrates the wake potential for the collimator with parameters $a = 30$ mm, $b = 2$ mm, $c = 50$ mm, $L = 200$ mm and the Gaussian bunch with RMS length $\sigma = 1$ mm. Again, the high accuracy of the suggested three-dimensional scheme can be seen.

## Acknowledgment

Work supported in part by the Deutsche Forschungsgemeinschaft, project WE1239/22-1.

## References

[1] A. Novokhatski, M. Timm, T. Weiland, Transition dynamics of the wake fields of ultra short bunches, TESLA Report 2000-03, DESY, 2000.
[2] I. Zagorodnov, R. Schumann, T. Weiland, J. Comput. Phys. 191 (2003) 525.
[3] K.S. Yee, IEEE Trans. Antennas Propag. 14 (1966) 302.
[4] MAFIA Collaboration, MAFIA Manual, CST GmbH, Darmstadt, 1997.
[5] T. Weiland, Int. J. Numer. Model. 9 (1996) 295.
[6] A.A. Samarskii, The Theory of Difference Schemes, Marcel Dekker, Inc., New York, 2001.
[7] I. Zagorodnov, R. Schuhmann, T. Weiland, Int. J. Numer. Model. 16 (2003) 127.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 100–105
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Precision modeling of grid fields and beam dynamics in a large-aperture accelerator for an electroionization laser

Yuri V. Zuev

*D.V. Efremov Institute of Electrical and Physical Apparatus (NIIEFA), RU-196641 Saint Petersburg, Russian Federation*

Available online 22 November 2005

**Abstract**

The accelerator considered is a planar tetrode-type system. It contains a grid of lengthy glowers emitting beamlets, which cover an anode exit window of 6600 $cm^2$ in area. Key parameters of the system, i.e. maximal extracted current, continuous service life and reliability, depend directly on the physical properties of the exit window foil, as well as non-uniformity in density of the accelerated beam. Adequate computer simulation is necessary in designing the latter. In this paper, careful modeling of electrostatic fields, which are highly nonlinear near the grid rods of small diameter, has been performed using the boundary collocation method. Stray magnetic cathode fields were found analytically. Electron ray tracing in these fields provided useful information on abnormal local heating of the foil due to aberrations in the beamlets.


*PACS:* 29.27.Eg; 41.20.Cv; 41.85.Ew

*Keywords:* Electron optics; Grid fields; Large-area beam; Electroionization laser

## 1. Introduction

Large-area electron beams are widely used in radiotechnology and plasmochemistry. The necessary cross-section size of the beams is usually the result of numerous beamlet overlapping. The accelerator considered in this paper is involved in the ionization device of a mobile powered $CO_2$ laser intended for remote metal construction cutting in accident emergencies [1]. It is an electrostatic three-stage accelerator. Fig. 1 shows schematically the accelerator optics with calculated trajectories of the electrons produced by one emitter (i.e. with a beamlet).

In general, the accelerator cathode contains a flat grid of lengthy glowers emitting beamlets. The beamlets combine to form an integral electron flow with a typical spread of transverse velocities. The cathode operates in the space-charge-limited mode. The beamlet current is controlled by the low voltage of the first thin grid. The second thicker grid screens the cathode from the main accelerating field of 160–180 keV applied between the grid and the anode. The anode simultaneously serves as a beam output device. It consists of a foil support window through which the electrons are injected into the laser gas medium. Not far from the glowers, a reflecting screen (spreader) is mounted at the cathode potential. All the grids take the form of rods parallel to the cathode filaments. Rod diameter and spacing are chosen with regard to breakdown strength.

Destination and service conditions of the laser place a high requirement on impulse density of the excitation beam with consequent restrictions on the accelerator parameters, namely, maximal extracted current, continuous service life and reliability. These parameters directly depend on the physical properties of the exit window foil, as well as divergence and spatial non-uniformity of the accelerated beam. Adequate computer simulation of the latter is necessary for design study.

## 2. Electric field

Grid field calculations in the accelerator are reduced to the solution of the first boundary problem in the Laplace equation. Applying common numerical methods, such as finite difference or finite elements, to the problem failed for

*E-mail address:* yuri_zuev@mail.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.020

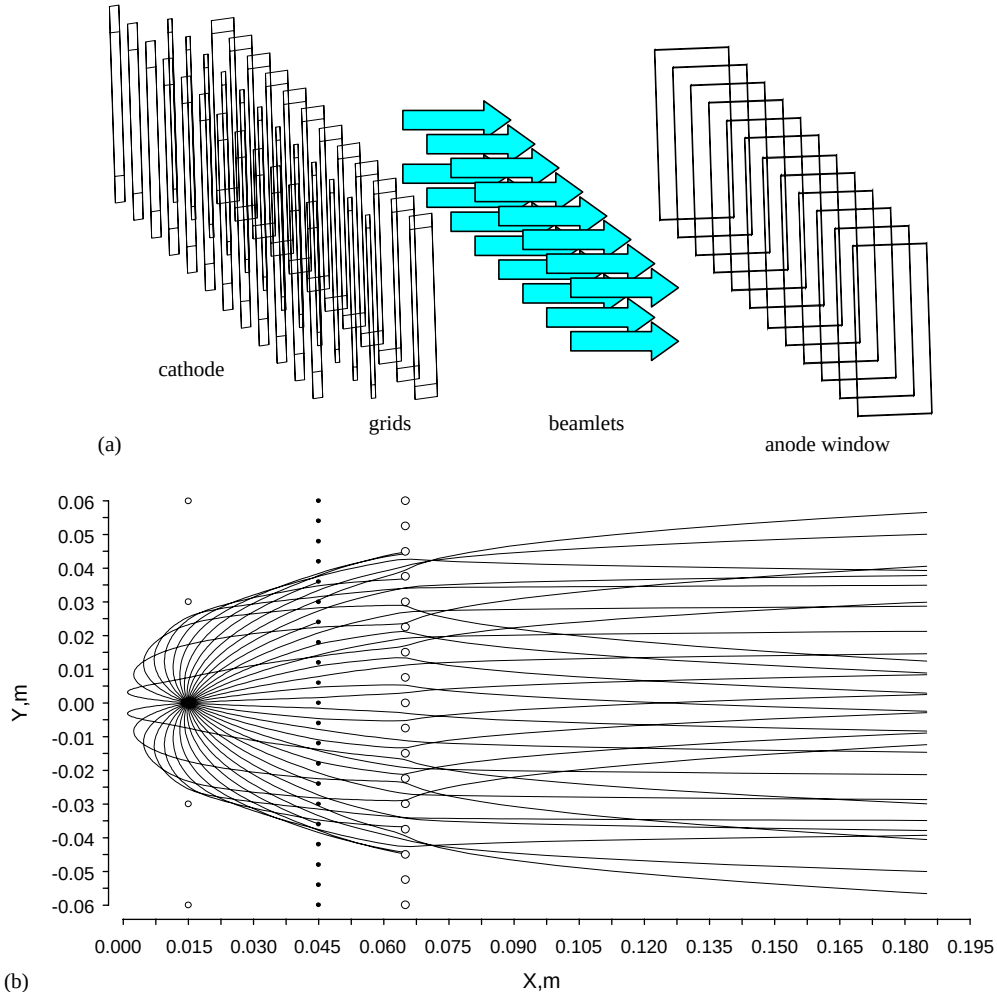



Fig. 1. Layout of accelerator optics and calculated trajectories of one beamlet.

the following two reasons: the large disproportion of the rod cross-section size to the accelerator volume, the aliquant rod spacing in the grids.

In practice, variously approximating solutions to the problem are used [3,4]. As a rule, for example, see Ref. [5], the planar grid field is replaced by the potential field of a row of infinitely long line charges of charge density $q$ per meter:

$$V(x,y) = \frac{q}{4\pi\varepsilon_0} \ln\left(2\cosh\left(\frac{2\pi x}{p}\right) - 2\cos\left(\frac{2\pi y}{p}\right)\right), \qquad (1)$$

where $p$ is the line charge spacing. For validity of Eq. (1) in a real situation, the grid rod diameter must be small and the flat electrodes, such as the anode or spreader, must be parallel to and distant from the grid. Otherwise, the rod contours and the electrode profiles mismatch to the respective equipotential surfaces of Eq. (1). Moreover, the approximation neglects the edge effect caused by the finite length and number of rods. Nevertheless, the effect can be compensated constructively within certain limits.

The approximation is improved by replacing every grid with a system of dummy line charges [6]. In the simplest case it takes the form

$$U(x,y) = A + Bx + \sum_{i=1}^{N_G} \sum_{j=1}^{N_C} V_{i,j}(x,y), \qquad (2)$$

where

$$V_{i,j}(x,y) = \left(\frac{q_{i,j}}{4\pi\varepsilon_0}\right) \ln\left(2\cosh\left(\frac{2\pi(x-x_{i,j})}{p_i}\right) - 2\cos\left(\frac{2\pi(y-y_{i,j})}{p_i}\right)\right)$$

is the potential filed produced by the $j$th infinite row of dummy line charges replacing the $i$th grid. The coordinates $(x_{i,j}, y_{i,j})$ comply with singular points of the field and specify a spatial arrangement of the line charges. It is self-evident that dummy charges have to be placed inside the grid rods. To provide proper equipotential contours of plane electrodes, part of the line charges may be located out of the computed area, following the mirror-image method.

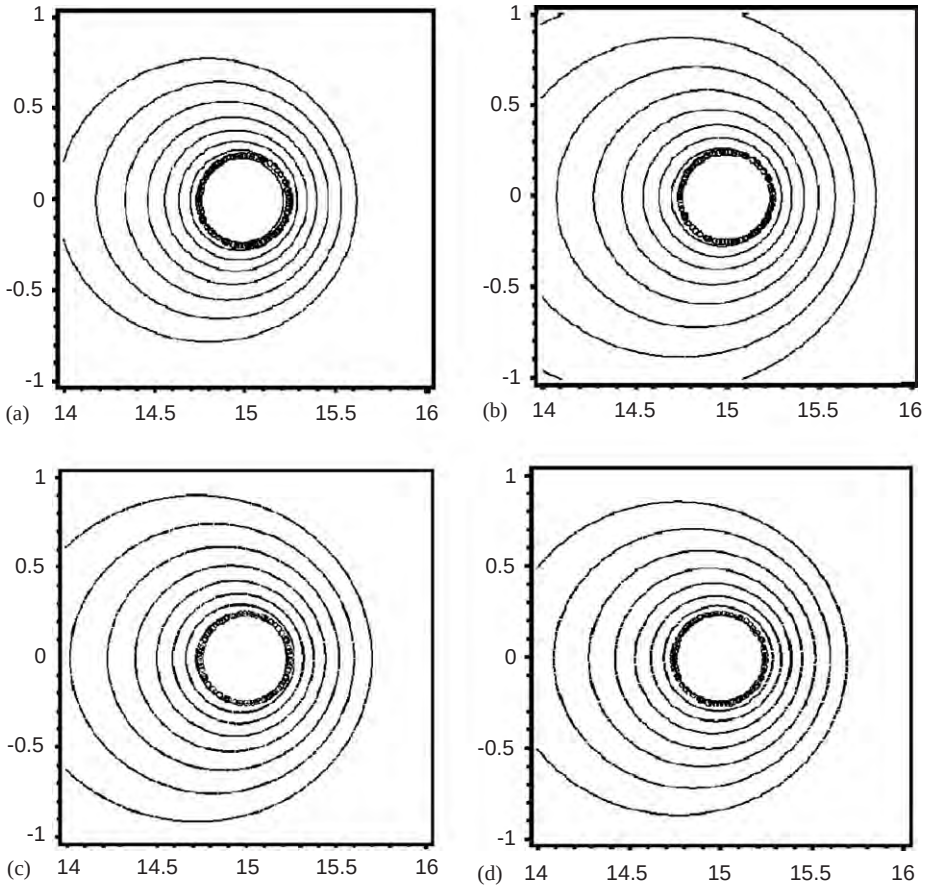



Fig. 2. Variation of the same equipotential lines near the cathode filament with number of dummy charges and position of collocation points. Hollow circles show real contour of filament.

Otherwise, the dummy charge arrangement is arbitrary, while influencing the accuracy of the approximation used. If $(x_{i,j}, y_{i,j})$ have been chosen, $U(x,y)$ is fitted to the boundary condition without difficulty by means of the other free parameters: $A$, $B$ and $q_{i,j}$. Substituting the coordinates of collocation points $(x_k, y_k)$ and their associated values of potential $U_k$ in Eq. (2) gives the mandatory system of linear equations:

$$U(x_k, y_k) = U_k, \quad k = 1 \ldots N_G N_C + 2.$$

The system is solvable by conventional methods.

Figs. 2 and 3 present the results of such an operating sequence. The figures depict the field maps, calculated by Eq. (2), by varying the collocation points and the number of line charges. Variants (a)–(c) correspond to one dummy charge being at the center of a real rod. Collocation points lie on the rod contour and have the coordinates: (a) $x_1 = x_0 - R_0$, $y_1 = y_0$; (b) $x_1 = x_0 - R_0$, $y_1 = y_0$; (c) $x_1 = x_0$, $y_1 = y_0 \pm R_0$, where $(x_0, y_0)$ is the center of the rod, $R_0$ is rod radius. Variant (d) corresponds to two dummy charges at $x_{1,2} = x_0 \pm \delta x$; , $y_{1,2} = y_0$, and the collocation points at $x_{1,2} = x_0 \pm R_0$, $y_{1,2} = y_0$. Comparing Figs. 2 and 3 reveals a strong dependence of approximation accuracy on differences in electric strength at the front and rear of the grid. Having added terms, such as

$$u_n(x,y) = \frac{q_n}{4\pi\varepsilon_0} \ln\left[\frac{\cosh(\pi(y-y_n)/h) + \cos(\pi(x+x_n)/h)}{\cosh(\pi(y-y_n)/h) - \cos(\pi(x-x_n)/h)}\right], \quad (3)$$

Eq. (2) offers a simple way of accounting for edge effects and grid defects (Fig. 4). Eq. (3) is the potential field produced by one line charge $q_n$ placed at the point $(x_n, y_n)$ between two parallel metal plates at a distance $h$ [2,4].

## 3. Magnetic field

The magnetic field $\bar{B}$ induced by the cathode filament current may be calculated in quasistationary approximation through the vector magnetic potential $\bar{A}$. In plane-parallel fields, vector $\bar{A}$ has a single component $A_z(x,y)$, which satisfies the Poisson equation:

$$\frac{\partial^2 A_z}{\partial x^2} + \frac{\partial^2 A_z}{\partial y^2} = -\mu_0 j_z(x,y), \quad (4)$$

where $j_z$ is filament current density. A solution to Eq. (4) with allowances for translation symmetry (equispaced cathode filaments and oppositely directed currents in

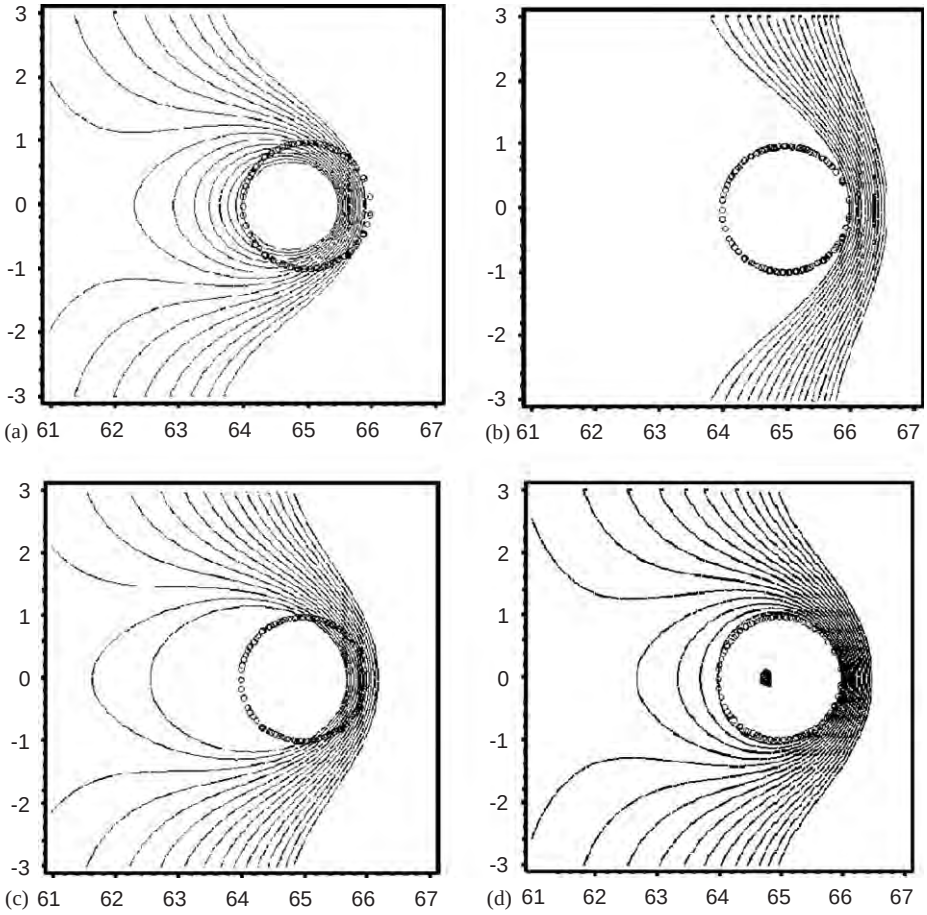



Fig. 3. Variation of the same equipotential lines near the screening grid rod with number of dummy charges and position of collocation points. Hollow circles show real contour of rod.

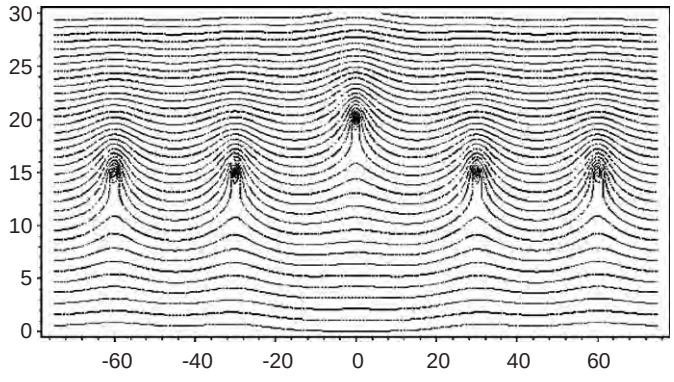


Fig. 4. Example of overstrained electric field in the system with defective grid.

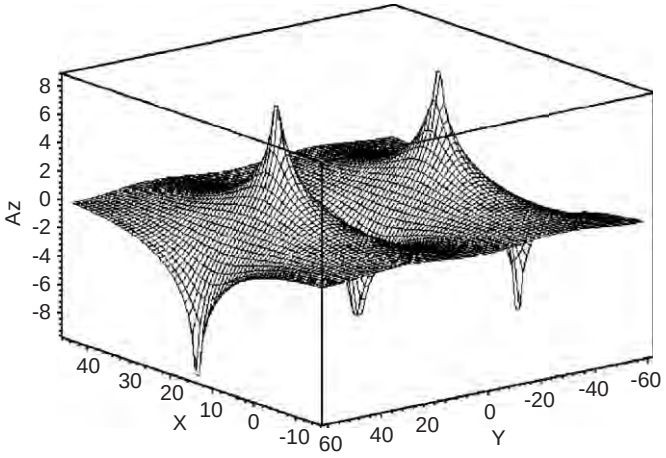



Fig. 5. General view of component $A_z(x, y)$.

adjacent filament) can be obtained in the same manner as the solution to the problem of an infinitely long line charge placed between parallel metal plates [2]:

$$A_z(x, y) = A_0 \ln\left[\frac{\cosh(\pi(x - x_0)/p) + \cos(\pi y/p)}{\cosh(\pi(x - x_0)/p) - \cos(\pi y/p)}\right]. \quad (5)$$

Here, $p$ is the spacing between the cathode filaments and $x_0$ is the coordinate of the central filament. The integration constant $A_0$ is calculated from the condition

$$\oint_l \bar{B}\,\mathrm{d}l = \mu_0 I,$$

where $l$ is an arbitrary contour that encloses the filament with a current $I$. The general view of component $A_z(x, y)$ is given in Fig. 5.

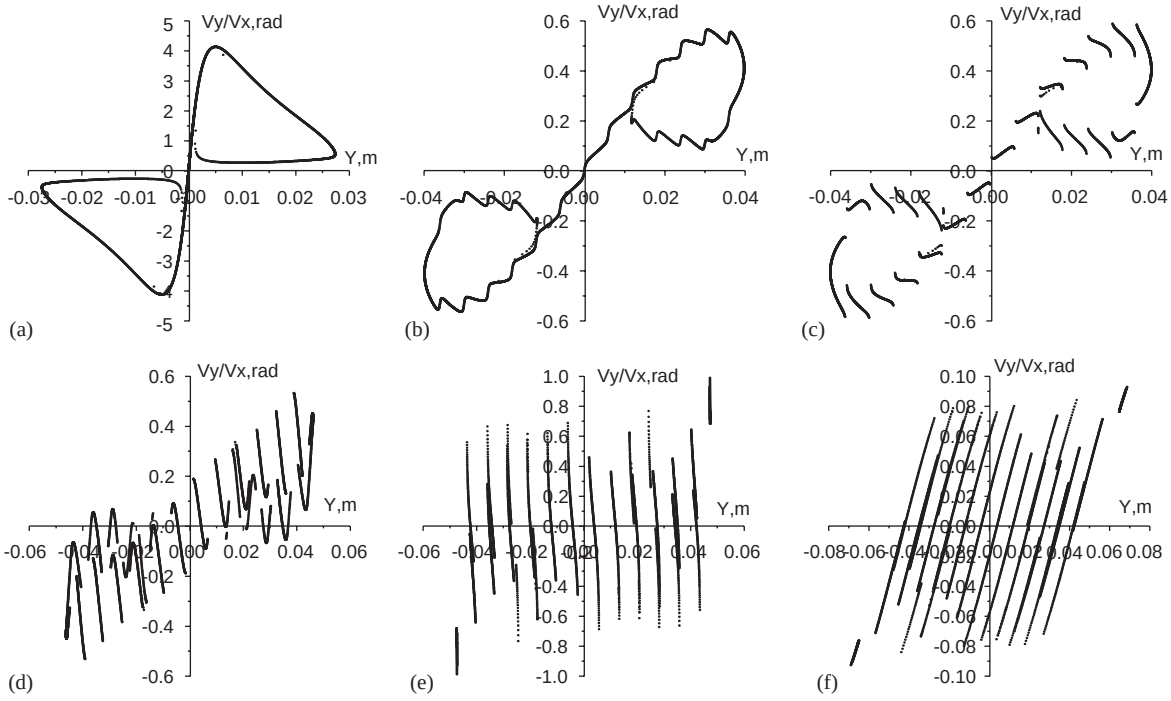



Fig. 6. Phase portrait of particles on plane $yy'$ at $x = 15.5$, 44.8, 45.2, 64, 66 and 185 mm. The particles are emitted by one cathode filament located at $x = 15$ mm.

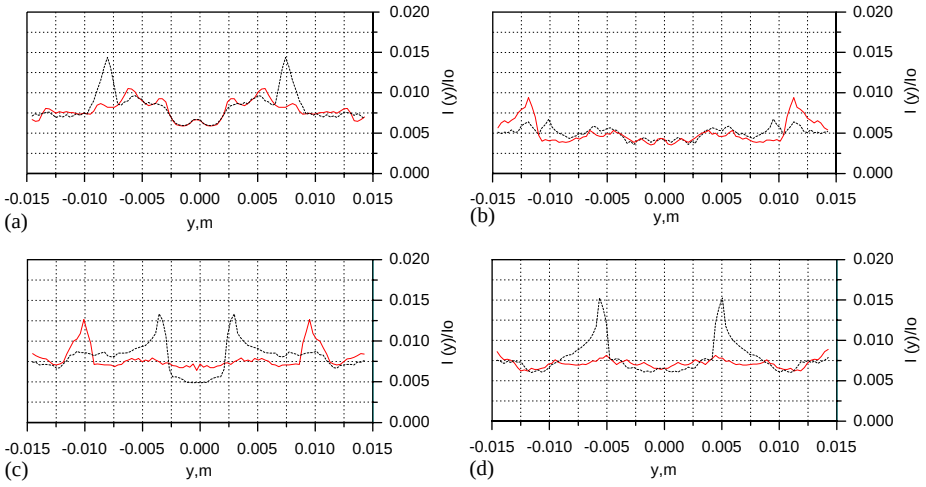



Fig. 7. Beam current distribution over an elementary optics cell near the foil support window at zero (dash) and maximal (solid) magnetic field intensity. Cases (a)–(d) correspond to the electric field approximation of cases (a)–(d) in Figs. 2 and 3, respectively. Operating mode: grid potentials are 800 and 1500 V; anode potential is 200 kV.

## 4. Ray tracing

Ray tracing was performed by integrating a relativistic form of the Newton–Lorentz equation in the fields of Eqs. (2) and (5). It was assumed that, at the initial moment, all test particles are uniformly distributed over the surface of the cathode filament. Integration was carried out numerically by the fourth-order Runge–Kutta method with automatic step selection. The solution accuracy was checked with the integrals of motion:

$$m_0\gamma c^2 = m_0c^2 - e(U(x,y) - U(x_S,y_S)),$$
$$m_0\gamma\dot{z} = -e(A_z(x,y) - A_z(x_S,y_S)),$$

the relative deviation from which was no greater than 0.1%. Here, $U(x_S,y_S)$ and $A_z(x_S,y_S)$ are the electrostatic and vector potentials at the start point, and are the same for all particles from any emitter because $x_S^2 + y_S^2 = r_C^2$,

where $r_C$ is the radius of the cathode filament. Since particle travel time from cathode to anode is much shorter than filament current alternation, effects associated with magnetic field variation in time are neglected.

Detailed analysis of tracing can be found in Ref. [7]. Fig. 6 shows the evolution of a beamlet on the $yy'$ plane as the particles travel through the accelerator optics. The beam-cutting and beam-focusing effects due to grid rods placed in the planes $x = 45$ and 65 mm are clearly seen. Fig. 7 demonstrates the extent to which computed spatial non-uniformity in the beam near the exit window depends on the accuracy of the electric field approximation, as well as magnetic field effects.

The accelerator optics inevitably causes an intersection of the adjacent trajectories. The beamlet envelope is found to be a caustic surface (Figs. 1 and 6). The resultant density distribution is most uniform when the beamlet edges are intercepted by the second grid. Otherwise, the efficiency of the accelerator is higher but the beam distribution contains regions of anomalous density. To preserve the foil from dangerous local overheating, the accelerating voltage pulse must be phase-synchronized with the sinusoidal filament current.

## References

[1] M.A. Abroyan, et al., in: Proceeding of the 10th International Conference on Applied Charged Particle Accelerators in Medicine and Industry, St. Petersburg, 1–4 Oct., 2001, pp. 422–424.
[2] G.A. Grinberg, Selected Topics of Mathematical Theory of Electrical and Magnetic Phenomena, Akad. Nauk SSSR, Moscow, 1948.
[3] W.R. Smythe, Static and Dynamic Electricity, McGraw-Hill, New York, 1950.
[4] A.M. Strashkevich, Electron Optics of Electrostatic Systems, Energiya, Moscow, 1966.
[5] D.A. Huchital, J.D. Rigden, J. Appl. Phys. 43 (5) (1972) 2291.
[6] M.S. Epshtein, in: Proceedings of the Third All-union Workshop on Computation Methods for Electron-Optical Systems, Novosibirsk, 1970, pp. 104–112.
[7] M.A. Abroyan, Yu.V. Zuev, S.L. Kosogorov, V.Ya. Shvedyuk, Zh. Tekh. Fiz. 73 (8) (2003) 98 (Tech. Phys. 48 (8) (2003) 1035).

Available online at www.sciencedirect.com
SCIENCE DIRECT
Nuclear Instruments and Methods in Physics Research A 558 (2006) 106–111
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# High-order representation of Poincaré maps

Johannes Grote*, Martin Berz, Kyoko Makino

*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*

Available online 23 January 2006

**Abstract**

A method to obtain DA approximations of Poincaré maps directly from a DA approximation of the flow of the differential equation for certain types of flows and Poincaré sections is presented. Examples of the performance of the method, its computational implementation, and its use for problems in beam physics are given.


*PACS:* 41.85.−p; 05.10.−a; 02.60.Jh; 02.60.Lj

*Keywords:* Poincaré map; Differential algebra

## 1. Introduction

Poincaré maps are a standard tool in general dynamical systems theory for the study of properties of a system under consideration, e.g. the flow generated by an ordinary differential equation. A Poincaré map essentially describes how points on a plane $S_i$ (a Poincaré section) which is transversed by such an orbit $\mathcal{O}$ (the reference orbit) and which are sufficiently close to $\mathcal{O}$ get mapped onto another plane $S_f$ by the flow. A frequent application is the case where $S_i = S_f$, and one of the most prominent applications is the study of asymptotic stability of periodic or almost periodic orbits.

For applications in Accelerator Physics, Poincaré maps are important because the dynamics is usually described in terms of so-called curvilinear coordinates, i.e. an orthogonal coordinate system that is attached to a reference orbit such that one of its axes points in the direction of its velocity, another one in the direction of its acceleration component perpendicular to the velocity, and so on; for details see for example Refs. [1–4]. Instead of solving the ODEs of the original system under consideration with time as the independent variable, the ODEs are transformed such that the new independent variable is the arclength along the reference orbit, and everything is described in terms of the coordinates in the plane perpendicular to the reference orbit.

The benefits of this approach are many: for imaging systems like electron microscopes or other optical systems, the method directly describes how particles originating in the object plane $S_i$ are mapped into the image plane $S_f$, since what matters is the position on this image plane where a detector or, in earlier days, photographic paper or plates are located. Likewise, for dispersive systems, it describes how particles of different energies are mapped into different locations in a detector plane $S_f$. Finally, for large storage rings and circular accelerators, one usually picks one plane $S$ in the ring and assesses long-term stability by studying the Poincaré map for $S = S_i = S_f$.

For many of the conventional particle optical systems, the reference orbit and the dynamics in the corresponding curvilinear coordinates is well-known [1,5]. However, for some of the modern particle optical elements this is not the case, and the mere formulation of the ODEs describing the system under consideration represents a significant problem. Two of the prominent examples of such cases are the dynamics in muon accelerators and storage rings, which are characterized by very large emittances and unusual field arrangements [6–8], and the analysis of modern high-resolution large acceptance particle spectrographs, where the details of the orbits in the fringe field regions of magnets

*Corresponding author.
*E-mail addresses:* grotejoh@msu.edu (J. Grote), berz@msu.edu (M. Berz), makino@msu.edu (K. Makino).
*URL:* http://bt.pa.msu.edu.

0168-9002/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2006.01.001

play a prominent role [9–11]. In these cases it is still possible to solve the underlying ODEs in conventional Cartesian coordinates with time as an independent variable; but for the connection with subsequent analysis, it is necessary to transform the results to the form of a Poincaré map.

In the following sections, we will show how this can be obtained using differential algebraic (DA) tools as in Refs. [2,12]. We remark that the proposed algorithm is a part of an extended method which allows the computation of rigorous interval enclosures of the polynomial approximation of the Poincaré map discussed here.

## 2. Review: Essential DA-tools

The DA tools which are necessary to appreciate the method are described in detail in Ref. [2]. However, we wish to review briefly the two most important applications of DA-methods as far as they relate to the problem which is discussed here: the DA-integration method which is employed to obtain high-order polynomial approximations of the flow $\varphi(x_0,t)$ and the functional inversion tools which are necessary in later steps of the algorithm.

### 2.1. DA-integration of ODEs

First, we tackle the problem of obtaining a polynomial approximation of the dependence on initial conditions of the solution of the initial value problem

$$\dot{x}(t)=f(x(t),t),\quad x(0)=X_0+x_0 \tag{2.1}$$

where $f:\mathbb{R}^v\supset U^{\text{open}}\longrightarrow\mathbb{R}^v$ is given as a composition of intrinsic functions which have been defined in DA-arithmetic. As a byproduct this also entails that $f$ exhibits sufficient smoothness to guarantee existence and uniqueness of solutions for all initial conditions. The vector $X_0\in\mathbb{R}^v$ is constant and the midpoint of the domain box $D=[-d_i,d_i]^v$, $i\in\{1,\ldots,v\}$, for the small relative initial conditions $x_0\in D$. Typical box widths $d_i$ are of the order $10^{-2}$–$10^{-8}$. The polynomial approximation $\varphi(x_0,t)$ of the flow of Eq. (2.1) we desire is an expansion in terms of the independent time coordinate $t$ and the relative initial conditions $x_0$, and the representation of this approximation is a so-called DA-vector which stores the expansion coefficients up to a desired order $n$ in a structured fashion.

To achieve the aforementioned goal, we proceed by recalling that the standard procedure of a Picard-iteration yields a polynomial approximation of the solution of (2.1) after repeated application of a Picard-operator on the initial conditions. The iteration in general increases the order of the expansion by one in every step, and since a DA-vector can store coefficients up to a prespecified order $n$, we expect that the iteration converges after finitely many steps in the DA-case.

Accordingly, the Picard-operator in the DA-computation is defined by

$$\mathscr{C}(.):=(X_0+x_0)+\partial_{v+1}^{-1}f(.)$$

where $f$ is computed in DA-arithmetic and $\partial_{v+1}^{-1}$ is the *antiderivation operator*, essentially the integration with respect to the $(v+1)$st variable $t$. It can now be shown that $\mathscr{C}$ is a contracting operator (with a suitable definition of a contraction) and fixed-point theorems exist which guarantee that repeated application of $\mathscr{C}$ on the initial condition $x(0)=X_0+x_0$ will converge to the DA-vector representation of the solution $\varphi(x_0,t)$ of (2.1) in finitely many steps.

### 2.2. Functional inversion using DA-arithmetic

Next we wish to review the actual functional inversion which is employed to obtain the inverse $\mathscr{M}^{-1}$ of a function $\mathscr{M}$, or rather a DA-vector which stores the expansion coefficients of $\mathscr{M}^{-1}$ up to the desired order. Assume we are given a smooth map $\mathscr{M}:\mathbb{R}^v\longrightarrow\mathbb{R}^v$ s.t. $\mathscr{M}(0)=0$ and its linearization $M$ is invertible at the origin. This assures the existence of a smooth inverse $\mathscr{M}^{-1}$ in a neighborhood of the origin. If we write $\mathscr{M}=M+\mathscr{N}$, where $\mathscr{N}$ is the nonlinear part and insert this into the fundamental condition $\mathscr{M}\circ\mathscr{M}^{-1}=\mathscr{I}$, we easily obtain the relation

$$\mathscr{M}^{-1}=M^{-1}\circ(\mathscr{I}-\mathscr{N}\circ\mathscr{M}^{-1})$$

and see that the desired inverse $\mathscr{M}^{-1}$ is a fixed point of the operator $\mathscr{C}(.):=M^{-1}\circ(\mathscr{I}-\mathscr{N}\circ.)$, which proves to be a contraction using a suitable definition of a contracting operator in the DA-picture. Hence the existence of the fixed point $\mathscr{M}^{-1}$ of $\mathscr{C}$ is verified and $\mathscr{M}^{-1}$ can be obtained through repeated iteration of $\mathscr{C}$, beginning with the identity $\mathscr{I}$. Also in this case the iteration converges to $\mathscr{M}^{-1}$ in finitely many steps.

## 3. Description of the method

### 3.1. Preliminary remarks

We begin our discussion by the assumption that the ODE under consideration exhibits a periodic or almost periodic solution $\varphi(X_0,t)$ which starts on a suitable Poincaré section and returns after a period $T$, which has been determined approximately e.g. by a high-order Runge–Kutta-integration. Once such a periodic orbit $\varphi(X_0,t)$ has been identified, we proceed by performing the DA-integration of Eq. (2.1) for one cycle as described in the last section. The goal is to use the new found local dependence on the relative initial conditions $x_0$ to make statements about the qualitative properties of the periodic orbit.

As Poincaré sections, we want to be able to consider as large a class of surfaces as possible. A suitable assumption is that the Poincaré section $S\subset\mathbb{R}^v$ is given in terms of a function $\sigma:\mathbb{R}^v\longrightarrow\mathbb{R}$ as $S:=\{x\in\mathbb{R}^v:\sigma(x)=0\}$. Since the function $\sigma$ also needs to be expressed in terms of elementary functions available in the computer environment for DA arithmetic, it is necessarily smooth, and hence also the surface $S$. This should contain most types of

surfaces which might be of practical interest, in particular, the most common case where $S$ is just an affine plane of the form $S:=\{x \in \mathbb{R}^v : x_1 = c\}$ for some $c \in \mathbb{R}$; here we would have $\sigma(x) = x_1 - c$.

Another condition which needs to be met by $S$ is that the flow is transversal to it for all possible initial conditions $x_0 \in D$. Without this assumption a Poincaré map cannot be defined meaningfully, and we check this condition in the following way: at any point $s \in S$ the gradient $\nabla\sigma(s)$ is perpendicular to $S$, and the direction of the flow is $f(s)$. So we need to ensure that the scalar product $\langle \nabla\sigma(s), f(s)\rangle$ does not change its sign $\forall\, s \in S \cap \varphi(D, T)$. In fact, we will even demand the more stringent condition that $0 \notin [\langle \nabla\sigma(\varphi(x_0, T)), f(\varphi(x_0, T))\rangle]_{x_0 \in D}$, where $[g(x)]_{x \in D}$ denotes a rigorous interval enclosure of a function $g$ over $D$.

### 3.2. Outline of the method

The general goal of the method is to define a Poincaré map which is acting on a suitably chosen section $S$ using only information already available from the DA vector representation of the flow. For every possible initial condition, we wish to derive an expression of the crossing time $t_c(x_0)$ at which the trajectory originating at the said initial value traverses the section $S$, and then reinsert this time $t_c(x_0)$ back into the DA vector $\varphi(x_0, t)$ describing the flow. This yields a polynomial $\varphi(x_0, t_c(x_0))$ only depending on the initial conditions $x_0$ which projects these values almost exactly onto the Poincaré section, up to accuracy restrictions depending on the approximation order. The information about this crossing time is contained in the flow and the geometry of $S$ in an implicit way, hence we need to use suitable tools for functional inversion in the DA context as has been described above. The function $\varphi(x_0, t)$ as such cannot be invertible, since the dimensionality of its domain and range do not even agree. Instead, we will introduce an auxiliary function $\psi(x, t)$ which is substantially easier to handle and yields all relevant results. For $\psi(x, t)$ to be invertible in the first place we need $\psi$ to map into $\mathbb{R}^{v+1}$. This motivates the following:

**Definition 1** (*Auxiliary function for* $\varphi$). Let $\varphi$ be a polynomial representation for the flow under consideration, and let $S$ be a Poincaré section as described above which is traversed by the flow. We then define the *auxiliary function* $\psi : \mathbb{R}^{v+1} \supset D \times I \longrightarrow \mathbb{R}^{v+1}$ by

$$\psi_k(x, t) := x_k \quad \forall k \in \{1, \ldots, v\}$$
$$\psi_{v+1}(x, t) := \sigma(\varphi(x, t))$$

where $I \subset \mathbb{R}$ is open interval s.t. $T \in I$ and $\psi_{v+1}$ is well-defined.

Essentially, $\psi$ contains the crucial part of the flow and is "filled up" with trivial identities in order to achieve invertibility. It can be shown that the map $\psi$ is indeed invertible at the points $\{(x, T) \in \mathbb{R}^{v+1} : x \in D\}$. We can now employ DA inversion tools to manipulate $\psi$ and obtain the inverse $\psi^{-1}(x, t)$. Naturally, because of the identities in $\psi$, also $\psi^{-1}$ will preserve these identities and hence only the component $\psi^{-1}_{v+1}(x, t)$ is nontrivial. Once we have established $\psi^{-1}(x, t)$ we evaluate it at the point $y := (x_0, 0)$ to solve for the crossing time as a function of $x_0$ and set

$$t_c(x_0) := \psi^{-1}_{v+1}(x_0, 0).$$

However, $\varphi(x_0, t_c(x_0))$ still depends on all components of $x_0$, since the crossing time can be specified for the whole domain box $D$. But the Poincaré map $\mathscr{P}$ is supposed to be defined on the surface $S$, a $v-1$-dimensional smooth submanifold of $\mathbb{R}^v$, so one of the coordinates should be redundant. We can perform this restriction of $\mathscr{P}$ to $S$ in the following way:

We assume that $\forall\, x \in \varphi(D, T) \cap S$ the implicit condition $\sigma(x) = 0$ can be stated explicitly as $x_j = \tilde{\sigma}(x_1, \ldots, x_{j-1}, x_{j+1}, x_v)$ for some $j \in \{1, \ldots, v\}$. This can always be done locally, but it is not a very strict requirement that this be true globally in the set $\varphi(D, T) \cap S$. From now on, we will assume wlog that $j = 1$, i.e. $x_1 = \tilde{\sigma}(x_2, \ldots, x_v)$. Then we define the Poincaré map $\mathscr{P}$ by setting $x_{0,1} = \tilde{\sigma}(x_{0,2}, \ldots, x_{0,v})$ and $t = t_c(x_0)$ in $\varphi(x_0, t)$:

$$\mathscr{P}(x_0) := [\varphi(\tilde{\sigma}(.), ., t_c(\tilde{\sigma}(.), .))](x_{0,2}, \ldots, x_{0,v}).$$

As a special case again consider the common instance $S := \{x \in \mathbb{R}^v : x_1 = c\}$ for some $c \in \mathbb{R}$. Then we have $\sigma(x) = x_1 - c$ and thus $\tilde{\sigma}(x_2, \ldots, x_v) = c$ and $\mathscr{P}(x_0) := \varphi(c, x_{0,2}, \ldots, x_{0,v}, t_c(c, x_{0,2}, \ldots, x_{0,v}))$.

### 3.3. Summary of the algorithm

We conclude the presentation of the method by summarizing the algorithmic steps:

(1) Obtain a DA-vector representation of the solution $\varphi(x_0, t)$ for one cycle.
(2) Verify that $0 \notin [\langle \nabla\sigma(\varphi(x_0, T)), f(\varphi(x_0, T))\rangle]_{x_0 \in D}$.
(3) Set up and invert the auxiliary function $\psi$ using DA functional inversion to obtain a DA-vector representation of $\psi^{-1}$.
(4) Evaluate $t_c(x_0) := \psi^{-1}(x_0, 0)$.
(5) Evaluate $\mathscr{P}(x_0) := \varphi(x_0, t_c(x_0))$.
(6) Restrict $\mathscr{P}(x_0)$ to $S$ by replacing $x_{0,1}$ by $\tilde{\sigma}(x_{0,2}, \ldots, x_{0,v})$.

## 4. Example: A muon cooling ring

The method as described above has been implemented in the COSY Infinity environment [13], which offers support for the DA-vector data type and its operations. In fact, COSY even supports the remainder-enhanced DA-vector or Taylor Model data type, which offers rigorous error estimates on the polynomial DA-representation, and we hope to extend the method described above in such a way that also validated error bounds for the Poincaré map can be obtained.

The system we wish to analyze is a problem from accelerator physics, a simple muon cooling ring based on

continuous gas-based cooling and continuous re-acceleration. Focusing is provided by a quadrupole-based FODO system; for details refer to Refs. [7,8,14]. The ODEs governing the motion are described as

$$\dot{x}_1 = x_3$$
$$\dot{x}_2 = x_4$$
$$\dot{x}_3 = x_4 \cdot (1 + k_q r_{xy} + k_h r_{xy}^2) - a \cdot \frac{x_3}{\sqrt{x_3^2 + x_4^2}} + a \cdot \frac{x_2}{\sqrt{x_1^2 + x_2^2}}$$
$$\dot{x}_4 = x_3 \cdot (1 + k_q r_{xy} + k_h r_{xy}^2) - a \cdot \frac{x_4}{\sqrt{x_3^2 + x_4^2}} - a \cdot \frac{x_1}{\sqrt{x_1^2 + x_2^2}}$$

where $a$ describes the cooling and re-acceleration strength, $k_q$ is the quadrupole focusing strength, and $k_h$ is the sextupole strength. In the following simulations, we choose the specific values $a = 0.1$, a focusing scheme based on a 12-fold FODO structure described by $k_q = 10\sin(12t)$, and a sextupole strength of $k_h = 3$.

We consider the initial values $x(0) := X_0 + x_0$ with $X_0 := (0, 1, 1, 0)^T$ and $D := [-10^{-4}, 10^{-4}]^4$. The centerpoint $X_0$ lies on a periodic orbit of the form $\varphi(X_0, t) = (\cos(t), -\sin(t), -\sin(t), -\cos(t))$ and hence has a period of $T = 2\pi$. However, no other orbit originating in the box $X_0 + (D \backslash \{0\})$ is periodic, but instead is slowly pulled towards the invariant solution $\varphi(X_0, t)$ with an asymptotic phase, i.e. $\varphi(X_0 + x_0, t) \to_{t\to\infty} \varphi(X_0, t - \theta(x_0))$ for some phase $\theta(x_0)$.

For this example, we compute the Poincaré map for two planes $S_1$ and $S_2$ before and after one of the 12 FODO cells, so that they form an angle of $2\pi/12$. This map is suitable to be iterated by the COSY beam tracking routines to produce graphics output.

First, we show the results for the crossing time $t_c$ and the components of $\mathscr{P}$ after an 14th order computation with a choice of $a = 0.1$, $k_h = 3$ and $k_q = 10$. The polynomial coefficients are scaled to the phase-space coordinate widths of $10^{-4}$; as a result, each coefficient directly shows the maximum contribution that the corresponding term can make, which helps readability. For the crossing time $t_c$ we obtain

| I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|
| 1 | −0.2107121099493938E−01 | 1 | 1 | 0 | 0 | 0 | 0 |
| 2 | 0.9347612653241022E−02 | 1 | 0 | 1 | 0 | 0 | 0 |
| 3 | −0.1201628825073957E−01 | 1 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0.2946295382772426E−02 | 1 | 0 | 0 | 0 | 1 | 0 |
| 5 | −0.1066867771975123E−06 | 2 | 2 | 0 | 0 | 0 | 0 |
| 6 | 0.6346330104788356E−07 | 2 | 1 | 1 | 0 | 0 | 0 |
| 7 | 0.5162610528656688E−06 | 2 | 0 | 2 | 0 | 0 | 0 |
| 8 | 0.2072768191336032E−05 | 2 | 1 | 0 | 1 | 0 | 0 |
| 9 | −0.1221608485172613E−05 | 2 | 0 | 1 | 1 | 0 | 0 |
| 10 | 0.1203023359953680E−05 | 2 | 0 | 0 | 2 | 0 | 0 |
| 11 | 0.9954413063005544E−08 | 2 | 1 | 0 | 0 | 1 | 0 |
| 12 | −0.1069046574180368E−05 | 2 | 0 | 1 | 0 | 1 | 0 |
| 13 | −0.2865490547207094E−06 | 2 | 0 | 0 | 1 | 1 | 0 |
| 14 | −0.1679258946667011E−07 | 2 | 0 | 0 | 0 | 2 | 0 |
| ⋮ | | | | | | | |
| 95 | 0.1412525551591218E−16 | 5 | 1 | 0 | 2 | 2 | 0 |
| 96 | −0.8773781577679656E−17 | 5 | 0 | 1 | 2 | 2 | 0 |
| 97 | −0.8503558663031080E−17 | 5 | 2 | 0 | 0 | 3 | 0 |
| 98 | 0.7544699027997683E−17 | 5 | 1 | 1 | 0 | 3 | 0 |
| 99 | −0.9698655865255742E−17 | 5 | 1 | 0 | 1 | 3 | 0 |

Inserting this into the flow $\varphi(x_0, t)$ and restricting $\varphi(x_0, t)$ to $S$ yields that $\mathscr{P}_1(x_0)$ is given by

| I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 0.5551115123125783E−16 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | −0.2475735600348810E−19 | 4 | 0 | 0 | 0 | 4 | 0 |
| 3 | −0.3388692518831476E−19 | 5 | 0 | 3 | 2 | 0 | 0 |
| 4 | 0.3333120608373854E−19 | 5 | 0 | 2 | 3 | 0 | 0 |
| 5 | −0.3552515768008433E−19 | 5 | 0 | 2 | 2 | 1 | 0 |
| 6 | 0.3530916534253391E−19 | 5 | 0 | 1 | 3 | 1 | 0 |

which is indeed zero up to roundoff error, as expected. For the component $\mathscr{P}_2(x_0)$ we get

| I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 1.000000000000000 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0.4057920121836106E−04 | 1 | 0 | 1 | 0 | 0 | 0 |
| 3 | 0.1403405185907984E−04 | 1 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0.4821854587969232E−04 | 1 | 0 | 0 | 0 | 1 | 0 |
| 5 | −0.6880188908023518E−08 | 2 | 0 | 2 | 0 | 0 | 0 |
| 6 | 0.2218250197573597E−09 | 2 | 0 | 1 | 1 | 0 | 0 |
| 7 | 0.1569399973211117E−08 | 2 | 0 | 0 | 2 | 0 | 0 |
| 8 | 0.3959964998800090E−08 | 2 | 0 | 1 | 0 | 1 | 0 |
| 9 | −0.5900424606672487E−08 | 2 | 0 | 0 | 1 | 1 | 0 |
| 10 | 0.1484090992045488E−08 | 2 | 0 | 0 | 0 | 2 | 0 |
| ⋮ | | | | | | | |
| 36 | 0.2905642684509707E−19 | 5 | 0 | 4 | 1 | 0 | 0 |
| 37 | 0.3555459693963146E−19 | 5 | 0 | 3 | 1 | 1 | 0 |
| 38 | −0.9121533983931736E−19 | 5 | 0 | 2 | 2 | 1 | 0 |
| 39 | 0.8209709912306776E−19 | 5 | 0 | 1 | 3 | 1 | 0 |
| 40 | −0.2773032070261215E−19 | 5 | 0 | 0 | 4 | 1 | 0 |

for $\mathscr{P}_3(x_0)$:

| I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 1.000000000000000 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0.3916009498009659E−14 | 1 | 0 | 1 | 0 | 0 | 0 |
| 3 | 0.9999999999976929E−04 | 1 | 0 | 0 | 1 | 0 | 0 |
| 4 | −0.6735131658115734E−15 | 1 | 0 | 0 | 0 | 1 | 0 |
| 5 | −0.5073800545630164E−08 | 2 | 0 | 2 | 0 | 0 | 0 |
| 6 | 0.5939776450415390E−08 | 2 | 0 | 1 | 1 | 0 | 0 |
| 7 | −0.1806310552295999E−08 | 2 | 0 | 0 | 2 | 0 | 0 |
| 8 | 0.1202564475352457E−07 | 2 | 0 | 1 | 0 | 1 | 0 |
| 9 | −0.7241099329147064E−08 | 2 | 0 | 0 | 1 | 1 | 0 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 10 | −0.2320522246878191E−08 | 2 | 0 | 0 | 0 | 2 | 0 |
| ⋮ | | | | | | | |
| 47 | 0.4870521605534692E−18 | 5 | 0 | 1 | 2 | 2 | 0 |
| 48 | −0.1330388468272154E−18 | 5 | 0 | 0 | 3 | 2 | 0 |
| 49 | −0.8160539983180998E−19 | 5 | 0 | 2 | 0 | 3 | 0 |
| 50 | 0.9361435018707453E−19 | 5 | 0 | 0 | 2 | 3 | 0 |
| 51 | −0.3132073512809148E−19 | 5 | 0 | 1 | 0 | 4 | 0 |

and for $\mathscr{P}_4(x_0)$:

| I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|
| 1 | −0.5551115123125783E−16 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | −0.9657136347505140E−04 | 1 | 0 | 1 | 0 | 0 | 0 |
| 3 | 0.5966243086151465E−04 | 1 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0.1191088537008158E−03 | 1 | 0 | 0 | 0 | 1 | 0 |
| 5 | −0.1332746698395360E−07 | 2 | 0 | 2 | 0 | 0 | 0 |
| 6 | −0.1368532466582279E−07 | 2 | 0 | 1 | 1 | 0 | 0 |
| 7 | 0.2399464571937943E−08 | 2 | 0 | 0 | 2 | 0 | 0 |
| 8 | −0.2076327780857213E−09 | 2 | 0 | 1 | 0 | 1 | 0 |
| 9 | 0.3328742850177919E−08 | 2 | 0 | 0 | 1 | 1 | 0 |
| 10 | −0.1163278287959000E−08 | 2 | 0 | 0 | 0 | 2 | 0 |
| ⋮ | | | | | | | |
| 48 | 0.7596735610932742E−18 | 5 | 0 | 1 | 2 | 2 | 0 |
| 49 | −0.7681001402837021E−19 | 5 | 0 | 0 | 3 | 2 | 0 |
| 50 | 0.9040407096077550E−19 | 5 | 0 | 2 | 0 | 3 | 0 |
| 51 | −0.2515146267314533E−18 | 5 | 0 | 1 | 1 | 3 | 0 |
| 52 | 0.5404827028051936E−19 | 5 | 0 | 0 | 2 | 3 | 0 |

In the following we utilize the Poincaré maps just obtained to perform a tracking analysis of the system. Specifically, we use the rescaled maps $\mathscr{P}_2(x_0)$ and $\mathscr{P}_3(x_0)$ and perform beam tracking using the COSY Infinity TR-routine. A total of $n = 6$ particles are launched on the $x$-axis at the positions $n \cdot 4\,\mathrm{cm}$. Figs. (1–3) show the evolution of the motion over 10 turns, 20 turns, and 50 turns, respectively. The cooling action of the system is clearly visible, resulting in the apparent collapse towards the origin. Fig. 4 shows the dynamics displayed in normal form coordinates [2,15], which decouples horizontal and vertical motion and, in the case of the damping, leads to a motion that follows a perfect logarithmic spiral.

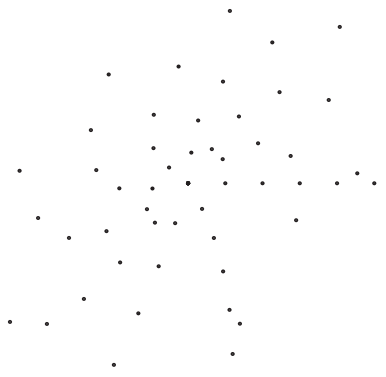


Fig. 1. Tracking of six particles for the first 10 turns in the muon cooling ring.

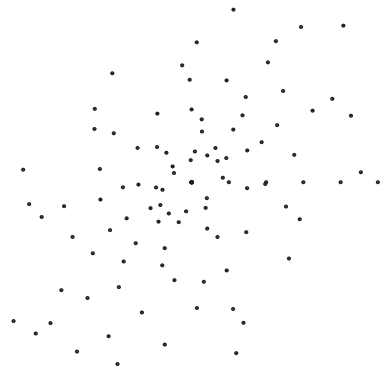


Fig. 2. Tracking of six particles for the first 20 turns in the muon cooling ring.

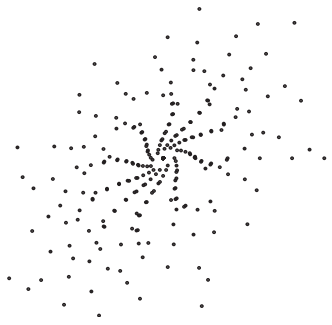


Fig. 3. Tracking of six particles for the first 50 turns in the muon cooling ring.

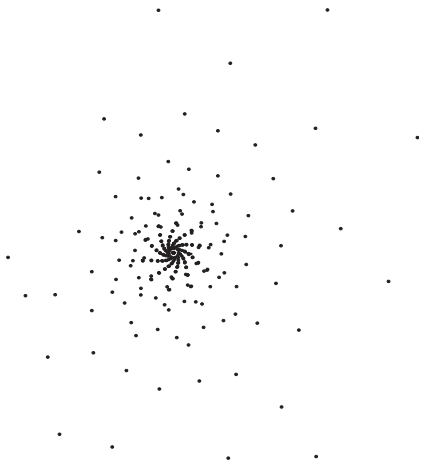


Fig. 4. Tracking of six particles for the first 50 turns in the muon cooling ring, displayed in normal form coordinates.

## References

[1] K.L. Brown, R. Belbeoch, P. Bounin, Rev. Sci. Instr. 35 (1964) 481.
[2] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999, also available at ⟨ http://bt.pa.msu.edu/pub⟩.
[3] K. Makino, M. Berz, Int. J. Appl. Math. 3 (4) (2000) 421.
[4] M. Berz, K. Makino, Int. J. Appl. Math. 3 (4) (2000) 401.

[5] H. Wollnik, Optics of Charged Particles, Academic Press, Orlando, FL, 1987.
[6] $\mu^+\mu^-$ Collider Collaboration, $\mu^+\mu^-$ collider: a feasibility study, Technical Report BNL-52503, Fermilab-Conf-96/092, LBNL-38946, 1996.
[7] S. Ozaki, et al., for the Muon Collaboration, Feasibility study-II of a muon-based neutrino source, Technical Report 52623, Muon Collider Collaboration, BNL, 2001.
[8] M.M. Alsharo'a, et al., Phys. Rev. ST-AB 6 (2003) 081001.
[9] N. Anantaraman, B. Sherrill (Eds.), Proceedings of the International Conference on Heavy Ion Research with Magnetic Spectrographs, Technical Report MSUCL-685, National Superconducting Cyclotron Laboratory, 1989.
[10] J. Nolen, A.F. Zeller, B. Sherrill, J.C. DeKamp, J. Yurkon, A proposal for construction of the S800 spectrograph, Technical Report MSUCL-694, National Superconducting Cyclotron Laboratory, 1989.
[11] M. Berz, K. Joh, J.A. Nolen, B.M. Sherrill, A.F. Zeller, Phys. Rev. C 47 (2) (1993) 537.
[12] M. Berz, K. Makino, Reliab. Comput. 4 (4) (1998) 361.
[13] M. Berz, J. Hoefkens, K. Makino, COSY INFINITY Version 8.1—programming manual, Technical Report MSUHEP-20703, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2001, see also ⟨http://cosy.pa.msu.edu⟩.
[14] K. Makino, M. Berz, C.J. Johnstone, D. Errede, Nucl. Instr. and Meth. A 519 (2004) 162.
[15] M. Berz, Differential algebraic formulation of normal form theory, in: M. Berz, S. Martin, K. Ziegler (Eds.), Proceedings of the Nonlinear Effects in Accelerators, London, 1992, p. 77, IOP Publishing, Bristol.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 112–118
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Modeling of self-consistent distributions for longitudinally non-uniform beams ☆

O.I. Drivotin*, D.A. Ovsyannikov

*Applied Mathematics and Control Processes Department, St. Petersburg State University, St. Petersburg 198504, Russia*

Available online 1 December 2005

**Abstract**

Self-consistent distributions for beams in longitudinal magnetic and electric fields are considered. The magnetic and electric fields are assumed to be uniform in the beam cross-section, but can slowly vary along the beam axis. As a consequence, the beam cross-section radius and the longitudinal velocities of particles can also vary in the longitudinal direction. New classes of self-consistent distributions are found as analytical solutions of the Vlasov equation. Some known distributions are particular cases of these distributions. The analytically described distributions were also numerically simulated.


*PACS:* 41.75.–i; 52.27.Jt; 52.65.Ff

*Keywords:* Charged particle beam; Vlasov equation; Self-consistent distributions; Phase density

## 1. Introduction

Consider an axially symmetric beam propagating through longitudinal magnetic and electric fields. Assume that all particles of any beam cross-section have the same longitudinal velocity, which can vary along the beam axis. It is usually assumed that the distribution density function in the four-dimensional phase space of transverse motion satisfies the non-linear Vlasov equation.

For longitudinally uniform beams, some stationary solutions of the Vlasov equation have been found. The most well-known distribution is the Kapchinsky–Vladimirsky (KV) distribution [1], where the particles have the same energy for transverse motion. The dimension of the support of the distribution density function in four-dimensional phase space for transverse motion is equal to 3. The simplest and best-known case is Brillouin flow [2], in which all particles rotate around the beam axis with the same angular velocity. In this case, the dimension of the support of the distribution density function is equal to 2. Another distribution known is the water bag distribution [1,3,4]. Its phase density is constant in that region of the phase space for transverse motion where the energy of the transverse motion of a particle is less than some given value. A wide class of distributions can be obtained using the density inversion theorem [5]. Distributions were also found [4,6] for which the phase density depends on a linear combination of the energy of the transverse motion and the angular momentum of a particle. Wide classes of self-consistent distributions have also been reported in the literature [7,8].

For longitudinally non-uniform beam, distributions have been investigated for the case of a periodic focusing magnetic field and new distributions that are similar to the KV distribution and rigid rotor distributions have been found [4,9–12].

This paper offers new stationary self-consistent distributions for an axially symmetric beam propagating through a longitudinal magnetic and electric field. Our approach is based on analysis of the distribution density in the space of integrals of motion. These distributions were found analytically and were also simulated numerically.

☆This work was supported by the Russian Foundation for Basic Research Grant # 03-01-0726.
*Corresponding author.
*E-mail address:* drivotin@yandex.ru (O.I. Drivotin).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.082

## 2. Integrals of motion

Assume that longitudinal components of the external electric field $E_z$ and of the external magnetic field $B_z$ do not depend on $r$ and $\varphi$ (where $r, \varphi, z$ are cylindrical coordinates). Then $E_r = -(r/2)\,\partial E_z/\partial z$, and the four-dimensional vector potential of the external magnetic field can be taken in the form $A_0 = A_r = A_z = 0$, $A_\varphi = B_z(z)r^2/2$.

Consider the case when longitudinal motion of the particles is relativistic, and the transverse motion is non-relativistic. Then the equation for particle dynamics, which is also the equation for the characteristic lines of the Vlasov equation, is

$$\frac{\mathrm{d}p}{\mathrm{d}s} = \frac{\partial L}{\partial x}, \quad L = -mc\sqrt{g_{ik}u^i u^k} + eA_i u^i \tag{1}$$

where $p$ is the particle canonical momentum, $s$ is proper time, $L$ is the Lagrangian of the particle, $u = \mathrm{d}x/\mathrm{d}s$ is the four-dimensional velocity of the particle, and $g_{ik}$ are components of the metric tensor in Minkowski space.

The equation for the azimuthal motion is $p'_\varphi = 0$ (where the stroke denotes differentiation on $s$). Introducing $M = p_\varphi/mc$, the first integral of this equation is of the form

$$M = r^2(\varphi' + \omega_0) \tag{2}$$

where $\omega_0 = eB_z/2m$.

Transforming Eq. (1) for longitudinal components taking into account $u^0 = \gamma$, $u^\varphi = \varphi'$, $\partial A_0/\partial z = E_z/c$, $\partial A_\varphi/\partial z = (r^2/2)\partial B_z/\partial z$, and neglecting the terms arising from the longitudinal components of the self-field, we have

$$\beta' = \frac{2w}{\gamma^2} + \frac{r^2\varphi'}{\gamma^3}\frac{\mathrm{d}\omega_0}{\mathrm{d}z} \tag{3}$$

where $\beta, \gamma$ are the reduced longitudinal velocity and reduced energy, $w = eE_z/2mc^2$. Assume that the magnetic field varies sufficiently slowly and that the second term in Eq. (3) can also be neglected. This means that the longitudinal motion does not depend on the transverse motion. Then the above assumption that all particles of any beam cross-section have the same longitudinal velocity holds, if all particles have the same energy in the initial cross-section. In this case, the four-dimensional vector potential of the self-field of the beam is $A_{\mathrm{self}} = (-U(r,z)/c, 0, 0, \beta U(r,z)/c)$. For distributions which are considered later it can be shown that the condition of slow variation of the magnetic field takes the form $\mathrm{d}\omega_0/\mathrm{d}z \ll \omega_0/R$, where $R = R(z)$ is the radius of a beam cross-section.

The equation for radial motion can be reduced to the form

$$r'' = -\omega_0^2 r + \frac{M^2}{r^3} - \frac{e}{m\gamma c^2}\frac{\partial U}{\partial r} - \gamma r\frac{\partial w}{\partial z}. \tag{4}$$

We search for distributions in which the particle density in the configuration space $\rho(r,z)$ is a step-function of the radial coordinate $r$

$$\rho(r,z) = \begin{cases} \rho_0(z), & r \leqslant R(z) \\ 0, & r > R(z) \end{cases} \tag{5}$$

where $\rho_0(z) = J/(\pi R(z)^2 e\beta c) = \lambda/R^2$ and $\lambda = J/(\pi e\beta c)$, $J$ is the beam current.

Assume that the radius of the beam cross-section $R$ significantly varies only at distances that are much greater than $R$. Then the Poisson equation and the boundary conditions for the potential of the self-electric field of the beam $U(r,z)$ can be written as

$$\frac{1}{r}\frac{\partial}{\partial r}\,r\,\frac{\partial U}{\partial r} = \frac{e\rho_0(z)}{\varepsilon_0}$$
$$r < R(z), \quad U(a) = 0, \quad \mathrm{d}U/\mathrm{d}r|_{r=0} = 0 \tag{6}$$

where $a$ is the channel aperture.

Substituting the solution to the problem (Eq. (6) into Eq. (4)) we obtain an equation for the radial motion of particles in the form

$$r'' = -\omega^2 r + M^2/r^3 \tag{7}$$

where $\omega^2 = \omega_0^2 - \lambda/R(z)^2 + \gamma\,\mathrm{d}w/\mathrm{d}z$, $\lambda = J/J_0\beta\gamma$, $J_0 = 2\pi\varepsilon_0 mc^3/e$.

For a longitudinally uniform beam, the variable $H$ given by the expression $H = r'^2 + \omega^2 r^2 + M^2/r^2$ is the integral of motion. However, in our case $H$ is not an integral. Let us formulate the conditions under which another integral of Eq. (7) exists. Assume that the beam envelope $R(z)$ can be defined as an envelope only for particles with $M = 0$. For such particles the equation for radial motion is linear

$$X' = AX, \quad X = \begin{pmatrix} r \\ r' \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}. \tag{8}$$

If in the initial cross-section at $s = s_0$ the particles fill the ellipse

$$X_0^* B_0 X_0 \leqslant 1, \quad B_0 = \begin{pmatrix} a_0^{-2} & 0 \\ 0 & c_0^{-2} \end{pmatrix}$$

then at $s \geqslant s_0$ they will fill the ellipses

$$X^* B X \leqslant 1$$

where $B = F^{*-1} B_0 F^{-1}$, and $F$ is the matrizant of the system (Eq. (8)). It can be readily shown that $R^2$ is the first diagonal element of the matrix $B^{-1}$, and $R$ satisfies the equation

$$R'' = -\omega^2 R + \frac{a_0^2 c_0^2}{R^3}. \tag{9}$$

The system of Eqs. (7) and (9) can be reduced to the known Ermakov system [13,14] if the variable $\omega$, which depends on $s$ and $R$, is regarded as a function of $s$.

Using the known expression for the integral of the Ermakov system [14], we obtain the value

$$I = (Rr' - rR')^2 + \frac{M^2R^2}{r^2} + \frac{a_0^2c_0^2r^2}{R^2} = \left(\frac{dq}{d\tau}\right)^2 + \frac{M^2}{q^2} + a_0^2c_0^2q^2 \quad (10)$$

which is the integral of motion, where $q = r/R$, $d\tau = ds/R^2$. Another integral of motion $M$ can be written as

$$M = q^2\left(\frac{d\varphi}{d\tau} + R^2\omega_0\right). \quad (11)$$

Note that when $M = 0, \omega = \omega(s)$, the integral (Eq. (10)) is identical to the well-known Courant–Snyder invariant [15].

Let us find a set $\tilde{\Omega}_1$ in the space of variables $M, I$ such that the condition $q \leqslant 1, \forall s \geqslant s_0$ holds for all particles. It follows from Eq. (10) that

$$I \leqslant M^2 + a_0^2c_0^2. \quad (12)$$

Furthermore, $I \geqslant \min_q((M^2/q^2) + a_0^2c_0^2q^2) = 2a_0c_0|M|$. Excluding particles corresponding to the lower boundary of the set $\tilde{\Omega}_1$, we obtain

$$I > 2|M|a_0c_0. \quad (13)$$

The set $\tilde{\Omega}_1$ defined by the conditions (Eqs. (12) and (13)) is shown in Fig. 1.

We also consider a set $\tilde{\Omega}(q)$ of $M$ and $I$ such that the particles possessing these $M$ and $I$ pass through a point corresponding to the given $q$. It follows from the equality (Eq. (10)) that if the inequality

$$I \geqslant \frac{M^2}{q^2} + a_0^2c_0^2q^2 \quad (14)$$

is satisfied, then the particle can move at a point corresponding to this $q$. Furthermore, the value of the integral $I$ is bound by the inequality (Eq. (12)). Therefore, the set $\tilde{\Omega}(q)$ is defined by Eqs. (12) and (14).

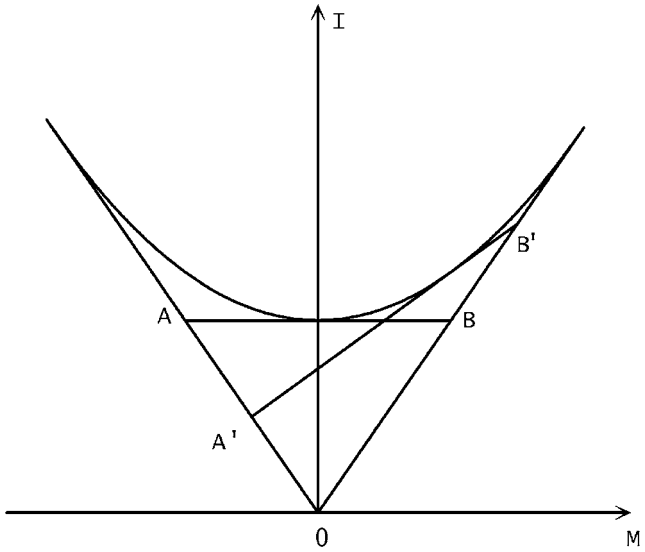



Fig. 1. The set $\tilde{\Omega}_1$.

## 3. Particle distributions in the space of integrals of motion

Consider the phase distribution of the particles of some infinitely thin layer bounded by two parallel planes moving along the $z$-axis with velocity $\beta(z)c$. The stationarity of the distribution can be ensured by the independence of the time of the distribution in the initial cross-section ($z = z_0$). Since the velocity $\beta c$ can vary along the beam axis, the thickness of the layer can vary correspondingly. Therefore, we normalize all densities by dividing them by the thickness of the layer $dz/\beta c$.

Assume that the particles are uniformly distributed on the angles $\varphi$ and on the phases of their trajectories. The first assumption implies that the beam has axial symmetry. Then the phase density in four-dimensional phase space of the transverse motion $n$ depends only on two phase variables. The integrals of motion $M$, $I$ can be taken as these two phase variables. Thus, let $n = n(M(r, \varphi'), I(r, r', \varphi'))$ where $n(M, I)$ denotes some function of $M$ and $I$.

Let us introduce the density distribution function in the space of integrals $M, I$: $f(M, I)$. The domain of this function is the set $\tilde{\Omega}_1$.

The simplest distribution is given by $f(M, I) = f_B\delta(M, I)$. For this distribution, all particles have the same values of $M, I$: $M = 0$, $I = 0$. It follows from Eq. (10) that $q' = 0$, $c_0 = 0$, and the envelope equation can be written as

$$R'' = -\omega_0^2R + \frac{\lambda}{R} - \gamma\frac{dw}{dz}R. \quad (15)$$

The density in configuration space is $\varrho_0(z) = J/\pi R(z)^2 e\beta(z)c$. Normalizing as described above, we obtain $f_B = J/e\beta(z_0)c$. If $\omega_0 = \text{const}$, $w = 0$, and $R = \text{const}$, then $R^2 = 2mJ/\pi\varepsilon_0eB_z^2\beta\gamma c$, and the density of such a distribution in configuration space is equal to the density of the Brillouin flow. Thus, the Brillouin flow is a particular case of such a distribution. Therefore, it can be called the generalized Brillouin flow. As distinct from the Brillouin flow, all particles can move in the radial direction, with their normalized radial coordinate $q$ constant for each particle.

It is easy to show for a uniform beam cross-section distribution (Eq. (5)) that

$$f(M, I) = \int_{q_{\min}(M,I)}^{q_{\max}(M,I)} \frac{2\pi}{|q'|}n(M, I)\,dq = \pi^2n(M, I)/a_0c_0. \quad (16)$$

Expressing the particle density in the configuration space $\rho(r)$ in terms of $f(M, I)$, we obtain

$$\rho(r) = \frac{1}{rR}\int_{\tilde{\Omega}(q)} \frac{n(M, I)\,dM\,dI}{|q'|} = \frac{a_0c_0}{\pi^2rR}\int_{\tilde{\Omega}(q)} \frac{f(M, I)\,dM\,dI}{(I - M^2/q^2 - a_0^2c_0^2q^2)^{1/2}}. \quad (17)$$

In the expression (Eq. (17)) we exclude the particles for which $q' \equiv 0$ in accordance with (Eq. (13)). To allow for these particles requires an additional term in expression (17).

Let us take $f(M,I)$ as a simple layer with density $f_0$ with respect to the variable $I$ on the segment belonging to the set $\tilde{\Omega}_1$ and the tangent to the upper bound of this set

$$f(M,I) = f_0 \delta_{I=I_0(k)+kM}, \quad f_0 > 0,$$
$$I_0(k) = a_0^2 c_0^2 - k^2/4, \ (M,I) \in \tilde{\Omega}_1. \tag{18}$$

Substituting Eq. (18) into Eq. (17) we have

$$\rho(r) = \frac{a_0 c_0 f_0}{\pi^2 r R} \int_{\tilde{\Omega}(q)} \frac{\delta_{I=I_0(k)+kM}\, \mathrm{d}M\, \mathrm{d}I}{(I - M^2/q^2 - a_0^2 c_0^2 q^2)^{1/2}} = \frac{a_0 c_0 f_0}{\pi}.$$

Hence, the density of the distribution (Eq. (18)) in the configuration space is constant throughout the beam cross-section, and therefore it is the solution to the problem. Normalizing as described above, we have $\rho_0 = J/\pi R^2$, so that $f_0 = J/a_0 c_0$.

It can easily be shown that the support of the distribution density function (Eq. (18)) is a segment of the straight line that is tangential to the upper boundary of the set $\tilde{\Omega}_1$. The segment is bounded by the lines $I = \pm 2a_0 c_0 M$ (segment A′B′ in Fig. 1). When $k = 0$, this segment is parallel to the axis $M$ (segment AB in Fig. 1). If $R, \lambda, \omega_0$ are constant, the distribution (Eq. (18)) at $k = 0$ is identical to the KV distribution, and the distributions at $k \neq 0$ are identical to the rigid rotor distributions described in Ref. [6]. Therefore, the distribution (Eq. (18)) at $k = 0$ can be called the generalized KV distribution.

Furthermore, each linear combination of the distributions (Eq. (18)) will also be uniform in the beam cross-section

$$f(M,I) = \sum_{k\in K} \alpha_k \delta_{I=I_0(k)+kM}, \quad \rho = \frac{a_0 c_0}{\pi R^2} \sum_{k\in K} \alpha_k$$

where $K$ is a finite set of real numbers, $K \subset (-2a_0 c_0, 2a_0 c_0)$, or

$$f(M,I) = \int_{-2a_0 c_0}^{2a_0 c_0} \alpha(k) \delta_{I=I_0(k)+kM}\, \mathrm{d}k$$
$$\rho = \frac{a_0 c_0}{\pi R^2} \int_{-2a_0 c_0}^{2a_0 c_0} \alpha(k)\, \mathrm{d}k.$$

Here, $\alpha_k$ in the first case and $\alpha(k)$ in the second case are functions defined on the discrete set $K$ or on the interval $(-2a_0 c_0, 2a_0 c_0)$ such that there are, respectively,

$$\sum_{k\in K} \alpha_k \quad \text{or} \quad \int_{-2a_0 c_0}^{2a_0 c_0} \alpha(k)\, \mathrm{d}k.$$

## 4. Integral equation for uniform distributions in the beam cross-section

Another way to search for uniform self-consistent distributions is to consider the equality Eq. (17) as an integral equation for the density of the distribution. Introducing new integration variables, we obtain the integral equation

$$\frac{a_0^2 c_0^2}{2\pi} \int_0^{2\pi} \int_0^1 \frac{F(y\cos(\alpha-\theta), y\cos(\alpha+\theta))}{(1-y^2)^{1/2}}\, y\, \mathrm{d}y\, \mathrm{d}\alpha = J \tag{19}$$

where

$$F(k_1,k_2) = \begin{cases} f(M,I)(M^2 - I + a_0^2 c_0^2)^{1/2}, & k_1 \geqslant k_2 \\ F(k_2,k_1), & k_1 < k_2 \end{cases} \tag{20}$$

$k_{1,2} = 2(M \pm (M^2 - I + a_0^2 c_0^2)^{1/2})$ and $\theta = \arccos q$. This is exactly the integral equation for the function of two arguments $F(k_1,k_2)$. Both arguments depend on $q$. The problem is to find a function $F(k_1,k_2)$ such that the result of the integration does not depend on $q$.

Any non-negative solution of Eq. (19) satisfying the condition $F(k_1,k_2) = F(k_2,k_1)$ corresponds to some self-consistent particle distribution.

The simplest case is $F(k_1,k_2) = g_0 > 0$, where $g_0$ is constant. In this case, the density of the distribution in the space of the integrals of motion is

$$f(I,M) = g_0 (M^2 - I + a_0^2 c_0^2)^{-1/2}. \tag{21}$$

This distribution does not reduce to any distribution obtained before.

Another simple case is $F(k_1,k_2) = g_1(k_1) + g_2(k_2)$. Substituting this expression into Eq. (19), it is apparent that this is exactly the solution. In view of the condition (Eq. (20)), we have $g_1(k) = g_2(k) + g_0$. The term with $g_0$ corresponds to the previous case (Eq. (21)). Omitting this, we obtain

$$f(M,I) = \frac{g(k_1) + g(k_2)}{(M^2 - I + a_0^2 c_0^2)^{1/2}}, \quad g(k) \geqslant 0.$$

Other solutions can be sought in the form of a series

$$F(x,z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_{mn} x^m z^n \tag{22}$$

or a polynomial. For example, searching for the solution to Eq. (19) as a third-degree polynomial, we obtain

$$f(M,I) = \frac{-c\,(I - a_0^2 c_0^2)(10M^2 - 5I + 2a_0^2 c_0^2) + g_0}{(M^2 - I + a_0^2 c_0^2)^{1/2}}. \tag{23}$$

The constant values $c$ and $g_0$ are such that $f(M,I) \geqslant 0$ for all $(M,I) \in \tilde{\Omega}_1$. The values of $F(k_1,k_2)$, $J$, and $\lambda$ can be arbitrarily great. In this case, the radius of the beam cross-section $R(z)$ behaves according to the envelope equation (15) with large values of $\lambda$.

## 5. Results of the computer simulation

The distributions described in this paper were numerically simulated using the macroparticle method. Only the radial motion of particles was investigated. Accordingly, each macroparticle represents an infinitely thin ring with the varying radius $r$ centered on the beam axis and lying in

the plane orthogonal to the beam axis. The dynamics of this ring is described by the equation

$$r'' = -\omega_0^2 r + \frac{M^2}{r^3} + \frac{\lambda N_r}{rN}$$

where $N_r$ is the number of macroparticles (rings) with a radius less than the radius of the macroparticle under consideration, and $N$ is the total number of macroparticles. The dependence of $B_z$ and $R$ on $z$ is presented in Fig. 2. The values of $B_z$ are given in tesla, $z$ in centimeters, and $R$ is divided by $a_0$, which in this case is equal to 2 cm (the initial radius of the beam cross-section). It is assumed that in the initial cross-section $R' = 0$. The magnetic field linearly increases from 0.3 T at $z = 0$ to 1.5 T at $z = 20$ cm and remains constant at greater values of $z$; $E_z = 0$. The particle energy is $W = 60$ kEv and remains constant along the beam. The beam current is $J = 82$ mA. The initial divergence is 60 mrad, which corresponds to $c_0 = 2.0 \times 10^5$.

The results of the simulation of the distribution (Eq. (21)) are presented in Figs. 3–6 for approximately 32,000 macroparticles. The left-hand side of the figures shows the distribution of macroparticles in the phase space of the radial motion (in reduced units $r/a_0$ and $v/c_0$). The dependence of the index of uniformity of the distribution, $k = N_r/N/(r^2/a_0^2)$, on $r$ is shown in the middle. This index should be constant for an ideal uniform distribution. Variations of $k$ when $r$ increases are explained by the discretization. The right-hand side of the figures shows the distribution of macroparticles in the space of the integrals of motion $M$ and $I$ (in reduced units $I_0 = a_0^2 c_0^2$, $M_0 = a_0 c_0$). The particles fill a curvilinear triangle defined by the conditions (Eqs. (12) and (13)) in the space of the integrals $M$ and $I$, which remains unchanged.

## 6. Conclusion

The use of the integrals of motion to analyze charged particle dynamics is widely known. In this case, the phase density of a self-consistent distribution is specified as a function of the integrals of motion. Within the framework of the approach presented in this paper (see also Refs. [16,17]), the particle density in the space of the integrals of motion $M$ and $I$ is introduced. This approach allows the specification of naturally various self-consistent distributions, both those known before and new ones.

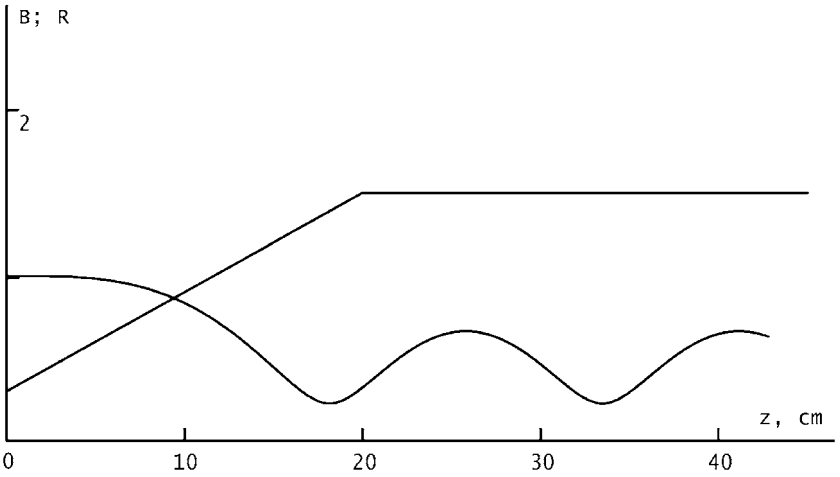



Fig. 2. The beam envelope $R(z)$ and the longitudinal component of the magnetic flux density $B_z(z)$ (T) versus the longitudinal coordinate $z$ (cm).

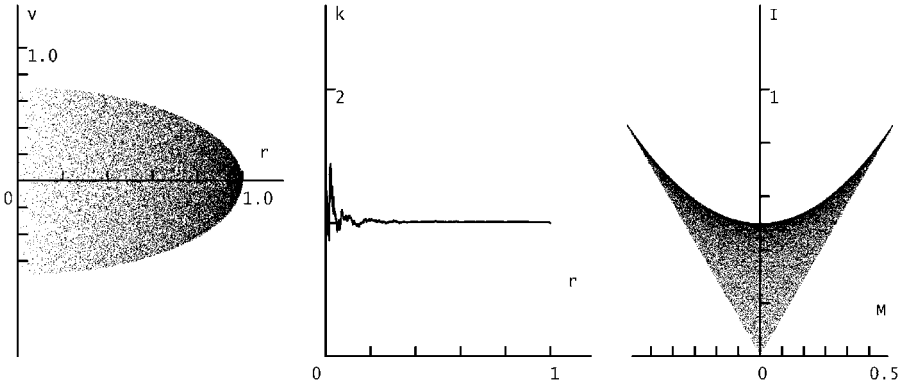



Fig. 3. The distribution of macroparticles in the phase space, the index of uniformity, and the distribution of macroparticles in the space of the integrals of motion $M$ and $I$ for the density function (20) at $z = 0$.

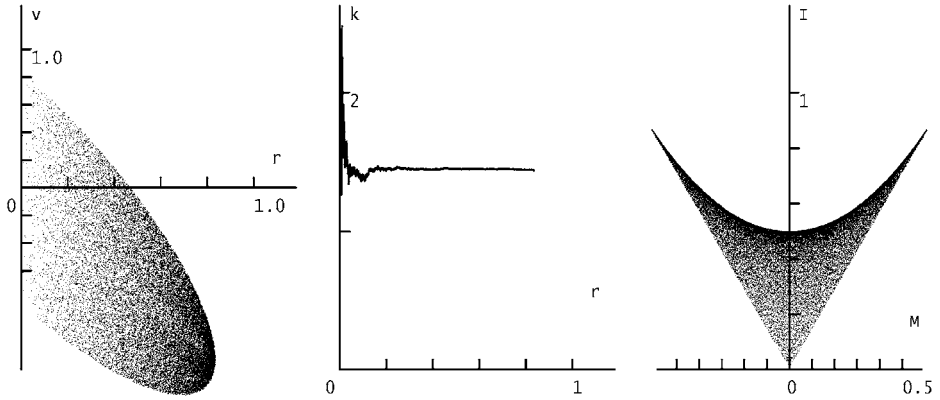



Fig. 4. The distribution of macroparticles in the phase space, the index of uniformity, and the distribution of macroparticles in the space of the integrals of motion $M$ and $I$ for the density function (20) at $z = 10\,\text{cm}$.

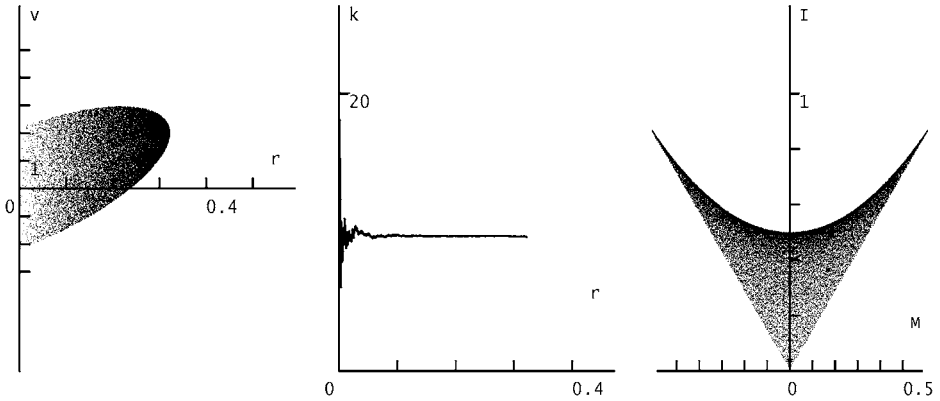



Fig. 5. The distribution of macroparticles in the phase space, the index of uniformity, and the distribution of macroparticles in the space of the integrals of motion $M$ and $I$ for the density function (20) at $z = 20\,\text{cm}$.

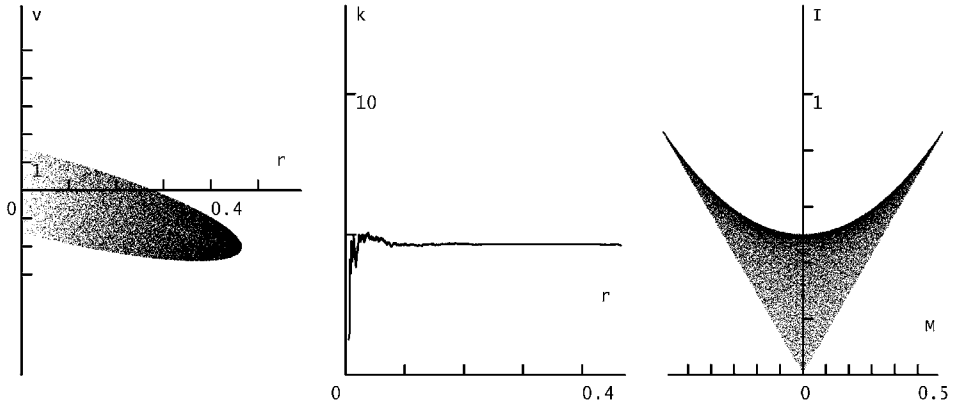



Fig. 6. The distribution of macroparticles in the phase space, the index of uniformity, and the distribution of macroparticles in the space of the integrals of motion $M$ and $I$ for the density function (20) at $z = 30\,\text{cm}$.

The proposed approach was applied to a beam non-uniform along its axis, for which wide classes of self-consistent distributions were found. In the particular instance of longitudinal uniformity, these are identical to the distributions for a uniform beam. Therefore, some of them can be regarded as generalizations of distributions, such as Brillouin flow and KV distribution. The important feature of these distributions is that the beam radius, the magnetic field, and the particle longitudinal velocity can vary along the beam axis, although this variation should be sufficiently slow. This means that they are more realistic than many of the previously considered distributions, and can be used for the solution of various problems of modeling and optimization of accelerating and focusing structures with high-density beams.

## References

[1] I.M. Kapchinsky, Particles Dynamics in Resonant Linear Accelerators, Atomizdat, Moscow, 1966.
[2] L. Brillouin, Phys. Rev. 67 (1945) 260.
[3] I. Hoffmann, in: Applied Charged Particle Optics, Part C, Academic Press, New York, 1983, p. 49.
[4] R.C. Davidson, C. Chen, Particle Accelerators 59 (1998) 175.
[5] R.C. Davidson, Phys. Plasmas 5 (1998) 3459.
[6] R.C. Davidson, Physics of Nonneutral Plasmas, Addison-Wesley, Reading, 1990.
[7] O.I. Drivotin, D.A. Ovsyannikov, Zh. Vychisl. Mat. Mat. Fiz. (USSR) 27 (1987) 416.
[8] O.I. Drivotin, D.A. Ovsyannikov, Zh. Vychisl. Mat. Mat. Fiz. (USSR) 29 (1989) 1245.
[9] R.L. Gluckstern, Phys. Rev. Lett. 73 (1994) 1247.
[10] R.L. Gluckstern, W.-H. Cheng, H. Ye, Phys. Rev. Lett. 75 (1995) 2835.
[11] C. Chen, R. Pakter, R.C. Davidson, Phys. Rev. Lett. 79 (1997) 225.
[12] R.L. Gluckstern, A.V. Fedotov, S. Kurennoy, R. Ryne, Phys. Rev. E 58 (1998) 4977.
[13] V.P. Ermakov, Univ. Izv. (Kiev) 20 (1880) 1.
[14] J.R. Ray, Adv. Nonlinear Waves 1 (1984) 230.
[15] E.D. Courant, H.S. Snyder, Ann. Phys. 3 (1958) 1.
[16] O.I. Drivotin, D.A. Ovsyannikov, Dokl. Russ. Acad. Nauk 334 (1994) 284.
[17] O.I. Drivotin, D.A. Ovsyannikov, in: Proceedings of the 1997 Particle Accelerators Conference, Vancouver, 1997, p. 1943.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 119–121
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# New vortices in axisymmetric beams in an inhomogeneous magnetic field

Yu.Ya. Golub

*Moscow Radiotechnical Institute, 132 Warshavskoye shosse, 113519 Moscow, Russia*

Available online 22 November 2005

**Abstract**

We analyzed localized vortices in non-neutral electron beams with inhomogeneous density and velocity propagating in vacuum along an inhomogeneous external magnetic field. These vortices are different to vortices in earlier work because of the inhomogeneous external magnetic field. New types of vortices were also obtained using a new solution for non-linear equations, different to the standard Larichev–Reznik or Reznik method. A new axisymmetric expression was found for the electric field potential of a vortex in a wave frame. The new vortices represent new solitons in an inhomogeneous external magnetic field.


*PACS:* 52.35.−g; 05.45.−a; 05.45.Yv

*Keywords:* Instability; Vortices; Solutions

## 1. Basic equations

We analyzed localized vortices in non-neutral electron beams with inhomogeneous density and velocity propagating in vacuum along an inhomogeneous external magnetic field. These vortices are different to those used previously [1–3] because of the inhomogeneous external magnetic field.

We investigated a non-relativistic electron beam propagating in vacuum along an external inhomogeneous magnetic field $B$ in the $z$-direction of a cylindrical coordinate system $(r, \theta, z)$:

$$\vec{B} = (B_0 + B_z(x))*\vec{e}_z + B_x(z)*\vec{e}_x$$

where $B$ satisfies the Maxwell equation:

$$\frac{\partial B_z(x)}{\partial x} = \frac{\partial B_x(z)}{\partial z}.$$

When the system is in equilibrium and homogeneous in $\theta$ and $z$, it is characterized by radial distributions of the electron density $n_0(r)$, velocity $v_0[0, v_{0\theta}(r), v_{0z}(r)]$ and electron field potential $\varphi_0(r)$. We assume $\omega_c^2 \gg \omega_p^2$, where $\omega_p$ is the plasma electron frequency and $\omega_c$ is the electron cyclotron frequency.

We investigated the non-steady state of the system characterized by the deviation of $n$, $v$ and $\varphi$ from equilibrium values of $n_0$, $v_0$ and $\varphi_0$. For a solution of the motion and continuity equations for the particles and the Poisson equation for the electric field potential, we chose a travelling wave for which all the parameters are functions of the variables $r$ and $\eta = \theta + k_z z - \omega t$, with constant wave number $k_z$ and frequency $\omega$. If we neglect the inertial drift of electrons due to a large value of $\omega_c$, we obtain the following equation [4]:

$$\left\{\Delta_\perp \varphi - \Lambda\varphi + S\varphi^2, \varphi - \frac{\omega_d B_0}{2c} r^2\right\}_{r,\eta} = 0 \tag{1}$$

where

$$\{f, g\}_{r,\eta} = \frac{1}{r}\left(\frac{\partial f}{\partial r}\frac{\partial g}{\partial \eta} - \frac{\partial f}{\partial \eta}\frac{\partial g}{\partial r}\right)$$

*E-mail address:* yurigolub@mtu-net.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.022

$$\Lambda = -\frac{k_z(k_z+k_v)\omega_p^2}{\omega_d^2} - \frac{(k_n+k_b)\omega_p^2}{v_0\omega_d}$$

$$S = \frac{k_z}{2}\left(\frac{(k_z+k_v)e}{m\omega_d^2}\right)^2$$

$$k_v = \frac{1}{\omega_c r}\frac{dv_{0z}}{dr}, \quad k_b = \frac{\partial B_z}{\partial x}\cdot\frac{v_0}{B_0\omega_c n_0}$$

$$k_n = \frac{v_0}{\omega_c r}\frac{dn_0}{dr}, \quad v_0 = v_{0z}(0)$$

$$\omega_d = \omega - k_z v_{0z} - \frac{v_{0\theta}}{r}$$

$m$ and $e$ are the electron mass and charge, $c$ is the speed of light, and $\Delta_\perp$ is the transverse part of the Laplace operator.

The new term in $\Lambda$ is $k_b$. The influence of the magnetic field inhomogeneity on the vortex is similar to the electron density inhomogeneity.

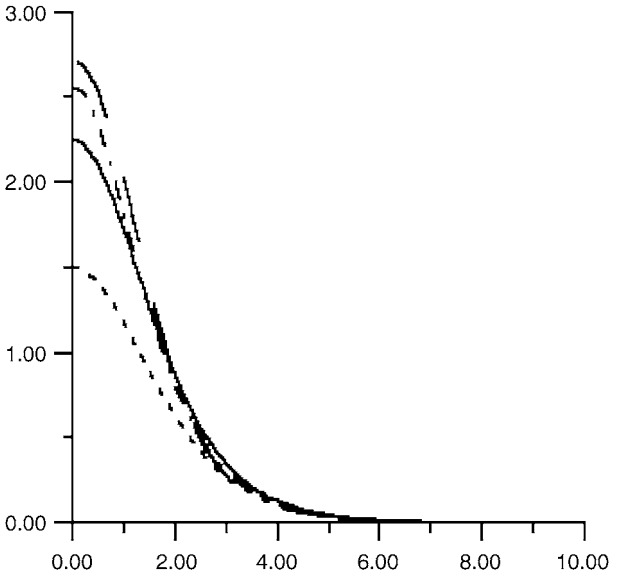



Fig. 1. Dependence of $\varphi^{(0)}$ (dotted line), $\varphi^{(1)}$ (solid line), $\varphi^{(2)}$ (dashed–dotted line), and $\varphi^{(3)}$ (dashed line) on the radius $r$.

## 2. Localized vortices

Larichev and Reznik [5,6] solved Eq. (1) only when they neglected the term $S\varphi^2$, leading to the so-called Larichev–Reznik solution. However, we did not neglect this non-linear term, but obtained a non-linear equation:

$$\frac{\partial^2\varphi}{\partial r^2} + \frac{1}{r}\frac{\partial\varphi}{\partial r} - \Lambda\varphi + S\varphi^2 = 0. \tag{2}$$

The non-linear Eq. (2) is distinguished from the Korteweg–de Vries (KdV) and Bessel equations. We obtain the approximate solution to Eq. (2) using an original method involving functional iteration. The next $(n+1)$th iteration is obtained from the equation:

$$\varphi^{(n+1)} = \varphi^{(n)} + \sin(\tau^{(n)}(0)) * \frac{1}{\Lambda}(\tau^{(m)}(r)) \tag{3}$$

where $\tau^{(n)}$ is the residual of $\varphi^{(n)}$ in Eq. (2):

$$\tau^{(n)} = \left(\frac{\partial^2\varphi^{(n)}}{\partial r^2} - \Lambda\varphi^{(n)} + \frac{1}{r}\frac{\partial\varphi^{(n)}}{\partial r} + S(\varphi^{(n)})^2\right)$$

and $\varphi^{(0)}$ is the solution for the KdV equation:

$$\varphi^{(0)} = \frac{3}{2}\frac{\Lambda}{S}\frac{1}{\left(\mathrm{ch}\left(\frac{\sqrt{\Lambda}}{2}r\right)\right)^2}.$$

The equation for the first iteration is

$$\varphi^{(1)} = \varphi^{(0)} - \frac{1}{\Lambda}\left(\frac{\partial^2\varphi^{(0)}}{\partial r^2} - \Lambda\varphi^{(0)} + \frac{1}{r}\frac{\partial\varphi^{(0)}}{\partial r} + S(\varphi^{(0)})^2\right)$$

with

$$\varphi^{(1)} = \frac{3\sqrt{\Lambda}\left(\mathrm{sech}\left(\frac{\sqrt{\Lambda}r}{2}\right)\right)^2\left(\sqrt{\Lambda}r + \tanh\left(\frac{\sqrt{\Lambda}r}{2}\right)\right)}{2Sr}$$

This iteration $\varphi^{(1)}$ is the approximate solution to Eq. (2). We can obtain $\varphi^{(2)}$, then $\varphi^{(3)}$, and so on. The iterations $\varphi^{(2)}$ and $\varphi^{(3)}$ are also approximate solutions to Eq. (2).

The equation for the second iteration $\varphi^{(2)}$ is

$$\varphi^{(2)} = \frac{1}{\Lambda}\left(\frac{\partial^2\varphi^{(1)}}{\partial r^2} + \frac{1}{r}\frac{\partial\varphi^{(1)}}{\partial r} + S(\varphi^{(1)})^2\right).$$

The dependence of $\varphi^{(0)}$, $\varphi^{(1)}$, $\varphi^{(2)}$ and $\varphi^{(3)}$ on the radius $r$ is shown in Fig. 1 for $\Lambda = 1\,\mathrm{cm}^{-2}$ and $S = 1\,\mathrm{cm}^{5/2}\,\mathrm{g}^{-1/2}\,\mathrm{s}$. We see that the maximum amplitude $\varphi^{(n)}$ approaches a constant value with increasing $n$.

We see that $\varphi^{(1)}$ and $\varphi^{(2)}$ are close to $\varphi^{(3)}$. Thus, the functional iteration method for the approximate solution shows convergence.

Thus, we obtain an approximate solution that exponentially decreases with radius $r$. This approximate solution is a continuous function in the first differential, in contrast to the Larichev–Reznik solution. That approximate solution is near the KdV solution at large $r$. A new expression has been found for the electric field potential of a vortex in a wave frame. The expression is axisymmetric in the wave frame. New vortices are the result of external disturbances or the appearance and development of instabilities, such as a diocotron instability in a hollow beam or a slipping instability in a solid beam. The influence of the magnetic field inhomogeneity on the vortex is similar to the electron density inhomogeneity.

## References

[1] Yu.Ya. Golub, M.G. Nikulin, N.E. Rozanov, in: V.G. Bar'yakhtar, et al. (Eds.), Nonlinear World: IVth International Workshop on

Nonlinear and Turbulent Processes in Physics, vol. 2 World Scientific Publishing, Singapore, 1990, p. 857.
[2] Yu.Ya. Golub, in: Proceedings of EPAC 2002, Paris, France, p. 1253.
[3] Yu.Ya. Golub, in: Proceedings of PAC 2003, Portland, OR, USA, p. 3103.
[4] G.D. Aburdzhaniya, F.F. Kamenetz, V.P. Lakhin, A.B. Mikhailovskii, O.G. Onishchenko, Phys. Lett. A 105 (1984) 48.
[5] V.D. Larichev, G.M. Reznik, Dokl. Akad. Nauk SSSR 231 (1976) 1077.
[6] G.M. Reznik, Dokl. Akad. Nauk SSSR 282 (1985) 981.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 122–126
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Beam dynamics simulation during final bunching and transport for heavy ion inertial fusion

T. Kikuchi[a,*], T. Someya[b], S. Kawata[b], M. Nakajima[c], K. Horioka[c], T. Katayama[d]

[a]Department of Electrical and Electronic Engineering, Utsunomiya University, Yohtoh 7-1-2, Utsunomiya 321-8585, Japan
[b]Department of Energy and Environmental Science, Utsunomiya University, Yohtoh 7-1-2, Utsunomiya 321-8585, Japan
[c]Department of Energy Sciences, Tokyo Institute of Technology, Nagatsuta 4259, Kanagawa 226-8502, Japan
[d]Center for Nuclear Study, University of Tokyo, Hirosawa 2-1, Saitama 351-0198, Japan

Available online 29 November 2005

**Abstract**

In a particle accelerator system for heavy ion inertial fusion, space-charge-dominated beam dynamics is investigated during a final beam bunching. While the space charge effect undertakes an important role for the beam dynamics in the regime, the beam instability induced by space charge oscillation may occur, depending on the particle distribution of the beam. Particle simulations present that the instability due to the strong space charge effect may not be serious for realistic particle distributions during the final beam bunching.


*PACS:* 52.58.Hm; 52.59.Sa; 52.65.Rr; 29.27.Bd

*Keywords:* Final beam bunching; Space-charge-dominated beams; Beam dynamics; Particle simulation; Heavy ion inertial fusion

## 1. Introduction

In a heavy ion inertial fusion (HIF), research key issues include a development of high flux ion source, an accelerator design, final focusing to a target, implosion dynamics of the fuel pellet, a reactor system design, an energy conversion system, and so on. Especially, the generation and transport of intense heavy-ion beam (HIB) are important in the HIF study. The required beam parameter values are ~GeV particle energy, ~100 kA total current and ~10 ns short pulse for an effective pellet implosion [1,2]. A particle accelerator can produce the intense HIBs, but these parameters and accelerator system have not been yet determined. Because the intense-heavy-ion-beam parameters are far from those of conventional particle accelerators, and there are no sufficient data for a system design. For this reason, the study of space-charge-dominated beam in the HIF accelerator system is performed in this paper.

In space-charge-dominated beams, theoretical analyses and numerical simulations have been performed to investigate the stability of the beam transport [3–5]. Although for the distribution of $\delta$ function in the four-dimensional phase space the Kapchinskij–Vladimirskij (KV) distributed beam [6] is useful for the analytical treatment, the analysis tends to overestimate the instability effect in the comparison with more realistic beams. For this reason, particle simulations with various distributions are also important for the stability evaluation.

In the final stage of HIF driver, the beam pulse must be longitudinally compressed from ~100 to ~10 ns. Induction voltage modulators, which have a precise waveform controllability and repetition capability, are proposed for this purpose [7]. A beam buncher by the induction modulator applies a bunching voltage so as to make a considerable velocity tilt between the head and the tail of the beam bunch. A final focus and beam illumination are crucial, and a large emittance interferes the focusing to the small fuel pellet [8]. The final beam bunching is an important technology in the heavy ion driver system. For the effective pellet implosion, we should transport and compress the bunch of HIB with a low emittance growth.

*Corresponding author.
*E-mail address:* tkikuchi@cc.utsunomiya-u.ac.jp (T. Kikuchi).

0168-9002/$ - see front matter 
doi:10.1016/j.nima.2005.11.057

We investigate the beam dynamics under a strong space charge effect during the longitudinal bunch compression in the final beam bunching. The growth of the beam instability depends on the initial particle distributions. It is summarized that the beam instability induced by the space charge effect may not contribute to the emittance growth for the beam transport with the realistic particle distributions.

## 2. Simulation model and beam parameters

The high-energy particle beam is transported by using a magnetic quadrupole focusing channel as a unit of focus-drift–defocus-drift (FODO) lattice [9]. Generally, the beam transport by the FODO lattice causes a non-axisymmetric behavior in the beam cross-section. We are interested in such beam dynamics with longitudinal bunch compression. For the above reasons, the fully three-dimensional numerical scheme is essentially required by the beam dynamics simulation. From the viewpoint of the computational cost, the full calculations are difficult. While the longitudinal bunch length is of the order of meter, the scale of the transverse cross-section is only a few centimeter in the stage of final beam bunching. Consequently, the small-scale phenomena by the space charge structure will be dominated by the transverse beam dynamics. We deal with the particle dynamics in the transverse cross-section of the beam by multi-particle simulations, and the effect of longitudinal compression is introduced as the beam current increases. The current lump-up history is not linear during the beam transport, and the schedule model can affect the transverse beam dynamics [10]. The linear current increase model causes the most serious influence in the particle dilution in the phase space. From the viewpoint of the beam physics study, the history of linear current increase is assumed as model of the longitudinal bunch compression.

We use the particle-in-cell (PIC) method for descriptions of the transverse behavior with the longitudinal compression, the effect of which was evaluated by assuming linear increase in the beam current. The charge and mass of the super particles are re-weighted with the beam transport [11]. The quadrupole occupancy is fixed at 0.5, and the one lattice period is 3 m. The beam parameters are assumed as Table 1 [2]. The initial generalized perveance is assumed to $3.58 \times 10^{-6}$. The initial undepressed and depressed phase advances are $\sigma_0 = 72^\circ$ and $\sigma = 65.2^\circ$. The transverse calculation region is fixed at the square of 10 cm × 10 cm, and the outer boundary condition is given as a conductor wall. The initially rms matched KV, waterbag (WB), Gaussian (GA), Parabolic (PA) [12], and semi-Gaussian (SG) [11] beams are chosen as the initial particle (non-stationary) distribution.

The ratio $\Delta/\lambda_D$ of the grid zone length $\Delta$ to the Debye length $\lambda_D$ is evaluated by rms emittance $\varepsilon_f$ after the final beam bunching. The unnormalized rms emittance $\varepsilon_i$ at initial condition is assumed as 10 mm mrad. The beam is a space-charge-dominated one, and the behavior likes a quasi-neutral plasma, because of the Debye shielding is important issue in this region. The grid size of $0.23\lambda_D$ is used to reproduce the Debye shielding effect and the calculation cost [13]. For the validation of the numerical convergence, we also tried to check the calculation results using the different numbers of grids and super particles. The mesh number is varied from $64 \times 64$ to $1024 \times 1024$, and the super particle number is changed from $1 \times 10^5$ to $3 \times 10^6$. Figs. 1 and 2 show the ratio of the final emittance

Table 1
Beam parameters for final beam bunching in HIF

| | |
|---|---|
| Ion species | $Pb^{1+}$ |
| Number of ions | $6.25 \times 10^{14}$ |
| Particle energy (GeV) | 10 |
| Initial beam current (A) | 400 |
| Final beam current (kA) | 10 |
| Initial pulse duration (ns) | 250 |
| Final pulse duration (ns) | 10 |

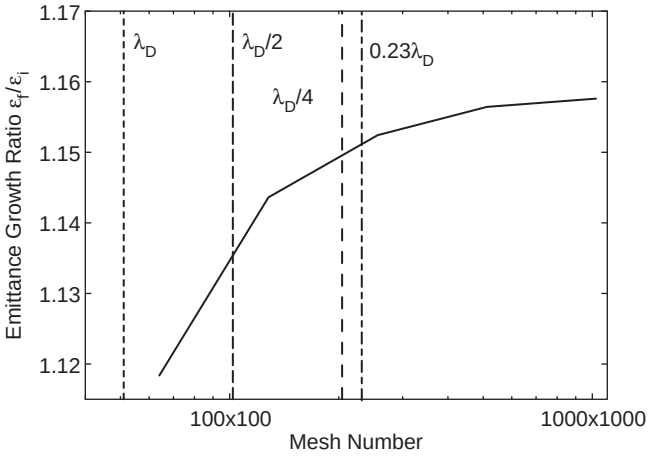



Fig. 1. Ratio of final rms emittance to initial one as a function of mesh number used by particle simulation in the case of initial WB beam.

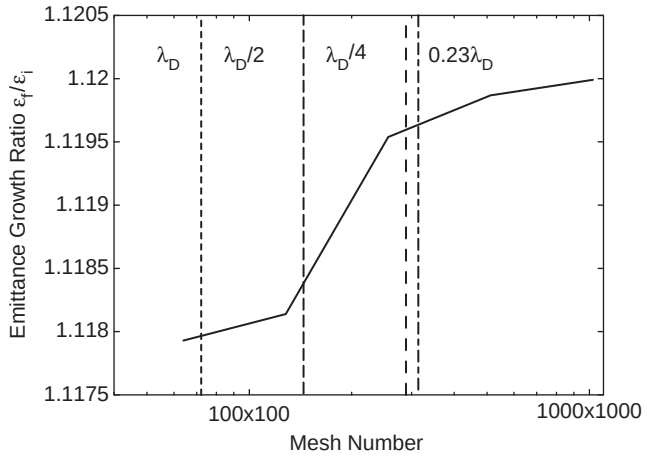



Fig. 2. Ratio of final rms emittance to initial one as a function of mesh number used by particle simulation in the case of initial GA beam.

to initial one as a function of the calculation mesh number at the initial WB and GA beams. As mentioned earlier, we tried to check the many test calculations to change the numbers of particles and cells used. These figures show typical results in such test calculations. In the space-charge-dominated beams, when the transverse particle distribution is assumed to be WB, the Debye length is approximated by [14]

$$\lambda_D \sim \frac{1}{2}\frac{v}{v_0} r_b \tag{1}$$

where $r_b$ is the beam radius and $v/v_0$ $(= \sigma/\sigma_0)$ is the tune depression, respectively. Also in the case of GA distribution, the Debye length can be assumed as [14]

$$\lambda_D \sim \frac{1}{\sqrt{8}}\frac{v}{v_0} r_b. \tag{2}$$

The density of the beam at the center is used for the evaluation of the tune depression. The final tune depression and beam radius are estimated as 0.16 and 2.45 cm. As shown in Fig. 1, the emittance growth is saturated if the mesh number is increased over $256 \times 256$. For this reason, it is found that we should consider the Debye shielding effect for the investigation of beam dynamics during the final beam bunching. However, we could not find significant changes from Fig. 2. As a result, we expected that the cell size does not contribute the emittance growth in initial GA beam. Consequently, we use the mesh number of $512 \times 512$ for the accurate calculation in this study.

## 3. Charge distribution and beam instability induced by space charge effect

Fig. 3 shows the charge density maps in physical space at each lattice period for the initial KV distributed beam. In the real space maps, the horizontal and vertical scales correspond to the full calculation region. The charge density is normalized by the maximum value at each map. The beam radius is extended with the beam current increase due to the longitudinal bunch compression. As shown in Fig. 3, the localized charge distribution in real space is indicated during the final beam bunching. The localized density profile can cause the non-linear space charge effect. As a result, the particle distribution is changed from the initial KV beam during the longitudinal compression. For this reason, the disappearance of the localized density distribution is observed in the near end of the KV beam simulation.

Fig. 4 shows the charge density maps in real space for the initial WB beam. In the WB beam, qualitatively similar behaviors with the KV beam are observed as shown in Fig. 4. From the previous study, it was reported that the localized charge distribution is caused by the instability due to the strong space charge effect [15].

As shown in Figs. 5, 6, and 7, the localized charge density cannot be observed in the initial GA, SG, and PA beams, respectively. The beam particles in the KV and WB beams are distributed into the narrow region in the four-dimensional phase space, e.g., the distribution of the KV beam is the $\delta$ function in the phase space. On the other hand, the GA, SG, and PA beams diffuse into the large area in the phase space. As a result, the KV and WB beams may be sensitive for the beam instability induced by space

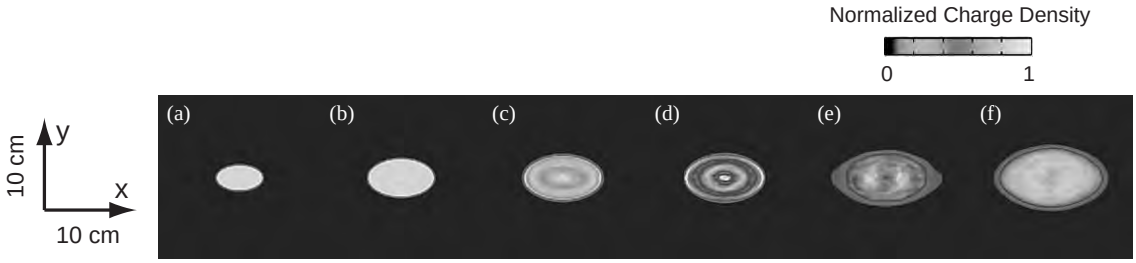



Fig. 3. Normalized charge distribution with initial KV beam, (a) at initial condition, (b) at 50 lattice periods, (c) at 82 lattice periods, (d) at 88 lattice periods, (e) at 108 lattice periods, and (f) at 150 lattice periods, respectively.

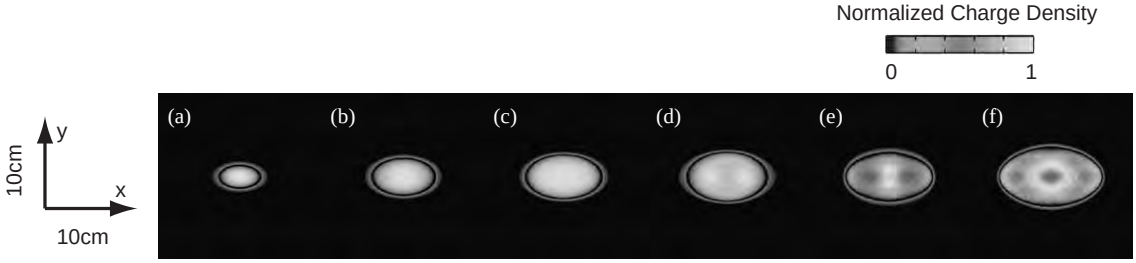



Fig. 4. Normalized charge distribution with initial WB beam, (a) at initial condition, (b) at 50 lattice periods, (c) at 77 lattice periods, (d) at 97 lattice periods, (e) at 106 lattice periods, and (f) at 150 lattice periods, respectively.

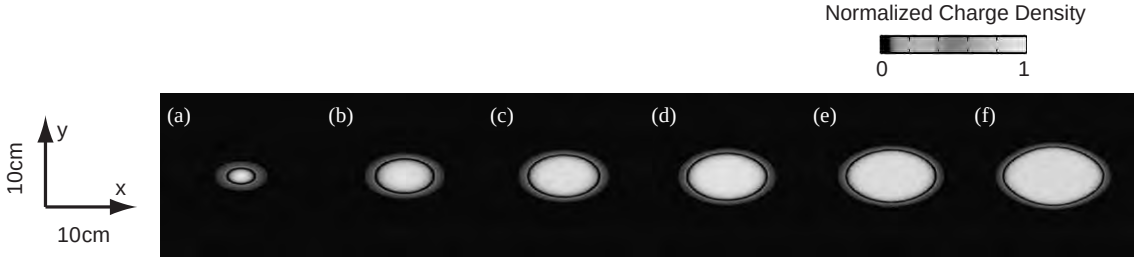



Fig. 5. Normalized charge distribution with initial GA beam, (a) at initial condition, (b) at 50 lattice periods, (c) at 77 lattice periods, (d) at 97 lattice periods, (e) at 120 lattice periods, and (f) at 150 lattice periods, respectively.

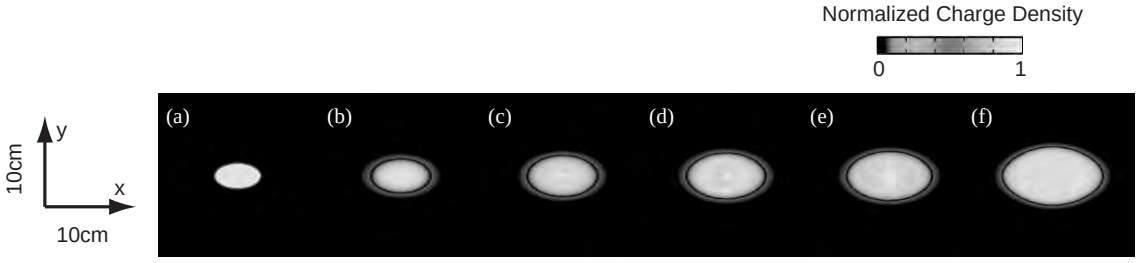



Fig. 6. Normalized charge distribution with initial SG beam, (a) at initial condition, (b) at 50 lattice periods, (c) at 77 lattice periods, (d) at 97 lattice periods, (e) at 120 lattice periods, and (f) at 150 lattice periods, respectively.

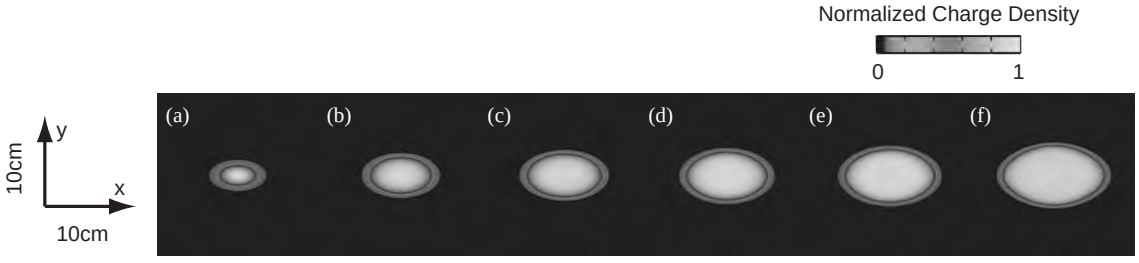



Fig. 7. Normalized charge distribution with initial PA beam, (a) at initial condition, (b) at 50 lattice periods, (c) at 77 lattice periods, (d) at 97 lattice periods, (e) at 120 lattice periods, and (f) at 150 lattice periods, respectively.

charge effect in comparison with the GA, SG, and PA distributions.

## 4. Conclusions

The transverse beam dynamics was investigated during the final beam bunching in HIF driver system. The PIC simulation with the beam current increase, as a model of the longitudinal bunch compression, was carried out for the study of beam transport under the strong effect of space charge oscillations. The initially KV and WB distributed beams caused the localized particle distribution due to the space charge effect during the final beam bunching. The GA, SG, and PA beams were transported without the notable effect by the space charge oscillation. It is expected from the PIC results that the realistic beam may cause the small effect of instability on HIB induced by the space charge oscillation during the final beam bunching in HIF.

## Acknowledgements

This work was partly supported by JSPS (Japan Society for the Promotion of Science) and MEXT. We would like to thank Drs. S.M. Lund, J.J. Barnard, E.A. Startsev, H. Qin, and M. Takanaka for their fruitful discussions and advises. We also thank to Ms. Y. Nakajima for the preparation of this manuscript.

## References

[1] J.J. Barnard, et al., Nucl. Instr. and Meth. A 415 (1998) 218.
[2] J.J. Barnard, et al., Phys. Fluids B 5 (1993) 2698.
[3] I. Hofmann, et al., Part. Acceler. 13 (1983) 145.
[4] S.M. Lund, R.C. Davidson, Phys. Plasmas 5 (1998) 3028.
[5] E.A. Startsev, R.C. Davidson, Phys. Rev. ST Acceler. Beams 6 (2003) 044401.
[6] I.M. Kapchinskij, V.V. Vladimirskij, Proceedings of the International Conference on High Energy Accelerators and Instrumentation (CERN, Geneva, 1959), p. 274.
[7] K. Horioka, et al., Laser Part. Beams 20 (2002) 609.

[8] T. Someya, A.I. Ogoyski, S. Kawata, T. Sasaki, Phys. Rev. ST Acceler. Beams 7 (2004) 044701.
[9] M. Reiser, Theory and Design of Charged Particle Beams, Wiley, New York, 1994, p. 153.
[10] T. Kikuchi, M. Nakajima, K. Horioka, J. Plasma Fusion Res. 79 (2003) 105.
[11] S.M. Lund, O. Boine-Frankenheim, G. Franchetti, I. Hofmann, P. Spiller, Proceedings of the 1999 Particle Accelerator Conference, New York, March 1999, p. 1785.
[12] Y.K. Batygin, Proceedings of the Computational Accelerator Physics Conference, Los Alamos, 1993, AIP Conference Proceedings No. 297, 1994, p. 419.
[13] R.W. Hockney, J.W. Eastwood, Computer Simulation Using Particles, McGraw-Hill, New York, 1981, p. 336.
[14] I. Hofmann, J. Struckmeier, Part. Acceler. 21 (1987) 69.
[15] T. Kikuchi, M. Nakajima, K. Horioka, T. Katayama, Phys. Rev. ST Acceler. Beams 7 (2004) 034201.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 127–130
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# An optimizer for high-current beamlines

S.V. Miginsky*

*Budker INP, Novosibirsk, Russia*

Available online 28 November 2005

**Abstract**

A novel approach to the problem of beamline development for high-current electron beam is put forward. A transient electron beam is considered as a set of independent stationary beams with different currents, energies, and initial conditions. In such a way, the overall acceptance of a beamline can be maximized effectively for some conditions and the loss can be predicted well. The simulation code implemented in this approach is described. Some examples of existing and designed beamlines are presented.


*PACS:* 41.75.−i; 41.85.Ja; 02.60.Cb; 02.60.Pn

*Keywords:* High-current beam; Beam transportation; Beamline optimization

## 1. The goal

An optimizer is intended for designing or an adjustment of a beamline carrying a high-current, transient electron beam. Additional conditions, such as $\eta$, $\eta'$ [1], rms size and divergence at the exit, can be set. Although a number of numerical simulators already exist, none fits the problem set. The problem is: (i) it seems to be impossible to optimize anything manually if the number of controlled parameters exceeds, say, five; and (ii) the parameters of the beam at the entrance are usually known only very inaccurately. Thus, an optimizer must minimize the loss of beam in a beamline, be very time-efficient and take into account main effects, but must not be very precise.

## 2. The method

First we define an appropriate model for the problem. It should take into account the main effects in a high-current beam and be computation conserving. It must not be too accurate, as the initial conditions are not known well.

### 2.1. Basic parametric models

An uncharged linear model [1] considers sin- and cos-like trajectories of particles in a beamline

$$\begin{Bmatrix} C & S & D \\ C' & S' & D' \\ 0 & 0 & 1 \end{Bmatrix}$$

$$\begin{matrix} C'' \\ S'' \end{matrix} = -\frac{e}{p} G \begin{matrix} C \\ S \end{matrix}, \quad D'' = \frac{1}{\rho} - \frac{e}{p} GD \tag{1}$$

where the second derivative is taken by the longitudinal coordinate $z$. Twiss parameters of the beam are converted with the matrix obtained, There is no way to take into account the space charge force in this case.

The KV approach [2] considers a charged beam with a special distribution in phase space. Thus, the type of distribution is preserved, and ordinary differential equations (ODEs) can be derived for its parameters (for rms values here):

$$x'' = \frac{\varepsilon_x^2}{x^3} + \frac{I}{I_0(\beta\gamma)^3} \frac{1}{x+y} - \frac{e}{p} G_x x$$

$$y'' = \frac{\varepsilon_y^2}{y^3} + \frac{I}{I_0(\beta\gamma)^3} \frac{1}{x+y} - \frac{e}{p} G_y y \tag{2}$$

*Tel.: +7 3833 394 443; fax: +7 3833 307 163.
*E-mail address:* s.v.miginsky@inp.nsk.su.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.031

where $x$ and $y$ are the horizontal and vertical coordinates, respectively. Emittance is preserved and included as a "force-like" term in each right-hand side. The ratio of the second term to the first gives a criterion for a high-current beam. Only if it is $\ll 1$ can the space charge effect be neglected. In reality, another criterion is typically much stronger (analogous to the betatron phase):

$$\int_L \sqrt{\frac{I}{I_0(\beta\gamma)^3}\frac{1}{(x+y)(x \text{ or } y)}}\,\mathrm{d}z \ll 1 \tag{3}$$

where $L$ is the total length of the beamline. If only the second criterion is not valid, the space charge effect can be considered as a small correction.

### 2.2. Proposed model

A transient charged beam is considered as a set of independent stationary charged beams (partial beams) with different currents and initial conditions, including energy. Each one has zero emittance and a uniform distribution of charge within some ellipse in the $xy$-plane. In $xx'$ and $yy'$ phase spaces, it looks like a line segment, with the middle at the point of origin. The rms size of each one is given by the KV Eq. (2) without the emittance term.

Passing through a limited aperture, each partial beam can lose a proportion of particles, depending on the oversize, i.e., the ratio of the current diameter to the aperture. The final oversize at the end is the maximum of local oversize values. The total loss is approximated as

$$\text{Loss} = \sum_i \text{Weight}_i(1 - 1/\text{Oversize}_i) \tag{4}$$

where $\text{Weight}_i$ and $\text{Oversize}_i$ are the weight coefficient (proportional to the carrying current) and the final oversize of a partial beam respectively, independently for $x$ and $y$.

The target function to be minimized is taken as

$$\sum_i \cosh\sqrt{2}\frac{\xi_i - \xi_0^i}{\Delta\xi_i} \tag{5}$$

where $\xi_i$ is the parameter obtained, $\xi_0^i$ is the desirable parameter, and $\Delta\xi_i$ is the admitted accuracy. Thus, the strongest effect in the transient beam is considered, while others are neglected. The latter are microscopic emittance and the non-uniform distribution of particles in each cross-section.

### 2.3. The code

The optimizer code consists of three parts: a beamline editor, a line geometry builder, and the optimizer itself. The beamline editor inserts a beamline element by element and edits it. It supports quadrupoles, solenoids, and bending magnets, including distributed ones. Step-like and linearly changed apertures are available. The geometry builder shows the trajectory in the inserted beamline and all the elements and apertures along it.

The optimizer receives the inserted beamline and optimizes the chosen elements to minimize the target function. This involves setting the optimized elements, the initial parameters, and the target parameters. The latter can include emittance and current at the entrance, $\eta$, $\eta'$, rms size and divergence at the exit, and always includes loss. The initial parameters are current, energy, energy spread, $\eta$, $\eta'$, rms size and divergence at the entrance.

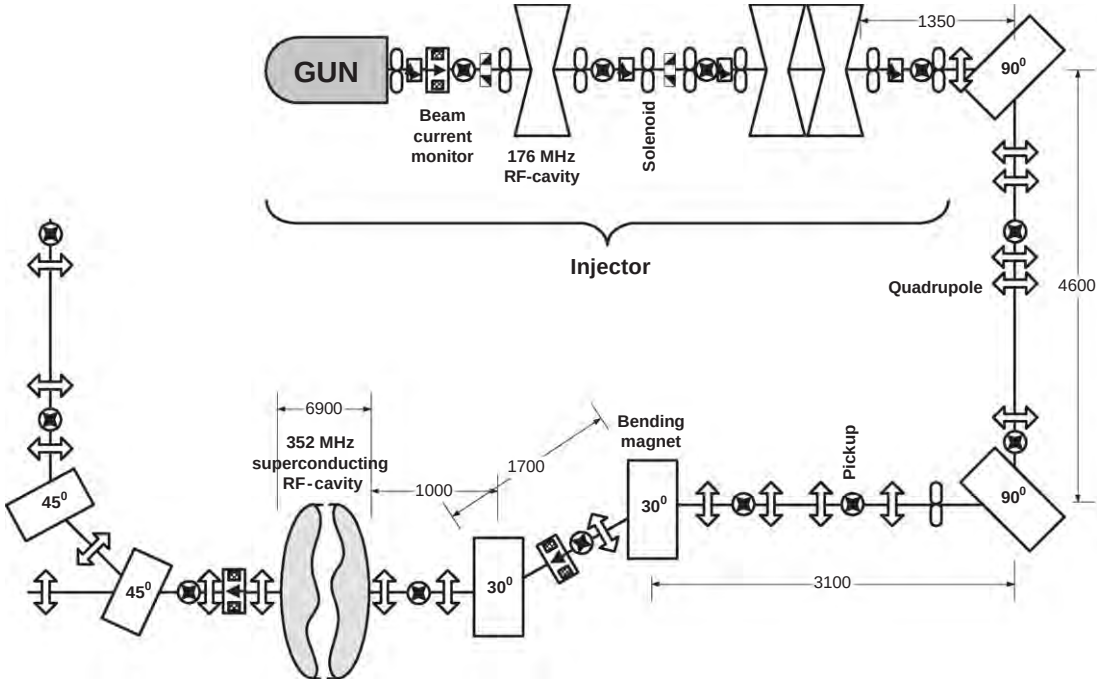



Fig. 1. Injection beamline at KAERI.

## 3. Optimized beamlines

### 3.1. Injection beamline at KAERI

The KAERI beamline [3] Fig. 1 is used to transport an electron beam from an injector to a superconducting RF cavity. Its regular aperture is 100 mm. Its basic parameters are listed in Table 1. The operation mode is continuous.

Originally the beamline was designed using the uncharged linear model. During commissioning, it was found that none of the regimes thus determined could be used due to great loss in the beamline. This is not surprising, as for a typical radius of 3 mm, both terms in Eq. (2) are $\sim 3 \times 10^{-3}\,\mathrm{m}^{-1}$. Simulations according to KV approach allowed the manual determination of new regimes that could be used as an initial approximation. After lengthy manual adjustment, acceptable loss was obtained. Nevertheless, significant discrepancy between the regimes installed and those after manual adjustment always occurred. Then the beamline was optimized with the code described, the regime found was installed, and further adjustment included only the first three lenses and the last two, i.e., only those lying outside the bends. Thus, they did not affect the chromatism of the beamline. Total loss of ~3% was easily obtained.

Table 1
Beamline at KAERI

| | |
|---|---|
| Kinetic energy of electrons, MeV | 1.5 |
| Peak current, A | 20 |
| Average current, mA | up to 10 |
| Initial emittance, $\pi$ mm mrad | 10 |
| Energy spread, relative | $3 \times 10{-}3$ |
| Acceptance at 4% loss, $\pi$ mm···mrad | 15.5 |

### 3.2. Beamline for beam splitting into three targets

The beamline Fig. 2 is used to split a short train of electron bunches (~200 ns) into three shorter ones and transport them to three separate bremsstrahlung converters. Thus, three X-ray sources are obtained that flash almost simultaneously. This is a prototype of a future diagnostic device. Its operation mode is one pulse for a long time. The BINP FEL injector [4] is used as a beam source; this is very similar to the one at KAERI.

The beamline was designed using the described code. The regular aperture was chosen as 40 mm. Additional conditions were: (i) achromatism of each branch; and (ii) small spots on the converters. It was found that this beamline has ~1.2-fold poorer acceptance than the one at KAERI, or 1.5-fold greater loss for the same beam. At the same time, it proved that the acceptance cannot be significantly improved and that the beamline cannot be

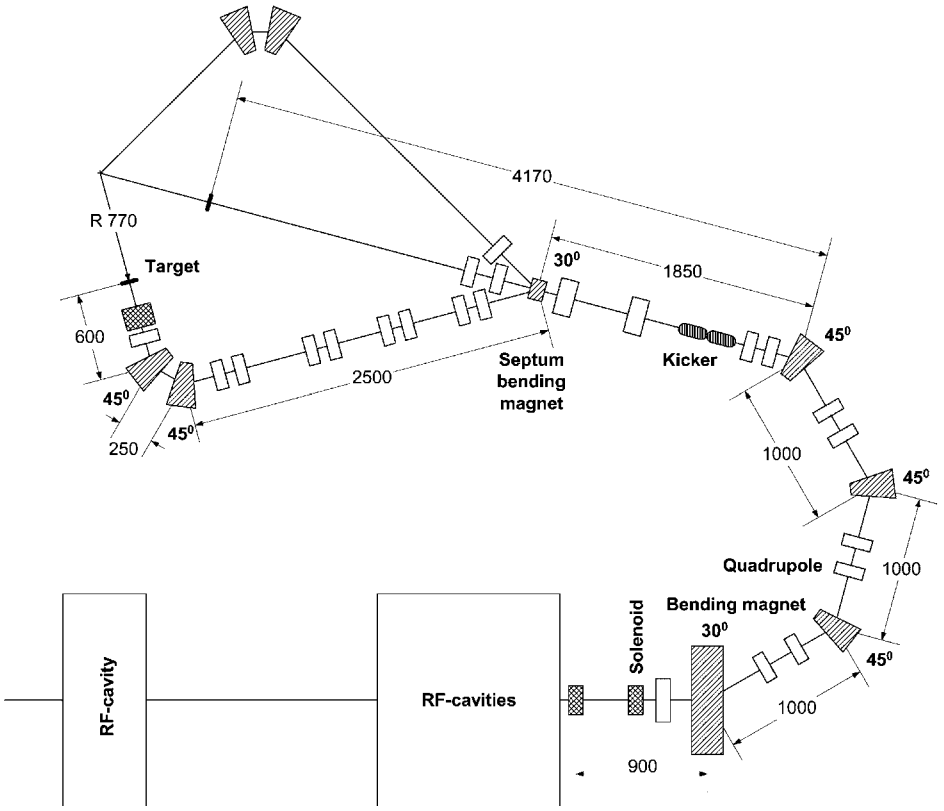



Fig. 2. Beamline for beam splitting into three targets at BINP.

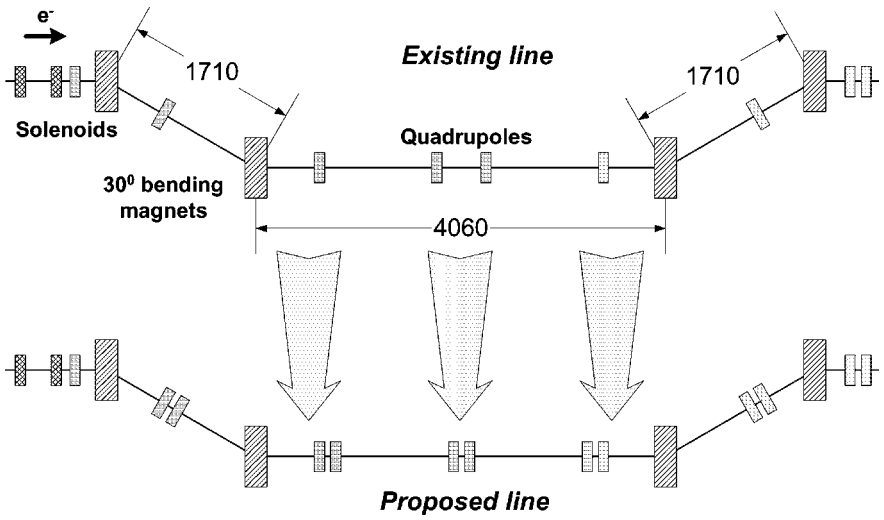



Fig. 3. Injection beamline of the BINP FEL.

simplified by extension of the aperture. The problem is that it is always limited in the kickers and the septum magnets, so there is almost no point in increasing the aperture in other parts. In any case, $<10\%$ loss is acceptable for this rare pulse machine.

### 3.3. BINP FEL injection beamline

The beamline Fig. 3 is used to transport an electron beam from the injector to a set of accelerating RF cavities of the BINP high-power FEL [4]. It contains four 30° bends and has to conduct a beam of 45 mA in continuous mode. The total length is ~8.5 m. The regular aperture is 100 mm. An additional condition is beam achromatism.

The beamline was originally designed using the uncharged linear model. It was later analyzed using the code described. No achromatic regime with acceptable loss was found, while in the absence of this condition, its acceptance is greater than that of the KAERI beamline. In situ, acceptable loss was obtained only by manual adjustment of all the lenses in the beamline, so the regime is not achromatic.

However, an achromatic regime with acceptable loss is possible if four additional quadrupole lenses are added, and some others are shifted. In this case, the acceptance is the same as in the original beamline, with no achromatism condition. This proposed modification is now being considered.

## 4. Conclusions

- A novel method for optimization of beamlines carrying a transient, high-current electron beam was proposed.
- A code implementing this method was developed.
- A number of beamlines were designed and analyzed with the code.
- Successful operation of the beamline at KAERI with the predicted regime verifies the reasonableness of the method.

## References

[1] J. Rossbach, P. Schmüser, Basic course on accelerator optics, CAS CERN Accelerator School, in: Fifth General Accelerator Physics Course Proceedings, Geneva, 1994, pp. 17–88.
[2] I.M. Kapchinskii, V.V. Vladimirskii, in: Proceedings of the International Conference on High-Energy Accelerators and Instruments, CERN, Geneva, 1959, p. 274.
[3] B.C. Lee, Y.U. Jeong, S.H. Park, et al., Nucl. Instr. and Meth. A 528 (2004) 106.
[4] E.I. Antokhin, R.R. Akberdin, et al., J. Synchrotr. Radiat. 10 (2003) 343.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 131–134
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Beam dynamics simulations based on the method of moments with polynomial approximation of the charge density

Elkuno A. Perelstein*, Liudmila V. Bobyleva

*Joint Institute for Nuclear Research, Dubna, Russia*

Available online 13 December 2005

**Abstract**

The problem of non-linear beam dynamics simulations based on the moment method is discussed. A new simulation method allows the study of non-linear effects, such as increasing emittance, a varying density profile and asymptotic behavior of the beam.


*PACS:* 02.60.CB; 41.75.−i

*Keywords:* Simulation; Non-linear beam dynamics; Charge particle

## 1. Introduction

At present, the beam space charge problem for low-energy ion beams is studied using particle-in-cell (PIC) methods [1] and root-mean squared (RMS) envelope equation (second-order moment equations) [2,3]. The second approach involves using an effective linear self-electric field [3]. In this report we discuss the problem of non-linear beam simulation based on the method of moments for a particle distribution function [4,5]. Polynomial approximation of the charge density of a beam with circular cross-section has been found using time-varying moments. The self-electric field was reconstructed with the help of the density approximation and used in the system of moment equations. Thus, approximate solutions of the Vlasov self-consistent problem were obtained. The new simulation method is demonstrated for examples of the free expansion of beams with non-uniform space charge density and beam propagation through a smooth non-linear focusing channel. Non-linear effects such as increasing emittance, varying density profile and asymptotic behavior of the beam are discussed.

## 2. Self-electric field reconstruction using polynomial density

Supposing that a beam with circular cross-section with radius $a$ has azimuthal symmetry charge density $\rho(r)$, then the self-electric field is

$$E_r = \frac{4\pi}{r}\int_0^r \rho(r) r\,\mathrm{d}r. \tag{1}$$

The $m$ lower density moments are known and defined as

$$\int_0^a r^n \rho(r) r\,\mathrm{d}r = M^{n,0}, \quad n = 0, 1, \ldots, m \tag{2}$$

under the condition $\int_0^a \rho(r) r\,\mathrm{d}r = 1$. The power expansion of the density is

$$\rho(r) = \sum_{i=0}^{m} \alpha_i r^i. \tag{3}$$

We use the accuracy measure

$$I_m(\alpha_m) = \int_0^a \left(\rho(r) - \sum_{i=0}^{m} \alpha_i r^i\right)^2 r\,\mathrm{d}r. \tag{4}$$

Using the minimum conditions for this function, we obtain a system of non-linear algebraic equations for the unknown

*Corresponding author.
*E-mail address:* perel@nusun.jinr.ru (E.A. Perelstein).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.097

coefficients $\alpha_i$, $i = 0, 1, \ldots, m$ and beam radius $a$:

$$\sum_{i=0}^{m} \alpha_i \frac{a^{i+j+2}}{i+j+2} = M^{j,0}, \quad j = 0, 1, \ldots, m. \tag{5}$$

Thus, we can reconstruct the beam density $\rho(r)$ and find the strength of the self-electric field using Eq. (1)

$$E_r = \frac{4\pi}{r} \int_0^r \left( \sum_{i=0}^{m} \alpha_i r^i \right) r \, \mathrm{d}r = 4\pi \sum_{i=0}^{m} \alpha_i \left( \frac{r^{i+1}}{i+2} \right). \tag{6}$$

The reconstruction procedure used may be demonstrated for two of the simplest cases. The first is the case of uniform density $\rho(r, z) = \alpha_0(z)$, i.e. $m = 0$. Then we have

$$\alpha_0 \frac{a^2}{2} = M^{0,0}, \quad \alpha_0 \frac{a^4}{4} = M^{2,0}. \tag{7}$$

Solving these equations gives $\alpha_0 = (M^{0,0})^2/M^{2,0}$, $a = \sqrt{2M^{2,0}}/M^{0,0}$, $\rho(r) = \alpha_0 = (M^{0,0})^2/M^{2,0}$. In a second case, we suppose that charge density has the form $\rho(r, z) = \alpha_0(z) + \alpha_2(z)r^2$, i.e. $m = 2$. Eq. (5) transforms to the following:

$$\alpha_0 \frac{a^2}{2} + \alpha_2 \frac{a^4}{4} = M^{0,0}, \quad \alpha_0 \frac{a^4}{4} + \alpha_2 \frac{a^6}{6} = M^{2,0},$$
$$\alpha_0 \frac{a^6}{6} + \alpha_2 \frac{a^8}{8} = M^{4,0}. \tag{8}$$

As a result of the solution of Eq. (8), we obtain

$$\alpha_0 = \frac{8M^{0,0}}{a^2} - \frac{12M^{2,0}}{a^4}, \quad \alpha_1 = -\frac{12M^{0,0}}{a^4} + \frac{24M^{2,0}}{a^6} \tag{9}$$

and the equation for the change in radius

$$-\frac{a^4 M^{0,0}}{6} + a^2 M^{2,0} = M^{4,0}. \tag{10}$$

Thus, we only need the first three moments to obtain the charge density approximation in this case.

## 3. Self-field polynomial approximation and moment evolution

The moment evolution during beam propagation is defined by the following system of equations [4]:

$$\frac{\mathrm{d}M^{pq}}{\mathrm{d}z} = pM^{p-1,q+1} + q \sum_k a_k(z) M^{p+k,q-1} \tag{11}$$

where $a_k(z) = \beta_k + \gamma_k$. An external field is represented as $E_{\mathrm{ext}} = \sum_{k=0}^{n} \gamma_k r^k$.

To integrate the system in Eq. (11), we use a symmetric Euler scheme. The scheme conserves emittance, whereby an exact solution corresponds to emittance conservation. Initial values of the problem include $a$, the beam radius, and the first $m$ moments of the charge density. Using these values we can find an approximate charge density and power expansion of the self-electric field strength. Then we integrate Eq. (11) to obtain the beam radius and moments at the next step. Using these values, the charge density is reconstructed in an iterative manner. Such an approach provides the numerical solution for a self-consistent problem of beam dynamics.

## 4. Solving the self-consistent problem of charge beam-free expansion in a polynomial density approximation model

Free expansion of a beam of uniform density is defined by the system in Eq. (7) for charge density and beam radius $a$ and the following equations:

$$\frac{\mathrm{d}M^{2,0}}{\mathrm{d}z} = 2M^{1,1}, \quad \frac{\mathrm{d}M^{1,1}}{\mathrm{d}z} = M^{0,2} + \beta_1 M^{2,0},$$
$$\frac{\mathrm{d}M^{0,2}}{\mathrm{d}z} = 2\beta_1 M^{1,1} \tag{12}$$

where $\beta_1 = 4\pi(\alpha_0/2)$, $E_r = \beta_1 r$ is the strength of the linear beam self-field. Let the initial conditions correspond to $a = 1$ and $M^{1,1}_{z=0} = 0$, $M^{0,2}_{z=0} = 0.000028$. Beam envelope evolution is defined by the well-known equation [6]:

$$\frac{\mathrm{d}^2 a}{\mathrm{d}z^2} = \frac{\varepsilon^2}{a^3} + \frac{2ZI_a}{A\beta^3 I_A} \frac{1}{a} \tag{13}$$

where $\varepsilon$ is beam emittance, $I_a(A)$ is the beam current, and $I_A = 17\,\mathrm{kA}$, $Z$ and $A$ atomic number of an ion or $Z = A = 1$ for an electron beam. The numerical results of the envelope evolution for the current parameters chosen according to Eqs. (7) and (12) (line D) and to Eq. (13) (line B) are presented in Fig. 1.

It is evident that the results coincide quite well. The emittance is conserved with an accuracy of approximately $10^{-6}$. The uniform density is also conserved and $M^{2,0}/a^2 = 0.5$.

The free expansion of a beam with non-uniform density is defined by the following system of equations:

$$\alpha_0 \frac{a^2}{2} + \alpha_2 \frac{a^4}{4} = M^{0,0}, \quad \alpha_0 \frac{a^4}{4} + \alpha_2 \frac{a^6}{6} = M^{2,0} \tag{14}$$

$$\alpha_0 \frac{a^6}{6} + \alpha_2 \frac{a^8}{8} = M^{4,0} \tag{15}$$

and

$$\frac{\mathrm{d}M^{2,0}}{\mathrm{d}z} = 2M^{1,1}, \quad \frac{\mathrm{d}M^{1,1}}{\mathrm{d}z} = M^{0,2} + \beta_1 M^{2,0} + \beta_3 M^{4,0} \tag{16}$$

$$\frac{\mathrm{d}M^{0,2}}{\mathrm{d}z} = 2\beta_1 M^{1,1} + \beta_3 M^{3,1} \tag{17}$$

where $\beta_k = \alpha_{k-1}/k + 1$ , $E_r = \beta_1 r + \beta_3 r^3$ is the strength of a non-linear beam self-field. Let the initial conditions correspond to non-uniform initial density $\rho(r) = 1 - 0.9r^2$, $a = 1$ and $M^{1,1}_{z=0} = 0$, $M^{0,2}_{z=0} = 0.000028$. To terminate the chain of moment equations we have used an equality that corresponds to known simulation and

experimental results [7–10]:

$$M^{4,0} = kM \cdot (M^{2,0})^2,$$
$$kM = kM_0 + (kM_\infty - kM_0) \cdot \left(1 - \mathrm{Exp}\left(-\frac{kE \cdot z}{z\,\mathrm{max}}\right)\right) \quad (18)$$

where the coefficient $kM_0 = M^{4,0}/(M^{2,0})^2$ contains the initial moment values. The coefficient $kM_\infty = 1.33$ corresponds to the asymptotic value of this ratio (here it corresponds to uniform density). The parameter $kE$ must be chosen from the problem conditions; here $kE = 10$. A moment $M^{3,1}$ is defined as $M^{3,1} = \mathrm{d}M^{4,0}/\mathrm{d}z$. The results of the numerical solution of the problem are presented in Figs. 1(line F)–3. Beam envelope evolution is shown in Fig. 1 (line F). Increasing emittance is demonstrated in Fig. 2. In Fig. 3, changes in charge density are shown for a step $\Delta z = 2.5$ cm along the beam propagation (lines B–P).

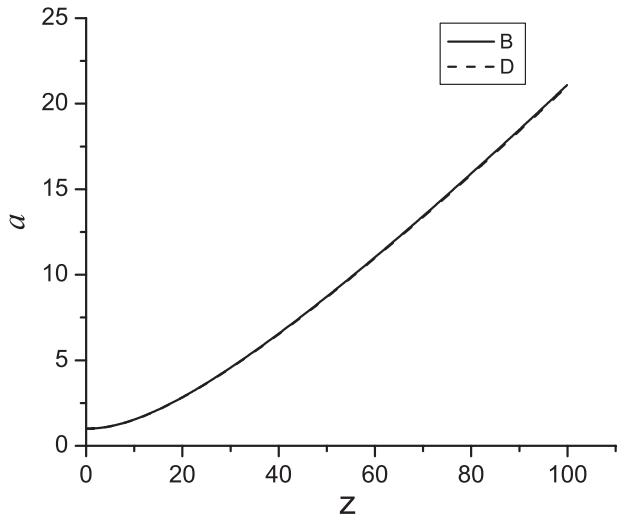



Fig. 1. The evolution of envelopes.

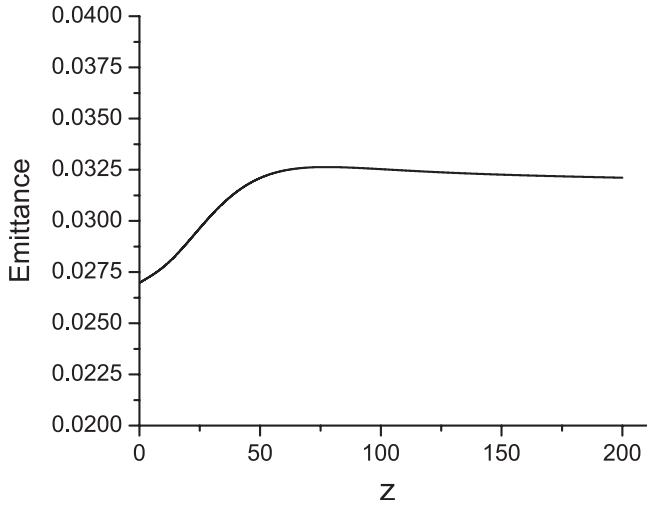



Fig. 2. Changing beam emittance.

## 5. Solution of the problem of beam dynamics in a smooth focusing channel in a polynomial density approximation model

Let us consider the dynamics of a beam with a circular cross-section in a smooth focusing channel. We suppose that the beam density has the form $\rho(r,z) = \alpha_0(z) + \alpha_2(z)r^2$ and non-linear focusing force $F_{\mathrm{ext}} = -r + 0.2r^3$. The beam dynamics is defined by the system of equations (16) and (17) with $\beta_1 = -1$, $\beta_3 = 0.2$ and the system in Eq. (8), which describes the beam radius and changes in charge density. The initial conditions correspond to the following: $\rho(r,z) = 3.6(1 - 0.9r^2)$ inside a circle with $a = 1$ and $M^{1,1}_{z=0} = 0$, $M^{0,2}_{z=0} = 0.000028$. To terminate the chain of equations we use Eq. (18). The main simulation results are presented in Figs. 4–6. Changes in RMS beam dimension is shown in Fig. 4. Fig. 5 demonstrates that emittance increases due to the non-linearity of the external force.

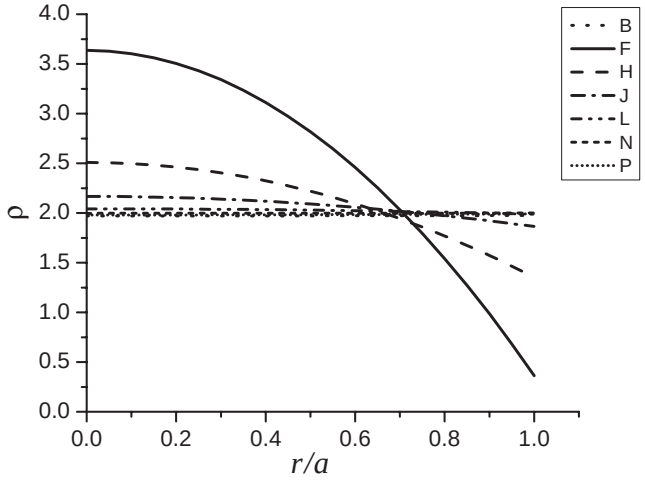



Fig. 3. Changes in charge density for a beam with non-uniform density.

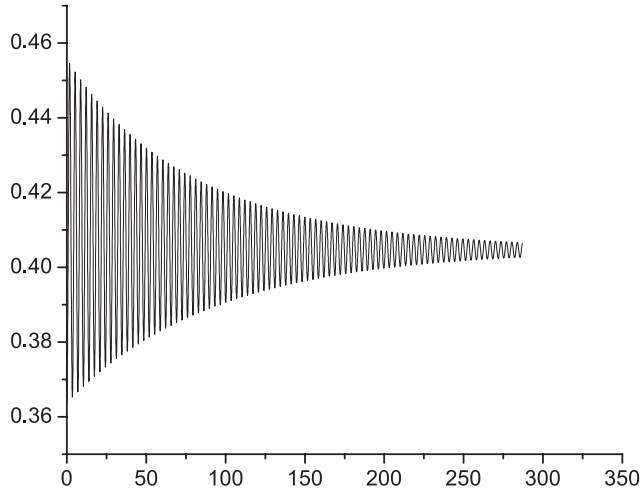



Fig. 4. Changes in of the RMS dimension of a non-uniform beam in a smooth non-linear focusing channel.

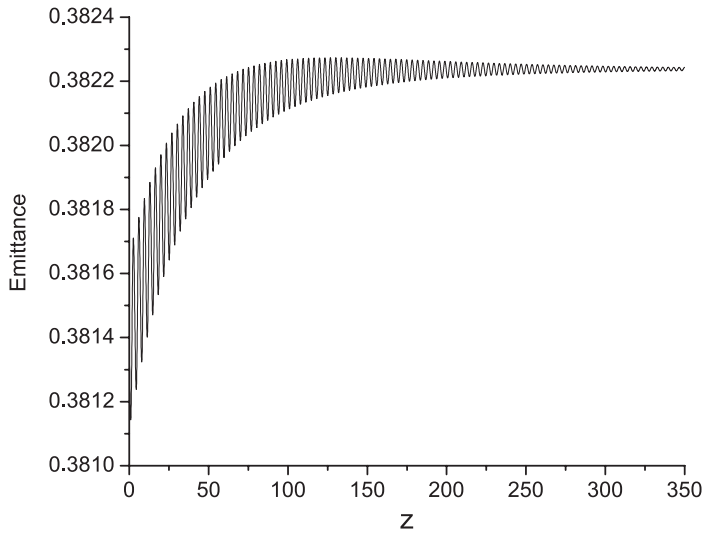



Fig. 5. Increasing emittance of a non-uniform beam in a smooth non-linear focusing channel.

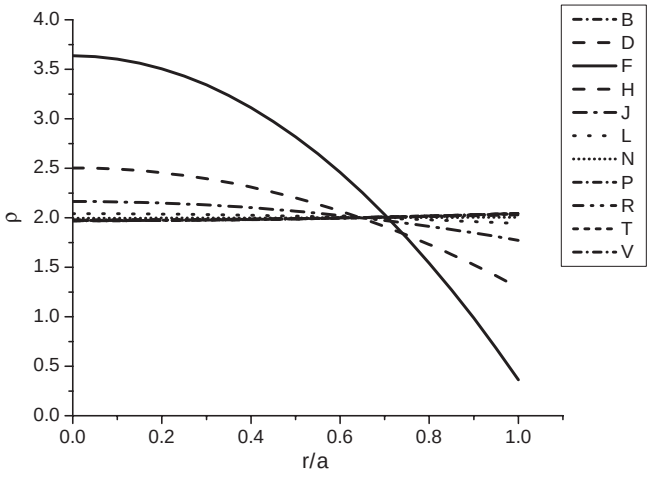



Fig. 6. Changes in the charge density of a non-uniform beam in a smooth non-linear focusing channel.

In Fig. 6, the changes in charge density are shown with a step $\Delta z = 2.5\,\text{cm}$ along the beam propagation (lines B–V).

## References

[1] A.S. Roshal, Charge Beam Simulation, Atomizdat, Moscow, 1979 (in Russian).
[2] P.M. Lapostolle, IEEE Trans. Nucl. Sci. NS-18 (3) (1971) 1101.
[3] F.J. Sacherer, IEEE Trans. Nucl. Sci. NS-18 (3) (1971) 1105 (CERN/SI/Int. 70-12).
[4] A.D. Dymnikov, E.A. Perelstein, Nucl. Instr. and Meth. 148 (1978) 567.
[5] N.Yu. Kazarinov, E.A. Perelshtein, V.F. Schevtsov, Part. Acceler. 10 (1980) 1.
[6] I.M. Kapchinsky, V.V. Vladimirsky, Proceedings of the International Conference on High-Energy Accelerators and Instruments, CERN, Geneva, 1959, p. 274.
[7] L.V. Bobyleva, E.A. Perelstein, Proceedings of the 1996 European Particle Accelerator Conference, Bristol, 1996, p. 1060.
[8] P.L. Bruhwiler, Y.K. Batygin, Proceedings of the 1995 Particle Accelerator Conference, Dallas, 1995, p. 3254.
[9] M. Raiser, Theory and Design of Charged Particle Beams, Wiley, New York, 1994.
[10] I. Hofmann, Proceedings of the CERN Accelerator School, Geneva, 1997, p. 327.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 135–141
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Optimizing the adiabatic buncher and phase-energy rotator for neutrino factories

Alexey A. Poklonskiy[a,*], David Neuffer[a], Carol J. Johnstone[a], Martin Berz[b], Kyoko Makino[b], Dmitriy A. Ovsyannikov[c], Alexandre D. Ovsyannikov[c]

[a]*MS221, Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510-500, USA*
[b]*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*
[c]*Universitestkiy Prospekt, 35, Petergof, Faculty of Applied Mathematics and Control Processes, Saint-Petersburg State University, St. Petersburg 198504, Russia*

Available online 13 December 2005

**Abstract**

In the US scenario for a Neutrino Factory presented in "A feasibility study of a neutrino source based on a muon storage ring", N. Holtkamp (Ed.), D. Finley (Ed.), Fermilab, April 15th, 2000), a large percentage of the cost is related to an induction linac for phase-energy rotation and bunching of the muon beam collected after the production target and decay channel. A more cost-effective adiabatic buncher and phase-energy rotator has been proposed to replace the induction linac system (D. Neuffer, A. Van Ginneken, High-frequency bunching and $(\phi - \delta E)$ rotation for a muon source, Proceedings of the 2001 Particle Accelerators Conference, Chicago, 2001, p. 2029). The new method uses consecutive RF cavities with differing frequencies. The frequencies are changed to enable bunching and phase-energy rotation. In this paper, the theoretical concept is developed and demonstrated with simulation results obtained with the map code *COSY Infinity* (http://cosy.pa.msu.edu). An optimization strategy is also explored.


*PACS:* 29.17.+w; 29.27.−a; 02.30.Yy; 07.05.Tp

*Keywords:* Neutrino factory; Buncher; Phase rotator; Map methods

## 1. Introduction

In various scenarios for muon-based accelerator projects such as a Muon Collider or a Neutrino Factory (Fig. 1), phase-energy $(\phi - \delta E)$ rotation is used in the μ beam exiting the $\pi$ production and decay channel, because this beam has not just a relatively small initial phase spread, but an energy spread that is much larger than the device acceptance [1–3]. In this process the beam is first allowed to lengthen and then the radio-frequency (RF) system is used to reduce the energy spread (by decelerating the high-energy "head" of the bunch and accelerating the low-energy "tail", so that the beam "rotates" in phase-energy space). The resulting beam has the energy spread reduced to a level where the majority of the beam particles is captured by a subsequent bunching and/or cooling system. The phase-energy rotation region is outlined by rectangle in Fig. 1.

The difficulty with the previously proposed $(\phi - \delta E)$ rotation systems is that they require either very low-frequency RF, or an induction linac, matched to the elongated bunch length of the $(\phi - \delta E)$ rotated system. This long-wavelength (or long rise-time) acceleration system would require new technology development and considerable expense. In this paper we present an approach which uses high-frequency RF systems for bunching the beam and reducing its overall energy spread [4]. With this approach it is possible to produce a particle distribution similar to that obtained in the induction linac and RF buncher system proposed in Study 2 [1–3], except that this

*Corresponding author.
*E-mail addresses:* poklonsk@fnal.gov (A.A. Poklonskiy), neuffer@fnal.gov (D. Neuffer), cjj@fnal.gov (C.J. Johnstone), berz@msu.edu (M. Berz), makino@msu.edu (K. Makino), dovs@apmath.spbu.ru (D.A. Ovsyannikov), ovs@apmath.spbu.ru (A.D. Ovsyannikov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.038

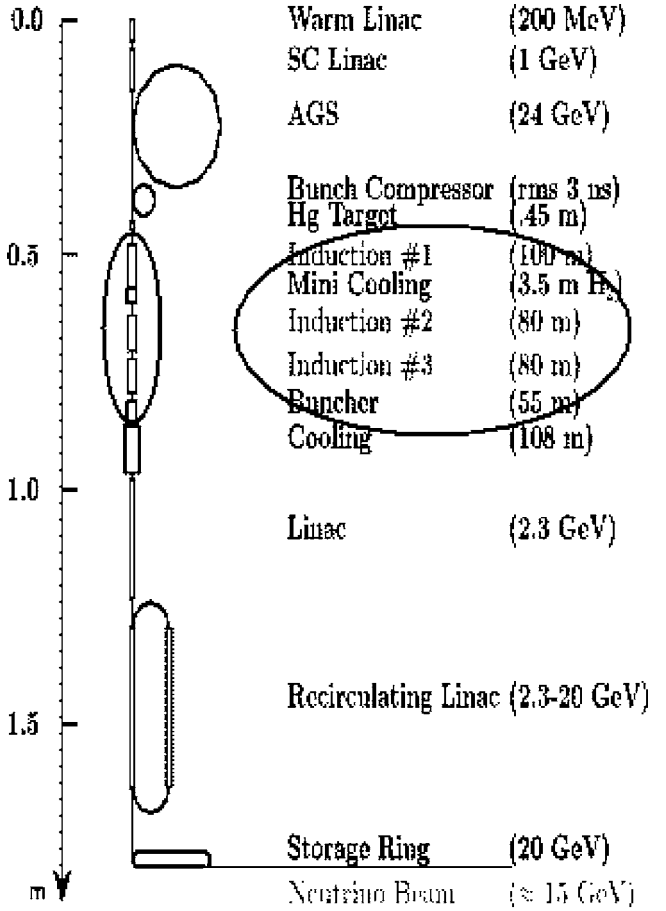



Fig. 1. Neutrino Factory schematics as proposed in Study 2 with outlined phase-energy rotation section.

system simultaneously captures both $\mu^+$ and $\mu^-$. The concept, key parameters, example simulations, and an optimization strategy based on the control theory approach are presented.

## 2. Concept

The initial μ beam with a small phase spread and a large energy spread from a $\pi \rightarrow \mu$ production target is allowed to drift in a solenoidal field. The *drift* section is followed by *adiabatic buncher* where beam is formed into a string of bunches and *phase-energy rotator* where its overall energy spread is reduced. Both buncher and rotator consists of RF cavities within a solenoidal (transversely focusing) field (Fig. 2).

To bunch particles we choose some particle to be the main central particle of the beam. We set all RF cavities parameters in such a way that this particle passes every one of them in the same phase of E field oscillations (0 in a buncher). By the virtue of the equations of motion in such a structure, particles close enough to this central one, will be formed into a stable group called "bunch". Because of the specific choice of this main central particle's phase and cavities parameters, we also have some other particles passing cavities in the same 0 phase and, by the same equations of motion, we will have bunching effect around those particles as well. In the following text we will call them "central particles" and the one chosen first "main central particle". Of course, all central particles are not real particles, they are just an idealization chosen to make equations of motion simpler.

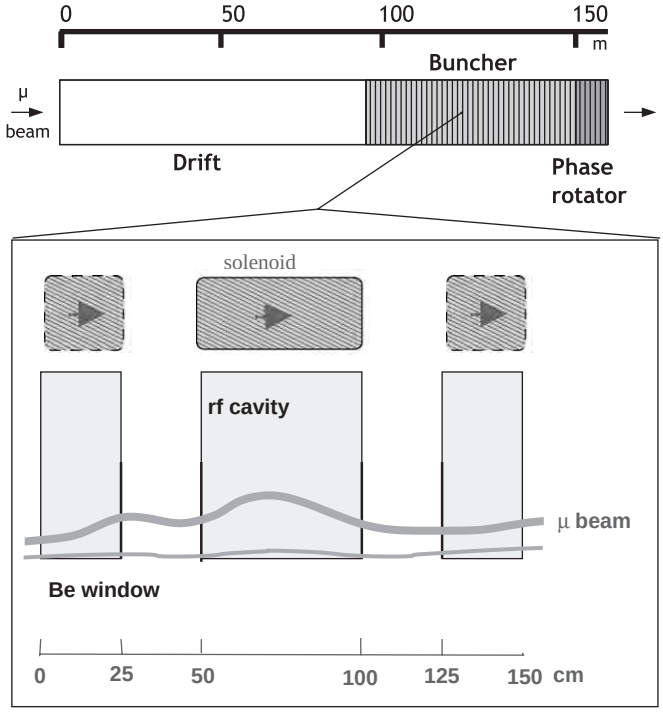



Fig. 2. Example simulation plots in ($T$,ct) phase space. Beam is shown after drift, buncher, rotator and cooler, respectively.

Each cavity in the buncher has its frequency set to maintain the following condition: the time of arrival difference between two central particles in a place of RF field application remains equal to a fixed integer number of RF oscillations periods so this condition holds true as the beam propagates through the buncher. As we set E field phase to be 0 for the main central particle, for other central particles it is also 0 so they gain no energy in each cavity and their energies stay constant through the buncher. We keep the final system frequency fixed because of matching into 201.25 MHz cooling and/or accelerating sections, so, using these two conditions we can find a frequency for each RF cavity as a function of its position in a buncher.

In the buncher, the RF gradient is increased as length increases. The goal here is to perform an *adiabatic capture*, in which the beam within each bunch is compressed in phase so as to be concentrated near the central particle. From these considerations we obtain the following relations for the lattice parameters and central particles of the bunches:

$$\Delta t = t_n - t_c = z\left(\frac{1}{v_n} - \frac{1}{v_c}\right) = nT_{RF} = n\frac{\lambda_{RF}}{c}, \quad n \in \mathbb{Z} \tag{1}$$

$$\delta\left(\frac{1}{\beta}\right) = \left(\frac{1}{\beta_1} - \frac{1}{\beta_c}\right) = \frac{\bar{\lambda}}{\bar{L}} \tag{2}$$

from which follows

$$\frac{1}{\beta_n} = \frac{1}{\beta_c} + n\delta\left(\frac{1}{\beta}\right) \Rightarrow T_n, \quad n \in \mathbb{Z} \tag{3}$$

$$\lambda_{RF}(z) = z \cdot \delta\left(\frac{1}{\beta}\right) \Rightarrow \nu_{RF} = \frac{c}{z \cdot \delta\left(\frac{1}{\beta}\right)} \tag{4}$$

$$V_{RF}(z) = B\frac{(z - z_D)}{L} + C\frac{(z - z_D)^2}{L} \tag{5}$$

where $n$ — denote the number of the bunch counted from the main central particles's one, $T_{RF}$ — period of RF field oscillations, $z$ — longitudinal coordinate counted from the beginning of the drift, $\beta_c$ and $\beta_n$ — normalized central particle and $n$-th reference particle's velocities, $t_c$ and $t_n$ — time of arrival of main central and $n$-th central particle (main central particle has $n = 0$); $\lambda_{RF}$, $\nu_{RF}$ and $V_{RF}$ — wavelength, frequency and gradient of the electric field in the cavity, $z_D$ — longitudinal coordinate of the beginning of the buncher (equals to the drift length); $\bar{\lambda}$ and $\bar{L}$ — wavelength of electric field and the longitudinal coordinate of the last RF in buncher; $c$ — speed of light, $B$ and $C$ — positive constants, defining RF gradients in a buncher. Note that, since each of the bunches is centered at different energy, they all have different longitudinal oscillation frequencies, and a simultaneously matched compression for all bunches is not possible. Instead a quasi-adiabatic capture resulting in an approximate bunch length minimization in each bunch is attempted.

Following the buncher is the so-called $(\phi - \delta E)$ *vernier rotation system* in which the RF frequency is almost fixed to the matched value at the end of the buncher and the RF voltage is constant. In this system the energies of the central particles of the low-energy bunches increase, while those of the high-energy bunches decrease. So the whole energy spread reduces to the point where beam is a string of similar-energy bunches and it could be captured into the ~200 MHz ionization cooling system matched to the central energy of the beam.

Let us describe the rotator parameters calculation in more detail. At the end of the buncher we choose a second central particle kept $N$ RF periods from main central one along the buncher and the *vernier offset* $\delta$. We then keep this second central particle at $(N + \delta)\lambda_{RF}$ wavelengths from the main one through the rotator. So now it passes all RFs in a constant accelerating phase $\phi_N$ having constant energy gain $\Delta T_N$, and after $|T_c - T_N|/\Delta T_N$ cavities, energies of the central particle and the chosen second reference one will be nearly equal. This process also aligns the energies of other reference particles and their bunches, hence at the end we have the beam rotated in $(\phi - \delta E)$ space with significantly reduced energy spread.

Example simulation of this process in 1D was developed in a Pascal code [5,6]. We take the main central particle's energy to be 125 MeV, beam's energy spread to be ±50 MeV and $(\phi - \delta E)$ coordinates distribution to be Gaussian. We arbitrarily set the initial drift length to 90 m and define buncher to consist of 60 pillbox cavities 1 m long each. With these numbers we get $\delta(1/\beta) = 1.5/150 = 0.01$, so, by plugging this values into Eq. (4) we obtain the RF frequency at the beginning of the buncher section ~333 MHz, and at the end ~200 Mhz (we match into a 201.25 Mhz cooling system). We choose the RF gradient to be quadratically increasing from 0 to 4.8 MV/m along the buncher, so, from Eq. (5) it follows that

$$V_{RF}(z) = 4.8\frac{(z - z_D)^2}{(\bar{L} - z_D)^2}\text{MV/m}. \tag{6}$$

In rotator we choose $N = 10 \Rightarrow T_N \approx 77.28$ MeV; $\delta = 0.1 \Rightarrow \phi_N = 36^\circ$ and the RF gradient to be 10 MV/m, so, after $\frac{1}{4}$ synchrotron oscillation (~8.4 m), the central energies spread becomes nearly 0. Coordinates of the particles in $(\phi - \delta E)$ phase space through the structure are shown in 2.

## 3. Problem description and key parameters/controls

The concept is defined in the previous section, but there are many variations in the structure parameters because of the minimal cost constraint, different possible final RF frequencies, reduced number of RF frequencies and gradients, etc., and in the final beam properties: shorter/longer bunch trains, constraints on the number of muons captured, desired central energy, etc. Matching into the accelerating/cooling structures following the buncher-rotator system and the transverse beam dynamics should

also be considered. The problem is in finding specific optimal parameters of the beam and/or the structure under imposed constraints.

In our example we use 60 RF cavities, each with different gradient and frequency, which is definitely makes structure too expensive to be built. Structure consisted of 10 cavities would be much more satisfactory in a cost sense. We could also try to combine the buncher and the $(\phi - \delta E)$ rotator into one structure for simultaneous bunching and rotation to reduce cost of the system. As could be seen from the concept and the relations given above, the control parameters of the structure are:

(1) *Drift*: the length of the section $L_{\mathrm{D}}$. Future studies, which include transverse motion, must also consider the apertures and focusing fields (this study uses fixed-field solenoids for transverse focusing). These focusing parameters are also critical for system performance.
(2) *Buncher*: the length of the section $L_{\mathrm{B}}$, RF voltages $V_{\mathrm{B}}^{i}$, $i = \overline{1, n_{\mathrm{RFs}}}$ or initial and final voltage and the law of voltage increase (linear, quadratic, etc.). Final frequency is usually strictly specified by the cooling/accelerating subsections of the accelerator, but could also be varied to find optimum.
(3) $(\phi - \delta E)$ *Rotator*: the length $L_{\phi\mathrm{R}}$, RF voltage $V_{\phi\mathrm{R}}$ of the phase-energy rotation section, number $N$ of RF field oscillation periods between chosen second central particle and the main central particle (with $n = 0$), and the vernier parameter $\delta$. Also the kinetic energy $T_{\mathrm{c}}$ of the main central particle could be changed (usually we take $T_{\mathrm{c}}$ to be the peak of energy distribution of beam's particles).

## 4. Simulations

Advanced particles dynamics simulation program was written in the *COSY Infinity* code [7,8] which uses a map-based approach of beam dynamics calculation. The dynamics is described in terms of a high-order Taylor expansion of the flow, i.e. the relationship that connects final particle coordinates to initial coordinates via

$$\vec{z_f} = \mathscr{M}(\vec{z_i}) \tag{7}$$

where the flow (also called "map") $\mathscr{M}$ is determined by either automatic differentiation of numerical integration algorithms, or by dedicated tools determining the flow of partial differential equations based on differential algebraic techniques. Depending on the complexity of the task, different orders of expansion from 1 to 15 are needed, and in our case, the necessary order is high ($\geqslant 7$). Apart from the powerful abilities to calculate high-order dynamics properties, *COSY Infinity* has its own programming language, which allows the construction of complicated optimization scenarios, it provides user with powerful DA (*differential algebra*) framework, and it has large built-in library of standard accelerator elements, so it fits well to our needs.

Because the beam has a very large energy spread and *COSY Infinity* calculates dynamics with the use of Taylor expansions on particle coordinates, a division of the initial coordinates domain into sub-domains with small energy coordinate range is required before tracking. The natural way of doing this is to divide the set of coordinates by the number of bunches in a beam. This division makes relative coordinates small enough to have Taylor expansions with acceptably small remainders at a reasonable order (order 7 or 8 is enough). The standard RF kick approximation for a pillbox RF cavity is used, i.e. we assume particles to pass the cavity instantly so that field does not change, having constant energy increment/decrement dependent on the particle coordinates and cavity parameters. This approximation is suitable for our simulations because our particles are fast enough (200 MeV μ has $\beta \approx 0.94$) and cavities are short enough to neglect any field change during particle transition. If deemed necessary, in future simulations a more realistic RF cavity model like the one developed in Ref. [4] could be used.

In our example structure we use 50 sub-domains for more realistic purely Gaussian (larger energy spread) and "almost real" distributions from *MARS* code [9]. Simulation results for these distributions are shown in Figs. 3–5.

## 5. Optimization problem formulation

The problem of simulation and search for optimal parameters naturally presents itself as one of the problems of control theory in beam physics [10]. The exact definition of the problem depends on what parameters of the structure lattice and/or the beam at the end of the structure are critical to achieve. We can consider an impulse effect model:

$$\vec{x}_k = \mathscr{A}^k(\vec{x}_{k-1}, \vec{u}_k) \tag{8}$$

where $\vec{x}_k$ is a vector of coordinates in phase space after the $k$-th lattice element represented by the $\mathscr{A}^k$ operator and the control $\vec{u}_k$. As could be seen from comparison with Eq. (7), this method is exactly the one used in *COSY Infinity* to calculate beam dynamics. We can as well use the continuous model

$$\dot{x} = f(t, x, u), \quad t \in [0, T], \ \ x \in \Omega \subset \mathbb{R}^{\mathrm{n}}, \ \ u \in U \subset \mathbb{R}^{\mathrm{r}} \tag{9}$$

where $x, u$ are coordinate and control vectors, respectively, $t$ is the time of flight, and $f$ is a continuous function, which describes particle's dynamics dependence on control functions representing the structure's parameters. On the trajectories obtained from any of these models, we define a quality functional as

$$I(u) = \int_0^T \int_{M_{t,u}} \phi(t, x_t, u(t))\, \mathrm{d}x_t\, \mathrm{d}t + \int_{M_{T,u}} g(x_T)\, \mathrm{d}x_T \tag{10}$$

where $x_t$, $x_T$ are the coordinates of the particle at the time $t$ and at the terminal time $T$ respectively, $M_{t,u}$, $M_{T,u}$ are the

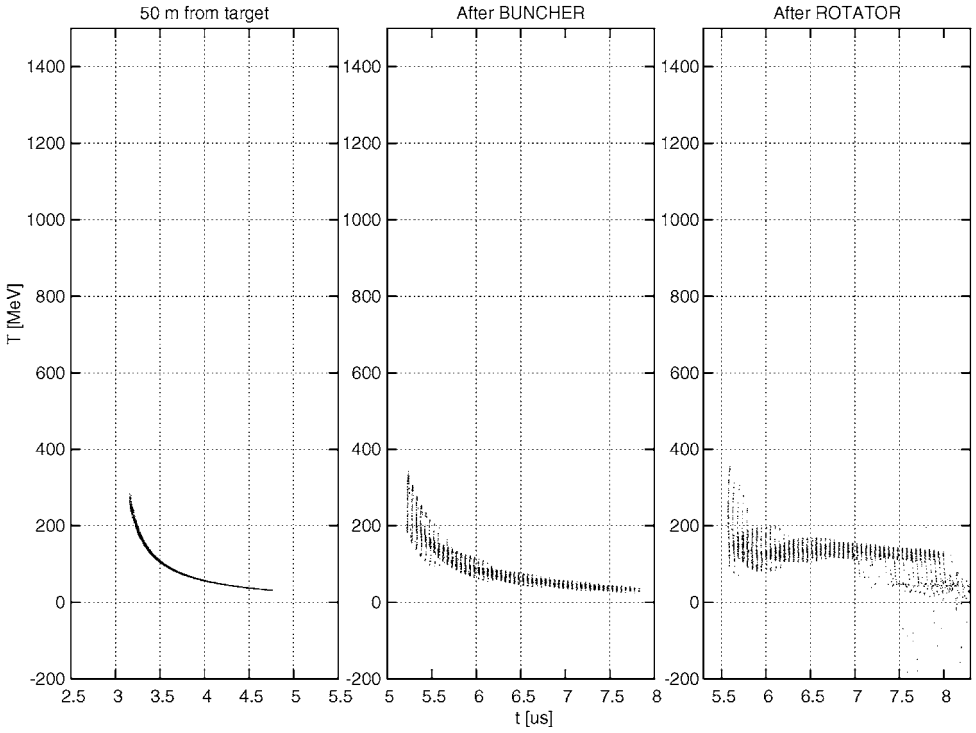



Fig. 3. *COSY Infinity* simulations plots in $(T, t)$ phase space. Initial distribution Gaussian, $T = 125\,\text{MeV} \pm 50\,\text{MeV}$).

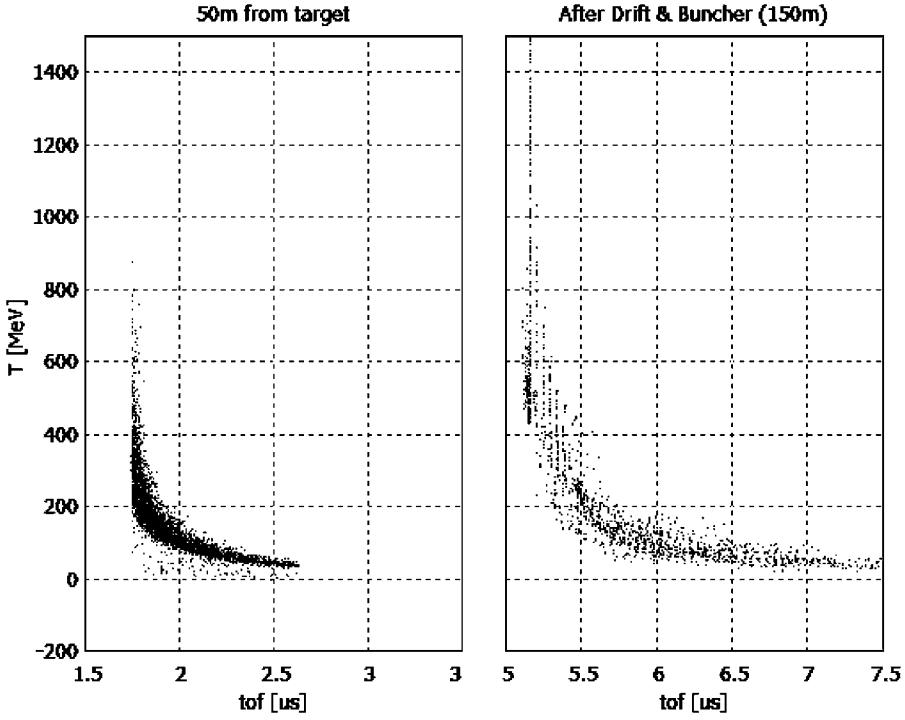



Fig. 4. *COSY Infinity* simulations plots in $(T, t)$ phase space. Initial distribution Gaussian $T = 10-1000\,\text{MeV}$.

sets of coordinates of beam particles at the time $t$ and at the terminal time $T$, respectively.

Functionals of the type (10) with problem-defined functions $\phi$ and $g$ allow one to evaluate any desired beam parameters throughout the whole structure with the first item and the terminal beam parameters with the second item. In search for optimal structure parameters (optimal in a sense they make dynamics of the particles and terminal beam parameters optimal) one then needs to find control functions that brings minimum/maximum to this functional.

As a variant of such formulation applicable to our problem we consider the problem of optimal transportation of the initial coordinates $\vec{x} = (x_1, x_2)^{\mathrm{T}}$ set $M_{0,u}$ into the

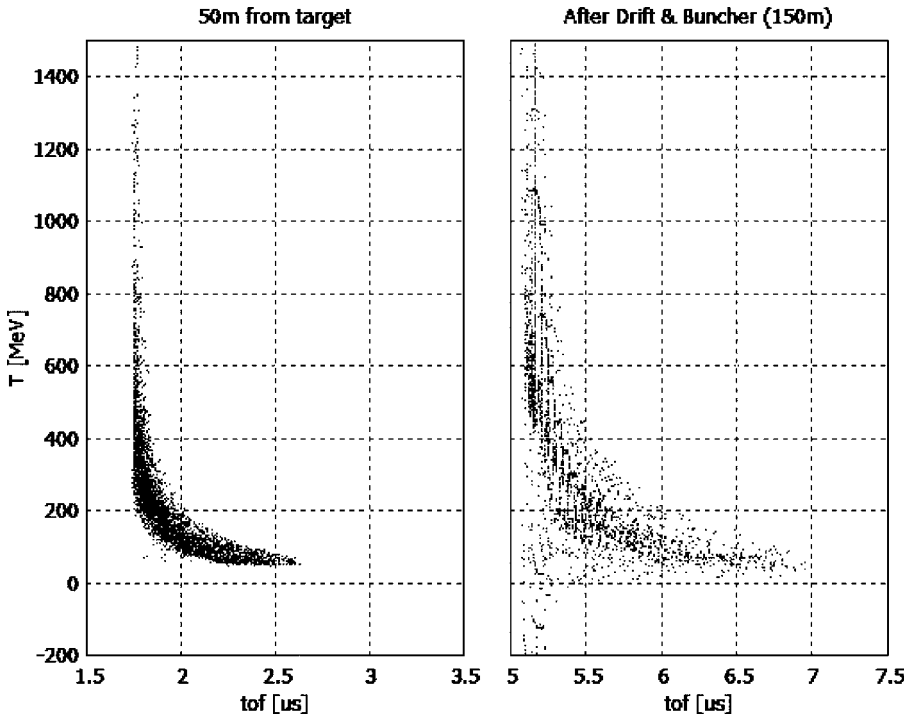



Fig. 5. *COSY Infinity* simulations plots in $(T, t)$ phase space. "Realistic" initial distribution obtained from *MARS* code [9].

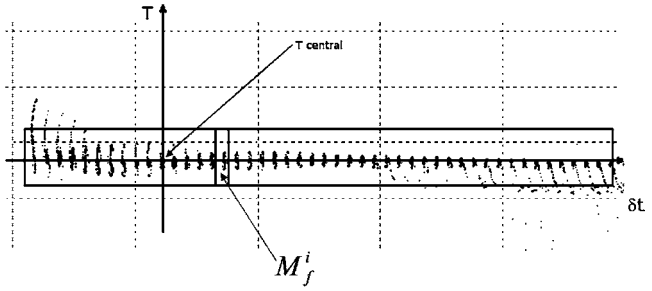



Fig. 6. Final beam structure and definition of bounding sets.

set of final coordinates $M_{T,u}$ with specified boundaries, i.e. we will base our evaluation on the terminal beam parameters. In our case $x_1 \leftrightarrow \phi$, $x_2 \leftrightarrow \delta$E. As an initial coordinate boundary we take a rectangle in phase-energy phase space, which encloses all or almost all particles of the beam. For the final boundary we may consider another rectangle with shorter length along energy coordinate as we are interested in reducing overall energy spread. But, in fact, we are also interested in enclosing the particles in each bunch to a so-called bucket area, because apart from the small energy spread we have the constraint of capturing, i.e. particles in each bunch should be in a stable area called "bucket" (in sense of equation of motion in resonant RF structure which has stable and unstable solutions). So we might divide this bounding rectangle into sub-rectangles (Fig. 6), the number of which is equal to the number of bunches in a beam, and define a penalty function to evaluate proximity of the terminal particles coordinates in each bunch to the rectangle corresponding to this bunch $[a_1, a_2] \times [b_1, b_2]$

$$\phi_1(\vec{x}) = \begin{cases} 0, & x_1 \in [a_1, a_2] \\ k_1(x_1 - a_2)^{2q_1}, & x_1 \geqslant a_2 \\ k_1(a_1 - x_1)^{2q_1}, & x_1 \leqslant a_1 \end{cases} \tag{11}$$

$$\phi_2(\vec{x}) = \begin{cases} 0, & x_2 \in [b_1, b_2] \\ k_2(x_2 - b_2)^{2q_2}, & x_2 \geqslant b_2 \\ k_2(b_1 - x_2)^{2q_2}, & x_2 \leqslant b_1 \end{cases} \tag{12}$$

where $k_1$, $k_2$, $q_1$, $q_2$ are arbitrary positive weight constants for selective optimization. Then, we define the quality functional as

$$I = \int_{M_{T,u}} (C_1\phi_1(\vec{x}) + C_2\phi_2(\vec{x}))\,\mathrm{d}\vec{x}. \tag{13}$$

With this functional we could perform optimization using some method of functional minimization (stochastic, gradient, etc.). This optimization is the main direction of the future research.

## References

[1] N. Holtkamp, D. Finley (Eds.), A feasibility study of a neutrino source based on a muon storage ring, Technical Report, FERMILAB-PUB-00-108-E, Fermi National Accelerator Laboratory, see ⟨http://library.fnal.gov/archive/test-preprint/fermilab-pub-00-108-e.shtml⟩ (April 15, 2000).

[2] S. Ozaki, R. Palmer, M. Zisman, J. Gallardo (Eds.), Feasibility study-II of a muon-based neutrino source, Technical Report, BNL-52623, Muon Collider Collaboration, see ⟨http://www.cap.bnl.gov/mumu/studyii/FS2-report.html⟩ (June 2001).

[3] M.C. Collaboration, $\mu^+\mu^-$ collider: a feasibility study, Technical Report BNL-52503, Fermilab-Conf-96/092, LBNL-38946, Muon Collider Collaboration, pres. at the Snowmass'96 workshop (July 1996).
[4] A. Geraci, T. Barlow, M. Portillo, J. Nolen, K. Shepard, M. Berz, K. Makino, Rev. Sci. Instrum. 73 (9) (2002) 3174.
[5] D. Neuffer, A.V. Ginneken, High-frequency bunching and $\phi - \delta E$ rotation for a muon source, in: Proceedings of 2001 Particle Accelerators Conference, Chicago, 2001, p. 2029.
[6] D. Neuffer, Exploration of the "high-frequency" buncher concept, Neutrino Factory/Muon Collider Notes MUC-NOTE-DECAY_-CHANNEL-0269, Fermi National Accelerator Laboratory, see ⟨http://www-mucool.fnal.gov/notes/notes.html⟩ 2003.
[7] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999 (also available at ⟨http://bt.pa.msu.edu/pub⟩).
[8] M. Berz, K. Makino, COSY INFINITY Version 8.1—user's guide and reference manual, Technical Report, MSUHEP-20704, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, see also ⟨http://cosy.pa.msu.edu⟩ 2001.
[9] N.V. Mokhov, A. van Ginneken, Pion production and targetry at $\mu^+\mu^-$ colliders, Technical Report, Fermilab-Conf-98/041, Fermi National Accelerator Laboratory, published Proceedings of the 4th International Conference on Physics Potential and Development of Muon Colliders, San Francisco, CA, December 10–12, 1997 (January 1998).
[10] D.A. Ovsyannikov, N.V. Egorov, Mathematical Modelling of Electron and Ion Beams Forming Systems, St.-Petersburg University Press, 1998.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 142–146
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Study and optimal correction of a systematic skew quadrupole field in the Tevatron

Pavel Snopok[a,b,*], Carol Johnstone[a], Martin Berz[b], Dmitry A. Ovsyannikov[c], Alexander D. Ovsyannikov[c]

[a]MS 221 Fermilab, P.O. Box 500, Batavia, IL 60510, USA
[b]Dept. of Physics and Astronomy, MSU, East Lansing, MI 48824, USA
[c]Dept. of Applied Mathematics and Control Processes, Universitetskii prospekt 35, Petergof, Saint-Petersburg 198504, Russia

Available online 5 December 2005

**Abstract**

Increasing demands for luminosity in existing and future colliders have made lattice design and error tolerance and correction critical to achieving performance goals. The current state of the Tevatron collider is an example, with a strong skew quadrupole error present in the operational lattice. This work studies the high-order performance of the Tevatron and the strong nonlinear behavior introduced when a significant skew quadrupole error is combined with conventional sextupole correction, a behavior still clearly evident after optimal tuning of available skew quadrupole circuits. An optimization study is performed using different skew quadrupole families, and, importantly, local and global correction of the linear skew terms in maps generated by the code COSY INFINITY [M. Berz, COSY INFINITY version 8.1 user's guide and reference manual, Department of Physics and Astronomy MSUHEP-20704, Michigan State University (2002). URL http://cosy.pa.msu.edu/cosymanu/index.html]. Two correction schemes with one family locally correcting each arc and eight independent correctors in the straight sections for global correction are proposed and shown to dramatically improve linearity and performance of the baseline Tevatron lattice.
Published by Elsevier B.V.

*PACS:* 02.60.Pn; 29.20.−c; 29.27.−a; 29.85.+c

*Keywords:* Tevatron; Optimization; Dipole error; Skew quadrupole correction; COSY INFINITY; Tracking; Transfer map

## 1. Introduction

Increasing the luminosity reach of existing and future colliders demand considered and precise optical design and predictability in operation. Driven by nonlinear fields, "high-order" beam dynamics are generally difficult to control, calculate, and can severely limit a machine's region of stable operation. An approximately linear lattice is desirable for operational simplicity and understanding; it also generally exhibits more robust, broader-range performance. Nonlinear sources arising from field and alignment errors, and the required correction elements are unavoidable. Successful management of nonlinear sources, however, depends on the linear lattice. Attributes of the linear lattice and relative locations of sources generate interference, constructive or destructive, between the nonlinear terms depending on their periodicity. In a highly effective linear lattice design, the strongest nonlinear amplitudes can be mitigated passively by intelligently exploiting periodicity, phase advance and optimal placement of nonlinear correctors. Such a lattice enhances precision and predictability in the machine optics.

Passive cancellation, however, is generally not sufficient to address certain systematics or widespread field errors; active correction in the form of added corrector elements is usually required. The overall lattice approach must be evaluated not only by its tolerance of errors, nonlinearities and natural aberrations, but also by its potential for active correction. Such correction may be "global" in the sense

*Corresponding author.
*E-mail addresses:* snopok@pa.msu.edu (P. Snopok), cjj@fnal.gov (C. Johnstone), berz@msu.edu (M. Berz), dovs@compmath.spbu.ru (D.A. Ovsyannikov), ovs@compmath.spbu.ru (A.D. Ovsyannikov).

0168-9002/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.083

that an error or aberration is corrected over one-turn optics. Global correction is not always adequate to maintain sensitive collider optics. Immediate—or "local" correction—of source terms, particularly if such terms propagate through the delicate optics of the interaction regions, may be an additional requirement for stability and linearity. A case addressed in this work is the current state of the Tevatron collider, where a strong, systematic, skew quadrupole error is present in the operational lattice as a result of a coil shift in the superconducting arc dipoles.

With increasing demands for luminosity, optimal performance must be extracted from the existing Tevatron optics. Local correction of errors and other strong sources of aberrations is necessary to achieve the desired optical performance and luminosity. We have, therefore, initiated a high-order dynamical study of the Tevatron to assess the performance, functionality and potential of the baseline lattice. For this study, we are concerned only with the baseline Tevatron lattice which we consider to be simply the linear lattice (quadrupoles and dipoles) combined with the strongest low-order nonlinearities. The strongest sources of nonlinearities are first, the chromatic correction and feed-down sextupoles and, second, strong sextupole and skew quadrupole error fields found in the arc dipoles. Skew quadrupole errors are very important because they change the linear lattice. This work describes the nonlinear performance of the Tevatron lattice with emphasis on the coupled and increased nonlinear behavior introduced by the significant skew quadrupole error in combination with conventional sextupole correction, a behavior still clearly evident after optimal tuning of available skew quadrupole circuits. An optimization study is then performed using available skew quadrupole circuits, and, importantly, local and global correction of the linear skew terms in maps generated by the code COSY INFINITY (COSY) [1].

Two correction schemes with one skew quadrupole family locally correcting each arc and eight independent correctors in the straight sections for global correction proved themselves to give the best results and dramatically improve the linear performance of the baseline Tevatron lattice. In both schemes, the source of the skew error is corrected in such a way as to allow the single-family circuit available to complete the correction and decoupling of the base lattice, which is technically achieved by fixing the coil shift in part of the Tevatron dipoles.

## 2. Tevatron lattice description

The Tevatron lattice [2] is comprised of 6 arcs and 6 straight sections with interaction regions CDF and D0, occupying two of the straights. The lattice has a simple periodicity of one, but with no reflective symmetry. Even the arcs are not perfectly regular, but remain adequately described by a FODO cell with 72° of phase advance in each plane. The global tunes are 20.585 and 20.575, in the horizontal and vertical, respectively, and clearly not split by an integer as is common in current lattice design.

## 3. Lattice data and method

First a high-order Taylor series one-turn map of the Tevatron is generated using the differential algebra code COSY with the baseline lattice described above. The different baseline components of the lattice: the chromatic correction and feed-down families of sextupoles, the skew quadrupole correctors, the strong skew and sextupole systematic errors are implemented in such a way that they could be turned on and off to study individual and correlated effects on performance and effectively trouble-shoot the lattice. Initial and updated Tevatron lattice data plus component strengths were obtained from the input deck for the code OptiM [3]. An automated converter has been written to transcribe the OptiM input format to the language of COSY [4]. The converter itself is written in PHP [5], so that it is straightforward to perform online updates or entire conversions of lattices from OptiM to COSY. For now a conversion exists for the following sets of elements: dipoles, dipole kicks, pure and skew quadrupoles, quadrupole kicks, pure and skew sextupoles, sextupole kicks, solenoids and electric separators. The generated code is ready-to-use by COSY.

## 4. Checks of the linear lattice

Linear maps without the skew quadrupole correctors and errors and linear parameters such as tunes have been verified and cross-checked with both OptiM and an independent COSY implementation [6]. The checks on the proper conversion of the lattice are as follows. First, beta functions [7] for closed orbit were compared with OptiM. Quantitative comparison showed less than the percent level difference. Slight differences are due to a more realistic implementation of the detector solenoids in COSY.

## 5. Simulation details and tracking results

Typically an 11th to 15th order map was required for complete convergence of nonlinear effects, but lower orders (7th, for example) provided a quicker check on the direction of results and optimizations. Particles were launched at the CDF interaction point in steps of one sigma (for normalized emittance of $10\pi$ at injection $\sigma = 1.2 \times 10^{-4}$ m). Particles were tracked in COSY by applying the map repetitively for typically 10,000 turns. Only the injection optics was being studied. It is important to note that the study is not, per se, a dynamic aperture one for which particles are launched along phase space vectors scaled to the linear injection ellipse and the transmitted transverse phase is mapped. Dynamic aperture studies are not always informative as to beam dynamics. In a predominately linear lattice, tracking along a single vector in one plane of phase space and then the other is sufficient to trace out the matched ellipse. Particles can be simply launched along the $x$- or $y$-axis, for example. We are

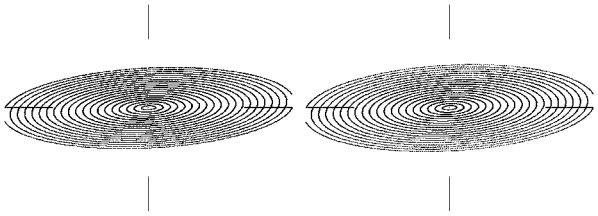


Fig. 1. $x$- and $y$-plane phase portraits, only dipoles and pure quadrupoles active, particles launched along $x$- and $y$-axis respectively.

Fig. 2. $x$- and $y$-plane phase portraits before the optimization, particles launched along $x$- and $y$-axis, respectively. The phase portraits include all sextupole and skew quadrupole fields (correctors plus errors) in addition to quadrupoles and dipoles where 15% of the skew quadrupole errors have been removed in specific dipoles.

looking for degradation of linear motion as evidenced by dissolution or distortion of the linear invariant ellipses. Since the current study is directed at optimizing linear performance, this is the approach used for tracking and the criterion for improvement.

The tracking results presented in this and subsequent sections are obtained for 10,000 turns with points plotted every 10th turn, and the scales are $2.4 \times 10^{-3}$ m for $x, y$ axis and $4.0 \times 10^{-3}$ for $a, b$ axis ($a = p_x/p_0$, $b = p_y/p_0$). Tracking is performed with a symplectification algorithm written by Bela Erdélyi [8–11] and calculation order 7. All the particles are launched either along $x$- or $y$-axis, which is explicitly mentioned in each figure caption. To start comparing the impacts of different sets of nonlinear elements, in Fig. 1 phase portraits for linear motion are shown. This includes only pure dipoles and pure quadrupoles.

According to the status of the Tevatron before August 2004, With 15% of the skew quadrupole errors removed in selected dipoles, the otherwise unchanged lattice shows significantly reduced stability (Fig. 2). The regular structures disappear and most of the particles can be considered lost in just 10,000 turns. These phase portraits can be considered a starting point of the study of different schemes of the skew quadrupole correction.

## 6. Skew quadrupole circuits optimization proposals

Each of the Tevatron arcs has 15 FODO cells with skew quadrupole correctors in every odd-numbered cell, which means one corrector every two FODO cells. The skew correctors are placed next to horizontally focusing quadrupoles only.

One family of skew quadrupole correctors is not sufficient to correct all the skew quadrupole errors in dipoles along the ring. But during the next shutdown approximately 50% of the coil shift errors in dipoles can be fixed and in this case one circuit of skew quadrupole correctors is capable of removing the coupling in the arcs. The problem here is to discover both the optimal dipole pattern for error correction and the new strength for the skew quadrupole correctors.

An optimization where all the strengths of the skew correctors are different is not practical. All the correctors in each arc have the same power supply, so it is more realistic to use one strength for all the correctors arc-wise or even ring-wise.

The optimization process itself consists of two steps. First, the optimization of each arc is performed using skew quadrupole corrector strengths as control parameters. This optimization would be close to optimal if no skew quadrupole components existed in the straight sections of the Tevatron, but there are skew errors and correctors for the interaction regions. Because of these components and the residual skew terms from the arcs since the arcs are not perfectly regular, the skew terms of the one-turn transfer map has nonzero skew quadrupole terms which require correction also. To remove this smaller, final stage of coupling requires a second step to the optimization. In four of the six straight sections there exist eight skew quadrupole correctors and the strengths of these correctors were used to finish the skew-quadrupole term cancellation in the one-turn map.

Two optimization schemes were considered which differed in the dipole pattern used for correcting the skew quadrupole error. The first scheme, proposed by us, attempts the elimination of the skew quadrupole error source predominately in the vertical plane by fixing the two dipoles flanking each vertically focusing arc quadrupole. With one degree of freedom corrected, the remaining predominately horizontal sources can be corrected with the existing single family of skew quadrupole correctors. The layout for scheme I is shown in Fig. 3. One arc of the Tevatron is shown and the others are similar. The dipoles

Sector E scheme I

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| FODO 1 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 2 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 3 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 4 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 5 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 6 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 7 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 8 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 9 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 10 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 11 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 12 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 13 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 14 | D*2 | FQ | | D*2 | D*2 FIX | DQ | D*2 FIX |
| FODO 15 | D*2 | FQ | SQC | D*2 | D*2 FIX | DQ | D*2 FIX |

Fig. 3. Correction scheme I, the skew quadrupole error is removed in dipoles in each cell surrounding the defocusing quadrupole.

with skew quadrupole errors are marked with the "D * 2" symbol, fixed dipoles—with the "D * 2 FIX" symbol. The vertically focusing main quadrupole is marked with "DQ" as it is defocusing in the horizontal plane.

The results of the two-stage optimization are shown in Fig. 4. Phase portraits show much more stability, though this scheme is not perfect due to the deviation from completely periodicity in the arcs. However, this scheme was also found to be the most robust to any lattice alterations than the second one which is described next.

The second approach proposed by Michael Syphers [12] is to correct the skew quadrupole errors in each FODO cell missing a skew quadrupole corrector. This scheme was further improved upon by removing specific correctors from the single family to provide more consistent correction in each arc as a function of fractional phase advance. The underlying idea is to correct the error locally at the source and remove it from cells without local correction. This scheme gives improved performance, particularly using one corrector strength across all arcs. Phase portraits for scheme II are shown in Fig. 5. The layout for the scheme is given in Fig. 6. Skew quadrupole errors reside in odd cells with skew quadrupole correctors marked with "SQC". Removing part of the correctors improved performance further and these removed correctors are marked with "SQC RMV".

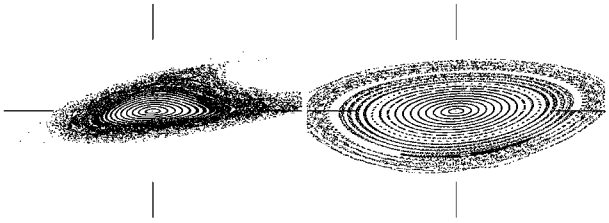


Fig. 4. $x$- and $y$-plane phase portraits after the optimization with 50% skew quadrupole errors in dipoles, errors fixed around each defocusing quadrupole, particles launched along $x$- and $y$-axis, respectively.

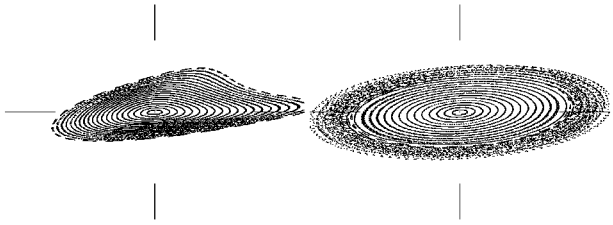


Fig. 5. $x$- and $y$-plane phase portrait after the optimization with the skew quadrupole errors removed in the two dipoles flanking each horizontally defocusing quadrupole, particles launched along $x$- and $y$-axis, respectively.

Sector E scheme II

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| FODO 1 | D*2 | FQ | SQC RMV | D*2 | D*2 | DQ | D*2 |
| FODO 2 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 3 | D*2 | FQ | SQC RMV | D*2 | D*2 | DQ | D*2 |
| FODO 4 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 5 | D*2 | FQ | SQC | D*2 | D*2 | DQ | D*2 |
| FODO 6 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 7 | D*2 | FQ | SQC | D*2 | D*2 | DQ | D*2 |
| FODO 8 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 9 | D*2 | FQ | SQC | D*2 | D*2 | DQ | D*2 |
| FODO 10 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 11 | D*2 | FQ | SQC | D*2 | D*2 | DQ | D*2 |
| FODO 12 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 13 | D*2 | FQ | SQC | D*2 | D*2 | DQ | D*2 |
| FODO 14 | D*2 FIX | FQ | | D*2 FIX | D*2 FIX | DQ | D*2 FIX |
| FODO 15 | D*2 | FQ | SQC RMV | D*2 | D*2 | DQ | D*2 |

Fig. 6. Correction scheme II, skew quadrupole errors in dipoles fixed in each even cell.

## 7. Conclusions

With all the skew quadrupole errors in dipoles, one circuit of skew quadrupole correctors is not sufficient or requires moving correctors, which is disruptive and expensive. With half of the errors fixed in the dipoles, one set is sufficient to achieve far reaching correction, and with all correctors in the ring set to equal strengths. One of these two schemes is to be implemented during the Tevatron shutdown in August 2004.

The code used, COSY INFINITY allowed tracking to be performed for a very large number of revolutions in a very short time. For example, the whole cycle of calculations, including one-turn transfer map calculation, skew quadrupole correctors circuits optimization and 10,000 revolution high-order (7th order) tracking, runs for 5–10 min on Intel Celeron 1.5 GHz with 256 Mb of RAM. For 11th order, the same task requires 8–12 h. However, 7th order was sufficient to judge improvements and optimization in preparation for the longer, final simulations.

With the skew quadrupole corrector circuits optimized and the dynamics uncoupled, the next study step might be to address the sextupole families and optimization of conflicts between the chromaticity and feed-down sextupoles which was also observed in this study.

Further study should also address possible effects of fringe fields. At this point, such simulations are not easily possible since detailed shapes of fringe field fall-off are not readily available.

Also, in the future normal form methods [13] and methods of control theory [14] might be employed for solving such problems. Normal form methods are effective for long-term stability study, and variational methods along with automatic differentiation algorithms implemented in COSY make it easy to study the behavior of the sets of trajectories of the perturbed motion with respect to the reference particle dynamics. This perturbed motion is then minimized using control theory techniques.

## References

[1] M. Berz, COSY INFINITY version 8.1 user's guide and reference manual, Department of Physics and Astronomy, MSUHEP-20704, Michigan State University (2002). URL http://cosy.pa.msu.edu/cosymanu/index.html.

[2] M.A. Martens, Tevatron Lattice Page. URL http://www-ap.fnal.gov/~martens/tev_lattice/tev_lattice.html.

[3] V. Lebedev, private communication.

[4] P. Snopok, A converter program for Tevatron lattices from OptiM to COSY INFINITY, Technical Report MSUHEP-40909, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2004.
[5] PHP: Hypertext Preprocessor. URL http://www.php.net.
[6] B. Erdélyi, private communication.
[7] S. Turner (Ed.), CERN Accelerator School, Fifth General Accelerator Physics Course, vol. 1, CERN, Geneva, 1994.
[8] B. Erdélyi, M. Berz, Phys. Rev. Lett. 87 (11) (2001) 114302.
[9] B. Erdélyi, M. Berz, Int. J. Pure Appl. Math. 11 (3) (2004) 241, available at http://bt.pa.msu.edu/pub.
[10] B. Erdélyi, Symplectic approximation of Hamiltonian flows and accurate simulation of fringe field effects, Ph.D. Thesis, Michigan State University, East Lansing, Michigan, USA, 2001.
[11] M.L. Shashikant, M. Berz, B. Erdélyi, IOP CP 175 (2004) 299.
[12] M. Syphers, private communication.
[13] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999, also available at http://bt.pa.msu.edu/pub.
[14] D.A. Ovsyannikov, N.V. Egorov, Mathematical Modelling of Electron and Ion Beams Forming Systems, St.-Petersburg University Press, 1998.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 147–149
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Emittance measurements of low-energy beam line at KVI

Dragan Toprek[a,*], Iwo Formanoy[b]

[a]VINCA Institute of Nuclear Sciences, Laboratory for Nuclear and Plasma Physics, P.O. Box 522, 11001 Belgrade, Serbia and Montenegro
[b]Kernfysisch Versneller Instituut (KVI), Zernikelaan 25, 9747 AA Groningen, The Netherlands

Available online 28 November 2005

**Abstract**

In this paper is represented the results of beam profile measurements of $^3He^+$ beam delivered from ECR ion source at KVI. The beam emittance is estimated by varying quadrupole method. The estimated values for the beam emittance at the different profile grid locations along the transport beam line shows that beam emittance in the both, horizontal and vertical planes is not conserved. That means there is a strong-coupling effect and/or the beam is losing through the transport line.


*PACS:* 07.77.Ka; 29.27.Ac; 29.27.Fh; 41.75.Ak

*Keywords:* ECR ion source; Emittance

## 1. Introduction

The beam current of $^3He^+$ beam at the beginning of the transport injection line is about 20 μA but at its ending (after matching section) is only 4 μA. To improve the transmission of the beam through the injection beam line the beam emittance should be known. Beam profile is measured at the three different positions along the beam transport line and in the both planes, horizontal and vertical, by using profile grids. At each position of the profile grids the beam profile is measured as a function of the strength of an upstream quadrupole. First, horizontal/vertical profile grid (marked by letters HH1/HV1 in Fig. 1) is located 7.17/7.27 m downstream from the ion source, second profile grid (HH2/HV2) is located 10.20/10.30 m and the third one (HH4/HV4) is located 15.40/15.41 m downstream from the ion source.

Data from the profile grids are digitized, acquired and processed by Control System. For the each $\mu$ setting of the strength of an upstream quadrupole the profile data are fitted to the Gauss distribution [1,2].

## 2. Results

The characteristics of the $^3He^+$ beam are: kinetic energy; $T_k = 8.33\,keV/n$, magnetic rigidity; $B\rho = 0.0394\,Tm$, electrostatic rigidity; $\chi_E = 0.05\,MV$.

The length of each electrostatic quadrupole is 12 cm and half of their aperture is 5 cm.

For each settings of the corresponding quadrupole it is measured the beam profile in the horizontal or vertical planes. Then each beam profile is fitted as the Gauss distribution. The estimated values for the emittance and its corresponding error at different profile grid position along the transport beam line are presented in Table 1. We consider the case when the half beam profile is equal $2.35\sigma_{ST}$, where $\sigma_{ST}$ is the standard deviations of the Gauss distribution. Due to the space limited we will here presented dependence, on the quadrupole strength, of the $^3He^+$ beam profile at the profile grid HH2 and HV2 (see Figs. 2 and 3).

In the case of the horizontal plane it is noticed slightly increasing of the beam emittance along the transport line. In the case of the vertical plane there is a suddenly take-off at the HV4 profile grid position. We can see from the Table 1 that the beam emittance in the vertical plane is slightly bigger than the beam emittance in the horizontal plane $\varepsilon_y > \varepsilon_x$.

*Corresponding author.
*E-mail address:* dragan.toprek@accel.de (D. Toprek).

0168-9002/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.032

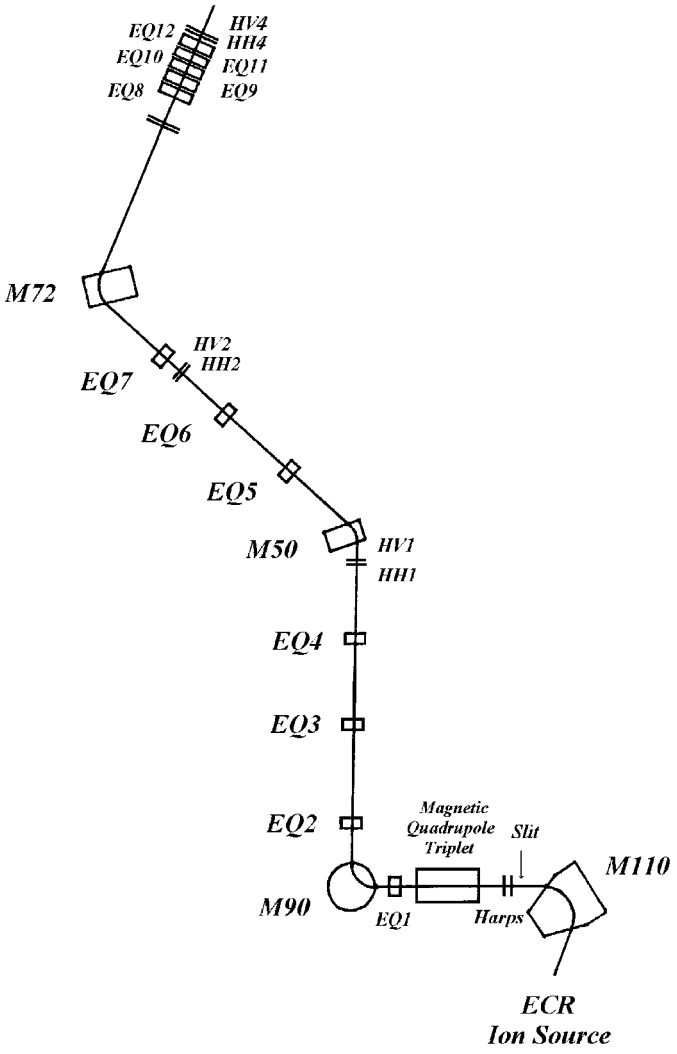



Fig. 1. Layout of the low-energy beam line at KVI.

Table 1

| $^3He^+$ | $V_{ex} = 25$ kV | |
|---|---|---|
| Profile's grid position | $\varepsilon_x$ [$\pi$ mm mrad] | $\varepsilon_y$ [$\pi$ mm mrad] |
| HH1 & HV1 | $28 \pm 2$ | $42 \pm 5$ |
| HH2 & HV2 | $37 \pm 1$ | $54 \pm 1$ |
| HH4 & HV4 | $52 \pm 1$ | $190 \pm 10$ |

The estimated values for the beam emittances at the different profile grid locations along the transport beam line shows that the beam emittance in the both planes is not conserved. The fact that the beam current decreases through the transport beam line means that the beam emittance is not conserved not only due to of a strong coupling effect but also due to the losing of the beam through the transport line.

The expression for the emittance's error is given by

$$\Delta\varepsilon_y = \frac{1}{2\varepsilon_y}\sqrt{\Delta\sigma_{33}^2\sigma_{44}^2 + \Delta\sigma_{44}^2\sigma_{33}^2 + 4\Delta\sigma_{34}^2\sigma_{34}^2 + \delta_y} \quad (1)$$

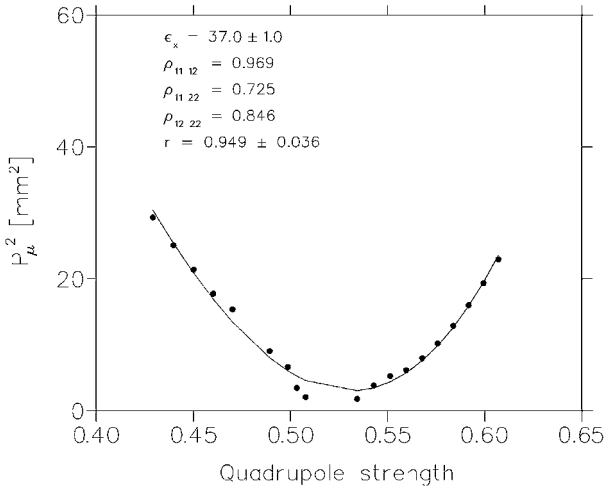



Fig. 2. Square of the horizontal beam size versus strength of the quadrupole 10.20 m upstream from the measurement point. The points are experimental data and the full line is the best fit curvature. $r$ and $\rho_{ij\ kl}$ are the corresponding correlation coefficients.

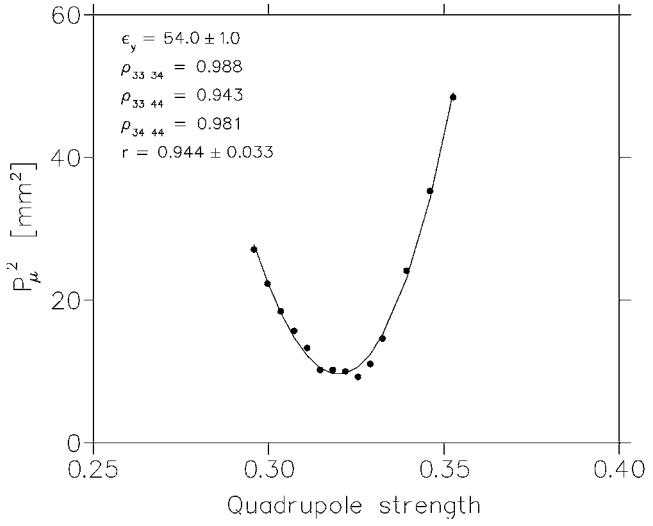



Fig. 3. Square of the vertical beam size versus strength of the quadrupole 10.30 m upstream from the measurement point. The points are experimental data and the full line is the best fit curvature. $r$ and $\rho_{ij\ kl}$ are the corresponding correlation coefficients.

where $\delta_y$ is the coefficient of interrelation between $\sigma_{33}$, $\sigma_{34}$ and $\sigma_{44}$ given by

$$\begin{aligned}\delta_y = &-4\sigma_{33}\sigma_{34}\rho_{3444}\Delta\sigma_{34}\Delta\sigma_{44} \\ &+2\sigma_{33}\sigma_{44}\rho_{3344}\Delta\sigma_{33}\Delta\sigma_{44} \\ &-4\sigma_{34}\sigma_{44}\rho_{3334}\Delta\sigma_{33}\Delta\sigma_{34}. \end{aligned} \quad (2)$$

In the upper equations $\rho_{ij\ kl}$ are the corresponding correlation coefficients between the elements of the beam matrix $\sigma_{ij}$ and $\sigma_{kl}$.

Similar equations as Eqs. (1) and (2) are valued for $\varepsilon_x$ and $\Delta\varepsilon_x$.

## 3. Conclusion

By beam profile measurements and prediction of the beam emittance at the different position along the transport beam line it is shown that the reason for the beam current dropping can be the over-focusing in the vertical plane produced by bending magnet M72. To eliminated this suspicious more diagnostic elements before and after M72 magnet should be placed.

## References

[1] D. Toprek, AC Code PROFILE for Fitting the Beam Profile's Data to Gaussian, private calculations.
[2] PHYSICA, TRIUMF Computer Document, TRI-CD-93-01, V1.3, Vancouver, Canada.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 150–153
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Algebraic modeling and parallel computing

S.N. Andrianov*, N.S. Edamenko, A.A. Dyatlov

*Saint-Petersburg State University, Saint-Petersburg, 198504, Russia*

Available online 28 November 2005

**Abstract**

In this paper, we present an algebraic approach for beam dynamics simulation in linear and circular accelerators. A two- and three-dimensional distribution function approach is employed within the matrix formalism for Lie algebraic methods and computer algebra codes. Implementing software design provides the code rather good maintainability, reusability, and extensibility. This approach is an alternative to well-known Particle-in-Cell approach. But it can be easily applied to the PIC-approach and thus to make better their effectiveness. As a required demand, the code also include symplectic integration methods (based on a correcting procedure for aberration matrices) up to an approximation order.


*PACS:* 41.85.−p; 02.20 Sv; 02.60.Pn; 02.70.Wz

*Keywords:* Beam optics; Lie methods; Numerical optimization; Symbolic computation

## 1. Introduction

Increasing interest in high intensity beams leads us to necessity more accurate selection of a mathematical background for created programming tools. This remark becomes actual especially for studying high intensity beams. The most popular method for similar problems is particle-in-cell (PIC) simulation method. But, as it is known, that parallelization procedure for the forces acting on particles due to the inter-particle Coulomb field is very intricate and complex (see, for example, Ref. [1]). In this approach the corresponding computing process is based on the well known architecture Single Instruction Multiple Date (SIMD). Indeed current phase data play a role of multiple date and methods of evolution equation solution (including evaluation of space charge forces) play a role of a single instruction. The main load on computational resources is defined by larger number of numerical particles and used grid resolution for realization of the PIC-scheme. It is obviously that the number of numerical particles is restricted by grid resolution, which for its part is restricted by number of downloaded processors. The main problems of similar approach concern first of all with used methods for numerical integration of motion equations. Besides the pointwise presentation of a beam phase portrait leads to difficulties of parallelization procedure realization.

In this report, we suggest an approach for beam evolution simulation based on Lie algebraic tools on the one hand and model distribution function for beam phase portrait description on the other hand. This approach allows us to use powerful mathematical tools—algebraic and group-theoretical methods. In particular, this permits to build any kind of special construction for the beam propagator $\mathscr{M}$—beam evolution operator. In the second, we are permitted to apply the algebra and group representation theories, first of all for Lie algebras and Lie groups. In particular this leads to a possibility of usage of the matrix formalism for Lie algebraic tools [2]. Here we note that the matrix formalism, based on usual matrix algebra, should be supplemented by Kronecker operations (product, sum and others). Just this allows us to realize parallel and distributed computing without difficulty.

## 2. Mathematical background

Let us describe some main features of the mathematical tools underlying in the suggested approach. This tools is

*Corresponding author.
*E-mail address:* sandrianov@yandex.ru (S.N. Andrianov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.036

based on the priority of map idea, which originate (at least in beam physics) in works of A.J. Dragt [3].

### 2.1. The lie algebraic tools for a beam propagator

Let

$$\frac{d\mathbf{X}}{ds} = \mathbf{F}(\mathbf{X}, \mathbf{U}, s) \tag{1}$$

be a motion equation for beam particles in external and space–charge fields, $s$—a length measured along some reference orbit. Here the vector $\mathbf{U}(s)$ describes control functions corresponding to guiding and focusing fields. According to the map technique any solution of Eq. (1) can be written in the form

$$\mathbf{X}(s) = \mathscr{M}_{\mathbf{F}}(\mathbf{U}; s|s_0) \circ \mathbf{X}_0, \quad \mathbf{X}_0 \in \mathfrak{M}_0$$

where $\mathscr{M} = \mathscr{M}_{\mathbf{F}}(\mathbf{U}; s|s_0)$ is so-called Lie map (transformation), associated with the vector function $\mathbf{F}$ from Eq. (1)—a beam propagator. This map satisfies the following operator equation:

$$\frac{d\mathscr{M}_{\mathbf{F}}(\mathbf{U}; s|s_0)}{ds} = \mathscr{L}_{\mathbf{F}}(s) \circ \mathscr{M}_{\mathbf{F}}(\mathbf{U}; s|s_0)$$

$$\mathscr{M}_{\mathbf{F}}(\mathbf{U}; s_0|s_0) = \mathscr{I}d$$

where $\mathscr{L}_{\mathbf{F}}(\mathbf{U}; s)$ is a Lie operator. In general case this map can be written in the form of so-called time chronological exponent operator (according to the chronological Volterrá series)

$$\mathscr{M}_{\mathbf{F}}(\mathbf{U}; s|s_0) = \mathrm{T}\exp\left(\int_{e_0}^{s} \mathscr{L}_{\mathbf{F}(\mathbf{U};\tau)}\, d\tau\right).$$

For non-autonomous systems the so-called Magnus's representation [4] can be used. This approach allows to pass to a routine exponential operator.

$$\mathscr{M}_{\mathbf{F}}(\mathbf{U}; s|s_0) = \exp\left(\int_{e_0}^{s} \mathscr{L}_{\mathbf{G}(\mathbf{U};s|\tau)}\, d\tau\right)$$

where $\mathbf{G}$ is a new vector function, generated by the initial function $\mathbf{F}$. The connection between these function can be evaluated using the techniques presented in Refs. [4,2]. In particular, the expansion of the function $\mathbf{F}(\mathbf{X}, s)$

$$\mathbf{F}(\mathbf{X}, s) = \sum_{k=0}^{\infty} \mathbb{P}^{1k}(s)\mathbf{X}^{[k]}$$

generates an expansion of the function

$$\mathbf{G}(\mathbf{X}; s|s_0) = \sum_{k=0}^{\infty} \mathbb{G}_k(s|s_0)\mathbf{X}^{[k]},$$

which appears in the Magnus's representation and one can write

$$\mathscr{M}(s|s_0) = \exp\left\{\sum_{k=0}^{\infty} \mathscr{L}_{\mathbf{G}_k(\mathbf{X},\mathbf{U};s|s_0)}\right\} \tag{2}$$

where $\mathbf{G}_k(\mathbf{X}, \mathbf{U}; s|s_0) = \mathbb{G}_k(\mathbf{U}; s|s_0)\mathbf{X}^{[k]}$. The similar to the Dragt–Finn factorization for the Lie transformations allows to rewrite the exponential operator Eq. (2) as an infinite product of exponential operators of Lie operators

$$\begin{aligned}\mathscr{M} &= \ldots \circ \exp\{\mathscr{L}_{\mathbf{H}_2}\} \circ \exp\{\mathscr{L}_{\mathbf{H}_1}\} \\ &= \exp\{\mathscr{L}_{\mathbf{V}_1}\} \circ \exp\{\mathscr{L}_{\mathbf{V}_2}\} \circ \ldots\end{aligned}$$

where $\mathbf{H}_k = \mathbb{H}_k\mathbf{X}^{[k]}$, $\mathbf{V}_k = \mathbb{V}_k\mathbf{X}^{[k]}$ are homogeneous polynomials of $k$th order. The matrices $\mathbb{H}_k$ or $\mathbf{V}_k$ can be calculated with the help of the continuous analogue of the CBH and Zassenhauss formulae and by using the Kronecker product and Kronecker sum technique for matrices [2]. Moreover, using the matrix representation for the Lie operators one can write a matrix representation for the Lie map generated by these Lie maps (transformations)

$$\begin{aligned}\mathscr{M} \circ \mathbf{X} &= \mathbb{M}\ \mathbf{X}^{\infty} \\ &= (\mathbb{M}^{10}\mathbb{M}^{11}\mathbb{M}^{12}\ldots\mathbb{M}^{1k}\ldots)\mathbf{X}^{\infty} \\ &= \sum_{k=0}^{\infty} \mathbb{M}^{1k}\mathbf{X}^{[k]}\end{aligned}$$

where $\mathbf{X}^{\infty} = (1\mathbf{X}\mathbf{X}^{[2]}\ldots\mathbf{X}^{[k]}\ldots)^*$, and matrices $\mathbb{M}^{1k}$ (*aberration matrices*) can be calculated according to a recurrent sequence of formulae of the following types:

$$\begin{aligned}\mathscr{M}_k \circ \mathbf{X}^{[l]} &= \exp\{\mathscr{L}_{\mathbf{G}_k}\} \circ \mathbf{X}^{[l]} \\ &= \mathbf{X}^{[l]} + \sum_{m=1}^{\infty} \frac{1}{m!} \prod_{j=1}^{m} \mathbb{G}_m^{\oplus((j-1)(k-1)+l)}\mathbf{X}^{[m(k-1)+l]}.\end{aligned}$$

Here $\mathbb{G}^{\oplus l} = \mathbb{G}^{\oplus(l-1)} \otimes \mathbb{E} + \mathbb{E}^{[l-1]} \otimes \mathbb{G}$ denotes the Kronecker sum of the $l$th order.

It is well known that Lie algebraic method preserves qualitative symplectic property residing to any Hamiltonian system. But in practice we have to truncate series presenting Lie maps. This procedure leads to symplectic property loss, and as a consequence to loss of occurrence one or another effect. The used matrix formalism allows to correct aberration matrices $\mathbb{M}^{1k}$, $\forall k \leqslant 2$ [5]. This procedure adds up to some linear algebraic equations with a simple structure. As a result we obtain a truncated beam propagator, which guarantees symplecticity up to some desired approximation order $N$ (desired precision). Moreover, corresponding solutions can be easy solved using computer algebra codes, such as Maple or Mathematica.

### 2.2. Particle beam distribution function

For the distribution function describing the beam particle at some moment one can use different kinds of presentation types. One of them is based on a symbolic form for distribution function (phenomenological function type), generated from experimental data or some another information. As an example we can use the following pseudo-normal distribution (in the case of a symmetric character of particle distribution):

$$f_0(\mathbf{X}) = Q_{2m}(\mathbf{X})\mathrm{e}^{-P_{2n}(\mathbf{X})} \tag{3}$$

where $Q_{2m}(\mathbf{X})$, $P_{2n}(\mathbf{X})$ are polynomials of the $2m$th and $2n$th order correspondingly. Coefficients of these poly-

nomials are determined from experimental data fully or partly. In the last case the remainder coefficients can be determined from other information or used as control parameters.

The other function type uses two types of following Taylor series expansions. Here we have two forms of such presentation. In general case one can write

$$f_0(\mathbf{X}) = f_0^0 + \sum_{k=1}^{\infty} \mathbf{F}_{0k}^{*} \mathbf{X}^{[k]}, \tag{4}$$

with $f_0^0$ and $\mathbf{F}_{0\,k}$ as a scalar coefficient and vector coefficients. In the case of elliptical symmetry of the beam phase portrait has the form

$$f_0(\mathbf{X}) = f_0^0 + \sum_{k=1}^{\infty} a_k \varkappa_0^{2k}(\mathbf{X}) \tag{5}$$

where $\varkappa^2(\mathbf{X})$ is a quadratic form, describing the phase ellipsoid (in 4D- or 6D-phase space). According to the algebraic Lie methods (see, for example, Ref. [3]) one can write for the current distribution function

$$f(\mathbf{X}, s) = f_0(\mathscr{M}^{-1}(\mathbf{F}; s|s_0) \circ \mathbf{X}). \tag{6}$$

For the inverse map $\mathscr{M}^{-1}$: $\mathbf{X} \to \mathbf{X}_0 = \mathscr{M}^{-1} \circ \mathbf{X}$ one can compute the corresponding block-matrices $\mathbb{T}^{ik}$ using the generalized Gauss's algorithm. Here it is necessary to note that only $\mathbb{M}^{11}$ should be inverted. The rest block matrices are evaluated using usual matrix operations: multiplication and summation. So one can write the following matrix presentation for Eq. (6):

$$f(\mathbf{X}, s) = f_0\left(\sum_{k=0}^{\infty} \mathbb{T}^{1k}(\mathbf{F}; s|s_0)\mathbf{X}^{[k]}\right).$$

So, similar formulae allow us to evaluate current distribution function for any moment $s$. In the case, when the space charge forces may be neglected, this approach provides effective tools for practically all problems of beam physics.

### 2.3. The case of space–charge dominated beam

The physical system for beam dynamics investigations consists of the beam (the subject of control) and the control system (the object of control), which in turn consists of accelerating and focusing elements (in our terms—control elements). So all types of forces acting on particles can be separated in externally applied fields and the inter-particle Coulomb field.

In the case, when the second type of forces have an influence upon the beam dynamics the described approach should be added by some procedures, see, for example, Refs. [2,6]. In this paper, we suggest to combine the matrix formalism for Lie algebraic methods with the well-proved PIC tools, which has the form of 2D- and 3D-dimensional versions.

Usually the particles dynamics is written with the help the Poisson-Vlasov system of equations. Let us remember that in our case the longitudinal independent variable $s$ is used rather than the time $t$. The Vlasov equation for the distribution function $f(\mathbf{X}, s) = f(\mathbf{Q}, \mathbf{P}, s)$ can be written as

$$\left\{\frac{\partial}{\partial s} + \mathbf{V}\frac{\partial}{\partial \mathbf{Q}} + \mathbf{F}\frac{\partial}{\partial \mathbf{P}}\right\} f(\mathbf{Q}, \mathbf{P}, s) = 0$$

where $\mathbf{X} = \{\mathbf{Q}, \mathbf{P}\}^*$ and $\mathbf{F} = \mathbf{F}(\mathbf{E}, \mathbf{B}, s)$ is the Lorentz force, written in the chosen coordinates system [2]. For space charge fields we use the Poisson equation, written in the form

$$\nabla\phi(\mathbf{Q}) = -\frac{\rho(\mathbf{Q})}{\varepsilon},$$

$$\rho(\mathbf{Q}) = \int_{\mathfrak{M}(s)} f(\mathbf{Q}, \mathbf{P})\, \mathrm{d}\mathbf{P} \tag{7}$$

where $\mathfrak{M} = \mathfrak{M}(s)$ is a current manifold occupied beam particles at the moment $s$. We suppose that for a current distribution function $f(\mathbf{X}, s)$ we can apply one of the presentations Eqs. (3)–(5) (in the case of Eq. (3) we also use series presentation like Eqs. (4) or (5) up to some approximation order $N$). This presentation allows us to find solution of Eq. (7) also in the form of series on $\mathbf{Q}$ (see, for example, Ref. [6]). We suppose that the beam size is much smaller than the inside wall radius of the machine. This permits to treat the beam as an isolated system. In such a case, the necessary concatenation procedure leads us to an algebraic equations for determination of the above mentioned indeterminate coefficients of corresponding truncated series.

Similar approach leads us to a polynomial presentation in a symbolic mode for self-fields with some indeterminate coefficients, which have to determine using boundary conditions for every cell of our particle grid. A two-dimensional schema of beam particle decomposition in the transverse configuration subspace is shown in Fig. 1. The solid lines define the domain boundaries. Each processor contains one rectangular block domain which contain equal number of particles. This domain separation allows us providing each processor with an equal computational load. If a particle drops out of the local current domain, the necessary information is sent to the corresponding processor where the particle is located. The transfer of this information is managed using corresponding MPI commands.

In this paper, we calculate the self-beam-field in a different way, using numerical procedure for a variant for the PIC method. But such computing has a special form meant for the matrix presentation of the beam propagators. In other words the aberration matrices $\mathbb{M}_{\mathrm{ext}}^{1k}$ corresponding to external (control) fields are evaluated up to some approximation order $N$ using symbolic formulae (see above), while the $\mathbb{M}_{\mathrm{self}}^{1k}$ are computed on the base of the PIC codes. We keep the matrix form for the beam propagator $\mathscr{M}$: $\mathbb{M}^N = (\mathbb{M}^{11}, \ldots, \mathbb{M}^{1N})$ as this allows us to use all advantages of the matrix algebra in parallel computing. Here we should note that the PIC codes admit enough simple parallelization procedure, which can be

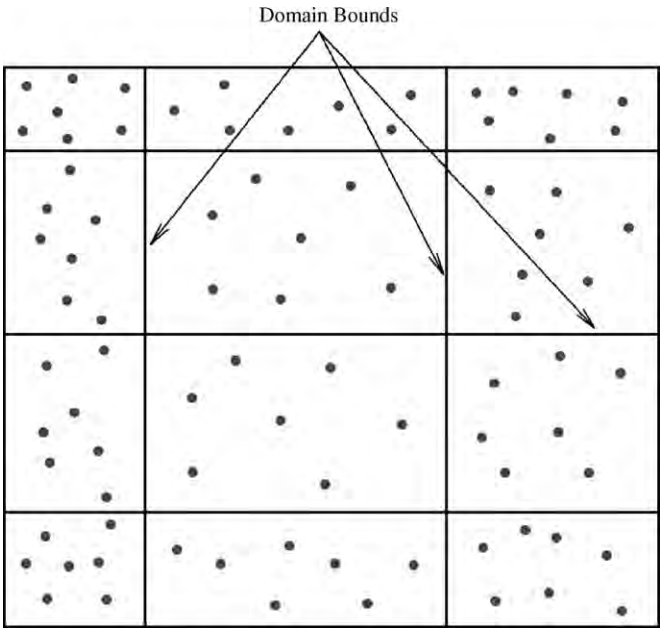



Fig. 1. A schematic 2D-plot of decomposition on $x$–$y$ domain. The number of particles is identical for each domain.

come to an agreement about matrix form for the beam propagator.

## 3. Computer model generation

The described approach of the problem solution are based on some linear algebra operations, which applied to any kind of objects having a vector or matrix nature. These operations form a mathematical core of our computer codes (which are adequate to using computer architecture), and just these operations are parallelized in the first place. Here one can use usual parallel tools employed for parallel computers.

### 3.1. Computer algebra and knowledge bases

The desired solutions are created in the form of power series up to an approximation order $N$. For most practical problems it is enough to evaluate the matrices $\mathbb{P}^{1k}$, $\mathbb{G}_k$, $\mathbb{H}_k$, $\mathbb{V}_k$ and $\mathbb{M}^{1k}$ up to third order in symbolic forms using the computer algebra codes (in our case Reduce and Maple). For some other types of ion-particle machines these matrices are computed up to fifth order, in particular for the nuclear microprobe (see, for example, Ref. [7]). These tools have algebraic character and can be easily realized on parallel computers.

### 3.2. The map creation

Knowledge of the Lie map $\mathscr{M}(s|s_0)$ in the matrix form (see Eq. (2)) allows to create necessary criteria for beam line working and current images $\mathfrak{M}(s) = \mathscr{M}(s|s_0) \circ \mathfrak{M}_0$, where $\mathfrak{M}_0$ is an initial beam phase portrait. We should note that the approximation order $N$ has to make consistent with the order of approximation for the external and certainly for space charge fields. In this approach one follows current phase beam portraits. These portraits allow to calculate additional criteria of beam evolution. From computing point of view this approach has two levels of realization too. The map computing is based on a set of ready matrices $\mathbb{M}^{1k}_{\mathrm{ext}}(s|s_0)$ for different accelerator elements calculated in symbolic forms (using computer algebra codes Reduce and Maple). These matrices are the fulfilment of the corresponding database. The self-field matrices are calculated numerically.

As the matrix algebra admits the parallelization in a natural way, then one can calculate necessary criteria using only matrix elements of current matrices $\mathbb{M}^{1k}$, $k \leqslant N$ (see, for example, Ref. [8]).

### 3.3. Additional modules

The above described approach has an intermediated character. Indeed the block matrices for external and self-consistent fields ($\mathbb{M}^{1k}_{\mathrm{ext}}(s|s_0)$ and $\mathbb{M}^{1k}_{\mathrm{self}}(s|s_0)$) are calculated using sufficient different approaches. Here we waste an advantages of the matrix formalism. Therefore it will be very useful to constraint an approach permitting to evaluate $\mathbb{M}^{1k}_{\mathrm{self}}(s|s_0)$ using the matrix formalism too, for example based on the approach suggested in Refs. [2,6]. For this purpose the PIC codes are suggested for a distribution function approximation in some class of model distribution function. This allows us in future to pass to matrix presentation for all block matrices, which correspond both self and external fields. This approach was approved for some test problems [7].

## References

[1] M. Giovannozzi, in: Proceedings of the 1998 European Particle Accelerators Conference, Stockholm, Sweden, 1998, pp. 1189–1191.
[2] S.N. Andrianov, AIP Conference Proceedings 1997, No. 391, New York, pp. 355–360, Dynamical Modelling of Beam Particle Control Systems, SPbSU, SPb, 2004 (in Russian).
[3] A.J. Dragt, No. 87, New York, 1982, pp. 147–313.
[4] S. Klarsfeld, J.O. Oteo, J. Phys. A 22 (21) (1989) 4565.
[5] S.N. Andrianov, Math. Comput. Simul. 57 (3–5) (2001) 139.
[6] S.N. Andrianov, in: Proceedings of Sixth European Particle Accelerator Conference, Stockholm, Bristol (UK), 1998, pp. 1091–1093.
[7] S.N. Andrianov, Nucl. Instr. and Meth. A 519 (2–3) (2004) 37.
[8] S.N. Andrianov, N.S. Edamenko, in: Proceedings of the 11th International IFAC Workshop Control Application of Optimization—CAO'2000, vol. 1, Pergamon Press, Oxford, UK, 2001, pp. 11–16.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 154–158
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# A massively parallel particle-in-cell code for the simulation of field-emitter based electron sources

Arno E. Candel[a,*], Micha M. Dehler[b], Matthias Troyer[c]

[a]*Theoretische Physik, Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland*
[b]*Paul Scherrer Institut, Villigen PSI, Switzerland*
[c]*Theoretische Physik, Eidgenössische Technische Hochschule, CH-8093 Zürich, Switzerland*

Available online 29 November 2005

**Abstract**

For the realistic simulation of electron sources using field emitter arrays, the sub-micron resolution required for the emitters leads to simulation models not suitable for current serial codes. Thus, a parallel high-performance 3D Particle-In-Cell code, called *Capone*, has been implemented in C++ using the POOMA II framework on the Linux platform. Sophisticated C++ expression templates techniques deliver Fortran performance combined with high-level programming and development comfort. For the computation of external fields, matching parallel field solvers are in development with the electrostatic one being completed.

The Maxwell field solver is based on the Finite Integration Algorithm on a non-uniform rectilinear grid. Anisotropic $\varepsilon$ and $\mu$ constants and perfect electric/magnetic materials stored in triangulated grid cells are supported as well as open, electric and magnetic boundary conditions. Self-consistent macro-particle pushing is accomplished by integrating the classical relativistic equations of motion in combination with charge-conserving current scattering onto the computational grid.

Parallelization is performed by partitioning the calculation domain into patches associated to individual processors. Fields are statically distributed while Particles are concurrently distributed to processors according to their position to allow fast local interpolation.



*PACS:* 41.20.Jb; 41.75.Fr; 41.75.Ht; 41.60.−m; 41.85.Ar; 41.85.Ne; 41.75.−i; 41.60.Cr; 83.10.Rs

*Keywords:* Electromagnetism; Charged-particle beam; Free-electron lasers; Computer simulation; Maxwell; Self-consistent; Particle dynamics; Particle-In-Cell; Time-domain; Parallel computing; C++

## 1. Introduction

### 1.1. Physical background

After the successful completion of the synchrotron light source project SLS at the Paul Scherrer Institut (PSI), PSI is looking forward at options for a Free Electron Laser (FEL) used as a complementary research tool to the SLS.

Our focus is on compact Angstrom wavelength FEL sources with a modest beam energy and a short saturation length of the undulator. The challenge is the required quality of the generating $e^-$ beam, mainly determined by the electron source. The upper limit for the required transverse emittance scales with the beam energy. A compact FEL working at reduced beam energies needs emittances orders of magnitudes below current designs. Our design goal is in the order of a few times $10^{-2}$ mm mrad.

The current approach towards a suitable electron source consists of using field emitter arrays (FEA) as a cold, high intensity electron cathode [1]. This cathode is used in a DC gun, driven by a 1 MV, 200 ns FWHM pulse [2], which is currently under development.

An example for a field emitter structure suited to our application is shown in Fig. 1. Here, the required field is

*Corresponding author. Stanford Linear Accelerator Center, 2575 Sand Hill Road, M/S 27, Menlo Park CA 94025, USA.
*E-mail addresses:* arno@candel.org, candel@slac.stanford.edu (A.E. Candel).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.059

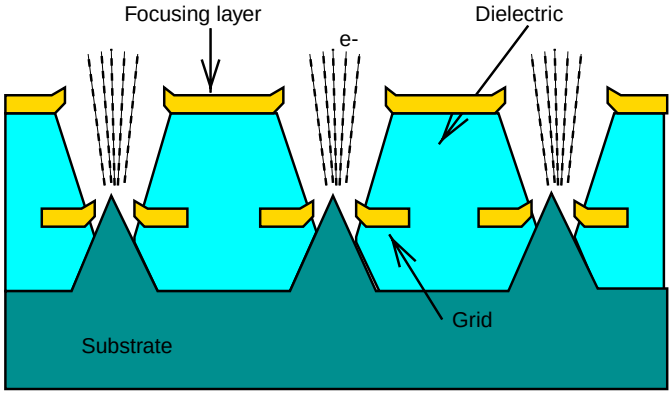



Fig. 1. Field emitter with additional grid and focusing layers.

applied via a conducting gate layer at a distance in the μm range from the substrate. The small distance yields relatively small turn on voltages in the order of 100 V, which in turn allows the use of fast pulse electronics to create short pulses. Having only a gate would lead to a divergent beamlet. In order to combine an array of these beamlets minimizing the emittance of the total beam, an additional focusing layer is mandatory, acting as a set of electrostatic lenses, thus creating an array of parallel beamlets.

### 1.2. Simulation strategy

Given the complexity of the problem, the dynamics of the emitted electrons in a given setup can only be simulated with a self-consistent, parallel 3D Maxwell Particle-In-Cell (PIC) code consisting of a Maxwell field dynamics solver, a static field solver and a relativistic particle pusher. Currently, a reasonable spatial resolution in 3D cannot be accomplished without parallelization due to high memory needs. Existing available codes like MAFIA [3] are either not fully 3D, not self-consistent, restricted to special cases or are only serial, and therefore not suited to our problem.

Over the last two years, a massively parallel high-performance 3D Maxwell PIC code has been implemented in C++ using the POOMA II framework for parallel computing on the Linux platform. Sophisticated C++ expression templates techniques deliver Fortran performance combined with high-level programming and development comfort. The simulation code is compatible to MAFIA's PIC modules TS2/TS3 which are used as reference. The code is called *Capone*, which stands for Charged Accelerated Particles Outta Next-generation Emitters.

## 2. Methods

### 2.1. Dynamic field solver

The numerical simulation of the electromagnetic field dynamics uses the Finite Integration Technique (FIT) [4].

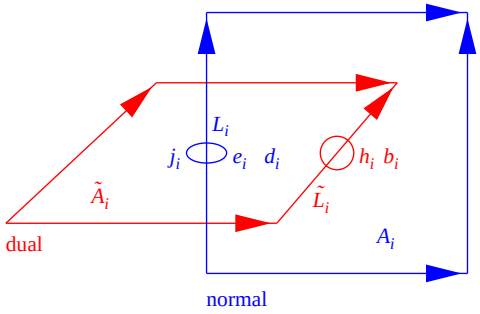



Fig. 2. Topological structure of the Finite Integration Algorithm.

A complete discretization of the calculation volume on two dual rectilinear grids $G$, $\tilde{G}$ is done, described by cells with volumes $V_i$, $\tilde{V}_i$, cell faces $A_i$, $\tilde{A}_i$ and grid lines $L_i$, $\tilde{L}_i$. Integrated field components $e_j = \int_{L_j} \vec{E}\,\mathrm{d}\vec{s}$ etc. are stored at positions indicated in Fig. 2.

Maxwell's equations are mapped to this topological space, using discrete curl and divergence operators $C$, $\tilde{C}$, $S$, $\tilde{S}$ as well as discrete material operators $D_\varepsilon$, $D_\mu$, resulting in the following set of discrete Maxwell's equations:

$$Ce = -\frac{\partial b}{\partial t} \tag{1}$$

$$\tilde{C}h = \frac{\partial d}{\partial t} + j \tag{2}$$

$$\tilde{S}d = q \tag{3}$$

$$Sb = 0 \tag{4}$$

$$d = D_\varepsilon e \tag{5}$$

$$b = D_\mu h. \tag{6}$$

Non-homogeneous distributions of $\varepsilon$ and $\mu$ are permitted, with material boundaries given by cell boundaries and cell diagonals. In addition to Dirichlet and Neumann boundary conditions, open boundaries have been implemented.

The system is solved in the time domain by integrating the two curl equations using the leap frog algorithm [13]. Due to the orthogonality of the discrete curl and div operators, the continuity equations remain fulfilled throughout the iteration process. There is no systematic accumulation of spurious space charges [5].

The discrete electric current density $j$ in (2) is obtained with a Nearest-Grid-Point (NGP) scatter interpolation in combination with a scheme fulfilling the continuity equation in discrete space [6]. A spatial filter is used to smoothen the electric current density at every time step.

### 2.2. Static field solver

For the generation of a consistent initial electrostatic field solution, exactly the same grid and the same material distribution as for the FIT solver are used. In our parallel approach, we are using an iterative conjugate gradient solver together with an incomplete Cholesky

preconditioner IC(0) with additional red/black checkerboard type domain decomposition. For details on these techniques, see e.g. Ref. [7].

### 2.3. Particle dynamics

Like in MAFIA, electron macro-particles are created inside of designated source materials and drift towards the emission surface with their initial velocity where they start to interact with the fields. Time distribution as well as the various initial position and momentum distributions are given by user input and are independent of local field strengths.

Adapting a leap-frog scheme for the numerical integration of the classical relativistic collisionless equations of motion, the following equations are obtained [8]:

$$\vec{r}^{n+1} = \vec{r}^{n} + \Delta t \frac{c}{\gamma^{n+1}} \vec{u}^{n+1} \tag{7}$$

$$\vec{u}^{n+1} = \vec{u}^{n} + \Delta t \frac{q}{m_0 c} \left( \vec{E}^{n+1/2} + \frac{c}{\gamma^{n+1/2}} \vec{u}^{n+1/2} \wedge \vec{B}^{n} \right) \tag{8}$$

where $\vec{u}$ is the normalized momentum, time steps are indicated by superscripts and $\vec{E}, \vec{B}$ are interpolated superpositions of static and dynamic fields at the particle positions using a first-order cloud-in-cell scheme.

The implicit momentum Eq. (8) is solved by using an explicit scheme [9], splitting the momentum update into an electric acceleration of error $O(\Delta t^2)$ and a magnetic rotation of error $O(\Delta t^3)$.

## 3. Implementation and parallelization

### 3.1. Program structure

*Capone* is based on the C++ POOMA II framework on the Linux platform and uses LAM MPI for parallelization [10,11]. A schematic of the program structure is shown in Fig. 3.

Efficient parallelization is performed by partitioning the calculation domain into patches associated to individual processors. Fields are statically distributed with an overlapping guard layer to optimize communication. Particles are concurrently distributed to processors according to their positions. The disadvantages of this approach are a strong dependency of the parallelization efficiency on the particle distribution and communication overhead. The advantage is having fast local interpolations for one of the most CPU intensive parts of the code.

### 3.2. Parallel scaling

One a single CPU, needed computational resources are comparable to MAFIA. With 64 CPUs, the Maxwell solver has a parallel scaling performance of about 70% and shows no saturation. The scaling performance of the PIC module

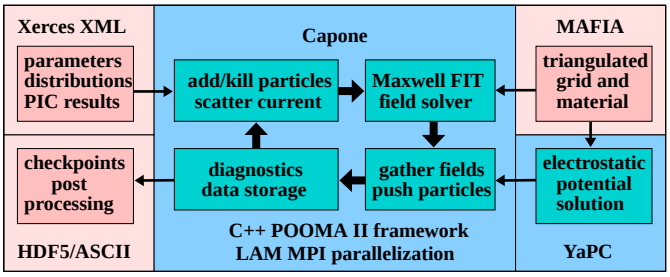



Fig. 3. Schematic program structure of the 3D PIC code *Capone*.

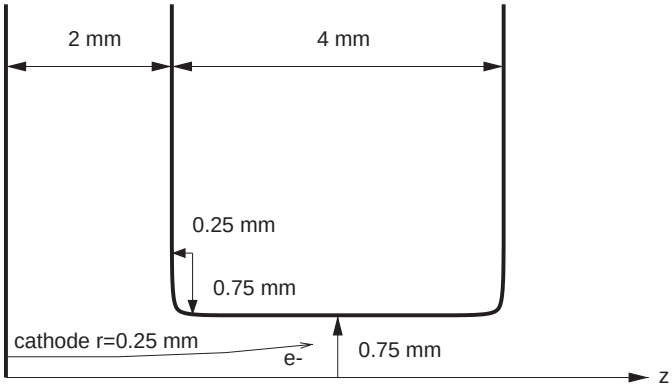



Fig. 4. Gun geometry used in the simulations.

is slightly less, depending strongly on domain partitioning, particle distribution and relative network speed.

## 4. Results

### 4.1. DC gun parameters

The DC gun geometry used in the simulations is shown in Fig. 4. It is a simple diode configuration with a gap of 2 mm and a gap voltage of 1 MV. For reduced surface fields, the iris opening has an elliptical rounding. The beam at the 0.5 mm diameter cathode surface is assumed to correspond to a parallel beam at the exit of the FEA focusing layer. The initial transverse momentum spread, which we would expect to be of an order of 200–300 meV, is assumed to be zero. Initial longitudinal electron momentum is set to about 100 eV, which is the value we expect at the exit of the focusing layer. In all simulations, the longitudinal distribution is assumed to be Gaussian with FWHM pulse durations of either 20 or 2 ps. The total charge is chosen such that the peak current was 5 A in all cases. Transversal distributions are assumed to be uniform and circular.

### 4.2. 3D PIC simulations

#### 4.2.1. Validation with MAFIA TS3

Validation of *Capone* against the MAFIA TS3 3D PIC module is presented in the following. A discretization of the full ("4/4") DC gun with 4.4 M grid points ($127 \times 127 \times$

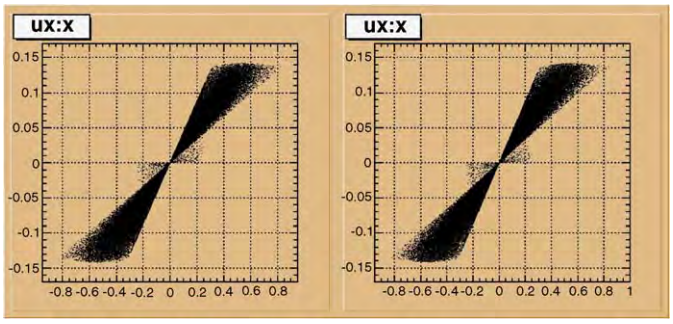



Fig. 5. Comparison of *Capone* (left) to MAFIA TS3 v4.106 (right) calculations. Shown here are transverse phase space plots, $u_x$ is normalized momentum, $x$-positions are indicated in mm.

Table 1
Diagnostics of particle phase space snapshots at mean longitudinal position $\langle z\rangle \simeq 7.671$ mm

| | *Capone* | MAFIA TS3 |
|---|---|---|
| # of macro-particles | 89734 | 89735 |
| $\langle z\rangle$/mm | 7.6712 | 7.6717 |
| $\langle E\rangle$/keV | 996.6 | 996.5 |
| $\sigma_E$/keV | 31.84 | 31.84 |
| $\sigma_r$/mm | 0.3906 | 0.3908 |
| $\sigma_{r'}$/mrad | 35.733 | 35.760 |
| $\varepsilon_x^n$/mm mrad | 3.9180 | 3.9236 |

Both *Capone* and MAFIA TS3 calculations are done with the same discretization of 4.4 M grid points and 90k macro-particles and show excellent agreement within statistical errors.

273) is used for both codes. 90k macro-particles are used to simulate a bunch of 20 ps (FWHM) duration. This problem size is near the limit of MAFIAs current capabilities, with a highest spatial resolution in the beam region of 30 μm.

A detailed comparison of the particle phase space distributions at a fixed point in time is done, as shown in Fig. 5 and Table 1.

For the given temporal snapshot, the bunch spans almost the entire longitudinal calculation domain and the first particles have already left the upper boundary at $z = 14$ mm. For a given MAFIA TS3 particle dump, the selected time step in the *Capone* calculation is chosen such that the corresponding mean longitudinal positions are as equal as possible ($\langle z\rangle \simeq 7.671$ mm). The difference is less than the change in one time step. Excellent agreement is found.

### 4.2.2. Simulation of FEAs

One motivation for the simulations presented in the following is to compute the effect of the granularity in the current density distribution due to an emitter array, as compared to a smooth density expected from e.g. a photocathode. A second question is the influence of imperfections in the homogeneity of real-life FEAs on the beam quality.

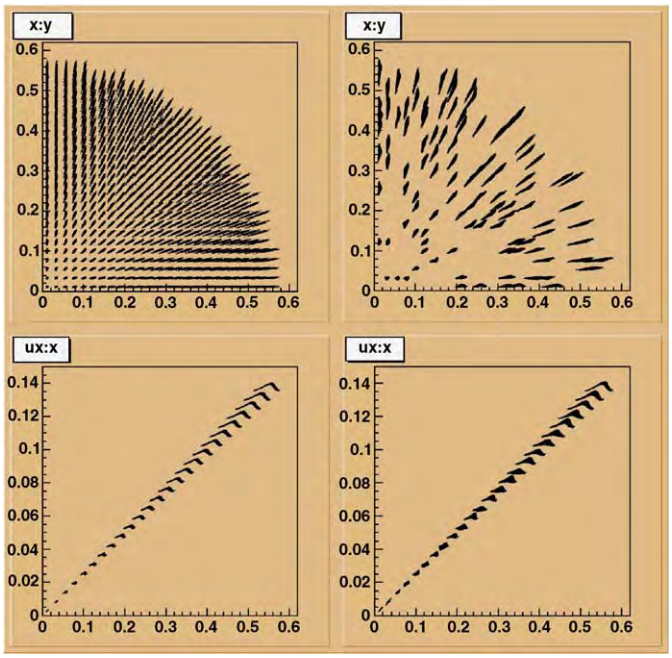



Fig. 6. Effect on transversal particle distribution and phase space for FEA simulations with 0% and 80% failure probability, at $z = 9$ mm. Positions are indicated in mm.

The performance of FEAs is typically degraded by several effects. First are variations in the emission of individual tips due to production and aging. Typically only a part of the emission sites contributes to the beam, with some global, highly correlated variation over the whole region.

Another effect inherent even in perfect field emitters is noise in the emission current due to adsorbate diffusion and adsorbates switching between emission states [12]. For low current emission at room temperature, the switching adsorbates jumping between emission sites leading to current variations is the dominant process. The resulting current fluctuations are in the millisecond range. Talking about the whole array, the noise effects from individual tips are uncorrelated.

In the following, we present simulation results showing the effect of uncorrelated tip-to-tip failure, where individual field emitters are switched off with a given probability.

A discretization of one symmetrical quarter of the gun with 8 M ($158 \times 158 \times 310$) grid points with focus on good transverse resolution of 4 μm in the beam region is used. Up to 2000 beamlets with FWHM pulse durations of 2 ps are emitted from a circular emitter array of $r = 0.25$ mm with a pitch of 10 μm and local emitter radii of 1 μm. Failure probabilities of the individual emitters are varied from 0% to 80%. Total peak current of 5 A and number of macro-particles of 200k are kept constant in all cases.

Fig. 6 shows the effect on transverse particle distribution. Some minor differences in transverse phase space can be observed, leading to an rms emittance growth as shown in Fig. 7. Growth of the normalized slice emittance at the

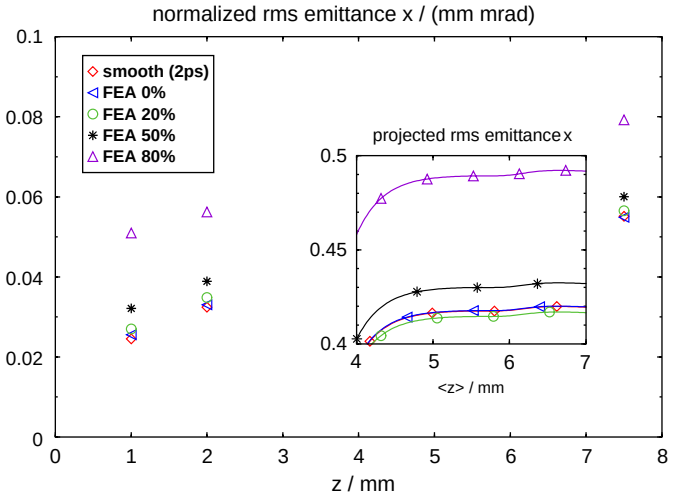



Fig. 7. Normalized rms emittance evaluated at three different longitudinal $z$-positions 1, 2 and 7.5 mm. Smooth distribution is compared to emission from an FEA with $x$% failure probability, denoted by FEA $x$%. The projected normalized rms emittance of the full bunch as a function of mean longitudinal position is shown in the box.

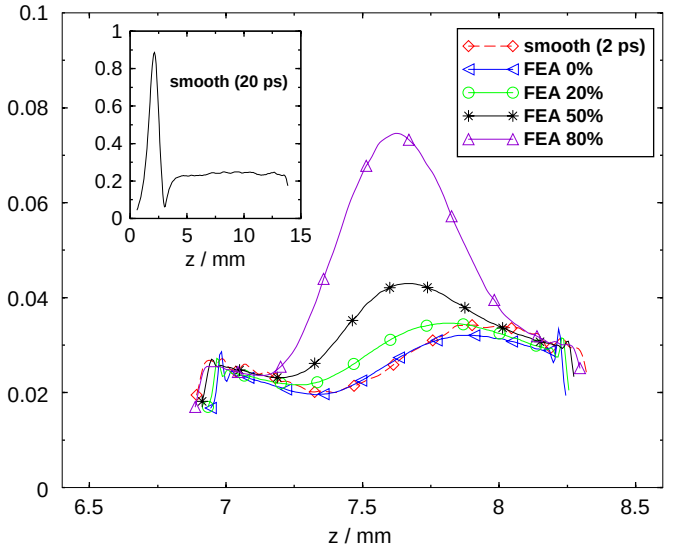



Fig. 8. Normalized slice emittances. Smooth distribution is compared to emission from an FEA with $x$% failure probability, denoted by FEA $x$%. Comparison with the slice emittance of a 20 ps FWHM bunch shows the influence of wake fields. All slice emittances have been calculated using a Gaussian weighting function with a $\sigma_z$ of $\frac{1}{100}$ of the bunch length.

center of the bunch from $2.5 \times 10^{-2}$ mm mrad (0% failure probability) to $4.1 \times 10^{-2}$ mm mrad (50%) and $7.7 \times 10^{-2}$ mm mrad (80%) is shown in Fig. 8.

## 5. Conclusion

A self-consistent, parallel high-performance 3D Maxwell PIC code, called *Capone*, has been implemented using the C++ POOMA II/MPI framework on Linux. The code has been designed for the simulation of DC guns and is validated with MAFIA. First results on simulations of field emitter arrays show the feasibility of detailed studies where high resolution accuracy is required to resolve highly non-linear space-charge effects.

## References

[1] R. Ganter, M. Dehler, J. Gobrecht, C. Gough, S. Leemann, K. Li, M. Paraliev, L. Rivkin, A. Wrulich (PSI, Villigen), A. Candel (ETH, Zürich), Preliminary results on a low emittance gun based on field emission EPAC'04, submitted for publication.
[2] C. Gough, PSI, Villigen personal communication.
[3] MAFIA 4, Computer simulation technology, Darmstadt, ⟨http://www.cst.de⟩.
[4] T. Weiland, A discretization method for the solution of Maxwell's equations for six-component fields, Electron. Commun. AEUE 31 (3) (1977) 116.
[5] M. Dehler, Numerische Loesung der Maxwellschen Gleichungen auf kreiszylindrischen Gittern, Darmstaedter Dissertationen, Technische Hochschule Darmstadt, 1993.
[6] O. Buneman, in: Relativistic Plasmas, Benjamin, New York, 1968, p. 205.
[7] Y. Saad, Iterative Methods for Sparse Linear Systems, first ed. ⟨http://www-users.cs.umn.edu/~saad/books.html⟩, second ed., in hard print, SIAM, Philadelphia ⟨http://ec-securehost.com/SIAM/ot82.html⟩.
[8] U. Becker, Berechnung und Darstellung von Bahnen geladener Teilchen in elektromagnetischen Feldern, Studienarbeit, Institut für Hochfrequenztechnik, Technische Hochschule Darmstadt, 1992.
[9] P. Schuett, Zur Dynamik eines Elektronen-Hohlstrahls, Dissertation Hamburg, DESY M-88-03 (1988).
[10] Parallel Object-Oriented Methods and Applications II, LANL, http://www.codesourcery.com/pooma/.
[11] LAM/MPI, Parallel computing, ⟨http://www.lam-mpi.org⟩.
[12] R.T. Olson, G.R. Condon, J.A. Panitz, R. Schwoebel, Analysis of bistable noise from microfabricated field emission cathodes, Appl. Phys. 87 (4) (2000) 2031.
[13] K.S. Yee, Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, IEEE AP-14 (1966) 302.

Available online at www.sciencedirect.com
SCIENCE DIRECT
Nuclear Instruments and Methods in Physics Research A 558 (2006) 159–162
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Moments conservation in adaptive Vlasov solver

M. Gutnic[a,c,*], M. Haefele[b,c], E. Sonnendrücker[a,c]

[a]IRMA, Université Louis Pasteur, Strasbourg, France
[b]LSIIT, Université Louis Pasteur, Strasbourg, France
[c]Projet CALVI, INRIA Lorraine, Nancy, France

Available online 27 December 2005

**Abstract**

We previously developed an adaptive semi-Lagrangian solver using a multiresolution analysis based on interpolets which are a kind of interpolating wavelets introduced by Deslauriers and Dubuc. This paper introduces a new multiresolution approximation for this solver which allows to conserve moments up to any order in the thresholding step by using the lifting method introduced by Sweldens.


*Keywords:* Vlasov; Phase-space grid; Adaptive; Multiresolution; Plasma physics; Beam physics

## 1. Introduction

The model we consider throughout this paper is the nonrelativistic Vlasov equation

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla_x f + \frac{q}{m}(\mathbf{E}+\mathbf{v}\times\mathbf{B})\cdot\nabla_v f = 0 \quad (1)$$

where the self-electric field $\mathbf{E}$ is computed from Poisson's equations. The magnetic field is external and considered to be known.

The numerical solution of the Vlasov equation is usually performed by particle-in-cell (PIC) methods, which are known to suffer from numerical noise. To remedy this problem and obtain a more accurate description of the distribution function, methods discretizing the Vlasov equation on a mesh of phase-space have been proposed [1–3]. In order to avoid the high numerical cost of such methods using a uniform and fixed mesh, we develop adaptive methods.

Our adaptive method is overlaid on a classical semi-Lagrangian method. Adaptivity, that is refinement or derefinement of the mesh, is based on a multiresolution analysis (Section 2, see Ref. [4] for more details). In the present work, we focus on moment conservation for the Vlasov equation during the thresholding step. This ensures in particular that the total number of particles is conserved. We explain (Section 3) how to use the lifting procedure introduced by Sweldens [5] in order to ensure this conservation of moments. Relevant numerical results are presented in Section 4.

## 2. Multiresolution analysis

The semi-Lagrangian method consists in computing point values of the distribution function on a grid of phase-space. It consist of two steps, an advection step needed to determine the origin of the characteristics ending at the grid points and an interpolation step which is used to compute the value of the distribution function at these points. The multiresolution approximation will be used to minimize the number of interpolation points for a given approximation error.

First, we define an infinite sequence of nested grids $(G_j)_{j\in\mathbb{Z}}$, where $j$ is called the level of the grid, the grid points being located at $x_k^j = k2^{-j}$. In order to define an adaptive grid, we want to compare a function $f$ defined by its values $(c_k^{j+1} = f(x_k^{j+1}))_{k\in\mathbb{Z}}$ on a finer grid $G_{j+1}$ to its restriction $(c_k^j = f(x_k^j) = f(x_{2k}^{j+1}) = c_{2k}^{j+1})_{k\in\mathbb{Z}}$ on a coarser grid $G_j$. For this purpose, we need a prediction operator to define an approximation of $f(x_{2k+1}^{j+1}) = c_{2k+1}^{j+1}$ from the $(c_k^j)$. Using an odd degree Lagrange interpolation polynomial $P_{2N-1}$, this approximation is given by $P_{2N-1}(x_{2k+1}^{j+1})$ and the

*Corresponding author. IRMA, Université Louis Pasteur, Strasbourg, France.
*E-mail address:* gutnic@math.u-strasbg.fr (M. Gutnic).

0168-9002/$ - see front matter © 2006 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.225

approximation errors, called details in wavelet terminology, are given by

$$d_k^j = c_{2k+1}^{j+1} - P_{2N-1}(x_{2k+1}^{j+1}) = c_{2k+1}^{j+1} - \sum_{n=1-N}^{N} a_n c_{2k+2,}^{j+1}. \qquad (2)$$

This can be formulated in the framework of biorthogonal wavelets introduces by Cohen et al. [6]. The above decomposition of $f$ can also be expressed using basis functions $\varphi$, named scaling function, and $\psi$, named wavelet, such that $\psi(x) = \phi(2x-1)$:

$$f(x) = \sum_k c_k^{j+1}\varphi_k^{j+1}(x) = \sum_k c_k^j \varphi_k^j(x) + \sum_k d_k^j \psi_k^j(x) \qquad (3)$$

where $\varphi_k^j(x) = \varphi(2^j x - k)$ and $\psi_k^j(x) = \psi(2^j x - k)$. For interpolating polynomials $P_{2N-1}$, the values $a_n$ defined in Eq. (2) are $a_0 = a_1 = \frac{1}{2}$ and else $a_n = 0$ for degree 1, and $a_{-1} = a_2 = -\frac{1}{16}$, $a_0 = a_1 = \frac{9}{16}$ and else $a_n = 0$ for degree 3. The scaling function $\varphi$ and wavelet $\psi$ for the corresponding values of $N$ are displayed in Figs. 1 and 2.

We now consider the thresholding. From Eq. (2), we note that the improvement of approximation is locally important around grid point $2^{-j}(2k+1)$ when the detail $d_k^j$ is large, and conversely small when the detail is small. Hence the representation of a function at level $j+1$ can be compressed with a controlled approximation loss by setting to 0 the details with an absolute value less than some given threshold $\varepsilon_j > 0$ depending on the level $j$.

Thus, in the decomposition formula (3), we eliminate terms $d_k^j\psi_k^j$ such that $|d_k^j| < \varepsilon_j$ and we can bound the error committed because of this thresholding process as follows:

$$\|e_{j+1}^s\|_{L^p} = \left\| \sum_{k|d_k^j|<\varepsilon_j} d_k^j \psi_k^j \right\|_{L^p} \leqslant \sum_{k|\,|d_k^j|<\varepsilon_j} |d_k^j| \|\psi_k^j\|_{L^p}$$
$$\leqslant \varepsilon_j \sum_{k|\,|d_k^j|<\varepsilon_j} 2^{-j/p}\|\psi\|_{L^p} \qquad (4)$$

$$\leqslant \varepsilon_j 2^{-j/p}\|\psi\|_{L^p} \#(\{k|\,|d_k^j| < \varepsilon_j\}) \qquad (5)$$

since, assuming that $f_{j+1}$ has compact support, the number of removed terms is finite. In Eqs. (4)–(5), $\|\psi\|_{L^p} = (\int_{\mathbb{R}} |\psi(x)|^p \,\mathrm{d}x)^{1/p}$ for $p \geqslant 1$ is the norm in which the error is measured.

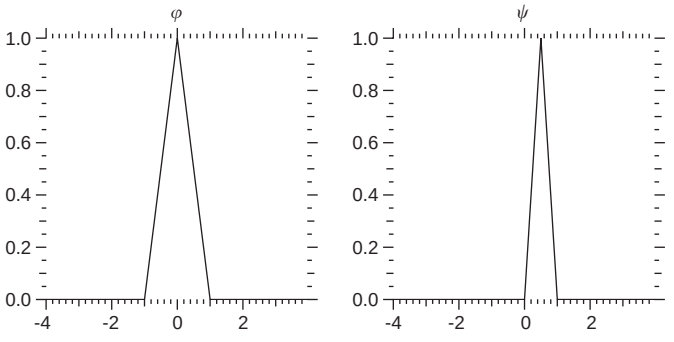



Fig. 1. Scaling functions and wavelets for linear interpolation polynomials ($N = 1$).

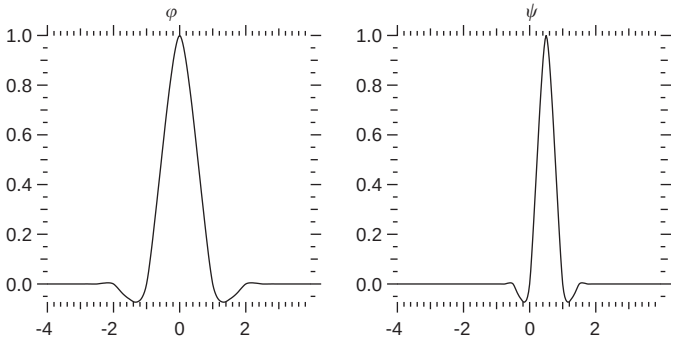



Fig. 2. Scaling functions and wavelets for cubic interpolation polynomials ($N = 2$).

## 3. Conservation of moments

When numerically solving the Vlasov equation, it is often essential to conserve the density of particles. Moreover, in order to get a better accuracy for the thresholded distribution function, it can be helpful to conserve higher order moments while performing adaptivity. Since the thresholding procedure consists in removing linear combinations of the $\psi_k^j$, this procedure conserves moments if the corresponding moments of $\psi$ that are $\int x^p\psi(x)\,\mathrm{d}x$ vanish.

We saw that the wavelet associated to the interpolating scaling function is $\psi(x) = \varphi(2x-1)$. Hence, as the moments of $\varphi$ do not vanish, the moments of $\psi$ do not vanish. However, the lifting procedure introduced by Sweldens [5] can be used to define a new set of biorthogonal scaling functions and wavelets from a given one so that some desired properties are satisfied.

Consider the decomposition formula (3). Following Swelden's construction, we can modify the basis $(\psi_j^k)$ without modifying the multiresolution approximation by taking a new wavelet $\bar{\psi}$ of the following form:

$$\bar{\psi}(x) = \psi(x) - \sum_k s_k \varphi(x-k) \qquad (6)$$

where the coefficients $(s_k)_k$ define the new wavelet.

To conserve density, one needs $\int \bar{\psi}(x)\,\mathrm{d}x = 0$ which yields

$$0 = \int \varphi(2x-1)\mathrm{d}x - \sum_k s_k \int \varphi(x-k)\,\mathrm{d}x$$
$$= \left(\frac{1}{2} - \sum_k s_k\right)\int \varphi(x)\,\mathrm{d}x. \qquad (7)$$

Thus, we need to ensure that $\sum_k s_k = \frac{1}{2}$. To keep the symmetry of $\bar{\psi}$, we also impose that $s_k = s_{1-k}$. Then the simplest choice for $(s_k)_k$ is $s_0 = s_1 = \frac{1}{4}$ and $s_k = 0$ else. With this choice of $s_0$ and $s_1$, $\bar{\psi}$ is an even function so that the first order moment is also conserved.

More generally, we can compute the $s_k$ in order to conserve the moment of any even order, always keeping the symmetry ensuring $s_k = s_{1-k}$. Then if all the moments of even order are conserved up to some even $n$, all the moments of any odd order are also conserved up to $n+1$.

In Fig. 3 and Table 1, we respectively construct the corrected wavelet in order to ensure density conservation and give the simplest choice of $s_k$ in order to ensure the conservation of the second order moment.

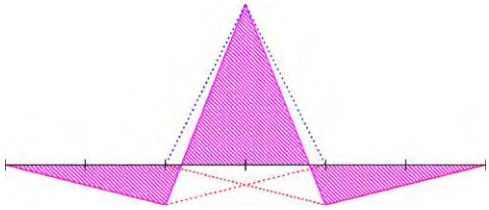


Fig. 3. Correction of linear wavelet in order to conserve density.

Table 1
Simplest choice of $s_k$ for second order moment conservation

| | | |
|---|---|---|
| Linear wavelets | $s_0 = s_1 = \frac{19}{64}$ | $s_{-1} = s_2 = \frac{-3}{64}$ |
| Cubic wavelets | $s_0 = s_1 = \frac{149}{512}$ | $s_{-1} = s_2 = \frac{-21}{512}$ |

## 4. Adaptive algorithm and numerical result

The wavelet decomposition we introduced yields an adaptive algorithm consisting of the following step: starting from a compressed distribution function $f^n$ defined on an adaptive grid, we predict the set of grid points that will contain the next adaptive grid by advecting the grid points along the characteristics. Then, we perform the semi-Lagrangian algorithm on the predicted points and finally compute the wavelet decomposition of $f^{n+1}$ to get the new set of important details and the corresponding

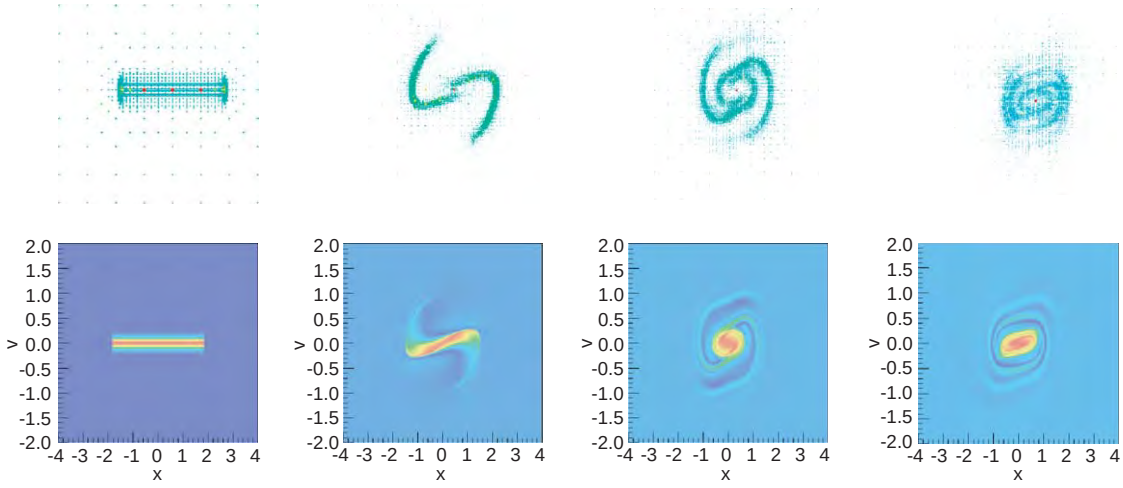



Fig. 4. Semi-Gaussian beam and the associated adaptive grid in a periodic focusing channel at 0, 15, 30 and $45\omega_p^{-1}$.

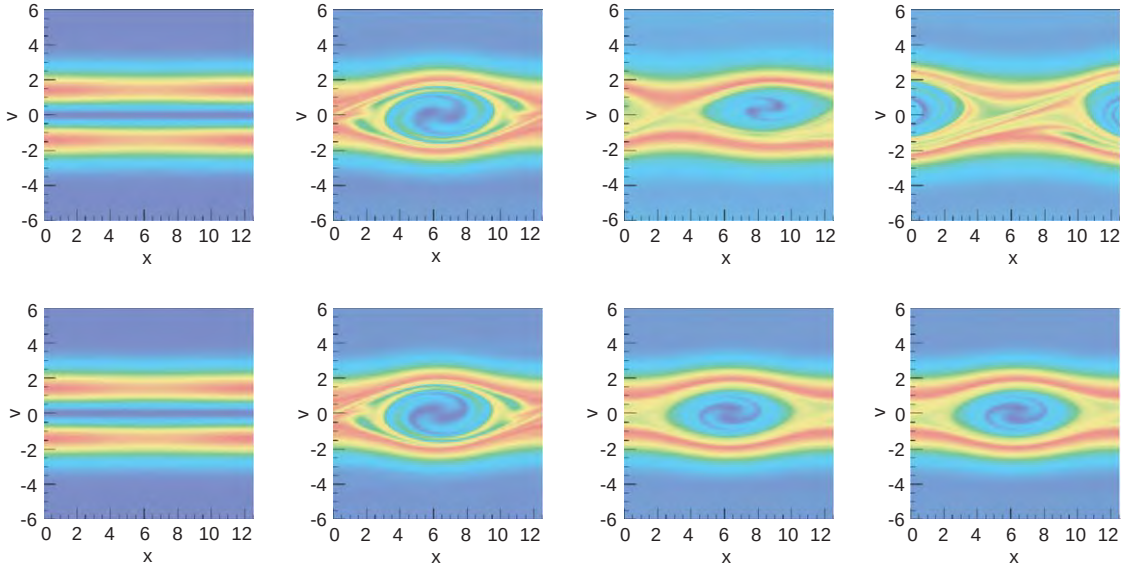



Fig. 5. Evolution of a two stream instability without (top) and with (bottom) density conservation at 0, 37.5, 75 and $112.5\omega_p^{-1}$.

adaptive grid (see Ref. [4] for more details on the algorithm).

In order to assess the benefits of the adaptive solver, we computed the transverse evolution of a semi-Gaussian beam in a periodic focusing field of the form $\alpha(1+\cos 2\pi z/S)$ for a tune depression $\sigma/\sigma_0$ of 0.17. The initial distribution function reads

$$f(r,v)=\frac{I}{\pi a^2\sqrt{2\pi}b}\mathrm{e}^{-1/2(v^2/b^2)} \quad \text{if } r<a \text{ and } f(r,v)=0 \text{ else.} \tag{8}$$

In order to show the importance of the moments conservation, we computed the evolution of a two-stream instability. The initial function is given by

$$f_0=\frac{1}{\sqrt{2\pi}}v^2\exp(-v^2/2)(1+\alpha\cos(kx)) \tag{9}$$

with $\alpha=0.05$ and $k=0.5$. The physical mesh is $x-$ periodic: $[0,4\pi[\times[-6,6[$.

We notice in Fig. 4 that the adaptive grid follows very well the evolution of the fine structures. In Fig. 5, we notice that ensuring density conservation avoided unstability problems at the expense of a slightly higher diffusion and computational cost when we use the lifting procedure.

## References

[1] C.Z. Cheng, G. Knorr, J. Comput. Phys. 22 (1976) 330.
[2] E. Sonnendrücker, J.J. Barnard, A. Friedman, D.P. Grote, S.M. Lund, Nucl. Instr. and Meth. A 464 (1–3) (2001) 653.
[3] E. Sonnendrücker, J. Roche, P. Bertrand, A. Ghizzo, J. Comput. Phys. 149 (1998) 201.
[4] M. Gutnic, M. Haefele, I. Paun, E. Sonnendrücker, Comput. Phys. Commun. 164 (2004) 214.
[5] W. Sweldens, SIAM J. Math. Anal. 29 (2) (1998) 511.
[6] A. Cohen, I. Daubechies, J.-C. Fauveau, Comm. Pure Appl. Math. 45 (5) (1992) 485.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 163–167
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Applications of parallel computational methods to charged-particle beam dynamics ☆

A. Kabel[a,*,1], Y. Cai[a,1], M. Dohlus[b], T. Sen[c], R. Uplenchwar[a,1,2]

[a] *Stanford Linear Accelerator Center, 2575 Sand Hill Road, Menlo Park, CA 94025, USA*
[b] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, D-22603 Hamburg, Germany*
[c] *Fermi National Accelerator Laboratory, Batavia, IL, USA*

Available online 1 December 2005

**Abstract**

The availability of parallel computation hardware and the advent of standardized programming interfaces has made a new class of beam dynamics problems accessible to numerical simulations. We describe recent progress in code development for simulations of coherent synchrotron radiation and the weak–strong and strong–strong beam–beam interaction. Parallelization schemes will be discussed, and typical results will be presented.


*PACS:* 02.70.−c; 29.27.Bd; 41.75.Ht

*Keywords:* Beam dynamics; Numerical simulation; Beam–beam effect; Coherent synchrotron radiation

## 1. Coherent Synchrotron Radiation: `TraFiC`[4]

Coherent Synchrotron Radiation (CSR) occurs when short bunches travel along strongly bent trajectories, leading to a tail-head interaction of the bunch with itself due to retardation effects. Such configurations typically occur in the bunch compression sections of Free Electron Laser Facilities; its impact can be macroscopic, leading to growth of the transverse projected or slice emittance, thus degrading FEL performance, or microscopic, leading to induced longitudinal short-wavelength density modulations in the beam, which may get amplified by the subsequent compression and transport mechanism.

The code `TraFiC`[4] was developed to handle the first class of problems; it models the bunch by a collection of spatially extended, non-compressible weighted macro-particles. These particles are tracked through the magnetic lattice in the laboratory frame of reference; CSR fields are calculated from first principles by storing the history of every macro-particle and using a retardation method on these histories. The fields of all particles are calculated for each particle applied at a single location for each time step, using a split-operator approach. Obviously, this algorithm is $O(N^2_{\text{particles}})$; in practice, in all but the simplest setups, parallelization is required to get manageable running times for simulations.

Among other things, `TraFiC`[4] has been used in the design of bunch compression sections for the DESY TESLA Test Facility [1], the LCLS [2], and for the simulation of a dedicated CSR experiment at the CERN CLIC Test Facility [3]. It has been benchmarked against other codes, using different approaches and/or simplified models for CSR.

Recent improvements of the code include a complete re-write of the tracking part in C++; extensive documentation and class structure documentation; a very flexible

☆ Invited talk presented at the 8th International Computational Accelerator Physics Conference, Saint Petersburg, Russia, June 29–July 2, 2004.

* Corresponding author.
*E-mail address:* akabel@slac.stanford.edu (A. Kabel).

[1] Work supported in part by the U.S. Department of Energy under contract number DE-AC02-76SF00515.
[2] Support by the U.S. Department of Energy's SciDAC program is gratefully acknowledged.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.079

mechanism of creating bunch populations by means of applying functional operators to pre-defined distributions or distributions read from files; MPI 'stub' libraries to allow for compilation and running of TraFiC$^4$ on single processor machines; more efficient storage of trajectory histories; dynamic load-balancing; and a basic checkpoint/restart mechanism.

Parallelization, so far, has been relying on a crude, but effective mechanism: the $n$th of $N$ nodes would run a complete replica of a single-node TraFiC$^4$ instance, reading the same input file. It would, however, only calculate the fields, due to *all* particles $P$, onto a subset of particles $P_n, P = \bigcup_i P_i, P_i \cap P_k = \{\}$. It would then gather the fields onto $P \backslash P_n$ from the other nodes in a collective synchronization step, apply them to the trajectories of $P$, and go to the next time step. This method, however, required storing the history of all particles on every node.

In our new approach, we only store the trajectories of particles $P_n$ on node $n$. For a given lab time $t$, it will broadcast a field calculation request for their positions $\vec{x}_i(t)$ to the other nodes, gather calculation requests for fields due to $P_n$ from the other nodes, process the requests, and then scatter its results and aggregate the other nodes' results, thus obtaining the total field. Most of these operations can be done asynchronously. This scheme involves more administrative overhead, however, there is no need to store any trajectories besides those of $P_n$. For typical problem and cluster sizes, the required memory per node could thus be reduced from 1 GB to some tens of MB.

An application of the improved TraFiC$^4$ code is a parameter study for the bunch compression section of the LCLS facility. In this study, more details of which can be found in Ref. [4], we vary the compression ratio of the bunch compressor to include bunch lengths well below the design values, while using nominal LCLS parameters for the other parameters. As the non-gaussian character of the initial distribution and the non-linear part of the initial energy distribution are of crucial importance, the new bunch population capabilities of TraFiC$^4$ were essential in this study.

The resulting final macro-particle distribution of each run was sorted into longitudinal bins; further post-processing removed the correlated energy spread. The resulting binned distributions are evaluated with respect to FEL figures of merit (namely, saturation length and saturation powers). The results show a gain in FEL performance with decreasing bunch length; even at 9.6 μm, the last bunch length investigated, we do not reach a break-even point.

The study was run on 512 processors per run on the NERSC facility. As the required longitudinal resolution was very high due to the high compression ratio and low natural energy spread, the number of macro-particles used was 7000, the highest number of macro-particles used so far in any TraFiC$^4$ run. A synopsis of resulting FEL figures of merit is shown in Figs. 1 and 2.

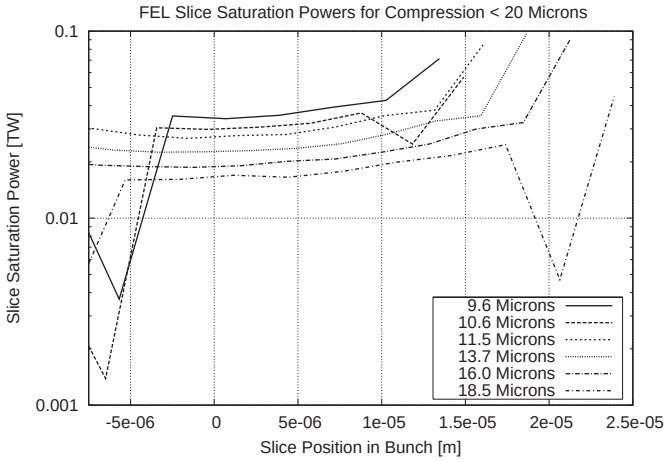



Fig. 1. Slice saturation powers for different bunch lengths.

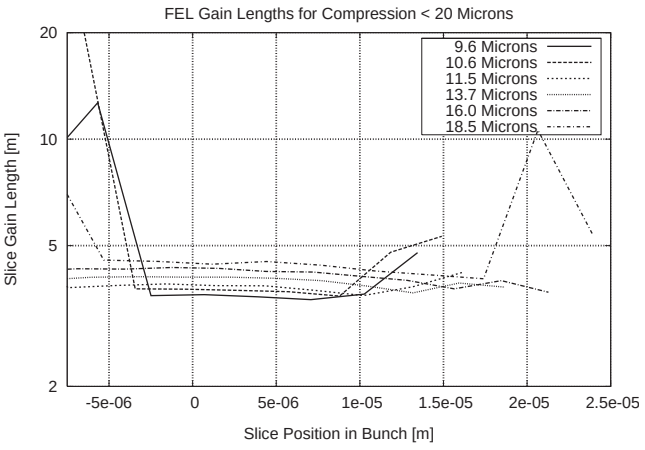



Fig. 2. Slice saturation length for different bunch lengths.

## 2. The weak–strong beam–beam effect: DUMBBB

The beam–beam interaction plays a crucial role in design and operation of colliding storage rings. It will limit luminosity, determine equilibrium emittance, and can affect beam lifetimes due to diffusion processes. For lack of an effective damping mechanism, the last item is especially important for hadron machines such as the Tevatron or the LHC. We can distinguish two realms of the beam–beam effect: strong–strong and weak–strong. In the former case, both beams' transverse fields affect each other significantly, while in the latter case the beams' charges differ strongly, so one ('strong') beam can be assumed to be unaffected by the other ('weak') beam. If the equilibrium distribution of the strong beam (as determined by the lattice and initial conditions) is known, the problem reduces to a single-particle dynamics problem in the presence of a highly non-linear force.

With current Tevatron operation parameters, there are 72 weak–strong beam–beam interactions affecting the dynamics of the weak (anti-proton) beam. It can be expected that the resulting, highly non-linear one-turn

map affects beam lifetime due to incoherent resonances or diffusion processes. To study these effects, we have developed a C++ code DUMBBB for fast tracking of single particles in the presence of weak–strong beam–beam interactions (for both parasitic and design interactions, i.e. off- or on-center). Being a single-particle code, parallelization reduces to the task of running many instances of the same code acting on different parts of a huge particle ensemble; communications is only required for calculation of collective quantities such as particle loss rates or emittances.

The code models the beam–beam interactions as a synchro-betatron mapping [5], the beam–beam kick itself is calculated from the Bassetti–Erskine [6] formula, using the Chiarella–Matta–Reichel approximation [7] for the evaluation of the complex error function. In a hadron machine, it is important to avoid all sources of numerical noise; the Chiarella algorithm is implemented as a templated C++ function with accuracy selectable at compile time. We find a $10^{-6}$ relative accuracy sufficient for turn numbers in the $10^5$ range.

The weak-beam part of the machine is modeled by a concatenation of beam–beam elements, linear $6\times 6$ transfer maps between non-linear elements (obtained by having a Perl script run MAD8 on a optics description file), a noise-inducing element to model emittance growth due to scattering processes, and a energy-dependent tune advance element to introduce total ring chromaticity.

The code allows for full coupling. All element transfer functions are templated with respect to the type of phase-space variables; in particular, they can operate on differential-algebraic quantities. This allows for finding exact solutions for the linear part of the one-turn map at start-up and constructing invariant initial weak and strong $6\times 6$ distributions, matched to measured emittances. Beam–beam elements have specialized functions depending on whether or not the strong beam shows hourglass effect, tilting, or position-depending tilting during an interaction. Also, they are templatized with respect to the number of slices used in the synchro-betatron mapping.

The aggregated lattice is repeatedly applied to real-value phase-space vectors of the initial weak distribution, which can be "de-cored" to remove particles from the core of the distribution, which are not expected to contribute to diffusive or resonant particle losses. Care must be taken to keep this operation invariant with respect to the one-turn map. The particles' excursions in action-angle space are recorded; once every few thousand turns, the $J_x, J_y$ space is swept and particles beyond a certain action aperture are counted. This way, we obtain a plot of particle loss vs. time for different assumptions about the limiting aperture of the machine. We typically run $10^{10}$ particle turns. The resulting dependencies are fitted with respect to $\tau$ against $\exp(-t/\tau)$ and $\exp(-\sqrt{t/\tau})$ particle loss behaviors, which are the limiting cases of solutions of the diffusion equation with absorbing boundary conditions for small and large-aperture boundaries. Due to the uncertainties of the diffusion model, the real aperture, and the simplifications in the model, the resulting lifetime should not be viewed as an absolute prediction, but as a figure of merit establishing signatures of the real lifetime of the machine.

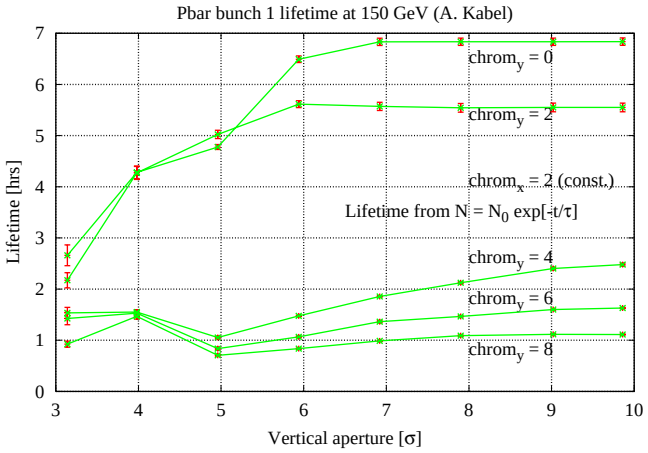



Fig. 3. Lifetime vs. vertical aperture and vertical chromaticity for Tevatron at injection.

We have done a series of parameter studies for the Tevatron at injection (150 GeV, 72 parasitic crossings, modeled as single-slice interactions). A typical result for varying chromaticity is shown in Fig. 3. Other parameter studies included sweeps of helix separations; weak-beam emittances, strong-beam charges, and two different bunch train schemes for 18 bunches on each of the bunch trains, resulting in lifetime differences of factors of two depending on deleting the odd- or even-numbered interactions; the latter result was checked independently with resonance strength studies.

To model non-linear effects due to the lattice, we have implemented a method for high-speed evaluation of multivariate polynomials. The method relies on the recursive definition of $P_n^v$, a $v$-variate homogeneous polynomial of degree $n$, as a direct sum $P_n^v = P_{n-1}^v \oplus P_n^{v-1}$ and a recursive evaluation algorithm $P_n^v(x_1,\ldots,x_v) = x_1 P_{n-1}^v(x_1,\ldots,x_v) + P_n^{v-1}(x_2,\ldots,x_v)$. Using C++'s templated data structure mechanisms for the definition of $P$ and inlining for the definition of the polynomial evaluation, the method effectively generates explicit expressions for Horner's scheme for any order at compile time. The method is easily generalized to inhomogeneous polynomials. We observe floating point efficiencies of $>0.85$ on Intel hardware and a speed gain of a factor of 4 as compared to standard implementations; still, we would need to gain another factor of 10 in speed to use 10th order polynomial transfer maps between beam–beam interactions.

## 3. The strong–strong beam–beam effect: NIMZOVICH

In the strong–strong realm, the colliding bunches influence each other substantially. Little is known analy-

tically about the resulting equilibrium distributions. Numerical methods used to determine them have converged on using PIC methods, modeling the collision process as a series of synchro-betatron mappings of test particles in the presence of two-dimensional field distributions (Ref. [8] and references therein). We have developed a code using this principle adapted to high longitudinal resolution, extreme aspect ratios, and the presence of parasitic crossings resulting in multi-bunch effects.

### 3.1. Parallelization

NIMZOVICH uses parallelization according to the SPMD (Single Program, Multiple Data) scheme. A cluster of processors is divided in two sections, called *Rings*. Each *Ring* is subdivided into several *Bunches*. Bunches within a Ring are completely independent. Bunches in opposing Rings are independent, except if they have a design or parasitic interaction point in common, i.e., if one of their two geometric interaction points falls into a section of the ring (the *Window*) shared by both beams.

Each Bunch is divided longitudinally into several Slices. For reasons of load balancing, the slicing scheme is chosen in such a way as to have the same number of particles within each slice, assuming an initial gaussian distribution of given length. Slice borders are, however, not dynamically adapted to changed longitudinal distributions.

Given enough available processors, each Slice's portion of particles can be further subdivided. Portions of a Bunch with the same subdivision index in each slice are called a *Slab*. They do not represent any geometric subdivision.

Each processor on each Ring runs through the following sequence of steps for each Turn:

(1) For each Bunch in the sequence of opposing Bunches:
   - For each Slice in the opposing Bunch:
     - Deposit particles onto grid in the center of gravity of my slice.
     - Solve Poisson's equation on that grid.
     - Calculate electric field.
     - Exchange electric field with opposing.
     - Slice in opposing Bunch.
     - Kick particles.
     - Advance particles to the next Slice.
   - Advance particles to next opposing Bunch.
(2) Advance particles according to one-turn map, possibly redistributing longitudinally.

We assume that the bunch is longitudinally frozen during interactions, so the slice-to-slice interactions are independent and can be done in parallel. Also, bunches in the same ring are independent, their mutual interaction can be handled in parallel. Synchronization is automatic, i.e., a Bunch will see the opposing Ring's bunches in the right order, as the slice-to-slice operation constitutes a barrier synchronizing the two Rings.

When a Slice has passed the last opposing Slice of its last opposing Bunch within a Window, it is transported back to the design IP, and the one-turn map is applied to its particles. After that, a particle may fall out of its current Slice. All particles with changed Slice numbers are moved to one out of a set of send queues, and an asynchronous send operation to its new Slice initiated. The leftover particles are deposited on the Grid. Then, the process opens a receive queue for particles from backward Slices, which might be moved onto this Slice by the action of the one-turn map. The process does not have to wait for all backward slices, as the synchrotron tune is usually small and a particle is extremely unlikely to pass distances of the order of a bunch length within a single turn. The actual number of backward slices a process will wait is dynamically adapted at run time; if the number of particles received after a Slice's first interaction with the next Bunch crosses a threshold (of the order of a few particles), the waiting period is increased.

A complication arises from the fact that the longitudinal resolution required is very different for parasitic and design interactions. Thus, a Bunch will have different slicing schemes, with $N_{\mathrm{Slabs}}N_{\mathrm{Slices}}$ constant, for different interaction points. It is easy to see that communications due to reassignment of slices by a change of resolution can be kept at its minimum by (1) letting the numbers of slices in adjacent IPs be integer multiples (provided the bunch length does not change between IPs) and (2) have formerly neighboring slices end up in the same new slice for a resolution decrease.

### 3.2. Field calculation

Point charges are deposited on a cartesian grid with typical dimensions of $N_x = 64 \ldots 512 \otimes N_y = 64 \ldots 512$, using a 9-site stencil. As the beam pipe is usually far away, Poisson's equation on the grid can be solved using free boundary conditions. This is done by convolving with an appropriately discretized and regularized version $\hat{G}_{ik}$ of the free Green's function $G(r) = 1/4\pi \log r^2$. The convolution is done by multiplication in momentum space; the transformation into momentum space is done by a two-dimensional Fast Fourier Transformation using the FFTW [9] package. Free boundary conditions are implemented by using the Hockney trick [10] of padding the array with zeroes to $2N_x \otimes 2N_y$.

The transformation is done by two sequences of one-dimensional transformations with a matrix transposition in between. In its parallel version, the transposition involves an expensive all-to-all communication, which might cancel the speed gains of parallelizing the transformation. In NIMZOVICH, the user has the choice of how finely to parallelize the solver. In our calculations, we find that the time spent in the solver equals the kick-deposit time at around $10^4$ particles.

Note that this is not the optimal solution; the fact that the array was zero-padded initially allows one to get rid of

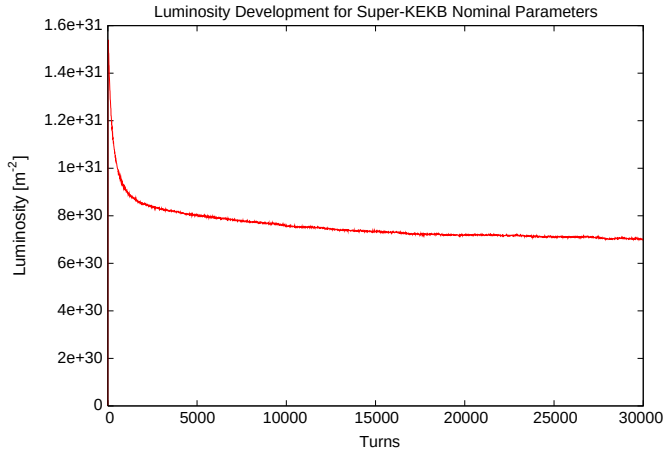



Fig. 4. Luminosity vs. time for NIMZOVICH SuperKEKB benchmark example.

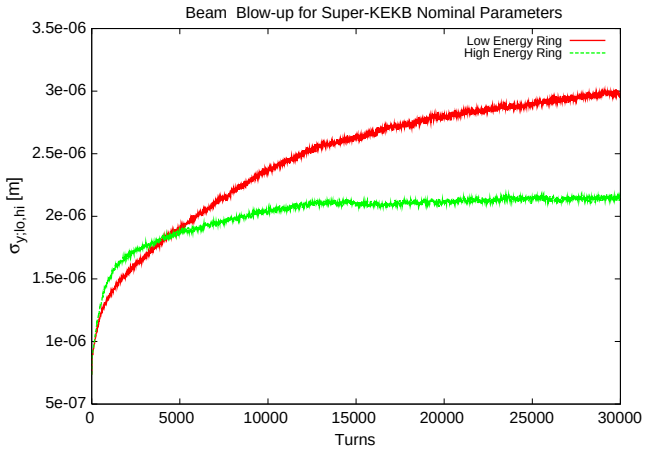



Fig. 5. Beam size vs. time for NIMZOVICH SuperKEKB benchmark example.

$2N_x$ of $2N_x + 2N_y$ FFT's right away. $2N_x$ other transformations can be done out-of-place. Also, the parallel transposition becomes simpler, as the padding space can serve as a scratch space, so send and receive operations can be done simultaneously and asynchronously, decreasing latency. We have implemented this scheme for the special case of symmetric $G$ functions and observe a speed gain of almost a factor of 2.

### 3.3. Slice-to-slice interaction and adaptive slices

The longitudinal domain decomposition makes use of Hirata slicing [5]. For field calculation, we make use of a convex interpolation method proposed by Ohmi [11] to avoid field discontinuities close to slice boundaries (at the cost of having to do two field calculations for each timestop). We use the same scheme to calculate the luminosity with high accuracy by sampling the opposing bunch's charge density with the macroparticles.

Using this scheme, each slice will execute grid operations (sampling fields or depositing particles) on four different temporal positions. In a beam with a pronounced hourglass effect, the transverse dimensions of the beam might vary substantially for these times. We adapt the transverse extensions of the grids to the expected extensions of the beam, calculated from the unperturbed Twiss functions. This way, we achieve constant effective resolution across the interaction process and can use a lower-resolution grid than codes with grids of constant absolute resolution. For each slice, we have to pre-calculate two $\hat{G}$ matrices for each opposing slice, as $\hat{G}$ does not follow a simple scaling law under temporal displacement for $\beta_x \neq \beta_y$. We are currently testing a dynamic scheme in which the grid sizes are adapted to the beam dimensions as measured during the course of the simulation, which would relieve the user of having to have an estimate of beam size increase.

### 3.4. Results

We present a typical result, a single-bunch luminosity simulation for Super-KEKB with parameters as given by Ohmi et al. [8]. We observe good agreement with the results in Ref. [8], obtained by other codes (Figs. 4 and 5).

## References

[1] TESLA Technical Design Report, 2001.
[2] Linac Coherent Light Source (LCLS) conceptual design report, SLAC-R-593.
[3] H.H. Braun, et al., Phys. Rev. ST Accel. Beams 3 (2000) 124402.
[4] A.C. Kabel, P. Emma, Peak current optimization for LCLS Bunch Compressor 2. Proceedings of the 9th European Particle Accelerator Conference (EPAC 2004).
[5] K. Hirata, H. Moshammer, F. Ruggiero, Part. Accel. 40 (1993) 205.
[6] M. Bassetti, G. Erskine, CERN ISR TH/80-06, 1980.
[7] F. Matta, A. Reichel, Math. Comp. 25 (1971) 339.
[8] K. Ohmi, M. Tawada, Y. Cai, S. Kamada, K. Oide, J. Qiang, Phys. Rev. Lett. 92 (2004) 214801-1.
[9] M. Frigo, S.G. Johnson, FFTW 2.15 User's Manual, ⟨http://www.fftw.org/fftw2_doc/⟩
[10] R.W. Hockney, J.W. Eastwood, Computer Simulation Using Particles, Bristol and Philadelphia, 1988.
[11] K. Ohmi, in: Proceedings of the 2003 IEEE Particle Accelerator Conference.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 168–174
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# High-performance computing in accelerating structure design and analysis

Zenghai Li[a,*], Nathan Folwell[a], Lixin Ge[a], Adam Guetz[a], Valentin Ivanov[a], Marc Kowalski[a], Lie-Quan Lee[a], Cho-Kuen Ng[a], Greg Schussman[a], Lukas Stingelin[b], Ravindra Uplenchwar[a], Michael Wolf[c], Liling Xiao[a], Kwok Ko[a]

[a]*Stanford Linear Accelerator Center, Stanford University, USA*
[b]*Paul Scherer Institute, Switzerland*
[c]*University of Illinois, Urbana-Champaign, USA*

Available online 1 December 2005

**Abstract**

Future high-energy accelerators such as the Next Linear Collider (NLC) will accelerate multi-bunch beams of high current and low emittance to obtain high luminosity, which put stringent requirements on the accelerating structures for efficiency and beam stability. While numerical modeling has been quite standard in accelerator R&D, designing the NLC accelerating structure required a new simulation capability because of the geometric complexity and level of accuracy involved. Under the US DOE Advanced Computing initiatives (first the Grand Challenge and now SciDAC), SLAC has developed a suite of electromagnetic codes based on unstructured grids and utilizing high-performance computing to provide an advanced tool for modeling structures at accuracies and scales previously not possible. This paper will discuss the code development and computational science research (e.g. domain decomposition, scalable eigensolvers, adaptive mesh refinement) that have enabled the large-scale simulations needed for meeting the computational challenges posed by the NLC as well as projects such as the PEP-II and RIA. Numerical results will be presented to show how high-performance computing has made a qualitative improvement in accelerator structure modeling for these accelerators, either at the component level (single cell optimization), or on the scale of an entire structure (beam heating and long-range wakefields).


*PACS:* 29.17.+w; 94.20.ws; 03.50.De; 07.05.Tp; 01.50.hv; 02.70.-c

*Keywords:* Acceleration; Linear collider; Cavity design; Wakefield; Parallel electromagnetic codes; Unstructured grid; High-performance computing; Large-scale simulation

## 1. Introduction

Particle accelerators are among the most important and most complex scientific instruments in use, and are critical to research in fields such as high-energy physics, nuclear physics, materials science, chemistry, and the biosciences. As new and existing facilities continually strive towards higher energy, higher beam current, and greater efficiency, accelerator physicists and engineers are faced with increasingly demanding specifications on the RF system to improve performance and reduce cost. As a result, the emphasis of designing the accelerating structures for these machines has been placed heavily on numerical modeling as the cost-saving approach to their R&D. This means the accuracy and reliability of the modeling software are becoming of paramount importance in order that the structures can meet the stringent design requirements.

The Damped, Detuned Structure (DDS) shown in Fig. 1 is the baseline linac design for the warm Linear Collider (NLC) [1] scheme. In the DDS, the frequency of the accelerating field must be accurate to within 1 part in 10,000 to maintain acceleration efficiency. This requirement has to be met in a complex cavity geometry that optimizes the accelerating field gradient while suppresses the long-range dipole wakefields. To provide the desired

*Corresponding author.
*E-mail address:* lizh@slac.stanford.edu (Z. Li).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.080

Fig. 1. Model of the 55-cell NLC DDS design.

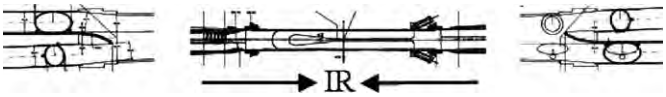



Fig. 2. Beamline complex of the PEP-II Interaction Region (IR).

accuracy in the DDS cell frequency and to verify the wakefield suppression by damping and detuning over the entire structure are modeling challenges that are beyond the capabilities of standard electromagnetic (EM) codes (e.g. MAFIA [2] and HFSS [3]) running on limited computing resources like desktop computers.

## 2. High-performance computing

Besides the NLC DDS design, the PEP-II [4] Interaction Region (IR), the RIA [5] RFQ cavity and the PSI [6] ring cyclotron are additional examples for which more advanced simulation tools other than available software are needed. In the PEP-II, beam heating in the IR due to trapped modes is an obstacle to high current operation, and thus high luminosity. Fig. 2 shows the IR beamline complex which consists of a central vacuum chamber of complicated, varying cross-sections (due to the synchrotron masks) that connects to the positron and electron beamlines via a crotch junction at both ends. Modeling the entire geometry is necessary to fully account for the beam heating effect. The proposed RIA plans to employ RFQ structures such as those shown in Fig. 3 in its low-energy linacs. Presently, provisions have to be made for tuners to compensate for frequency errors of about 1% from using existing design software. New modeling tools that can improve the accuracy by an order of magnitude would lead to a significant reduction in tuners required and a much simplified operation as well.

To address these modeling challenges, SLAC has embarked on a code development effort that was first initiated in 1997 under the support of the DOE Grand Challenge on Accelerator Physics, with the goal to implement high-performance computing (HPC) capabilities in accelerating modeling tools. This work has expanded with the follow-up DOE SciDAC Accelerator Simulation project in which SLAC leads the team that specifically targets high accuracy, large-scale electromagnetic applications.

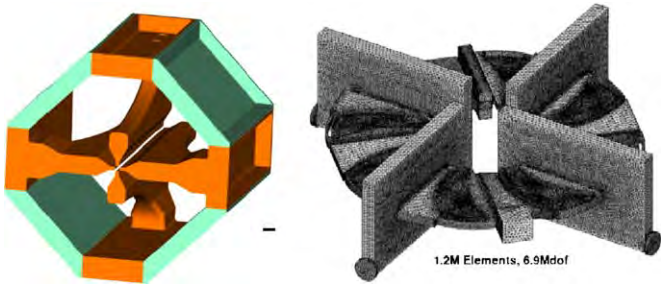



Fig. 3. The RIA RFQ and the PSI ring cyclotron.

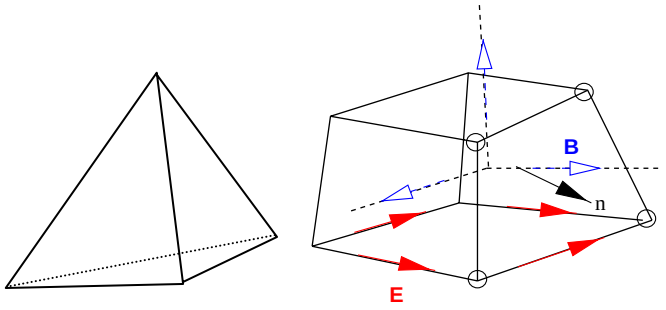



Fig. 4. Tetrahedral mesh for Omega3P/S3P/T3P and hexahedral mesh for Tau3P.

## 3. Parallel electromagnetic codes on unstructured grids

The suite of 3D, parallel electromagnetic codes that are finite element based consists of:

(1) Omega3P—eigenmode solver for finding normal modes in lossless and lossy cavities.
(2) S3P—solver in frequency domain to calculate S parameters of open structures.
(3) T3P—time-domain solver for modeling response due to beam, dipole and waveguide excitation, and it also includes:
(4) Tau3P—time-domain solver following the Discrete Surface Integral (DSI) formulation with same functionalities as T3P.
(5) Track3P—module for dark current simulation with surface physics using fields from solvers above.
(6) Viz3D—analysis and graphics package.

The finite element codes employ tetrahedral mesh elements while the DSI-based Tau3P uses hexahedral cells (Fig. 4), both unstructured grids able to conform to curved surfaces for very high accuracy modeling. The set of codes is developed under C++ with a unified data structure to facilitate geometry input and partitioning, and uses MPI for communication on distributed memory architectures.

The code development at SLAC is supported by an extensive and coordinated R&D program in computer science and applied mathematics that is sponsored by SciDAC and carried out in designated national laboratories and universities. Among these efforts include parallel meshing at Sandia and University of Wisconsin, partitioning at Sandia and Lawrence Berkeley Lab (LBL), linear solvers and eigensolvers at Stanford and LBL, adaptive refinement at RPI, and visualization at UC Davis. They contribute to the success of the large-scale simulations required for the challenging accelerator applications described above. We will next present two examples of

high resolution modeling and three examples of system scale simulations using the new HPC tools.

## 4. High-resolution design

The DDS cell is one of 206 cavities in the NLC first accelerating structure design whose cavity shape is optimized for high shunt impedance while connected to four damping manifolds that run the length of the structure for long-range dipole wakefield suppression. The wake is further reduced by a cavity-to-cavity variation that detunes the dipole modes in a Gaussian manner. As previously indicated, the fundamental mode frequency of the DDS cell has to be accurate to 0.01% to maintain acceleration efficiency. Applying Omega3P to the distributed model shown in Fig. 5 on NERSC's IBM/SP, a table of dimensions for all 206 different cavities along the DDS was generated for computerized machining based on calculations that met this accuracy requirement. Cold tests on fabricated cells based on this table indeed showed that their measured frequencies are within the targeted value (Fig. 5). This result demonstrates that high-resolution modeling utilizing unstructured grids and parallel computing is a powerful tool for designing the NLC structures which potentially could have saved the project more than $100 million in machine cost alone. Furthermore, the resolution and speedup provided by HPC together with high-precision machining have enabled simulation-based design to become a cheaper, faster alternative to the traditional expensive and time-consuming R&D process of repeated fabrication and testing.

### 4.1. Adaptive mesh refinement

Under SciDAC, SLAC and RPI are collaborating on developing an adaptive mesh refinement (AMR) capability in Omega3P to improve the accuracy and convergence of frequency and quality factor calculations for cavities of complex shapes. Accurate wall loss determinations are difficult when the wall currents are localized in narrow regions of the cavity, such as around the coupling iris. The increased wall loss reduces the cavity's quality factor, thus the shunt impedance, and can also lead to RF surface heating at high power. An adaptive mesh control loop based on error indication procedures has been implemented in Omega3P to provide increasingly refined meshes until a converged result is obtained. The adaptive procedure has been applied to the RIA RFQ cavity with the following results. Using measurement [7] as benchmark, the percentage deviations in frequency and Q from a standard code are 1.5 and 30, respectively, while the results from Omega3P using uniform refinement are 0.05 and 16.1 and with AMR the numbers are 0.11 and 16.2. The AMR case uses 0.4 million DOFs, 10 times less than the uniform refinement case to reach the same accuracy. We conclude that combining unstructured grid, third order elements and AMR, Omega3P's improvements over standard codes are an order of magnitude better in frequency and about a factor of two better in Q (Fig. 6).

## 5. System scale modeling (DDS)

The h60vg3 [8] X-band structure is a 55-cell DDS that is considered to be the baseline design for the NLC linac. Due to the complicated geometry, the long-range dipole wakefields in this structure have thus far been analyzed by an approximate equivalent circuit model representing only the lowest two dipole bands. It is of both theoretical and practical interests to be able to simulate the entire, realistic structure so that the effect of the higher bands and that due to the input/output couplers can be included. Modeling a problem of such size and complexity has not been tried before because the computational resources required were considered prohibitive. Using the parallel tools developed under SciDAC and the computing resources it provides on NERSC's IBM SP2 machine, such

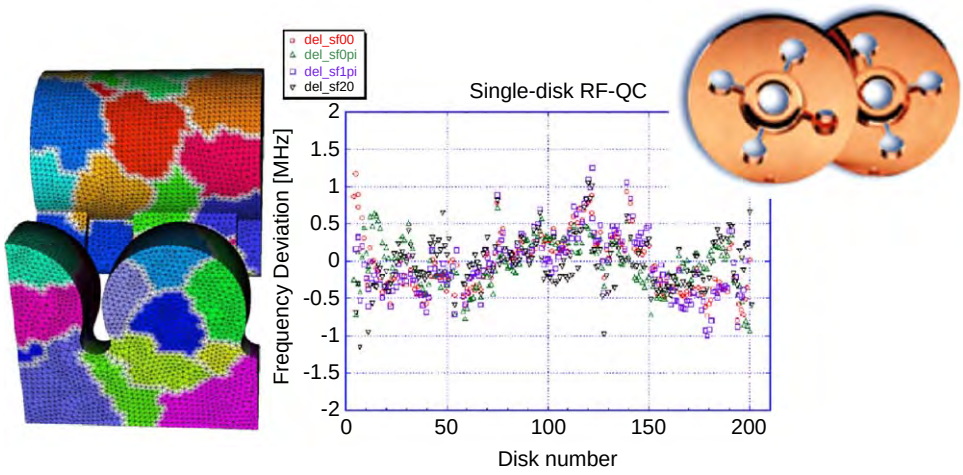



Fig. 5. (Left) A distributed Omega3P model of 1/8 of the DDS cell, (right) microwave QC of the fabricated cells (insert) showing measured frequencies within 0.01% of target value.

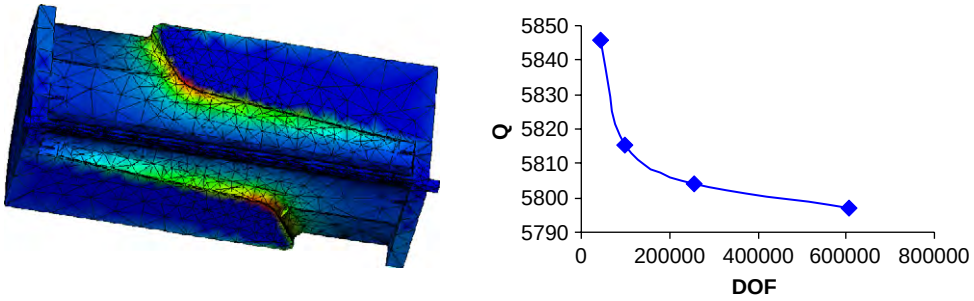



Fig. 6. Adaptively refined mesh and wall loss for the RFQ cavity and the convergence of Q vs. DOF.

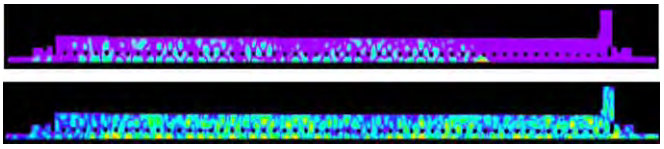


Fig. 7. Time domain Tau3P simulation of a rigid beam traversing the DDS structure, showing field distribution at two instances in time.

a system-scale simulation is within reach. We will next present the computations of the wakefields in the h60vg3 structure via time and frequency domain methods.

### 5.1. Time-domain simulation

The parallel time-domain solver Tau3P has been used to simulate the h60vg3 structure with a transit beam so that the wakefield can be found directly for the first time. Fig. 7 shows the beam-excited electric fields in the structure at two instances in time. The top snapshot shows the field excitation as the beam traverses the structure. The bottom one is the field distribution after the beam left the structure, clearly showing the coupling of the wakefield out to the manifold and subsequently out of the structure through the HOM coupler at the downstream end. The evolution of the wakefield from the Tau3P simulation provides insight into the HOM damping via the manifold.

The wakefield is plotted in Fig. 8 together with the results from the same structure but without the manifolds. The difference between the DS (Detuned Structure) and DDS wakefields due to manifold damping at long distances is clearly evident. Fig. 9 is the impedance spectrum showing the contributions from the higher bands up to 30 GHz which are absent from the two-band equivalent circuit model.

### 5.2. Frequency-domain calculation

The wakefield in the h60vg3 structure can also be found by summing the eigenmodes in the frequency domain. Because of the power loss through the fundamental and HOM couplers, a complex solver has been developed in Omega3P to calculate damped modes. In the computational model, lossy materials are used as matched terminations for the fundamental and HOM couplers in the frequency range of interest. About 400 complex eigen-

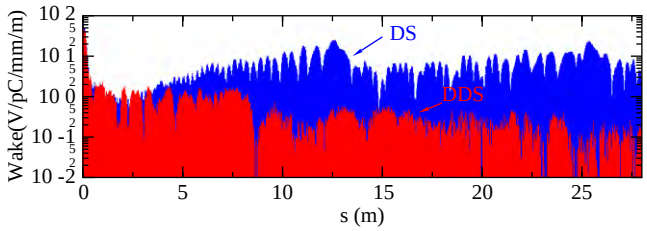



Fig. 8. Dipole wakefield in the 55-cell h60vg3 DDS in red as compared with that of the same structure but without the manifold in blue.

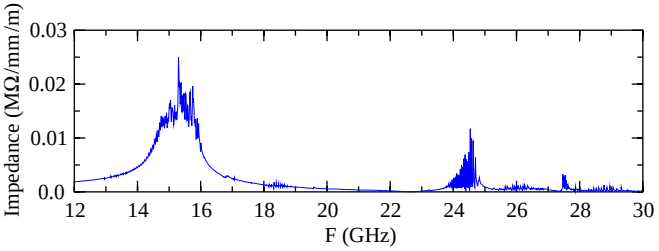



Fig. 9. Impedance spectrum of h60vg3 wakefield from Tau3P simulation covering up to 30 GHz.

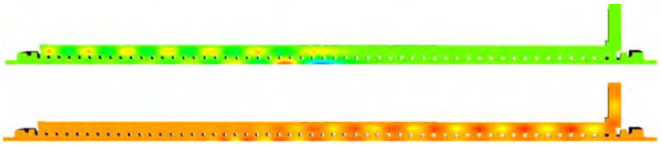


Fig. 10. (Top) Complex eigenmode in the h60vg3 structure that couples to a backward wave in the manifold; (bottom) mode couples to a forward wave that gets damped via the HOM coupler down stream.

modes covering the spectrum up to 30 GHz have been calculated to form the sum for the wakefield. In Fig. 10 are shown two typical modes in the DDS structure. The top plot is a mode that couples to the backward wave in the manifold, suggesting the need for a HOM load at the input end. The bottom plot shows a mode that couples to a forward wave in the manifold that is damped by the HOM load downstream. This is the first time these modes have been seen in a long, realistic structure.

The mode spectrum from Omega3P as depicted in Fig. 11 shows good agreement when compared with the

Tau3P result of Fig. 9. The agreement is even closer on the wakefields as shown in Fig. 12 which is remarkable considering the results have been obtained from two different methods each with its mesh model of the structure. This serves not only to validate the two codes but to also verify the effectiveness of the detuning and damping scheme for wakefield suppression. Furthermore, it points out the importance of high-performance computing and the SciDAC support in computational science research, without which such calculations would not have been possible.

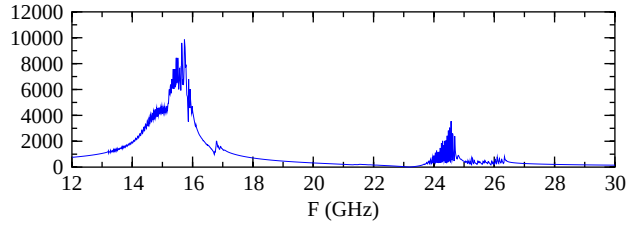



Fig. 11. Impedance spectrum of h60vg3 wakefield from Omega3P analysis.

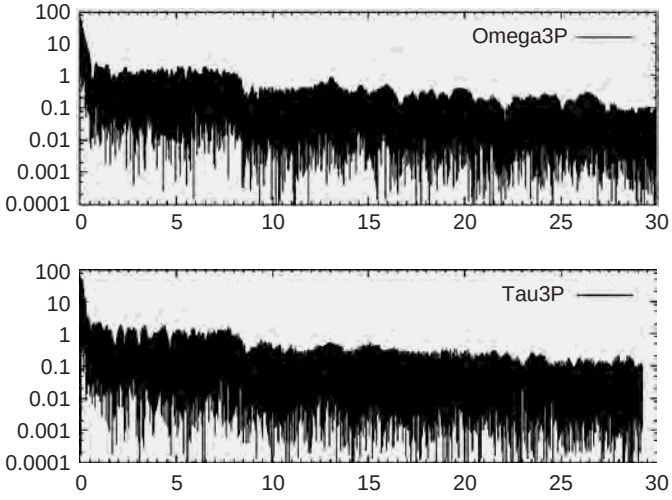



Fig. 12. Omega3P and Tau3P wakefields out to 30 m behind leading bunch, showing detuning effect at shot range, and damping effect at long range.

### 5.3. Advances in eigenmode solvers

The success of the large-scale simulations just presented relies heavily on the accuracy, convergence and scalability of the solvers used. SciDAC supports a strong effort consisting of SLAC, Stanford and LBL to develop new, efficient algorithms for solving the most challenging applications. They include a linear solver framework that interfaces Omega3P/S3P to direct solvers such as SuperLU and WSMP, leading to an improvement in speed by a factor of 50–100. Implementation of the AV FEM formulation adds a gain of 5–10 for using iterative solvers. In addition, higher-order hierarchical bases (up to $p = 6$) and preconditioners have been developed to increase accuracy and convergence. With these advances we have been able to solve the h60vg3 structure using up to 93 million DOFs, requiring about 800 GB of memory (Fig. 13 left), and they helped in reaching higher accuracy in AMR calculations. Furthermore, work on complex solvers has enabled the mode analysis of the h60vg3 and the PEP-II IR absorber design to be discussed later.

### 5.4. Research in domain decomposition

The scalability of parallel time-domain codes such as Tau3P is hampered by communication costs since the fields have to be updated at every time step, leading to poor parallel efficiency. Under SciDAC SLAC is working with SNL and LBL to improve Tau3P performance by exploring alternate schemes within the Zoltan partitioning library. For the Tau3P simulation of the h60vg3 55-cell structure, Fig. 13 (right) shows that the RCB1D partitioning provides better speedup over the existing partitioning using ParMETIS by reducing the number of neighboring MPI processes.

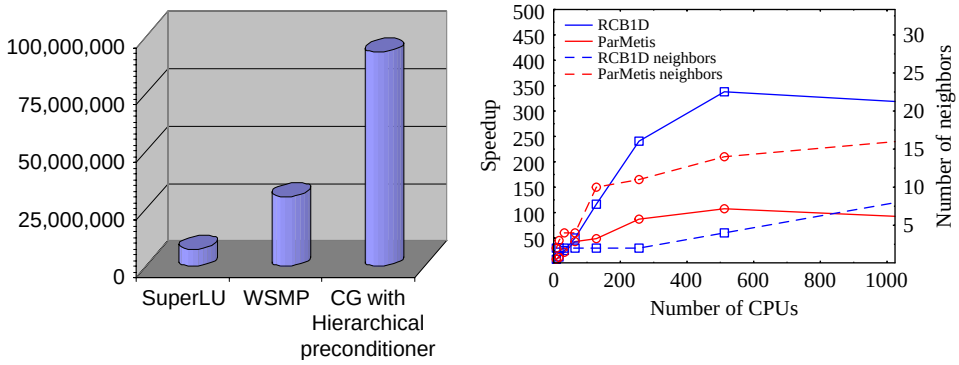



Fig. 13. (Left) Largest problem size in DOFs with three linear solvers. Maximum reached is 93 million; (right) Speedup gain using RCB1D over ParMETIS.

## 6. Simulating the PEP-II IR

A distributed model of the PEP-II IR shown in Fig. 14 is used in Tau3P simulation of two colliding beams traversing the beamline complex near the interaction point. Two snapshots of the simulation before and after the beams pass each other are shown in Fig. 15. The beam-excited fields are analyzed to find the trapped modes that contribute to local heating effects. These heating calculations were factored into the upgrade of the IR and the machine now is able to operate at 15% higher beam current without overheating.

In preparation for further increase in beam current a new IR design is under consideration that uses absorbers to damp trapped modes and reduce beam heating. The absorbers are located at the crotch region on both ends of the central beam pipe (see Fig. 16). Their effectiveness is studied using the complex eigensolver in Omega3P. Fig. 17 compares the quality factor of the trapped modes between the crotches with and without the absorbers. It is seen that most of the modes have been damped by two orders of magnitude. This simulation capability provides a powerful tool for optimizing the absorber design without extensive prototyping and testing.

Fig. 14. Distributed model of PEP-II IR for Tau3P simulation of beam heating.

Fig. 15. Beam-excited fields in PEP-II IR from Tau3P simulation before and after two beams passing each other at the collision point.

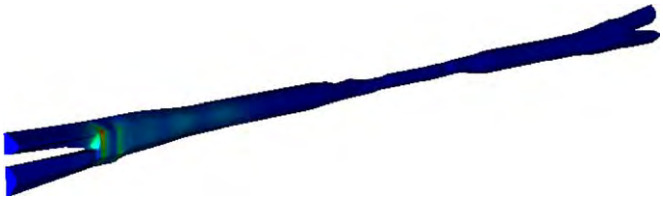


Fig. 16. Omega3P model of the PEP-II IR including placement of absorbers at the crotches.

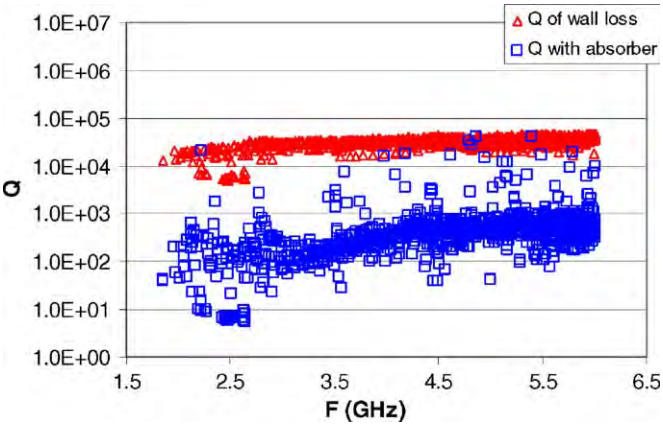



Fig. 17. Complex Omega3P calculation of damping effect on trapped modes in PEP-II IR using absorbers at the crotches (undamped modes in red; damped modes in blue).

## 7. Modeling entire ring cyclotron

The availability of new parallel codes such as Omega3P has generated interests in finding the HOM modes that exist in an entire ring cyclotron for better understanding of their effect on the beam dynamics in such a machine. PSI and SLAC are collaborating on modeling the PSI ring cyclotron with Omega3P and determining the HOM effects through an eigenmode analysis. The calculations have been performed using the 32 CPU IBM/SP4 at PSI. Using the ESIL solver in Omega3P, 280 modes with frequency close to a beam harmonic have been computed. They can be classified into three types. Forty four of them are cavity modes with low frequency and high gap voltage. There are 18 vacuum chamber modes having medium frequency and low gap voltage. The rest of the modes are hybrid modes that occupy both the cavity and the vacuum chamber, and these have high frequency and low gap voltage. In Fig 18, sample field pattern of the three types of modes in an entire ring cyclotron is shown for the first time.

## 8. Summary

A new suite of parallel electromagnetic codes based on unstructured grids has been developed under the DOE SciDAC project to model large, complex accelerator cavities and RF structures that have not been possible with standard software. It is demonstrated that the high performance computing capability in these codes has enabled the large-scale simulations that are necessary for high resolution cavity design and system-scale structure analysis. R&D in computer science and applied mathematics through SciDAC collaborations have shown to further improve their accuracy and efficiency. The codes have been applied to optimize existing accelerators and design planned facilities with considerable success.

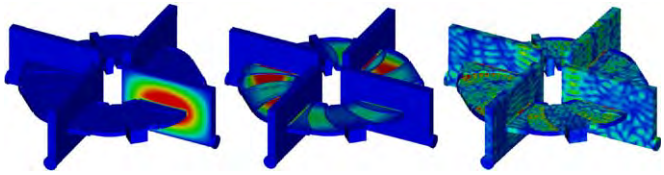


Fig. 18. Eigenmodes in the PSI Ring Cyclotron: cavity mode (left), vacuum chamber mode (middle), and hybrid mode (right).

## Acknowledgment

Work supported by the US DOE Contract no. DE-AC02-76SF00515.

## References

[1] ZDR report for the Next Linear Collider, LBNL-PUB-5424, SLAC Report 474, UCRL-ID-124161, 1996.
[2] MAFIA User's Guide, CST GmbH, Darmstadt, Germany.
[3] HFSS: http://www.ansoft.com/products/hf/hfss.
[4] PEP-II: http://www.slac.stanford.edu/accel/pepii.
[5] RIA: http://www.orau.org/ria.
[6] PSI: http://gfa.web.psi.ch.
[7] N. Vinogradov, et al., Progress with the room temperature structures for the RIA facility, RIA R&D Workshop, August 26–28, 2003.
[8] Z. Li, et al., Optimization of the X-band structure for the JLC/NLC, proceedings of PAC2003, Portland, 2003.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 175–183
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Multipole expansion solution of the Laplace equation using surface data

Shashikant Manikonda*, Martin Berz

*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48823, USA*

Available online 1 December 2005

**Abstract**

This paper provides a computational method to model a three-dimensional static electromagnetic field within a finite source free volume starting from discrete field information on its surface. The method uses the Helmholtz vector decomposition theorem and the differential algebraic framework of COSY INFINITY to determine a solution to the Laplace equation. The solution is locally expressed as a Taylor expansion of the field which can be computed to arbitrary order. It provides a natural multipole decomposition of the field which is required for the computation of transfer maps, and also allows to obtain very accurate finite element representations with very small numbers of cells.


*PACS:* 41.20.Gz; 41.20.Cv

*Keywords:* Laplace equation; COSY Infinity; Helmholtz vector decomposition theorem; Differential algebra; Multipole expansion; Surface data

## 1. Introduction

The detailed simulation of particle trajectories through magnets in spectrographs and other large acceptance devices requires the use of detailed field information obtained from measurements. Likewise, for high-energy accelerators like the LHC, higher order description of the beam dynamic via one-turn maps is required to study the long-term beam stability [1,2]. The construction of such high-order one-turn truncated Taylor maps [3] requires the precise information of the electromagnetic field in the individual electromagnetic components (quadrupoles, dipole, sextupoles, etc.) of the lattice.

It is commonly known that for a device that satisfies mid-plane symmetry, the entire field information can be extracted from the data in the mid-plane of the device [3]. However, it is well known that this method has limitations in accurately predicting nonlinear field information outside the immediate vicinity of the mid-plane because the extrapolation requires the computation of higher order derivatives of in-mid-plane data, which is difficult to do with accuracy if the data is based on measurements. Thus, it is particularly useful to employ techniques that rely on field measurements outside the mid-plane. In particular, in modern particle spectrographs it is common to measure the fields on a fine mesh on 2–4 planes outside the mid-plane. These data have frequently been used to model the overall field as a superposition of point-charge fields of so-called image charges [4,5]. However, the computational effort required for this approach is large, as it requires the inversion of a matrix with a dimension equal to that of the number of image charges.

However, the out-of-plane field measurements in essence provide field data on the top and bottom surfaces of a box containing the region of interest through which the beam passes. If the planes extend outward far enough to a region where the fringe field becomes very small, or can easily be modeled, and inwards far enough that the field becomes rather homogenous, field data are known on an entire surface enclosing the region of interest. The method we present in this paper can extract the field information as a multipole expansion in the volume of interest if a discrete set of field measurements are provided on a closed surface enclosing the volume of interest.

*Corresponding author. Tel.: +1 517 355 9200; fax: +1 360 294 3258.
*E-mail addresses:* manikond@msu.edu (S. Manikonda), berz@msu.edu (M. Berz).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.081

Thus, the field computation problem can be viewed as solving a boundary value problem for the three-dimensional (3D) Laplace equation for the field, i.e. to obtain the solution of the PDE

$$\nabla^2\psi(\vec{r}) = 0 \quad \text{in the volume } \Omega \subset \mathbb{R}^3$$

$$\text{where } \nabla\psi(\vec{r}) = \vec{f}(\vec{r}) \text{ is specified on the surface } \partial\Omega.$$

The existence and uniqueness of the solution for the 3D case can easily be shown through the application of Green's formulae. In particular, it is well known that the component of $\vec{\nabla}\psi$ normal to the surface is sufficient to specify the solution; however, in the following we will also make use of the additional field components available from measurements. It is well known that the analytic closed-form solution for the 3D case can be found for problems with certain regular geometries where a separation of variables can be performed. However, in most practical 3D cases, numerical methods are the only way to proceed. Usually the finite difference or finite element approach are used to find the numerical solution as data set in the region of interest. But because of their relatively low approximation order, for the problem of precise solution of PDEs, the methods have limited success because of the prohibitively large number of mesh points required. The method we present in this paper can determine local finite elements of in principle unlimited order.

In Section 2 we discuss the benefits of using the boundary data and present the analytic closed form solution for the 2D case that can be easily found by application of Cauchy's integral formula. We then use a 2D example to highlight the advantages of the methods that use the boundary data to compute the solution. In Section 3 we present the theory and the implementation of the new scheme to find the solution of the 3D Laplace equation when the gradient of the solution is specified on the surface enclosing the volume of interest . This scheme is based on the Helmholtz theorem and the tools of the code COSY Infinity [3,6,7]. In Section 4 we present an application of this new scheme to a theoretical bar magnet problem. We also address the results of an application to the study of a dipole magnet of the MAGNEX spectrograph.

## 2. Methods using boundary data

Boundary data methods such as those utilized below are based on a description of the interior field in terms of particular surface integrals involving the surface data. These approaches have various advantages. Firstly, the solution is analytic in terms of the interior variables, even if the boundary data fail to be differentiable or are even piecewise discontinuous; all such non-smoothness is removed after the integration is executed. Hence a Taylor polynomial approximation in terms of interior variables can be performed; and we expect that a Taylor approximation of a certain order will provide an accurate approximation over suitable domains.

Secondly, since for the PDEs under consideration here the solution functions are known to assume their extrema on the boundary because of analyticity or harmonicity, a method that uses boundary data is expected to be robust against errors in those boundary data with errors in the interior not exceeding the errors on the surface. Thirdly, if the boundary data given have statistical errors, such errors have a tendency to even average out in the integration process as long as the contributions of individual pieces of integration are of similar significance. Thus, we expect the error in the computed field in the interior to be generally much smaller than the error in the boundary data. This ensures that the methods using boundary data are computationally stable.

### 2.1. The 2D case

As an introduction to the general approach, we begin with the discussion of the 2D case, the theory of which can be fully developed in the framework of elementary complex analysis, and which also describes the situation of static electric or magnetic fields as long as no longitudinal field dependence is present. It is based on the use of Cauchy's integral formula stating that if the function $f$ is analytic in a region containing the closed path $C$, and if $\alpha$ is a point within $C$, then

$$f(\alpha) = \frac{1}{2\pi i}\oint_C \frac{f(z)}{z-\alpha}\,dz \tag{1}$$

where the integral denotes the path integral over $C$. Cauchy's formula is an integral representation of $f$ which permits us to compute $f$ anywhere in the interior of $C$, knowing only the value of $f$ on $C$. This integral representation of $f$ is also the solution of the 2D Laplace equation for the primitive of $(\mathrm{Re}(f), -\mathrm{Im}(f))$ with the function $f$ specified on the path $C$.

Now, suppose a random error of $\delta(z)$ is introduced in the measured data around the path $C$. Then by Eq. (1) we can compute the error $E(a)$ introduced in the computation of $f(\alpha)$ at some point $\alpha$ inside $C$ as

$$\begin{aligned} E(\alpha) &= \frac{1}{2\pi i}\oint_C \frac{f(z)+\delta(z)}{z-\alpha}\,dz - f(\alpha) \\ &= \frac{1}{2\pi i}\oint_C \frac{\delta(z)}{z-\alpha}\,dz \end{aligned} \tag{2}$$

We note that while $E(\alpha)$ is given by a Cauchy integral, $E$ need not be analytic since $\delta(z)$ need not assume the function values of an analytic function. In fact, if it would, then it already would be uniquely specified on any dense subset $S$ of $C$, which removes the freedom for all values of $E$ on points on $C$ that are not in $S$.

While the error $E$ itself may be bounded in magnitude, if the integral is approximated by one of the conventional numerical quadrature methods, the result can become singular as the point $\alpha$ approaches the boundary $C$. This case may limit the practical use of the method and needs to be studied carefully. As an example, we consider the case of

quadrature based on adding the terms of a Riemann sum, i.e. the approximation

$$\frac{1}{2\pi i}\oint_C \frac{\delta(z)}{z-\alpha}\,dz \approx \frac{1}{2\pi i}\sum_{j=1}^{N_z}\frac{\delta(z_j)}{(z_j-\alpha(r))}(z_j - z_{j-1}) = \tilde{E}(\alpha) \quad (3)$$

where the $N_z$ points $z_j$ are spaced equidistantly around $C$; since $C$ is closed, $z_0 = z_{N_z}$. By studying the approximation $\tilde{E}(a)$ as the point $\alpha$ approaches the boundary $C$, we can analyze the stability of the method with respect to the discretization of the path $C$.

As an example, we choose the path $C$ as a circle of radius $R$ enclosing the region of interest. We assume a random error of $\delta(z)(\pm 10^{-2})$ is introduced in the measured data around $C$. The point $\alpha$ is given by $r\exp(i\phi)$ and the points $z_j$ are given by $R\exp(i2\pi j/N_z)$ for $j = 0,\ldots,N_z$. Letting $\delta_m(z_j)$ denote the error assigned to point $z_j$ in error set $m$, for each of these error sets we express the Riemann sum $\zeta_m(\alpha)$ for point $\alpha$ by

$$\zeta_m(\alpha) = \frac{1}{2\pi i}\sum_{j=1}^{N_z}\frac{iz_j\delta_m(z_j)}{(z_j-\alpha(r))}\frac{2\pi}{N_z}.$$

We then form the average of the magnitude of the error over $N_e$ error sets to obtain

$$\eta(r) = \frac{1}{N_e}\sum_{m=1}^{N_e}|\zeta_m(\alpha)|.$$

Note that $\eta(r)$ still depends on the phase $\phi$. However, in the statistical limit there is apparently invariance under rotation by $\exp(i2\pi/N_z)$; and one quickly sees that there are two limiting cases for the choice of the phase. These are the case $\phi = 0$, where the $\alpha$ will eventually collide with the $z_j$ for $j = N_z$ as $r \to R$ and thus a "worst case" divergence will appear, and the case $\phi = 2\pi/2N_z$, in which case the $\alpha$ will approach the mid-point between $z_j$ for $j = N_z$ and $z_j$ for $j = 1$ as $r \to R$. Choosing sufficiently fine discretization of the path and sufficiently many error sets $\delta_m$, the quantity $\eta(r)$ for these two cases will be a good measure for the accuracy that can be achieved with the surface integral method.

For our specific example, we choose random errors of maximum magnitude $10^{-2}$ at $N_z = 10,000$ points on the circle of radius $R = 2$. For each value of $r$, we perform the computation for a total of $N_e = 10,000$ error sets. The results of this analysis are shown as plots in Figs. 1 and 2 for the two phases that represent the "worst case" and the "best case" situation.

We first observe that sufficiently away from the surface, the expected smoothing effect is happening, and the errors in the function values are indeed well below the errors assumed on the surface. A rough quantitative analysis shows that this error is about two orders of magnitude below the surface data error, corresponding well with the statistically expected decrease of the error by $1/\sqrt{N_z}$. As $\alpha$ approaches the curve closer than $10^{-3}$, in the "best case" situation, the error rises to about $10^{-2}$, which is because now only nearby grid points contribute to the sum and thus the smoothing effect disappears. In the "worst case" scenario, divergence actually happens; but the average error is still at the level of the original random error of $10^{-2}$ for values of $r$ that are only about $10^{-4}$ away from the radius 2.

So overall we see that the method performs significant smoothing, and even with the simplest discretization as a Riemann sum, good accuracy is maintained even as we approach $C$. We note in passing that with more sophisticated quadrature methods, for example, those based on Gaussian methods [3], the divergence effect can be significantly controlled.

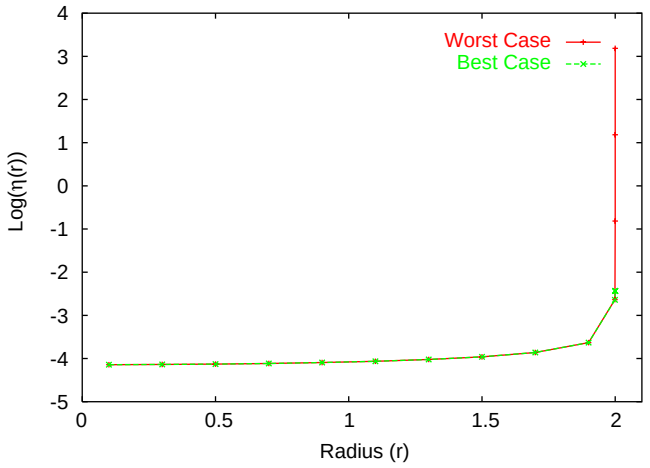



Fig. 1. The plot shows the dependency of $\eta(r)$ on the radius $r$. The $Y$-axis represents $\log(\eta(r))$ and the $X$-axis represents the radius $r$. Ten thousand error sets ($N_e$) around the circle or radius $R = 2$ were chosen for the analysis. We show the plot for both the best and the worst case scenario.

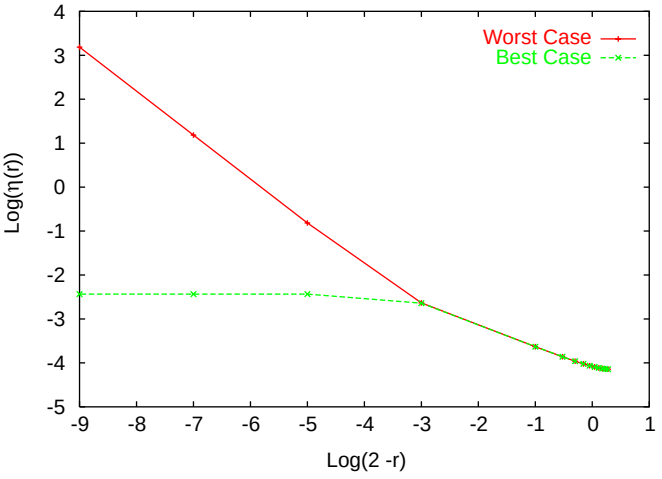



Fig. 2. The plot shows the dependency of $\eta(r)$ on the radius as the radius $r$ approaches the boundary. The $Y$-axis represents $\log(\eta(r))$ and the $X$-axis represents $\log(2-r)$. Ten thousand error sets ($N_e$) around the circle or radius $R = 2$ were chosen for the analysis. We show the plot for both the best and the worst case scenario.

## 3. The 3D case

The scheme we use for the 3D case is based on the Helmholtz vector decomposition theorem [8–14]. We begin by representing the solution of the PDE via Helmholtz' theorem, which states that any vector field $\overrightarrow{B}$ which vanishes at infinity can be written as the sum of two terms, one of which is irrotational and the other solenoidal as

$$\overrightarrow{B}(\vec{x}) = \vec{\nabla} \times \vec{A}_t(\vec{x}) + \vec{\nabla}\phi_n(\vec{x}) \tag{4}$$

$$\phi_n(\vec{x}) = \frac{1}{4\pi}\int_{\partial\Omega} \frac{\vec{n}(\vec{x}_s) \times \overrightarrow{B}(\vec{x}_s)}{|\vec{x}-\vec{x}_s|}\,\mathrm{d}s - \frac{1}{4\pi}\int_{\Omega} \frac{\vec{\nabla} \times \overrightarrow{B}(\vec{x}_v)}{|\vec{x}-\vec{x}_v|}\,\mathrm{d}V$$

$$\vec{A}_t(\vec{x}) = -\frac{1}{4\pi}\int_{\partial\Omega} \frac{\vec{n}(\vec{x}_s) \times \overrightarrow{B}(\vec{x}_s)}{|\vec{x}-\vec{x}_s|}\,\mathrm{d}s + \frac{1}{4\pi}\int_{\Omega} \frac{\vec{\nabla} \times \overrightarrow{B}(\vec{x}_v)}{|\vec{x}-\vec{x}_v|}\,\mathrm{d}V.$$

Here $\partial\Omega$ is the surface which bounds the volume $\Omega$. $\vec{x}_s$ denotes a point on the surface $\partial\Omega$, and $\vec{x}_v$ denotes a point within $\Omega$. $\vec{n}$ is the unit vector perpendicular to $\partial\Omega$ that points away from $\Omega$. $\vec{\nabla}$ denotes the gradient with respect to $\vec{x}_v$.

For the special case that $\vec{B} = \vec{\nabla}V$, we have $\vec{\nabla} \times \vec{B} = 0$; furthermore, if $V$ is a solution of the Laplace equation $\vec{\nabla}^2 V = 0$, we have $\vec{\nabla} \cdot \overrightarrow{B} = 0$. Thus in this case, all the volume integral terms vanish, and $\phi_n(\vec{x})$ and $\vec{A}_t(\vec{x})$ are completely determined from the normal and the tangential components of $\vec{B}$ on the surface $\partial\Omega$ via

$$\phi_n(\vec{x}) = \frac{1}{4\pi}\int_{\partial\Omega} \frac{\vec{n}(\vec{x}_s) \times \overrightarrow{B}(\vec{x}_s)}{|\vec{x}-\vec{x}_s|}\,\mathrm{d}s \tag{5}$$

$$\vec{A}_t(\vec{x}) = -\frac{1}{4\pi}\int_{\partial\Omega} \frac{\vec{n}(\vec{x}_s) \times \overrightarrow{B}(\vec{x}_s)}{|\vec{x}-\vec{x}_s|}\,\mathrm{d}s. \tag{6}$$

For static electric or magnetic fields without sources in $\Omega$, which are characterized by the Laplace problem that we are studying, the divergence and the curl of the field vanish and hence these fields can be decomposed into irrotational and solenoidal parts. For any point within the volume $\Omega$, the scalar and vector potentials depend only on the field on the surface $\partial\Omega$. And due to the smoothing properties of the integral kernel, the interior fields will be analytic even if the field on the surface data fails to be differentiable.

It is worth noting that there are also various higher dimensional extensions to the Helmholtz theorem [15,16] which may be useful to also solve certain 4D boundary value problems.

### 3.1. Surface integration and finite elements via DA

Since the expressions (5) and (6) are analytic, they can be expanded at least locally. The idea is now to expand them to higher orders in BOTH the two components of the surface variables $\vec{x}_s$ and the three components of the volume variables $\vec{x}$. The polynomial dependence on the surface variables will be integrated over surface sub-cells, which results in a highly accurate integration formula with an error order equal to that of the expansion. The dependence on the volume variables will be retained, which leads to a high-order finite element method. By using sufficiently high order, high accuracy can be achieved with a small number of surface elements, and more importantly, a small number of volume elements. We describe the details of the implementation in the following.

The volume $\Omega$ is subdivided into volume elements. Using the prescription for the surface field, the Taylor expansion of the field is computed at the center of each volume element. The final solution inside the overall volume is given as local expansions of the field in different volume elements.

To find the local expansions for each volume element, we first split the domain of integration $\partial\Omega$ into smaller elements $\Gamma_i$. From the surface field formula we extract an approximate Taylor expansion in the surface variables $\vec{x}_s$ about the center of the surface element. Then the integral kernel $1/|\vec{r}-\vec{r}_s|$ and the field $\vec{B}$ on the surface are Taylor expanded in the surface variables $\vec{r}_s$ about the center of each surface element. We also Taylor expand the kernel in the volume variables $\vec{r}$ about the center of the volume element. The final step is to integrate and sum the resulting Taylor expansions for all surface elements. Depending on the accuracy of the computation needed we choose step sizes, order of expansion in $r(x,y,z)$, and order of expansion in $r_s(x,y,z)$.

All the mathematical operations to perform the expansion, surface integration, curl and divergence were implemented using the high-order multivariate differential algebraic tools available in the code COSY Infinity [3,6,7] which automatically leads to the respective field representation to any order without any manual computations.

## 4. Applications

### 4.1. An analytical example: the bar magnet

As a reference problem we consider the magnetic field of an arrangement of the two rectangular iron bars with inner surfaces ($y = \pm y_0$) parallel to the mid-plane ($y = 0$) as shown in Fig. 3. The interior of these uniformly magnetized bars, which are assumed to be infinitely extended in the $\pm y$-directions is defined by: $x_1 \leqslant x \leqslant x_2$, $|y| \leqslant y_0$, and $z_1 \leqslant z \leqslant z_2$. From this bar magnet one can obtain an analytic solution for the magnetic field $\vec{B}(x,y,z)$—see for example

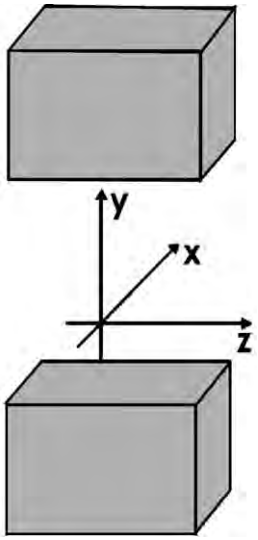



Fig. 3. Geometric layout of the bar magnet, consisting of two bars of magnetized material.

Ref. [17]—and the result is given by

$$B_y(x,y,z) = \frac{B_0}{4\pi}\sum_{i,j=1}^{2}(-1)^{i+j} \times\left[\arctan\left(\frac{X_i Z_j}{Y_+ R_{ij}^+}\right)+\arctan\left(\frac{X_i Z_j}{Y_- R_{ij}^-}\right)\right]$$

$$B_x(x,y,z) = \frac{B_0}{4\pi}\sum_{i,j=1}^{2}(-1)^{i+j}\left[\ln\left(\frac{Z_j+R_{ij}^-}{Z_j+R_{ij}^+}\right)\right]$$

$$B_z(x,y,z) = \frac{B_0}{4\pi}\sum_{i,j=1}^{2}(-1)^{i+j}\left[\ln\left(\frac{X_j+R_{ij}^-}{X_j+R_{ij}^+}\right)\right]$$

where $X_i = x - x_i$, $Y_\pm = y_0 \pm y$, $Z_i = z - z_i$, and $R_{ij}^\pm = (X_i^2 + Y_j^2 + Z_\pm^2)^{1/2}$. We note that because of the symmetry of the field around the mid-plane, only even order terms exist in the Taylor expansion of this field about the origin. The mid-plane field of such a magnet is shown in Fig. 4.

#### 4.1.1. Results and analysis

As a first step, we study the performance of the surface integration method. To this end, the six surfaces are each subdivided into a $44 \times 44$ mesh. The entire volume is considered as one volume element, which is characteristic of the situation in which the entire system has to be represented by merely a single multipole expansion. On each of the surface mesh cells, the contribution from the Helmholtz integral is Taylor expanded [3,18], and the resulting polynomial is integrated. Fig. 5 shows the accuracy of the predicted field, compared with the exact solution, as a function of the order of expansion within the surface mesh cells. Results are shown for the points $(0,0,0)$ and $(0.1,0.1,0.1)$. It can be seen that at order six, an

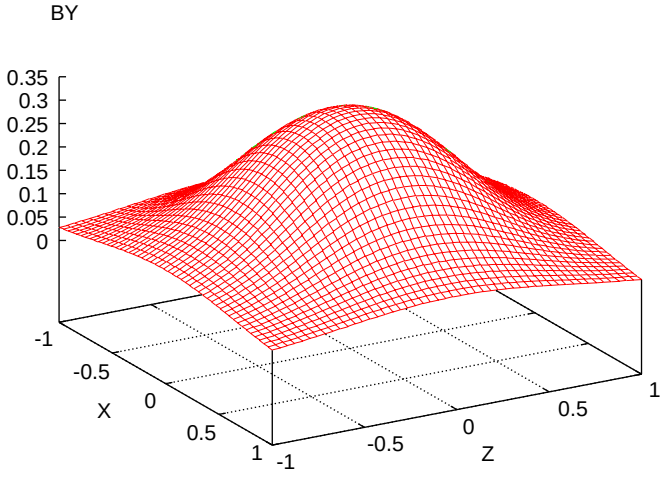



Fig. 4. Magnetic field $B_y$ on the center plane of the bar magnet. $B_0 = 1\,\mathrm{T}$ and the interior of this magnet is defined by $-0.5 \leqslant x \leqslant 0.5$, $|y| \leqslant 0.5$, and $-0.5 \leqslant z \leqslant 0.5$.

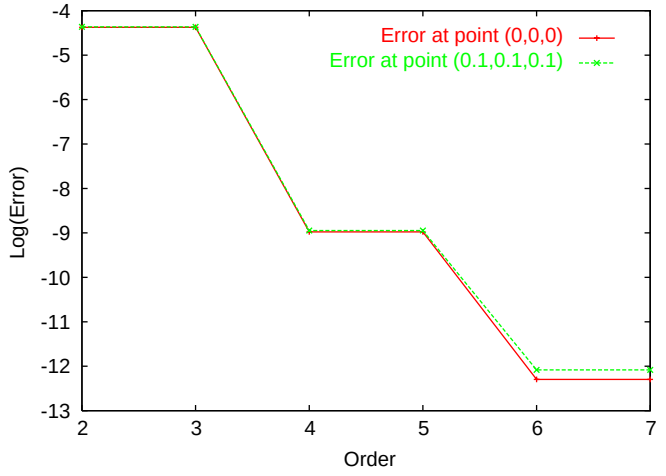



Fig. 5. Error for the field calculated for the bar magnet example for individual points $(0,0,0)$ and $(0.1,0.1,0.1)$.

accuracy of approximately $10^{-12}$ is reached, which is very high compared to conventional numerical field solvers.

We note that the change from order 2 to 3, 4 to 5 and 6 to 7 do not produce significant change in the error. This is due to the fact that odd order terms do not exist in the Taylor expansion of the field around the reference point $(0,0,0)$.

For the next example, we split the volume inside the bar magnet into $5 \times 5 \times 5$ finite elements of width $\pm 0.2$. Within each of the elements, a Taylor expansion in the three volume variables is carried out, resulting in a polynomial representation of the field within the finite element cell. The polynomial representation is used to evaluate the field at 1000 randomly chosen points within the cell, and comparing the result with the analytical answer. Fig. 6 shows the resulting RMS error for finite elements centered around $(0,0,0)$ and $(0.1,0.1,0.1)$. It can be seen that at order 7, an accuracy of approximately $10^{-6}$ is reached.

Overall, the method of simultaneous surface and volume expansion that can be carried out automatically with the

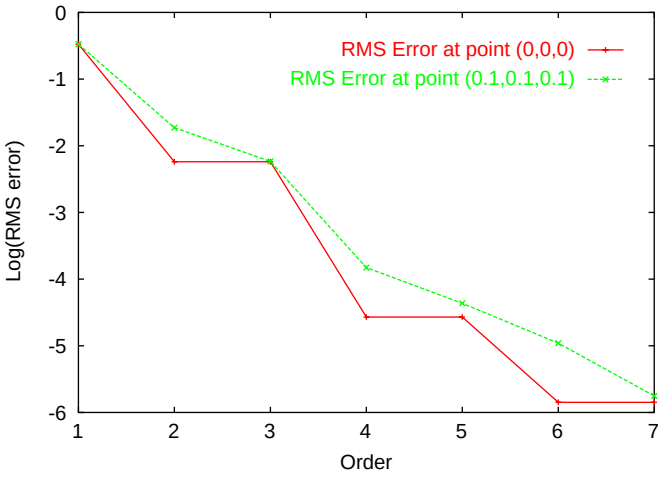



Fig. 6. Average error for the field calculated for the bar magnet example for finite elements of width 0.4 around points $(0,0,0)$ and $(0.1,0.1,0.1)$.

tools in the code COSY Infinity [3,6,7,18] leads to accuracies that are significantly higher than those of conventional finite element tools, even when unusually large finite elements are used.

For purposes of illustration, we now show the Taylor expansion of the field given by Eq. (5) and calculated using the DA tools of COSY over one surface element for a particular point frozen inside the volume of interest. The center of the surface element is at $(-0.39,-0.39,0.5)$ and the point is at $(0.1,0.1,0.1)$. The surface element is described by $(-0.39+0.5\lambda_x x_s,-0.39+0.5\lambda_y y_s,0.5)$, where $\lambda_x,\lambda_y$ represent the length and width of the surface element and $x_s,y_s\in[-1,1]$. In the representation of the Taylor expansion in $x_s$ and $y_s$ below, the entries in the first column provide the number assigned to each of the coefficients in the Taylor expansion to easily identify them. The entries in the second column provide the numerical value of the coefficients. The entries in the fourth, fifth and the sixth columns provide the expansion orders with respect to the volume variables $(x,y,z)$. And the entries in the seventh and eighth column provide the expansion orders with respect to the surface variables $(x_s,y_s)$. The total order for each coefficient is the sum of all the orders in columns four through eight, which is given in the third column. Since we compute the Taylor expansion about a particular point $(0.1,0.1,0.1)$ frozen in the volume of interest in two surface variables $(x_s,y_s)$, we notice that the entries in column four, five, six are all zero. It can be seen that in this expansion, the contributions of higher order terms depending on the surface variables decrease rapidly, and thus the expansion shown would lead to a result of very high accuracy.

Sample eighth order Taylor expansion in two surface variables

| | I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 0.1430015053365947E-01 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 2 | 0.6922600731781813E-03 | 1 | 0 | 0 | 0 | 1 | 0 |
| | 3 | -0.9437452710153340E-03 | 1 | 0 | 0 | 0 | 0 | 1 |
| | 4 | -0.1561210105220474E-04 | 2 | 0 | 0 | 0 | 2 | 0 |
| | 5 | -0.4471499751575185E-04 | 2 | 0 | 0 | 0 | 1 | 1 |
| ⋮ | | | | | | | | |
| | 20 | -0.3232493054085583E-07 | 5 | 0 | 0 | 0 | 1 | 4 |
| | 21 | 0.6156849473575023E-07 | 5 | 0 | 0 | 0 | 0 | 5 |
| | 22 | 0.8960505971632865E-10 | 6 | 0 | 0 | 0 | 6 | 0 |
| | 23 | 0.1890553337467643E-08 | 6 | 0 | 0 | 0 | 5 | 1 |
| | 24 | -0.9792219471281489E-09 | 6 | 0 | 0 | 0 | 4 | 2 |
| ⋮ | | | | | | | | |
| | 41 | -0.2417698920592542E-10 | 8 | 0 | 0 | 0 | 4 | 4 |
| | 42 | 0.7717865536738434E-10 | 8 | 0 | 0 | 0 | 3 | 5 |
| | 43 | -0.2649803372019223E-11 | 8 | 0 | 0 | 0 | 2 | 6 |
| | 44 | -0.2561415687161454E-10 | 8 | 0 | 0 | 0 | 1 | 7 |
| | 45 | 0.8506329051477273E-10 | 8 | 0 | 0 | 0 | 0 | 8 |

We now present the Taylor expansion of the contribution of (5) for one surface element and over one volume element inside the volume of interest. The center of the surface element is at $(-0.39,-0.39,0.5)$ and the center of the volume element is at $(0.1,0.1,0.1)$. The surface element and the volume element can be fully described by $(-0.39+0.5\lambda_x x_s,-0.39+0.5\lambda_y y_s,0.5)$ and $(0.1+0.5\rho_x x,0.1+0.5\rho_y y,0.1+0.5\rho_z z)$, respectively, where $\lambda_x,\lambda_y$ represent the length and width of the surface element, and $\rho_x,\rho_y,\rho_z$ represent the length, width and height of the volume element, and $x_s,y_s,x,y,z\in[-1,1]$. In this case the coefficients of the Taylor expansion depend on both the surface $(x_s,y_s)$ and the volume variables $(x,y,z)$. The coefficients

depending only on the surface variables and the coefficient of the zeroth order term are same as in the previous example of the expansion in just the surface variables. Once again we notice that the contributions of higher order terms decrease rapidly for higher order, showing that also the expansion in volume variables leads to a very accurate representation.

dependence of the average error on the length of the volume element and the total number of volume elements. As an example, for cell lengths of 0.1, which leads to a total number of only 550 finite elements, an accuracy of $10^{-10}$ can be reached with a ninth order method. Similarly, for a seventh order method with a cell length of 0.2, corresponding to 125 boxes, accuracies of about $10^{-6}$ can be reached.

Sample eighth order Taylor expansion in two surface variable and three volume variables

| | I | Coefficient | Order | Exponents | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 0.1430015055365947E-01 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 2 | -0.9590481459719686E-02 | 1 | 1 | 0 | 0 | 0 | 0 |
| | 3 | -0.9590481459719686E-02 | 1 | 0 | 1 | 0 | 0 | 0 |
| | 4 | -0.9768082968233012E-02 | 1 | 0 | 0 | 1 | 0 | 0 |
| | 5 | 0.6922600731781813E-03 | 1 | 0 | 0 | 0 | 1 | 0 |
| | 6 | -0.9437452710153340E-03 | 1 | 0 | 0 | 0 | 0 | 1 |
| ⋮ | | | | | | | | |
| | 454 | -0.4509222359486833E-07 | 6 | 0 | 1 | 0 | 0 | 5 |
| | 455 | -0.3067430813781439E-07 | 6 | 0 | 0 | 1 | 0 | 5 |
| | 456 | 0.8960505971632865E-10 | 6 | 0 | 0 | 0 | 6 | 0 |
| | 457 | 0.1890553337467643E-08 | 6 | 0 | 0 | 0 | 5 | 1 |
| | 458 | -0.9792219471281489E-09 | 6 | 0 | 0 | 0 | 4 | 2 |
| ⋮ | | | | | | | | |
| | 1283 | -0.2417698920592547E-10 | 8 | 0 | 0 | 0 | 4 | 4 |
| | 1284 | 0.7717865536738462E-10 | 8 | 0 | 0 | 0 | 3 | 5 |
| | 1285 | -0.2649803372019148E-11 | 8 | 0 | 0 | 0 | 2 | 6 |
| | 1286 | -0.2561415687161455E-10 | 8 | 0 | 0 | 0 | 1 | 7 |
| | 1287 | 0.8506329051477271E-10 | 8 | 0 | 0 | 0 | 0 | 8 |

We now study the error dependency on the size (length) of the volume element, or equivalently the number of volume elements chosen for the computation. For the order of computation 3,5,7 and 9, Figs. 7 and 8 provide the

Compared to conventional 3D Laplace solvers which typically utilize in the order of $10^6$ cells to achieve accuracies in the order of $10^{-3}$, these results are quite promising.

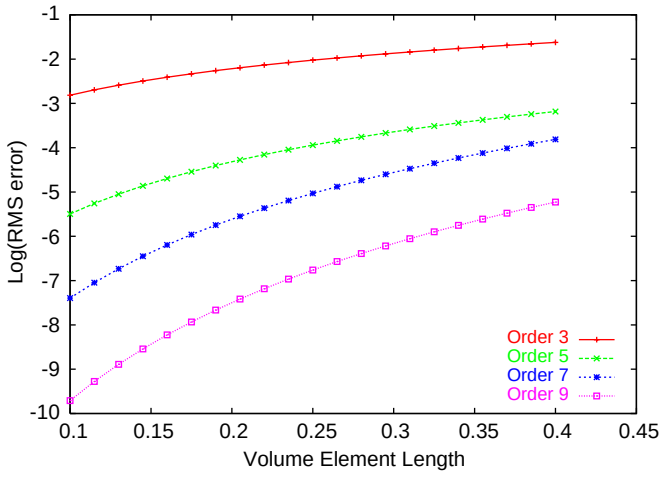



Fig. 7. The plot shows the dependency of the average error on the length of the volume element.

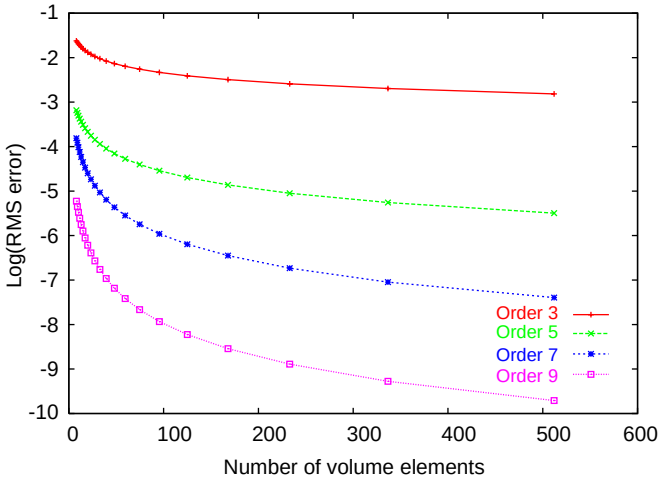



Fig. 8. The plot shows the dependency of the average error on the number of volume element.

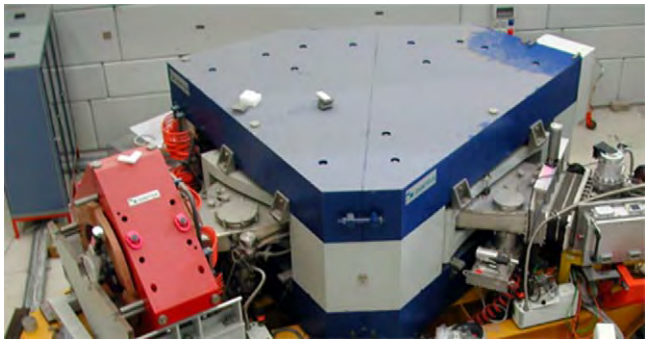


Fig. 9. The dipole magnet of the MAGNEX spectrometer; courtesy A. Cunsolo.


Fig. 10. Layout of the measurement grids in different regions of the dipole magnet. Courtesy A. Lazzaro.

### 4.2. The dipole magnet of the MAGNEX spectrometer

We now address a practical application of this method to magnetic spectrometers. The trajectory reconstruction method [19] is one of the important tools to study magnetic spectrometers. Good computational modeling of the dipole magnet is very important for this tool to work, and this is particularly so for modern large-aperture devices such as MAGNEX at INFN, Catania, Italy [20–22]. Fig. 9 shows the MAGNEX spectrometer configuration.

As mentioned above, for purposes of measurement economy, magnet builders usually provide the magnetic field only on few separate horizontal planes within the dipole, while the computational treatment of the device requires the knowledge of the field in all of space. The MAGNEX dipole was divided into a number of volumes defined by areas and planes as shown in Fig. 10. Four areas were mapped as indicated in Table 1; areas 1 and 4 comprise the effective field boundary regions of the magnet at the entrance and at the exit where the field undergoes a sudden variation due to the fringe field effects, whereas regions 2 and 3 represent the central region of the magnet. This subdivision is the result of the need of different grid sizes over the mapped area in order to limit the measurement time. For each of the regions, the measurements were taken on seven different planes as shown in Table 2.

Table 1
Areas mapped in the dipole

| Area | |
|---|---|
| 1 | EFB area at entrance |
| 2 | Central area entrance end |
| 3 | Central area exit end |
| 4 | EFB area at exit |

Table 2
Planes mapped in the dipole

| Planes | **Z** (cm) | |
|---|---|---|
| *a* | 0 | Mid-plane |
| *b* | 1.6 | Above mid-plane |
| *c* | 3.2 | Above mid-plane |
| *d* | 4.8 | Above mid-plane |
| *e* | 6.4 | Above mid-plane |
| *f* | 8.0 | Above mid-plane |
| *g* | −4.8 | Below mid-plane |

The magnetic measurement were organized so that the RMS error $\langle \Delta B_i/B \rangle$ $i = x, y, z$ at any mesh point inside the working volume of the magnet was not greater than $5 \times 10^{-4}$. The field measurement error due to the error of measuring the Hall probe voltage was $\Delta B = \pm 5 \times 10^{-5}$ G. The main source of the $B$ measurement error were assumed to be the errors of positioning the Hall probe [23–25].

Utilizing that sufficiently outside the dipole the fields will vanish, it is thus possible to provide field data over the surface of a finite box enclosing the region of interest, and thus to apply the methods described above to obtain a field representation everywhere. We use this method to compute the fields in region 1 and plane A of the dipole magnet. The contour plot of the resulting relative errors is plotted over region 1 in Fig. 11. The region where the sharp valley is observed coincides with the physical boundary of the dipole magnet.

## 5. Conclusion

A new technique for finding the multipole expansion solution of the 3D Laplace equation using surface data has been developed. Since this new technique uses the field information on the surface enclosing the volume of interest and is implemented using the high-order multivariate differential algebraic tools available in the code COSY Infinity [3,6,7], the accuracy achieved is much higher than that of conventional field solvers. If the data on the surface

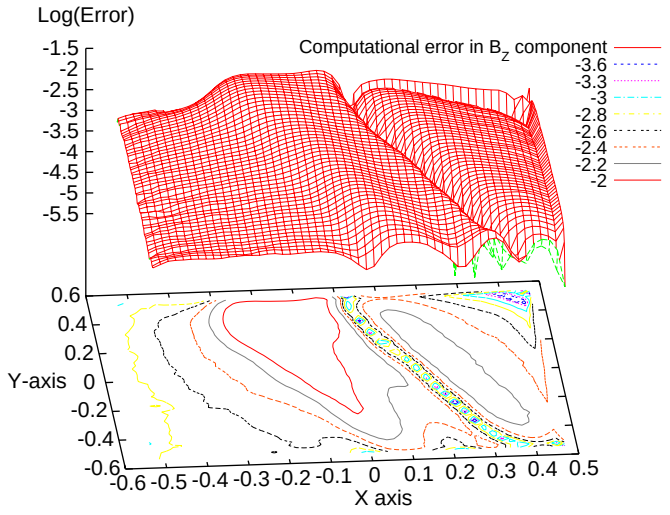



Fig. 11. Contour plot of the magnetic field errors for region 1 and plane A.

enclosing the volume of interest can be given exactly, then in principle arbitrarily high accuracy limited only by the computational resources available can be achieved by this new technique. In practical situations where the field data on the surface enclosing the volume of interest is experimentally measured, the discretization of the surface and the errors in the experimentally measured field data may limit the accuracy achieved, but because the method is naturally smoothing, the accuracy is expected to exceed that of the measurements. The use of this new technique has been tested rather successfully for an analytic example of the bar magnet. The result from a preliminary analysis of the data from the MAGNEX spectrograph dipole magnet is also presented.

## Acknowledgements

For various fruitful discussions, we are indebted to A. Lazzaro, A. Cunsolo, F. Cappuzzello, and P. Walstrom.

## References

[1] M.L. Shashikant, M. Berz, B. Erdélyi, COSY INFINITY's EXPO symplectic tracking for LHC, IOP CP, vol. 175, 2004, pp. 299–306.
[2] B. Erdélyi, M. Berz, K. Makino, Detailed analysis of fringe field effects in the large hadron collider, Technical Report MSUCL-1129, National Superconducting Cyclotron Laboratory, Michigan State University, East Lansing, MI 48824, 1999.
[3] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999 and Also available at ⟨http://bt.pa.msu.edu/pub⟩.
[4] R. Degenhardt, M. Berz, Nucl. Instr. and Meth. A 427 (1999) 151.
[5] M. Berz, B. Fawley, K. Hahn, Nucl. Instr. and Meth. A 307 (1991) 1.
[6] M. Berz, K. Makino, COSY INFINITY Version 8.1—user's guide and reference manual, Technical Report MSUHEP-20704, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2002, see also ⟨http://cosy.pa.msu.edu⟩.
[7] M. Berz, J. Hoefkens, K. Makino, COSY INFINITY Version 8.1—programming manual, Technical Report MSUHEP-20703, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2002. see also ⟨http://cosy.pa.msu.edu⟩.
[8] H.v. Helmholtz, J. Reine Angew. Math. Crelle's J. 55 (1858) 25.
[9] P.M. Morse, H. Feshbach, Meth. Theor. Phys. Part I and II, (1953).
[10] P.L. Walstrom, Nucl. Instr. and Meth. A 519 (1–2) (2004) 216.
[11] M. Venturini, A.J. Dragt, Nucl. Instr. and Meth. A 427 (1999) 387.
[12] M. Venturini, A. Dragt, Computing transfer maps from magnetic field data, in: IEEE Particle Accelerator Physics Conference, 1999.
[13] M. Venturini, D. Abell, A. Dragt, Map computation from magnetic field data and application to the lhc high-gradient quadrupoles, eConf C980914, 1998, pp. 184–188.
[14] P. Walstrom, A. Dragt, T. Stasevich, Computation of charged-particle transfer maps for general fields and geometries using electromagnetic boundary-value data, in: IEEE Particle Accelerator Conference 2001, 2001.
[15] D.A. Woodside, J. Math. Phys. 41 (7) (2000) 4622.
[16] D.A. Woodside, J. Math. Phys. 40 (10) (1999) 4911.
[17] R. Degenhardt, M. Berz, High accuracy description of the fringe fields of particle spectrographs, in: Proceedings 1993 Particle Accelerator Conference, Washington, DC, 1993.
[18] M. Berz, Forward algorithms for high orders and many variables, Automatic Differentiation of Algorithms: Theory, Implementation and Application, SIAM, Philadelphia, 1991.
[19] M. Berz, K. Joh, J.A. Nolen, B.M. Sherrill, A.F. Zeller, Phys. Rev. C 47 (2) (1993) 537.
[20] V.A. Shchepunov, A. Cunsolo, F. Cappuzzello, A. Foti, A. Lazzaro, A.L. Melita, C. Nociforo, J.S. Winfield, Nucl. Instr. and Meth. B 204 (2003) 447.
[21] A. Lazzaro, A. Cunsolo, F. Cappuzzello, A. Foti, C. Nociforo, S. Orrigo, V. Shchepunov, J.S. Winfield, M. Allia, Computational aspects of the trajectory reconstruction in the MAGNEX large acceptance spectrometer, in: Berz, Makino (Eds.), Computational Accelerator Physics 2002, Institute of Physics Conference Series, vol. 175, 2004.
[22] A. Cunsolo, F. Cappuzzello, A.V. Belozyorov, A. Elanique, A. Foti, A. Lazzaro, O. Malishev, L. Melita, W. Mittig, C. Nociforo, P. Roussel-Chomaz, V. Shchepunov, D. Vinciguerra, A. Yeremin, J.S. Winfield, Nucl. Condens. Matter Phys. 513 (1) (2000) 270.
[23] I. Vasserman, S. Sasaki, Comparison of different magnetic measurement techniques, in: Proceedings of 13th International Magnetic Measurement Workshop, 2003.
[24] S.I. Redin, N.M. Ryskulov, G.V. Fedotovich, B.I. Khazin, G.M. Bunce, G.T. Danby, J.W. Jackson, W.M. Morse, R. Prigl, Y.K. Semertzidis, et al., Nucl. Instr. and Meth. A 473 (3) (2001) 260.
[25] K.N. Henrichsen, Overview of magnet measurement methods, in: CERN Accelerator School 1997, 1997.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 184–187
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# BEAMDULAC code for numerical simulation of 3D beam dynamics in a high-intensity undulator linac

E.S. Masunov*, S.M. Polozov

*Moscow Engineering Physics Institute (State University), Kashirskoe shosse 31, Moscow 115409, Russian Federation*

Available online 28 November 2005

**Abstract**

The concept of a linear accelerator (linac) without a synchronous wave from an RF field has previously been suggested. This concept has been implemented in an innovative type of linac called a linear undulator accelerator. Comprehensive analytical and numerical studies of the beam dynamics resulted in the complete structure design for two versions of the undulator accelerator. The key results of these investigations are discussed in this paper.


*PACS:* 41.75.L; 41.85.E; 29.27.F

*Keywords:* Linear accelerator; RF field; RF undulator; Electrostatic undulator; Ribbon beam; High-intensity ion beam; Numerical simulation; BEAMDULAC code

## 1. Introduction

In a conventional radio frequency (RF) linear accelerator (linac) the beam is accelerated by a synchronous wave from an RF field. An alternative method of acceleration in fields without a synchronous wave has been suggested and analytically studied [1,2]. The physical acceleration mechanism is similar to that in an inverse free electron laser [3,4]. In our case, the accelerating force is driven by a combination of two non-synchronous waves provided by two undulators. We define an undulator as any structure producing a periodic-distribution electromagnetic field for which the phase velocity differs significantly from that of the beam. A new linac based on this approach has been termed a undulator linear accelerator (UNDULAC). There are three different types of undulator that can be used to design the required configuration of the accelerating fields—a magnetic, electrostatic or RF undulator. It has been shown that one of the undulators must be an RF type and the second can be a magnetic (UNDULAC-M), electrostatic (UNDULAC-E) or radio frequency (UNDULAC-RF) type [5]. The approach developed can be applied to the acceleration of axisymmetric beams. However, a more attractive application of the UNDULAC concept is ribbon beam acceleration. In this case, the total beam current significantly increases. UNDULAC-E (Fig. 1) and UNDULAC-RF (Fig. 2) devices can be used for ribbon ion-beam bunching and acceleration in the low-energy range. Acceleration and transverse focusing of the ribbon beam are realized in the longitudinal or transverse undulator field. The results of investigation of low-beta ribbon beam dynamics in UNDULAC-E and UNDULAC-RF accelerator are described.

## 2. Beam dynamics studies in the UNDULAC

UNDULAC-E and UNDULAC-RF accelerators can be designed as shown in Figs. 1 and 2. In the UNDULAC-E device, the RF field has a phase advance per period of $\mu = 0$ (zero mode) and the electrostatic field has $\mu = \pi$ ($\pi$ mode). In the UNDULAC-RF device, zero or $\pi$ modes of the electromagnetic field can be used. The beam particles interact with the zero RF field harmonic and the basic space harmonic of the electrostatic field in the

*Corresponding author. Tel.: +7 95 3239292, +7 95 3242995; fax: +7 95 3242111.
*E-mail address:* esmasunov@mephi.ru (E.S. Masunov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.037

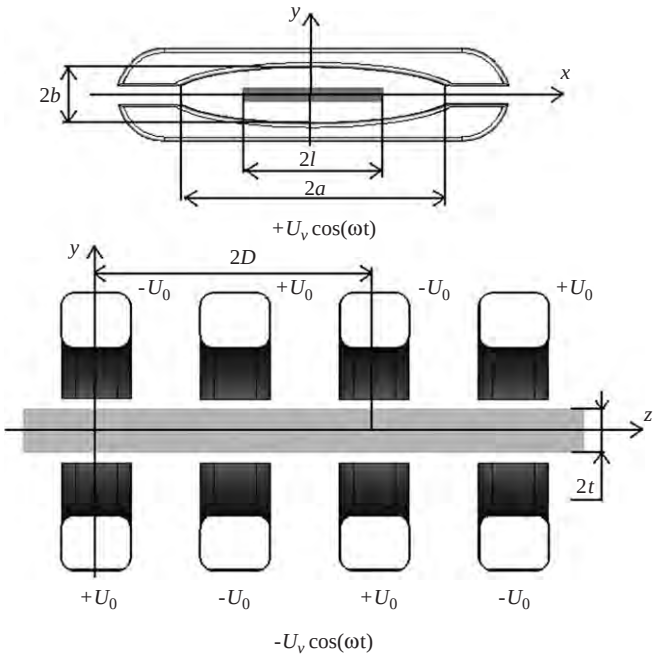



Fig. 1. UNDULAC-E with transverse field ($U_0$ is the electrostatic voltage, $U_v$ is the RF voltage amplitude and $D$ is the undulator period).

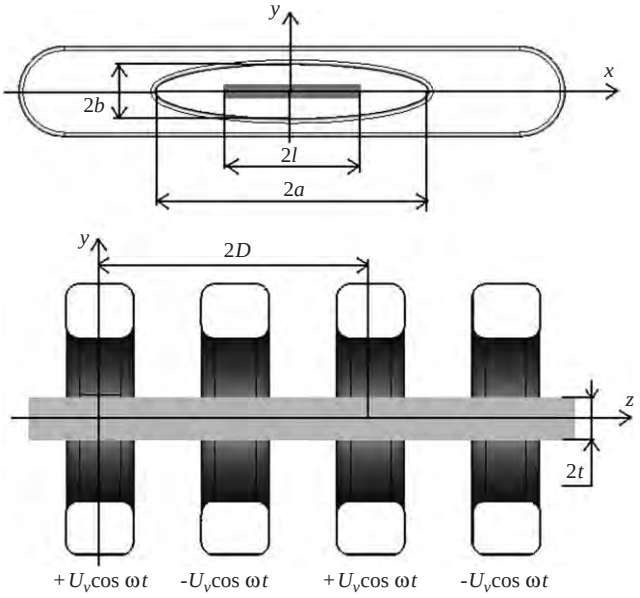



Fig. 2. UNDULAC-RF for the $\pi$ mode.

UNDULAC-E, or with two fundamental space RF field harmonics in the periodic RF structure in the UNDULAC-RF. For any UNDULAC, consideration of the beam dynamics within the scope of traditional theory is a challenging task due to absence of an accelerating wave synchronous with the beam. This difficulty can be overcome using the averaging technique. Beam acceleration is realized by a combined wave if the phase velocity of the combined wave is close to the beam velocity. In such an acceleration system, the momentum and coordinate of each particle can be presented as a sum of slow varying and fast oscillating components. Averaging over the rapid oscillations allows the motion equation to be obtained in Hamilton form. The results of analytical research of this equation are discussed elsewhere [1,5].

In an UNDULAC the ion bunch can lose particles due to a large phase and transverse non-linear oscillations. The choice of the undulator field amplitude is not arbitrary because, simultaneously to acceleration, it must ensure transverse beam focusing. Hamiltonian analysis of the motion equation was used for a complete 3D description of the beam dynamics in an UNDULAC. The optimal ratio of the amplitude of the zero ($E_0$) and first ($E_1$) harmonics of the RF field, $\chi = E_1/E_0$, must be equal to 0.3–0.4 in the UNDULAC-RF. Particle losses due to longitudinal sliding will be small in this case. Transverse beam focusing can be realized in the UNDULAC-RF for all values of $\chi$ if the $\pi$ mode is used and for $\chi > 1$ if the zero mode is used.

Hamiltonian analysis in a smooth approximation showed that the amplitude of the electrostatic $E_0^o$ and RF fields $E_0$ must be close ($\chi^o = E_0/E_0^o \approx 1$) for effective transverse beam focusing in the UNDULAC-E. The current transmission coefficient for the ion beam is 90–95% in this case [6,7].

## 3. BEAMDULAC code

The beam dynamics cannot be investigated completely using analytical methods only. Numerical simulation is necessary to verify the results of any analytical study. A space charge field can also be accurately treated by numerical simulation. The new BEAMDULAC code was developed especially for beam dynamics simulations in linear accelerators.

The BEAMDULAC code utilizes the cloud-in-cell (CIC) method for accurate treatment of the space charge effects that are especially important in the case of a high-intensity beam. The motion equation for each particle is solved in the external fields and the inter-particle Coulomb field. The charge density is deposited on the grid points using the CIC technique. To determine the potential of the Coulomb field, the Poisson equation is solved on the grid with periodic boundary conditions at both ends of the domain in the longitudinal direction. The aperture of the channel is represented as an ideally conducting surface of rectangular or circular cross-section. This allows consideration of the shielding effect, which is sufficiently important for transverse focusing of ribbon beams. The fast Fourier transform (FFT) algorithm is used to solve the Poisson equation on a 3D grid. The Fourier series for the space charge potential obtained can be analytically differentiated, and thus each component of the Coulomb electrical field can be found as a series with known coefficients. The Coulomb repulsion force is the main factor limiting the beam current in high-intensity linacs. In our code, the space charge field can be calculated with the same precision as the Coulomb potential without numerical differentiation. The external

Table 1

| | UNDULAC-RF | | UNDULAC-E | |
|---|---|---|---|---|
| | lg. | tr. | lg. | tr. |
| Accelerator length $L$ (m) | 2.5 | | | |
| Deuterium ion injection energy $W_{in}$ (keV) | 100 (0.01) | | 100–150 | |
| Injection velocity $\beta_{in}$ | | | (0.01–0.013) | |
| Amplitude of zero RF field harmonic $E_0$ (kV/cm) | 200 | 210 | 175 | |
| Amplitude of first RF field harmonic $E_1$ (kV/cm) | 80 | 70 | — | |
| Amplitude of electrostatic field $E_0^o$ (kV/cm) | — | — | 175 | |
| Bunching sub-section length $L_b/L$ | 0.5 | | 0.3 | |
| Accelerator channel cross-section $2a \times 2b$ (cm$^2$) | $10 \times 0.7$ | | $20 \times 0.8$ | |
| Optimal initial beam size $2l \times 2t$ (cm$^2$) | — | — | $12 \times 0.4$ | $12 \times 0.4$ |
| Limit initial beam size $2l \times 2t$ (cm$^2$) | $5 \times 0.3$ | $7 \times 0.3$ | $16 \times 0.4$ | $16 \times 0.5$ |
| Initial emittance | | | | |
| $E_x$ ($\pi$ mm mrad) | 30 | 30 | 20 | 20 |
| $E_y$ ($\pi$ mm mrad) | 0.7 | 0.06 | 0.8 | 0.9 |
| $E_\varphi$ (keV mrad) | 25 | 40 | 50 | 55 |
| Acceptance of accelerator channel | | | | |
| $A_x$ ($\pi$ mm mrad) | 60 | 60 | 20 | 20 |
| $A_y$ ($\pi$ mm mrad) | 2 | 2.5 | 1.5 | 1.5 |
| $A_\varphi$ (keV mrad) | 40 | 40 | 100–200 | 100–200 |
| Limit beam current $I_{max}$ (A) | 0.2–0.25 | 0.3–0.35 | 0.8–1.0 | 0.9–1.6 |
| Output beam energy $W_{max}$ (MeV) | 1.2–1.5 | | 1.0–1.2 | |
| Output velocity $\beta_{max}$ | (0.034–0.04) | | (0.03–0.034) | |

fields are represented as a series of space harmonics. Time is used as an independent variable and standard methods are applied for integration of the motion equation in the BEAMDULAC code.

## 4. Ribbon ion beam dynamics in UNDULAC-RF

The BEAMDULAC code was used for a deuterium ion-beam dynamics study in an UNDULAC-RF accelerator. Optimal linac parameters and the limit beam current were found for zero and $\pi$ modes of the RF field. Calculation showed that transverse focusing is not effective for the zero mode. The current transmission coefficient $K_T$ is smaller than 30–35%, and this linac version is not promising for practical use. The main parameters of the UNDULAC-RF for the $\pi$ mode are presented in Table 1, with output beam energy of $W = 1.3-1.5$ MeV for accelerator length $L = 2.5$ m. The optimal bunching sub-section length must be $L_b = 0.5L$ approximately. The rate of energy gain in the accelerating sub-section is 700–800 keV/m. The current transmission coefficient calculated will be smaller than in the smooth approximation. It is equal to 75–80% for low initial emittance and beam size of $2l \times 2t = 1 \times 0.04$ cm$^2$. Particle losses are connected with fast oscillations of the particle phases and longitudinal velocities in the UNDULAC-RF using a longitudinal field. Transverse oscillations of the particles cause enlargement of the beam envelope in an undulator with a transverse field. This effect increases particle losses in the transverse direction. Calculation showed that the optimum value for the amplitude ratio of the RF field harmonics is $\chi = 0.3-0.4$ (Fig. 3). This coincides with the result calculated analytically. The current transmission coefficient is smaller if the initial

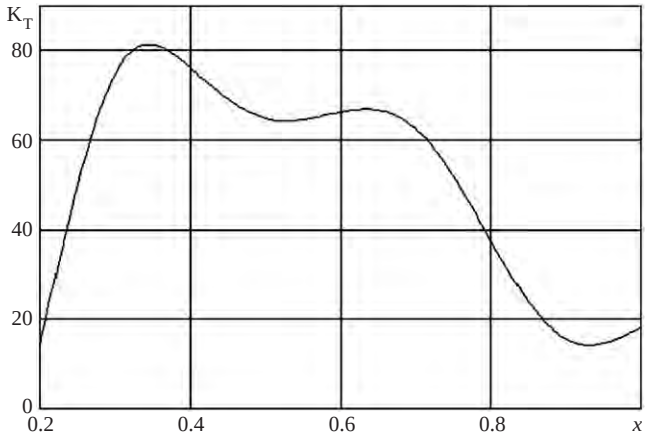



Fig. 3. Current transmission coefficient versus ratio of RF field harmonic amplitudes in the UNDULAC-RF.

emittance is $0.7\pi$ mm mrad. In this case the transmission coefficient is 65–70%.

The limit beam current is related to Coulomb repulsion in high-intensity ion linacs. Its value was found for a longitudinal and transverse UNDULAC-RF. The limit current is $I_{max} = 0.2-0.25$ A for a linac using a longitudinal RF field and $I_{max} = 0.3-0.35$ A for a transverse field.

## 5. UNDULAC-E beam dynamics

Numerical simulation of deuterium ion-beam dynamics shows that the limit current and current transmission coefficient for the UNDULAC-E are higher and the rate of energy gain is smaller than for the UNDULAC-RF.

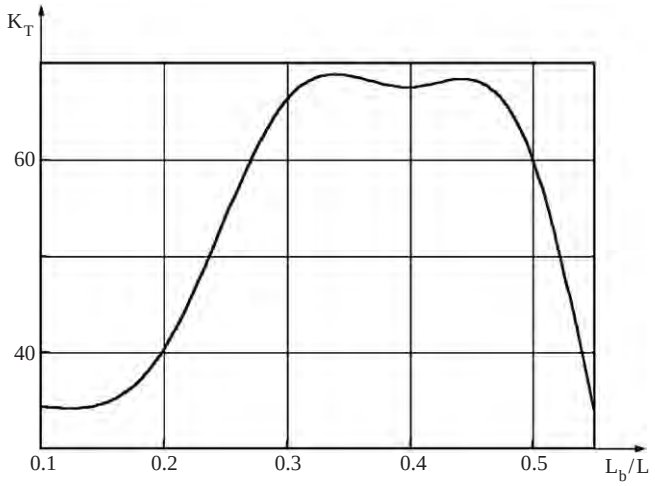



Fig. 4. Current transmission coefficient versus bunching sub-section length $L_b/L$ for the UNDULAC-E.

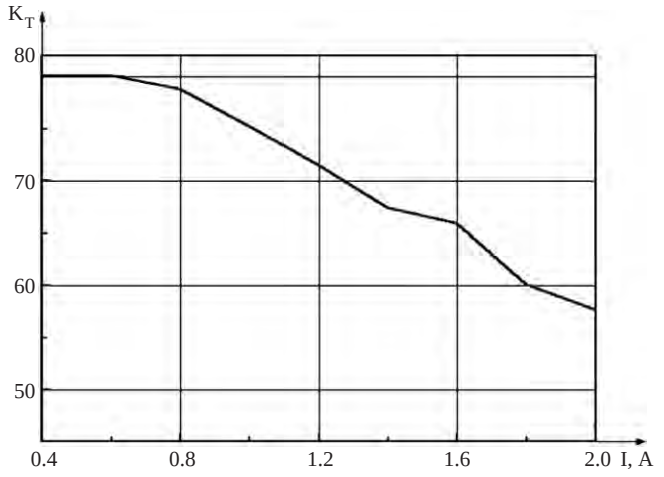



Fig. 5. Current transmission coefficient versus initial beam current $I$ for the UNDULAC-E.

The rate of energy gain in the accelerating sub-section of the UNDULAC-E is 500 keV/m. The main parameters of the UNDULAC are presented in Table 1. The optimal values of the undulator field amplitude and the RF field are $E_0^o = 120 - 180$ kV/cm and $E_0 = 150-200$ kV/cm. In this case the output beam energy is $W = 1$ MeV for accelerator length $L = 2.5$ m. The bunching sub-section length is $L_b = 0.3L$ (see Fig. 4). The current transmission coefficient is $K_T = 80\%$ for zero current beam. The limit beam current for the UNDULAC-E can be very high. In a transverse undulator field, the limit current is $I_{max} = 1.0$ A (initial beam size $l \times t = 6 \times 0.2\,cm^2$). The current transmission coefficient is $K_T = 75\%$ in this case. The current transmission coefficient versus the initial beam current is shown in Fig. 5. The particle loss coefficient $K_L$ versus the longitudinal coordinate $z$ is represented in Fig. 6. Curve 1 shows the total losses, curve 2, losses due to transverse motion, and curve 3, losses due to phase sliding. The major proportion of particle losses is observed in the bunching sub-section. The influence of the Coulomb field is the basic reason for ion losses in the accelerating part if the initial beam size is $l \times t = 8 \times 0.25\,cm^2$. The current transmission coefficient reduces to $K_T = 65-70\%$ in this case. The limit beam current in the UNDULAC-E device using a longitudinal field is 0.8–1.0 A, but $K_T$ is smaller than 65%.

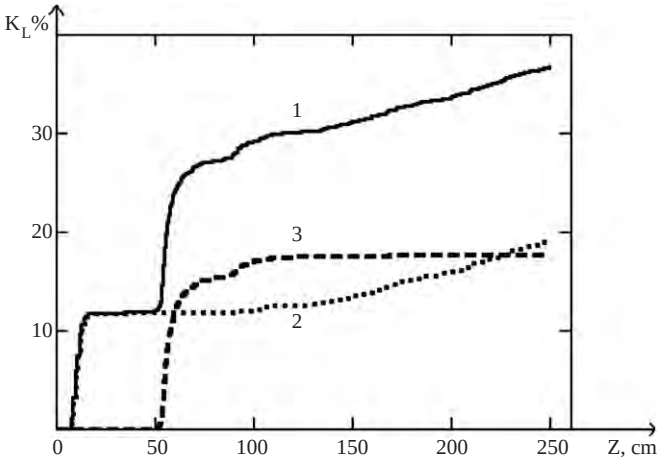



Fig. 6. Particle losses versus longitudinal coordinate.

## 6. Conclusion

Deuterium ion-beam dynamics was studied in two types of undulator linac. The new BEAMDULAC code has been used especially for optimization and calculation of the beam dynamics in UNDULAC devices. The numerical simulation results coincide with the analytical investigation. The limit beam current was also calculated in the UNDULAC. Its value can be increased up to 1.6 A. The numerical simulation shows that the current transmission coefficient and the limit current are larger in the UNDULAC-E, but the rate of energy gain is higher in the UNDULAC-RF when the $\pi$ mode of the RF field is used. All the results will be used for further design of the accelerator structure developed.

## Acknowledgement

The authors would like to thank Dr. Nikolay Vinogradov for her help in preparing the manuscript. The work was supported by RFBR: Grant #04-02-16667.

## References

[1] E.S. Masunov, Sov. Phys. Tech. Phys. 35 (8) (1990) 962.
[2] E.S. Masunov, Tech. Phys. 46 (11) (2001) 1433.
[3] R. Palmer, J. Appl. Phys. 43 (1972) 3014.
[4] E. Courant, C. Pellegrini, W. Zakowicz, Phys. Rev. A 32 (1985) 2813.
[5] E.S. Masunov, Beam Dyn. Newslett. 23 (2000) 23.
[6] E.S. Masunov, A.P. Novikov, in: Proceedings of PAC 1991, vol.5, pp. 3177–3179.
[7] E.S. Masunov, S.M. Polozov, A.S. Roshal, Radiat. Phys. Chem. 61 (2001) 491.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 188–191
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# A parallel adaptive Vlasov solver based on hierarchical finite element interpolation

M. Mehrenberger[a,*,1], E. Violard[b,1], O. Hoenen[b,1], M. Campos Pinto[c], E. Sonnendrücker[a,1]

[a]*IRMA, Université Louis Pasteur, Strasbourg, France*
[b]*LSIIT, Université Louis Pasteur, Strasbourg, France*
[c]*LJLL, Université Pierre et Marie Curie, Paris, France*

Available online 5 December 2005

**Abstract**

We present a parallel adaptive scheme for the Vlasov equation. Our method is based on a way of reducing dependencies between data, thanks to a hierarchical finite element interpolation approach. A specific data distribution pattern yields an efficient implementation. Numerical results are exhibited for a classical beam simulation in the 1D phase space.


*PACS:* 29.27.a; 52.65−*y*

*Keywords:* Adaptive; Parallel; Vlasov; Simulation; Beam; Plasma

## 1. Introduction

Thanks to the rapid increase of computing power in recent years, simulations of plasmas and particle beams based on direct solution of the Vlasov equation on a multi-dimensional phase-space grid are becoming attractive as an alternative to Particle-In-Cell (PIC) simulations. Their strength lies essentially in the fact that they are noiseless and that all parts of phase space, including the tail of the distribution, are equally well resolved. Their major drawback is that, for inhomogeneous systems, many of the grid points (where no particles are present) are wasted. This is especially the case for beam simulations where the beam moves rapidly through the phase space (due to varying alternating-gradient focusing forces, for example). In order to overcome this problem, adaptive methods can be used [1,2,6].

In this paper, we present an adaptive method based on bi-quadratic finite element interpolation as was first introduced in Ref. [2] and describe a new efficient parallel version of an adaptive Vlasov solver, which is also described in Ref. [3]. It turns out that adaptive numerical methods are often difficult to parallelize, because they introduce dependancies between data at different grid levels and it is then difficult to manage data locality. We have designed here a numerical method well fitted for parallelization where the underlying partitions of dyadic tensor-product cells offer a simple way to distribute data. In fact, with such a strategy, each data essentially depends on the neighbor data of the same level. After describing the numerical method in the adaptive context, we present its parallelization and exhibit an application to a classical beam simulation.

## 2. An adaptive resolution scheme for the Vlasov equation

Here is a brief description of the numerical method of resolution. We refer the reader to Ref. [2] for a more detailed presentation. For sake of conciseness, we give the scheme for a 2D phase space, but it generalizes to higher dimensions.

The adaptive method is based on a classical semi-Lagrangian method which takes benefit of the conservation of the distribution function along particle trajectories. This method requires to be able to localize a point in the

*Corresponding author. Tel.: +33 1 39 63 58 30; fax: +33 1 39 63 58 82.
*E-mail address:* Michel.Mehrenberger@inria.fr (M. Mehrenberger).
[1]Supported by CALVI, INRIA Lorraine project, Nancy, France.

0168-9002/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.094

phase-space and uses an interpolation operator. Therefore, we use a dyadic cutting of the phase space together with a finite stencil for the interpolation.

We represent our solution in a hierarchical way. Each mesh cell can be recursively subdivided into four subcells of same size. A function is approximated at a given level by a biquadratic interpolation using its value at nine nodes: the four vertices of the cell, the four midpoints of the edges and the midpoint of the cell. The biquadratic interpolation is a tensor product of quadratic 1D interpolation. Hence the interpolation procedure in 2D (or more dimensions) can be derived easily from the 1D procedure. When dealing, with the fine grid, instead of interpolating directly using the basis functions on the fine grid, we consider the three basis functions at the coarse grid and we add two others of the fine level. The basis functions for one coarse cell and its two subcells are represented in Fig. 1.

The solution at time $t^n$ is given by an adaptive mesh $\mathscr{M}^n$ consisting of a dyadic partition of cells and the level $j$ of a given cell will vary from a coarser level $j_0$ to a finer level $J$. Going from time step $t^n$ to $t^{n+1}$ consists in three steps:

(1) *Prediction of* $\mathscr{M}^{n+1}$ : For each cell $\alpha \in \mathscr{M}^n$, denoting $j$ its level, compute its center $c_\alpha$ and the forward advected point $\mathscr{A}(c_\alpha)$ by following the characteristics of the Vlasov equation (see [2] for more details about characteristics and advection operator $\mathscr{A}$). Then add to $\mathscr{M}^{n+1}$ the unique cell $\bar{\alpha}$ of level $j$ which fits at that place in $\mathscr{M}^{n+1}$ and all the necessary cells so that $\mathscr{M}^{n+1}$ is a dyadic adaptive mesh. Last, if $j<J$, refine $\bar{\alpha}$ of one level, that is, replace it by the four cells of level $j+1$ which cover the same surface.

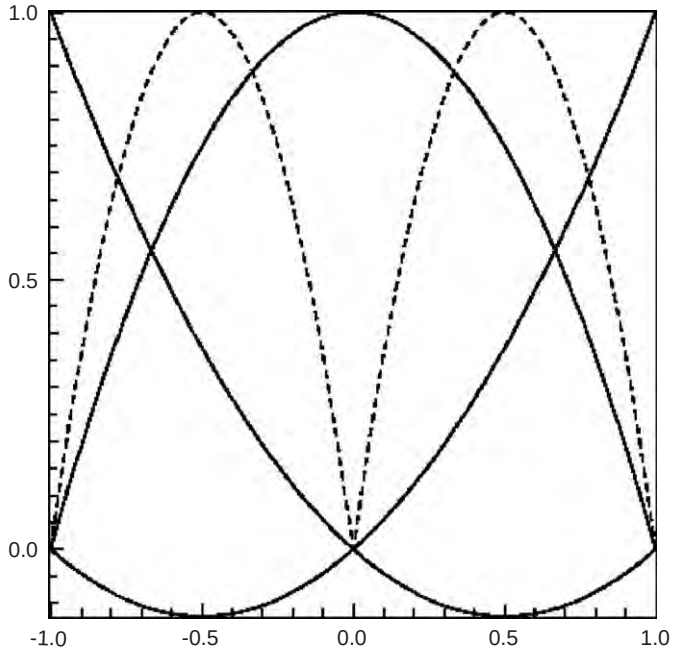



Fig. 1. Hierarchical quadratic basis functions. Basis functions on coarse cell (solid line). Additional basis functions for fine cells (dashed line).

(2) *Evaluation*: For each node $a$ of $\mathscr{M}^{n+1}$, compute the backward advected point $\mathscr{A}^{-1}(a)$ and set $f^{n+1}(a)$ to $f^n(\mathscr{A}^{-1}(a))$: the evaluation $f^n(c)$ of the solution at any point $c \in [0,1]^2$ is obtained by searching the unique cell $\alpha$ of the adaptive mesh $\mathscr{M}^n$ where the point is located, using the values at the nodes of that cell and computing the local biquadratic interpolation on that cell, say $I(c,\alpha,f^n(c))$.

(3) *Compression of* $\mathscr{M}^{n+1}$: From $j=J-1$ to $j_0$, replace four cells of level $j+1$ by a cell $\alpha$ of level $j$ (do the converse of refining $\alpha$) when the norm of the differences $f^{n+1}(a)-I(a,\alpha,f^n(a))$, for all node $a$ of $\alpha$, is small enough.

## 3. Parallel implementation

The computational domain is subdivided into *regions*. A *region* is a surface of the computational domain which is defined by an union of mesh cells. Regions are allocated to processors so that each processor owns and computes the mesh cells and nodes which are included in its own *region*. As the mesh adapts to the evolution in time of the physics, the number of cells within a region change and it is then necessary to include a load balancing mechanism which then consists in redefining regions for each processor. In order to minimize communications, we apply compression within the region limit only. So the compression phase does not require any communication in our implementation. This is an approximation of the numerical method since we eliminate fewer cells than in the original method, but it does not hazard convergence. Each processor owns a local representation of the mesh. The mesh is represented by two hash tables: the cell hash table stores a set of cells which forms a partition of the whole computational domain and associates each cell with its owner identity. The node hash table stores the value at each node within the region. This representation allows cells and nodes to be accessed in constant time while minimizing the memory usage.

As said previously, our load balancing mechanism consists in redefining regions for each processor, the number of cells in each region should be approximatively the same and each region should have a "good shape" to improve the compression. Moreover, every region should be connex in order to reduce the volume of communications. We use the Hilbert's curve [4] to achieve this last requirement.

We model the global load and its localization onto the computational domain by a quad-tree [5] whose nodes are weighted by the number of leaves in the subtree. Each leaf of this quad-tree identifies one cell of the mesh and the level of a leaf in the tree is the level of the corresponding cell in the mesh.

We then build the new regions by partitioning the quad-tree. Each region is the union of the cells corresponding to the leaves of each part of the quad-tree.

To obtain a good partition, we browse the quad-tree starting from its root to its leaves, and try to make a cut as soon as possible. A part, say $\mathscr{P}$, of the partition is such that $(1-\lambda)*I \leqslant \|\mathscr{P}\| \leqslant (1+\lambda)*I$, where $\|\mathscr{P}\|$ is the number of leaves of the part, $I$ equals to the total number of cells divided by the number of processors, and $\lambda \in [0,1]$ is an error factor that permits a certain degree of liberty for finding good parts.

We use this method at initialization, and a less expensive version to update regions at runtime without penalizing performance.

## 4. Numerical results

In order to assess the benefits of the adaptive solver we computed the transverse evolution of a semi-Gaussian beam in a uniform focusing channel. For such a beam, the initial distribution function reads

$$f(r,v) = \frac{1}{\pi a^2 \sqrt{2\pi} b} e^{-1/2(v^2/b^2)} \quad \text{if } r < a,$$

and $f(r,v) = 0$ else. Here, $a = 4/\sqrt{15}$, $b = 1/(2*\sqrt{15})$ and the time step is 1/32nd of period, that is $\Delta t = 2\pi/32$.

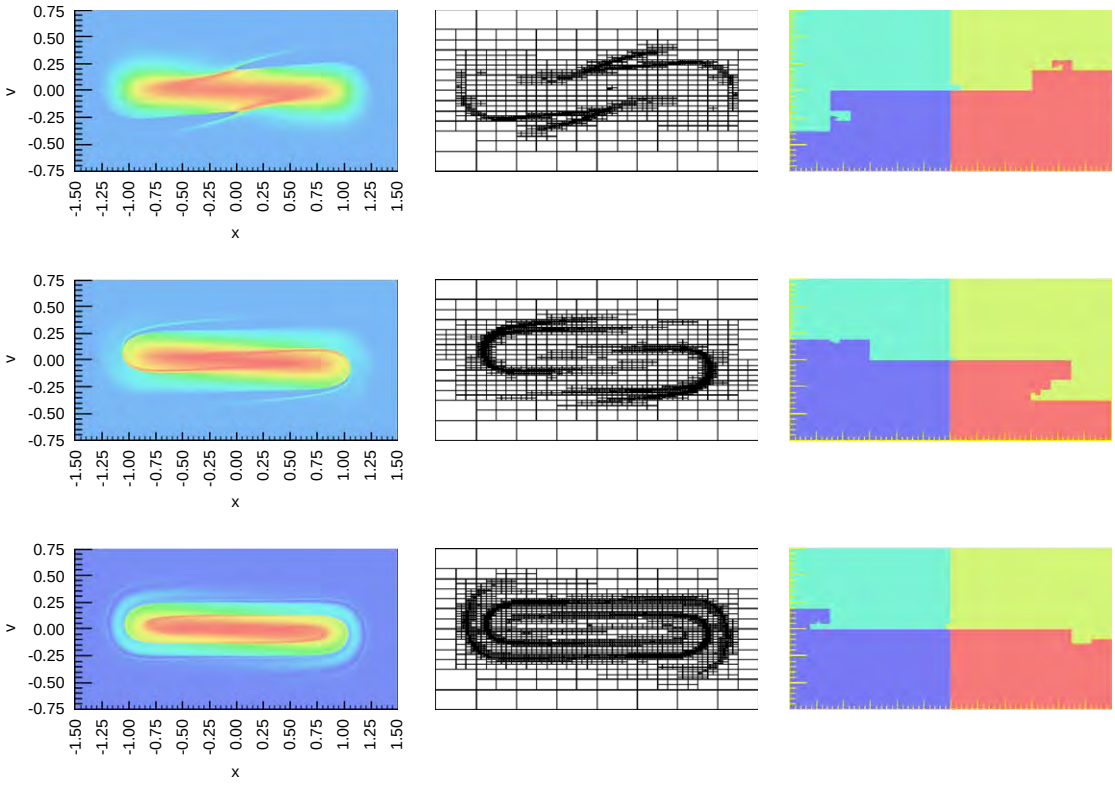



Fig. 2. Evolution of the distribution function together with the grid and the regions distribution of processors after one half, one and two periods (resp. 32, 64 and 128 iterations).

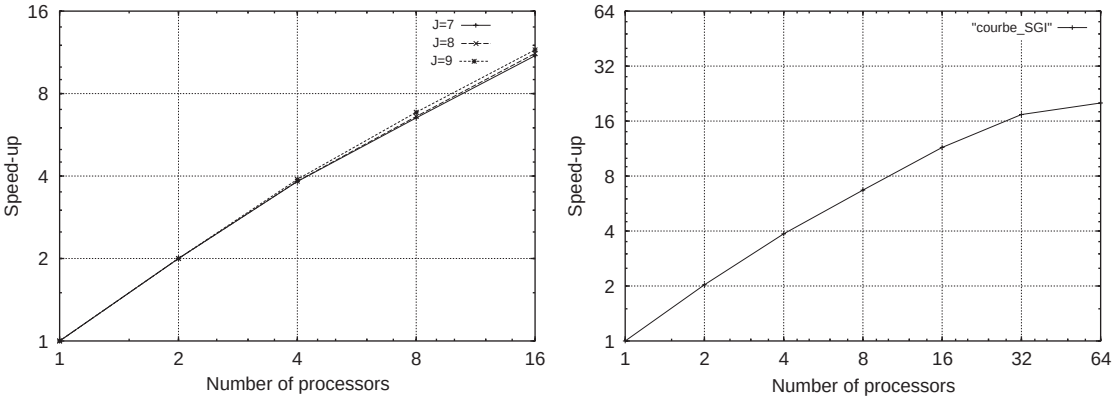



Fig. 3. Speed-up on a HP cluster (left) and on a SGI O3800 (right).

We have made a simulation of five periods (160 iterations). Our parallel code has been written in C++/MPI and tested (1) on a HP cluster, formed by 30 identical Itanium bi-processors nodes cadenced at 1.3 Ghz with 8 GB of main memory and interconnected through a switched 200 MB/s and (2) on a SGI Origin 3800, composed of $R14k$ running at 500 Mhz with 512 MB per node. The results are reported in Figs. 2 and 3. In Fig. 2, the finest level is $J = 8$, which corresponds to an underlying fine grid of $512 \times 512$ points. We observe that the adaptive grid follows very well the evolution of the fine structures. Fig. 3(left) shows the graphical representation of the speed-up on a logarithmic scale. We observe that the speed-up is approximatively constant as the level of details increases. Fig. 3(right) shows that the wall-clock time keeps decreasing as the number of processors increases up to 64 processors.

## 5. Conclusion and future work

In this paper, we have presented a Vlasov solver based on a hierarchical finite element interpolation. The inherent good localization of the cells has permitted to give an efficient parallel implementation. We have also shown numerical results which prove that our code works very well in the 2D phase space case. A 4D parallel code is being developed.

## References

[1] M. Gutnic, M. Haefele, I. Paun, E. Sonnendrücker, Vlasov simulations on an adaptive phase-space grid, Comput. Phys. Commun. 164 (2004) 214.
[2] M. Campos Pinto, Michel Mehrenberger, Adaptive numerical resolution of the Vlasov equation, IRMA Lectures in Mathematics and Theoretical Physics, 7 (2005), CEMRACS 2003, p 43.
[3] O. Hoenen, M. Mehrenberger, E. Violard, Parallelization of an Adaptive Solver, ParSim04, Special Session of EuroPVM/MPI2004, Lecture Notes in Computer Science, vol. 3241, September 2004, p 430.
[4] J.K. Lawder, P.J.H. King, Using Space-Filling Curves for Multi-dimensional Indexing, Lecture Notes in Computer Science, vol. 1832, Springer, Berlin, 2000.
[5] A. Patra, J.T. Oden, Problem decomposition for adaptive hp finite element methods, Comput. Syst. Eng. 6 (1995).
[6] E. Sonnendrücker, F. Filbet, A. Friedman, E. Oudet, J.L. Vay, Vlasov simulation of beams on a moving phase-space grid, Comput. Phys. Commun, 2004, vol. 164, p 390.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 192–195
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Simulating dark current in NLC structures ☆

Cho-Kuen Ng*, Nathan Folwell, Adam Guetz, Valentin Ivanov, Lie-Quan Lee, Zenghai Li, Greg Schussman, Kwok Ko

*Stanford Linear Accelerator Center, 2575 Sand Hill Road, Menlo Park, CA 94025, USA*

Available online 1 December 2005

**Abstract**

Dark current generation and capture are of great importance in high gradient accelerating structure R&D especially for the NLC which aims to operate at 65 MV/m with specific limits on dark current and RF breakdown rates. Although considerable effort has been devoted to building and testing various types of structures to meet these requirements, few theoretical studies have been done to understand these effects in actual structures. This paper focuses on the simulation of dark current in a NLC test structure for which experimental data are available. The parallel time-domain field solver Tau3P and the parallel particle tracking code Track3P are used together to simulate, for the first time, a dark current pulse to compare with the data measured downstream. Results from SLAC X-band 30-cell constant impedance structure for RF drive pulses with different rise times are presented and discussed.


*PACS:* 41.20.Cv; 02.70.Pt; 42.30.Va

*Keywords:* Particle accelerator; Electromagnetism; Dark current

## 1. Introduction

Next generation linear colliders strive to operate at high gradients to achieve greater efficiency. One limiting factor that prevents the accelerator from reaching high gradients is the generation of dark current. Dark current may lead to beam loading of the accelerator structure and, if captured, may also produce undesirable backgrounds downstream in the detector at the interaction point. Therefore, understanding the mechanism of dark current generation and capture is essential to the successful development of high gradient structures for future colliders. Thus far, dark current has been studied mostly through experiments in which measurements were made during high-power tests. There is growing interest in understanding the measured data using simulation in order to gain insight into the conditions and structure properties that lead to dark current generation.

Previous efforts in dark current simulation have focused on cylindrically symmetric structures and treated the accelerating fields in steady state. These assumptions ignore the 3D effects of the fundamental power couplers and the transient effects due to the finite pulse length of the accelerating fields. However, they greatly simplify the simulation and many results have been obtained with this approach [1–4]. This paper describes the simulation of dark current in a fully 3D model of the accelerating structure and includes the transient response to a realistic drive pulse. Specifically, we chose the SLAC X-band 30-cell constant impedance structure [5] for our end-to-end simulation which has been made possible by new parallel tools from the DOE SciDAC Accelerator Simulation Project.

Under SciDAC, SLAC has developed a new suite of 3D parallel electromagnetic codes based on unstructured grids for modeling large, complex accelerating cavities and RF structures to high accuracy using high-performance computers. The codes for use in dark current simulation include the parallel time-domain field solver, Tau3P [6], and the parallel particle-tracking module, Track3P [7]. Coupling Tau3P to Track3P provides the transient fields

☆Work supported by Department of Energy contract DE-AC03-76SF00515.
*Corresponding author.
*E-mail address:* cho@slac.stanford.edu (C.-K. Ng).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.072

needed for simulating the structure surface response to an RF drive pulse. Applying this simulation model to the 30-cell constant impedance structure allows us to study, for the first time, the dependence of dark current on pulse rise times so that comparison with available data can be made. At the same time, the transient fields from Tau3P are analyzed to understand the increase in peak surface fields due to shorter rise time and the correlation with enhancement in the dark current pulse.

## 2. Dark current simulation model

The simulation of dark current requires three components: a field solver for determining the electromagnetic fields in the accelerator structure, a physics model to govern the particle emission process when the fields interact with the structure surface, and a particle tracker that computes the trajectory of emitted particles under the influence of imposed fields.

(*a*) *Calculation of electromagnetic fields using Tau3P*: Tau3P is a 3D parallel time-domain solver which uses the discrete surface integrals method for unstructured grids to simulate large, complex accelerator structures. The use of unstructured grids allows for conformal meshing of curved surfaces which is particularly important for determining the peak surface fields to a high degree of accuracy. Parallelization provides the ability to handle large problem size and to speed up long simulation. Both capabilities are needed for modeling an entire structure over the duration of a finite RF pulse.

(*b*) *Particle tracking and surface physics in Track3P*: Track3P is a 3D parallel particle tracking code that computes particle motion through the Lorentz force equation using the Boris scheme [8] and with fields input from SLAC's parallel solvers such as Tau3P and Omega3P. On an unstructured grid, the localization of particle position and the interpolation of fields onto it require special treatment for good accuracy and computational efficiency. For example, we have implemented a uniform spatial subdivision scheme to facilitate the particle searching and work is in progress on developing a binned binary space partitioning scheme that can be an order of magnitude faster. The surface physics model in Track3P follows closely the emission processes described in the paper by Yamaguchi [1]. Field emitted or primary emissions are treated according to the standard Fowler–Nordheim formula [9] where the emission current is determined by the strength of the surface electric field. Secondary emissions are governed by the Lye and Dekker formula [10] which provides the yield of elastically as well as inelastically scattered electrons, and also true secondaries depending on the energy of the incident electron.

(*c*) *Coupling of Track3P to Tau3P*: In order that Track3P receives the instantaneous fields from Tau3P at each time step, the two codes have to be coupled directly. Although, in principle, one can compute and store the Tau3P fields (over a pulse length) beforehand, the data size would be too large for practical purposes. Our approach is to treat the computation like a PIC method but the fields are computed without charges and currents due to the emitted current, so the scheme is not self-consistent. With all parallel particle codes, the main computational issue is load balancing so that work is evenly divided among the processors at all times. Presently static load balancing is used based on the partitioning of the mesh. As the dark current evolves over time, the workload distribution between particles and fields can vary widely which reduces the parallel efficiency. Methods for improvement using dynamic load balancing schemes are under consideration.

## 3. Application to 30-cell constant impedance structure

(*a*) *Peak surface field increase during pulse rise time:* The SLAC X-band 30-cell constant impedance structure shown in Fig. 1 (iris opening same for all cavities) is a test structure for the NLC and suffered RF damage at the edge of the cell disks during high-power tests. It was suspected that RF breakdown was the cause for the damage and is due to extraordinarily high fields sustained at the disk surface. In an effort to understand the origin of these high fields, we drive a pulse with a given rise time and RF power through the structure using Tau3P and monitor the surface electric field around the iris edge of each cell. Fig. 2 (left) shows two instances in time of the pulse propagation through the structure. All surface emissions have been turned off. Fig. 2 (right) plots the surface electric field at a chosen disk location as a function of time for all 30 cells when excited by a pulse with a 10 ns rise time. It can be seen that, at each cell disk, the field overshoots during the pulse transient before settling down to its expected steady-state value. This overshoot is a dispersive effect due to the high-frequency content in the rise pulse. When adjusted for the wall loss (which is absent in this simulation) the maximum overshoot is about 17% over the steady state which is quite a significant increase. Simulation indicates that the increase gets smaller as the rise time gets longer. Since field emission

Fig. 1. (Left) The X-Band 30-cell constant impedance structure with fundamental couplers, and (right) the corresponding distributed mesh used in Tau3P simulation with different colors representing the portions of mesh being assigned to different processors.

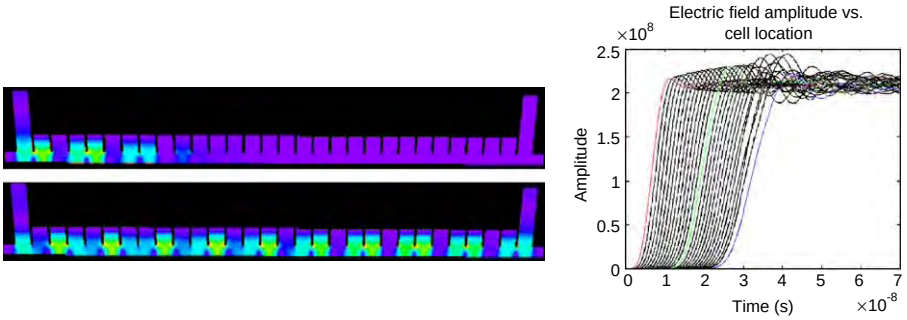



Fig. 2. (Left) Two snapshots in time of the pulse propagation in the 30-cell structure; (right) surface field monitored at cell disk in each of the 30 cavities as a function of time, showing overshoot due to the rising pulse (rise time = 10 ns).

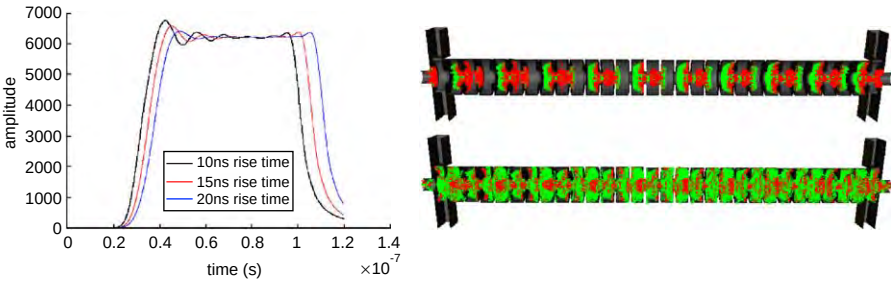



Fig. 3. (Left) Transmitted pulse shape monitored at the output waveguide for rise times of 10, 15 and 20 ns, respectively; (right) two instances in time of dark current generation with primary particles shown in red and secondaries in green.

depends on the surface field, this suggests some correlation with experimental data that the peak dark current also decreases with longer rise times.

(*b*) *Generation of dark current pulse:* To study dark current, the Tau3P simulation is now coupled to the particle tracking code Track3P which turns on the surface emission processes and tracks the emitted particles (as macro-particles) through the Tau3P fields. These fields are now adjusted by an attenuation factor to account for wall loss and Fig. 3 (left) shows the transmitted RF pulse (for three different rise times) monitored at the output waveguide. In Track3P's field emission model, the field enhancement factor $\beta$ is set to 40. The field emitted (or primary) particles, under the imposed fields, either hit the structure surface generating secondary particles or are captured and accelerated downstream. The simulated dark current is found by collecting all of the particles that exit the downstream beam pipe. A typical simulation using 30 Pentium-4 processors of a Linux cluster takes 12.6 h to advance 64,000 time steps for a mesh with about 500,000 hexahedral elements, and the average number of particles at each time step is about 150,000. Meeting the computational requirements and obtaining a fast turnaround time would not have been possible without parallel computing.

## 4. Comparison between simulations and experiments

One of the goals of high-power tests is the study of dark current generation because it is recognized that dark current may be the precursor to RF breakdown. In the high-power experiment, the dark current pulse is measured downstream of the structure for a range of pulse conditions that include different rise times, various field gradients and pulse lengths [5]. In this work, we focus on a set of data measured in the 30-cell structure that describes the dark current pulse as a function of three pulse rise times of 10, 15 and 20 ns, at a field gradient of 85 MV/m. The data plotted in Fig. 4b shows the input, output, and dark current pulses for each case. The same experiment is repeated using Tau3P and Track3P, and the results are summarized in Fig. 4a. It shows the computed dark current pulse from the 10 ns case in black, the 15 ns case in red, and the 20 ns case in blue, and that they are in reasonably good agreement with measurement.

## 5. Summary

This paper reports the first ever dark current simulation of an entire 3D realistic structure, with an actual RF pulse and including secondary emission, using newly developed parallel tools. Good agreement is found between numerical results and measured data on the dark current pulses generated for different rise times in a test structure. The shape of the dark current pulse is corroborated with the RF pulse shape which shows overshoot as a result of dispersive field effects during the rising pulse.

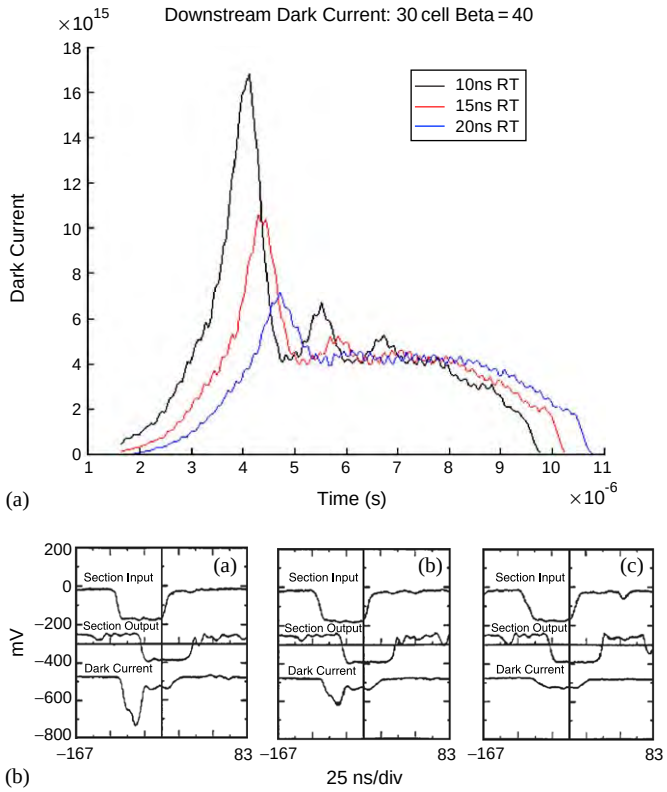



Fig. 4. (a) Dark current as a function of time for pulse rise times of 10 (black), 15 (red) and 20 (blue) ns; (b) experimental data for same set of rise times (10, 15, and 20 ns from left to right). From top to bottom are the input, output and dark current pulses, respectively.

## Acknowledgments

We thank Juwen Wang for many helpful discussions.

## References

[1] S. Yamaguchi, Simulation studies of high-gradient experiments, LAL/RT 92-18, December 1992.
[2] P.H. Stolz, M.A. Furman, J.-L. Vay, A.W. Nolvik, R.H. Cohen, Phys. Rev. 6 (2003) 054701.
[3] S. Setzer, W. Ackermann, T. Weiland, M. Krasilnilov, Proceedings of PAC-2003, p. 3566.
[4] K.L. Bane, V.A. Dolgashev, G.V. Stupakov, Simulation of dark currents in X-band accelerator structures, in: EPAC'04, Lucerne, Switzerland, 5–9 July 2004.
[5] J.W. Wang et al., High gradient studies on 11.4 GHz Copper Accelerator Structures, in: Proceedings of LINAC'92, Ottawa, Canada, 24–28 August 1992, pp. 716–718.
[6] M. Wolf, A. Guetz, C.-K. Ng, Modeling large accelerator structures with the parallel field solver Tau3P, in: Proceedings of 18th Annual Review of Progress in Applied Computational Electromagnetics, ACES, Monterey, CA, 2003.
[7] C. Ng, L. Ge, A. Guetz, V. Ivanov, Z. Li, G. Schussman, Y. Sun, M. Weiner, M. Wolf, Numerical studies of field gradient and dark currents in SLAC structures, in: ICAP-2002, East Lansing, MI, 15–18 October 2002.
[8] O. Buneman, Comput. Meth. Appl. Sci. Eng. 5 (1982) 529.
[9] R.H. Fowler, L. Nordheim, Proc. Roy. Soc. A 119 (1928) 173.
[10] R.G. Lye, A.J. Dekker, Phys. Rev. 107 (1957) 977.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 196–198
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Mathematical methods of data processing for high-energy introscopic examination of large-scale objects

Vladimir I. Petrunin

*D.V. Efremov Institute (NIIEFA), St. Petersburg, Russian Federation*

Available online 21 November 2005

**Abstract**

Optimized high-energy systems for introscopic examination of large-scale objects are considered. Advanced methods of mathematical data treatment for non-linear detector correction, image reconstruction and noise suppression based on the synthetic point spread function and continuous wavelet transformation are given. Experimental results for real, full-scale complexes are shown.


*PACS:* 07.05.Kf; 07.05.Pj; 07.05.Rm

*Keywords:* Introscopic examination; Detector correction; Image reconstruction; Noise suppression

## 1. Introduction

Data processing for high-energy introscopy and tomography (with photon boundary energy of 6–15 MeV) of large-scale objects (with integral absorption of approx. $10^4$) can be considered as three stages of data treatment. Routine operations in the first stage involve detector response leveling, dark current subtraction and dynamic range correction. This results in a so-called refined image suitable for the next processing step. In the second stage more powerful methods are used, in particular, detector response linearization, image reconstruction with point spread function (deconvolution), two-dimensional linear and non-linear trend elimination (global heterogeneity removal). In the third stage the most powerful mathematical methods are used—continuous wavelet transformation for effective noise suppression, cluster analysis for structure recognition and so called bio-correction (biological eyesight simulation). In this stage the primary image is intentionally changed according to our strategy for intellectual result reception (non-uniformity, emptiness, inclusions, suspicious substance identification).

## 2. Data correction

Two types of irregularity are apparent on the image, as vertical and horizontal stripes. Vertical stripes are due to photon flow non-stability; horizontal ones are the result of detector non-uniformity. There are also low-frequency variations due to the angular distribution of the photon beam and its temporal drift. Since the two aforementioned types of irregularity are statistically independent, they can easily be removed if found as vertical and horizontal stripes on an image without any objects (white field). In this case a two-dimensional corrected image $f^*(x,y)$ can be directly calculated from the primary image $f(x,y)$:

$$f^*(x,y) = C_1 \frac{f(x,y)}{\int_{x_1}^{x_2} f(x,y)\,\mathrm{d}x \int_{y_1}^{y_2} f(x,y)\,\mathrm{d}y},$$

where $x_1$ and $x_2$ define the two abscissa values and $y_1$ and $y_2$ the two ordinate values for image strips in which there are no objects (white field in the form of a cross with finite quantity of a surface).

In the corrected image, fast and slow fluctuations of the photon flow, the spread of the detector response, photon angular distribution and collimation system defects are eliminated. On the other hand, detector non-linearity is retained and cannot be corrected with a linear model. To eliminate image stripes at high absorption requires the use

*E-mail address:* petrunin@lek.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.004

of a moving ladder sample along the detector line to measure the detector response at each position along the sample thickness. Using this non-linearity matrix, the real image can be corrected.

The next problem is the elimination of two-dimensional non-linear brightness trends. Let us suppose that defects are found on a part of the image with a high local brightness gradient and thus cannot be seen. In this case leveling of the local averaged brightness can be achieved using wavelet theory. Decomposition of a two-dimensional image into a three-dimensional array of wavelet coefficients using continuous wavelet transformation allows realization of the image layer localization along a spatial frequency (see below). By removing layers with low spatial frequencies, the non-linear two-dimensional trend can be eliminated. The result is shown in Fig. 1.

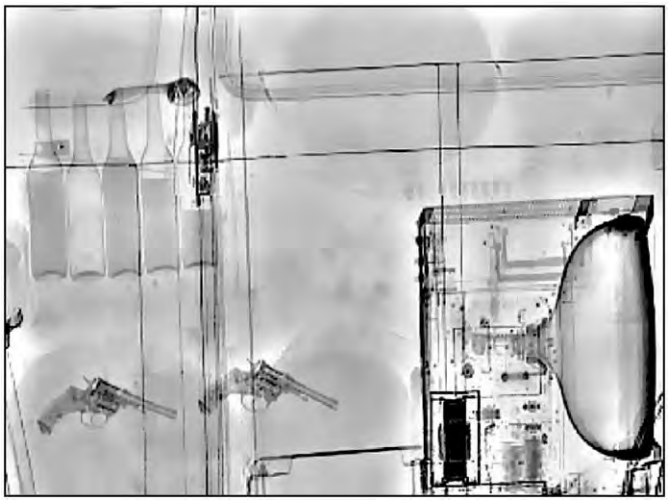


Fig. 1. Non-linear two-dimensional trend elimination.

## 3. Data reconstruction

X-ray images for object scanning are distorted due to the finite quantity of the linear accelerator focal spot, photon scattering and other factors. A direct method of true image revelation using Fourier transformation [1] has been used, but with the point spread function represented by an artificial synthetic function:

$$\Phi(x,y)=\left\{1-\sqrt{\left(\frac{x}{x_0}\right)^2+\left(\frac{y}{y_0}\right)^2}\right\}^N,$$

where $x_0,y_0$ denote the size of kernels along the $x$ and $y$ coordinates, and the exponent $N$ defines the shape of the kernel slopes (usually 8–16). The aforementioned parameters are experimentally defined according to the real point spread function of the linear accelerator. The sharp peak of the function leads to non-null Fourier coefficients and thus stability of the reconstruction process. A fragment of scanned container with hidden illegal objects in Fig. 1 has been reconstructed using a synthetic kernel.

## 4. Noise suppression

For noise filtration in introscopic images, the less popular continuous wavelet transformation (in contrast to discrete wavelet transformation) has been used. The main reason is its superfluity of image information, which is very useful for accurate processing of noisy objects.

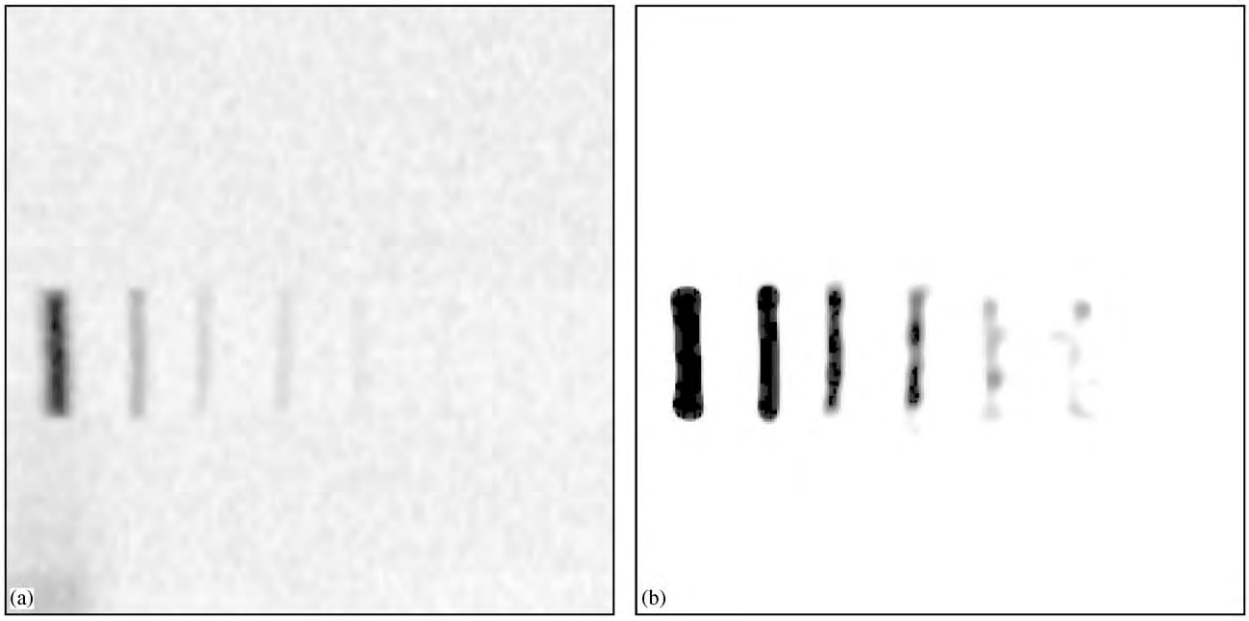



Fig. 2. (a) Primary image of wire samples. (b) CWT processed image.

Two-dimensional decomposition and subsequent image reconstruction are realized using the following expressions:

$$S(b,c,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\Psi\left(\frac{x-b}{a},\frac{y-c}{a}\right)\mathrm{d}x\,\mathrm{d}y,$$

$$f(x,y)=\frac{1}{c_{\mathrm{a}}}\int_{0}^{\infty}S(x,y,a)\,\frac{\mathrm{d}a}{a^{2}\sqrt{a}},$$

where $\Psi$ is the wavelet function, $a$ is a scale variable, the inverse value of which is spatial frequency, $b$ and $c$ are shift variables, $f$ is the primary two-dimensional image, and $S$ is the three-dimensional wavelet decomposition. It should be noted that in our case decomposition is performed by second derivation of a Gaussian function (Mexican hat), but reconstruction is by a $\delta$ function [2]. Flexibility of the continuous wavelet transformation can also be achieved using arbitrary non-integral or dyadic fractional scale variables:

$$a_j=a_0 2^{(j/j_{\max})\ln_2(a_n/a_0)},$$

where $a_0$ and $a_n$ are the minimum and maximum values of $a$, and $j$ and $j_{\max}$ are the current and maximum layer numbers. Proper choice of these parameters provides good accuracy of the decomposition–reconstruction cycle.

Further inquiry of the above-mentioned strategy shows the possibility of renormalization of the decomposition and reconstruction cycle. In other words, the weight function $1/a^2\sqrt{a}$ can be eliminated from reconstruction formulae. In this case, the decomposition formulae should also be changed:

$$S(b,c,a)=\frac{1}{a^{3}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)\Psi\left(\frac{x-b}{a},\frac{y-c}{a}\right)\mathrm{d}x\,\mathrm{d}y,$$

$$f(x,y)=\frac{1}{c_{\mathrm{a}}}\int_{0}^{\infty}S(x,y,a)\,\mathrm{d}a.$$

The amplitude of the wavelet coefficients decreases much sooner and reconstruction is a simple summation. At the same time the accuracy of the decomposition and reconstruction cycle is retained. An example using this strategy for a noisy image is shown in Fig. 2(a),(b).

## 5. Conclusion

Introscopic examination of large-scale objects with photon boundary energy of 15 MeV using the methods discussed above achieved highest resolution of 1-1T (according to the ASTM standard [3]) for a steel barrier of up to 400 mm in thickness.

## References

[1] T.J. Cornwell, Image restoration. In: Proceedings of the NATO Advanced Study Institute Conference on Diffraction-Limited Imaging with Very Large Telescopes, Cargèse, 1988, pp. 273–292.
[2] S.H. Maes, SIAM J. Appl. Math. 57 (6) (1997) 1763.
[3] ASTM E747-04. Standard Practice for Design, Manufacture and Material Grouping Classification of Wire Image Quality Indicators (IQI) Used for Radiology, ASTM International. http://www.astm.org

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 199–201
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Modeling and optimization of electron linac exit systems for nuclear technologies

A.N. Dovbnya, V.I. Nikiforov, V.L. Uvarov*

*NSC Kharkov Institute of Physics and Technology, 61108 Kharkov, Ukraine*

Available online 28 November 2005

**Abstract**

The possibility of producing linac-based high-intensity sources of high-energy bremsstrahlung ($>10^{15}$ photons/cm$^2$ s) and photoneutrons (up to $10^{14}$ neutrons/cm$^2$ s) holds much promise for the use of these sources in various branches of applied nuclear physics (medical isotope production, transmutation studies of nuclear-cycle long-lived by-products, testing materials for long-term immobilization of radioactive waste, etc.). An accelerator with an energy of $\geqslant 30$ MeV, beam power of $>10$ kW and the highest possible beam-current density appears appropriate for the generation of secondary radiation with the required parameters. However, the electron beam handling presents some thermophysical problems on account of high power absorption in the accelerator exit window, the bremsstrahlung converter and the target. Therefore, each of the technologies mentioned demands optimization of both the geometry and design of the exit systems and the accelerator beam parameters. This report offers some solutions to the problems mentioned by means of computer simulation based on GEANT and PENELOPE systems. The codes developed are supplemented with data on excitation functions of the corresponding photonuclear reactions.


*PACS:* 29.27.Ac; 25.20.Lj; 07.05.Tp

*Keywords:* Electron linac; Target; Secondary beams; Simulation; Isotope production

## 1. Introduction

The advantages of electron accelerators are their relative cheapness, the possibility of wide-range adjustment of particle energy and intensity, and control within certain limits over the spectral and geometrical parameters of secondary radiation (bremsstrahlung photons and photoneutrons). Therefore, in recent years increasing interest has been shown in the use of these facilities in different areas of nuclear technologies traditionally based on the use of reactors and heavy-particle accelerators. For instance, in a number of cases it appears possible to achieve the production of isotopes (e.g., the radionuclide $^{99}$Mo, popular in nuclear medicine) via the photonuclear channel with higher efficiency [1].

## 2. Photonuclear production of isotopes

To develop programs for isotope production in the photonuclear channel, a high-current linear accelerator was created, KUT-20 [2]. Its exit window and target station were investigated and optimized using the GEANT code in the geometry shown in Fig. 1 for $^{99}$Mo production.

Optimization was carried out for the following parameters:

- Total yield of the $^{100}$Mo$(\gamma, n)^{99}$Mo reaction per day of accelerator operation.
- Specific activity of the target.
- Thermal load at the exit window of the accelerator and at the target station elements.

Three versions of the target device were considered:

- A molybdenum metal cylinder, 10 mm in diameter and 18 mm in height (target 10-1).
- A Mo target as a truncated cone with a developed surface for efficient cooling (10-2).

*Corresponding author. Tel.: +380 57 335 63 32; fax: +380 57 335 37 31.
*E-mail addresses:* dan@kipt.kharkov.ua (A.N. Dovbnya), vinikiforov@kipt.kharkov.ua (V.I. Nikiforov), uvarov@kipt.kharkov.ua (V.L. Uvarov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.060

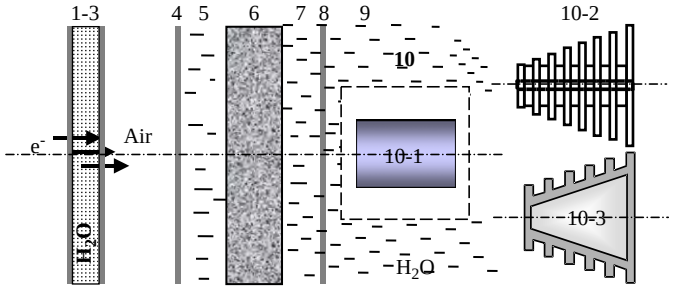



Fig. 1. Geometry of the exit devices. 1–3, exit window of the KUT-20 accelerator (two water-cooled Ti foils); 4, input foil of the target station; 6, Bremsstrahlung converter (tungsten); 5 and 7, water layers cooling the converter; 8, foil separating the converter chamber from the target device chamber; 9, water flow cooling the target; and 10, target.

Table 1
Yield of $^{99}$Mo for diverse target geometries ($10^{-4}$ nuclei/electron)

| Target | Energy (MeV) | | |
|---|---|---|---|
| | 25 | 30 | 40 |
| Solid $^{100}$Mo (10-2) | 4.60 | 7.99 | 13.93 |
| Powder $^{100}$Mo (10-3) | 3.70 | 5.75 | 10.95 |
| Oxide $^{100}MoO_3$ (10-3) | 2.51 | 4.24 | 7.28 |
| *Oxide $^{100}MoO_3$ (10-3) | 2.80 | 4.23 | 7.39 |

- An Al capsule containing $MoO_3$ powder (10-3).

Choice of the target version is determined by the conditions for particular isotope production and by the target treatment procedures after irradiation.

Considering that a divergent photon beam emerges from the converter, two more versions of the target as truncated cones (10-2 and 10-3) were studied. Table 1 gives data on $^{100}$Mo yield per accelerated electron for targets enriched in the $^{100}$Mo isotope. In the last line, the asterisk indicates data obtained for the real spectrum of the beam from the KUT-20 accelerator.

## 3. Transmutation

It is known that the nuclear energy cycle is accompanied by accumulation of a considerable amount of radioactive waste (RAW). In the initial stage of RAW disposal, the main dose load is mainly determined by the isotopes $^{137}$Cs and $^{90}$Sr.

Table 2 gives the calculations for transmutation of these isotopes in the irradiation geometry shown in Fig. 1 and with the version 10-1 target, 2 cm in diameter and 1 cm in height, at different electron energies. The simulation was performed based on the PENELOPE system.

## 4. Conclusion

The software packages developed allow computation and optimization of the radiation and thermophysical parameters for electron-accelerator exit systems to solve various problems in applied nuclear physics. Thus, for production of the medical isotope $^{99}$Mo, version 10-3 of the target (Fig. 1) appears to be the most acceptable. Experiments have corroborated the correctness of the simulation. Calculated data were also obtained for the production of some other isotopes [3].

The results of computer simulation of nuclear-cycle product transmutation make it possible to optimize the conditions for realization of this technology and to estimate its economic efficiency. In particular, computations show that the partition of elements is a necessary preliminary procedure for nuclear waste treatment.

Table 2
Yield from photonuclear transmutation of nuclear cycle products (10-1)

| Energy (MeV) | Isotope (nuclear transmutation coefficient, $10^{-4}$ nuclei/electron) | |
|---|---|---|
| | $^{137}$Cs | $^{90}$Sr |
| 30 | $2.16 \pm 0.44$ | $2.73 \pm 0.35$ |
| 40 | $3.78 \pm 0.58$ | $4.42 \pm 0.63$ |
| 70 | $7.39 \pm 0.82$ | $9.14 \pm 0.91$ |
| 100 | $9.63 \pm 0.93$ | $12.47 \pm 1.1$ |

Table 3
Simulation data for exit systems

| $\Delta W$ (cm) | $E_0$ (MeV) | $\Delta H_2O - 7 = 3$ mm | | | |
|---|---|---|---|---|---|
| | | ⊘beam = 5 mm | | | |
| | | $P$ (kW/10 kW) | | | $A_{Mo99}$, Ci/day × 10 kW |
| | | Ti-1 ($10^3$) | $W$ | Target 10-1 | |
| | 25 | 10.4 | 1.26 | 2.3 | 3.0 |
| 0.1 | 30 | 8.66 | 1.02 | 2.64 | 3.87 |
| | 35 | 7.42 | 0.87 | 2.89 | 4.44 |
| | 40 | 6.5 | 0.76 | 3.06 | 4.98 |
| | 25 | 10.4 | 2.98 | 1.18 | 2.84 |
| 0.2 | 30 | 8.66 | 2.38 | 1.48 | 3.59 |
| | 35 | 7.42 | 1.99 | 1.73 | 4.19 |
| | 40 | 6.5 | 1.71 | 1.94 | 4.69 |
| | 25 | 10.4 | 4.55 | 0.67 | 2.48 |
| 0.3 | 30 | 8.66 | 3.81 | 0.90 | 3.16 |
| | 35 | 7.42 | 3.25 | 1.10 | 3.73 |
| | 40 | 6.5 | 2.82 | 1.30 | 4.26 |
| | 25 | 10.4 | 5.46 | 0.43 | 2.06 |
| 0.4 | 30 | 8.66 | 4.85 | 0.59 | 2.75 |
| | 35 | 7.42 | 4.30 | 0.78 | 3.21 |
| | 40 | 6.5 | 3.83 | 0.92 | 3.69 |

## Appendix

Compared to reactors, electron accelerators are considerably cheaper, safer in service, and do not produce long-lived radioactive waste. Accelerator operation is not associated with the use or production of fissile substances.

On the other hand, the cross-sections for heavy-particle reactions are two or three orders of magnitude greater than those for photonuclear reactions. However, ionization losses for heavy particles in their interaction with the target increase to the same degree. Therefore, particles quickly leave the resonance region.

Table 3 presents the simulation results for target 10-1, which is made from natural molybdenum (the $^{100}$Mo isotope content is 9.62%). Here, $\Delta W$ is the $W$ converter thickness, $E_0$ is the accelerated electron energy, $\varnothing$ beam is the accelerated electron beam diameter, $\Delta H_2O - 7$ is the thickness of the cooling water layer in front of target 10, $P$ is the absorbed radiation power in exit systems elements for an electron beam of 10 kW power, Ti-1 is the Ti-foil exit window of the accelerator, $A_{Mo99}$ is the $^{99}$Mo activity induced in target 10-1 for 1 day at a beam power of 10 kW.

## References

[1] R.G. Bennett, J.D. Christian, D.A. Petti, W.K. Terry, S.B. Grower, Nucl. Technol. 126 (1999) 102.
[2] M.I. Ayzatsky, V.N. Boriskin, A.M. Dovbnya, V.A. Kushnir, Probl. Atom. Sci. Technol. Ser. Nucl. Phys. Invest. 2 (2003) 19.
[3] N.P. Dikiy, A.N. Dovbnya, V.L. Uvarov, Probl. Atom. Sci. Technol. Ser. Nucl. Phys. Invest. 2 (2003) 99.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 202–204
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# A parallel 3D particle-in-cell code with dynamic load balancing

Felix Wolfheimer*, Erion Gjonaj, Thomas Weiland

*Technische Universität Darmstadt, Institut für Theorie Elektromagnetischer Felder, Schloßgartenstr.8, 64283 Darmstadt, Germany*

Available online 28 November 2005

**Abstract**

A parallel 3D electrostatic Particle-In-Cell (PIC) code including an algorithm for modelling Space Charge Limited (SCL) emission [E. Gjonaj, T. Weiland, 3D-modeling of space-charge-limited electron emission. A charge conserving algorithm, Proceedings of the 11th Biennial IEEE Conference on Electromagnetic Field Computation, 2004] is presented. A domain decomposition technique based on orthogonal recursive bisection is used to parallelize the computation on a distributed memory environment of clustered workstations. For problems with a highly nonuniform and time dependent distribution of particles, e.g., bunch dynamics, a dynamic load balancing between the processes is needed to preserve the parallel performance. The algorithm for the detection of a load imbalance and the redistribution of the tasks among the processes is based on a weight function criterion, where the weight of a cell measures the computational load associated with it. The algorithm is studied with two examples. In the first example, multiple electron bunches as occurring in the S-DALINAC [A. Richter, Operational experience at the S-DALINAC, Proceedings of the Fifth European Particle Accelerator Conference, 1996] accelerator are simulated in the absence of space charge fields. In the second example, the SCL emission and electron trajectories in an electron gun are simulated.


*PACS:* 02.60.Cb; 41.75.−i; 89.20.Ff

*Keywords:* Computational electromagnetics; Particle-in-cell; Parallelization; Dynamic load balancing

## 1. PIC algorithm and SCL emission

The PIC algorithm [3] is commonly used to simulate electromagnetic field problems involving charged particles. In the electrostatic approximation, the electric field strength $\vec{E}$ is found by the solution of the Poisson equation $\nabla\varepsilon\nabla\phi = -\rho, \vec{E} = -\nabla\phi$. The equation of motion for the $i$th particle is given by the Newton-Lorentz equation $\mathrm{d}/\mathrm{d}t(\gamma\vec{v}_i) = q/m_0[\vec{E}(\vec{r}_i) + \vec{v}_i \times \vec{B}(\vec{r}_i)]$. For the discretisation of the problem the staggered grid approach of the Finite Integration Technique (FIT) [4] is used. The discrete Poisson equation in FIT notation reads

$$\widetilde{\mathbf{S}}\mathbf{M}_\varepsilon\widetilde{\mathbf{S}}^{\mathrm{T}}\mathbf{\Phi} = \mathbf{q}, \quad \widehat{\mathbf{e}} = \widetilde{\mathbf{S}}^{\mathrm{T}}\mathbf{\Phi} \tag{1}$$

where $\widetilde{\mathbf{S}}$ and $\widetilde{\mathbf{S}}^{\mathrm{T}}$ denote the discrete divergence and gradient operators and $\widehat{\mathbf{e}}$ is the vector of the electric voltages on the edges of the primary grid.

The algorithm used for the SCL particle emission was proposed in [1]. The algorithm is based on imposing exact charge conservation along emission surfaces of arbitrary curvature. In particular, the SCL condition of vanishing electric fields on the emission surface is fulfilled in every time step even for transient external fields and boundary conditions.

## 2. Parallelization of the PIC algorithm

One problem when parallelizing an algorithm concerns the decomposition of computational tasks on the processes. In this paper, an approach which couples the computational tasks with the grid cells by a weight function is proposed. The weight $w_{i,j,k}$ associated to a cell $(i,j,k)$ is the measure of the local computational cost for the update of fields and particles. Assuming that the simulation should be distributed among $N_p$ processes the total computational load is said to be ideally balanced between the processes, if the weight associated to each process is

*Corresponding author.
*E-mail address:* wolfheimer@temf.de (F. Wolfheimer).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.003

$w^{\text{ideal}} := (1/N_p)\sum w_{i,j,k}$. This optimal weight distribution can be approximately achieved for an arbitrary number of processes by a slightly modified orthogonal recursive bisection (ORB) of the computational domain. A domain, which has assigned $N_p$ processes and holds the weight $w$, is divided into two subdomains with weights $w^{\text{sub}_1}, w^{\text{sub}_2}$ in such a way that the constraint $N_1/N_2 = w^{\text{sub}_1}/w^{\text{sub}_2}$ is fulfilled, where $N_1 := \lceil N/2 \rceil$ and $N_2 := \lfloor N/2 \rfloor$ denote the number of processes assigned to the resulting subdomains. In Fig. 1a, a sample decomposition resulting from a simulation involving three processes is shown. In the parallel PIC algorithm the slowest of the processes determines the time needed for one PIC step, i.e., highly nonuniform distribution of weights due to the particle motion leads to a poorly balanced load. Thus the weights have to be rebalanced to preserve the parallel performance, i.e., the borders of the domains are adjusted as shown in the example of Fig. 1a. The decision, if a new partitioning of the computational tasks should occur, is based upon the value of the maximum normed deviation of the weights assigned to the processes from the ideal value $\sigma := \max_{i=1\ldots N_p}\{(w^{\text{sub}_i} - w^{\text{ideal}})/w^{\text{ideal}}\}$. The flowchart for the parallel PIC algorithm is shown in Fig. 1b.

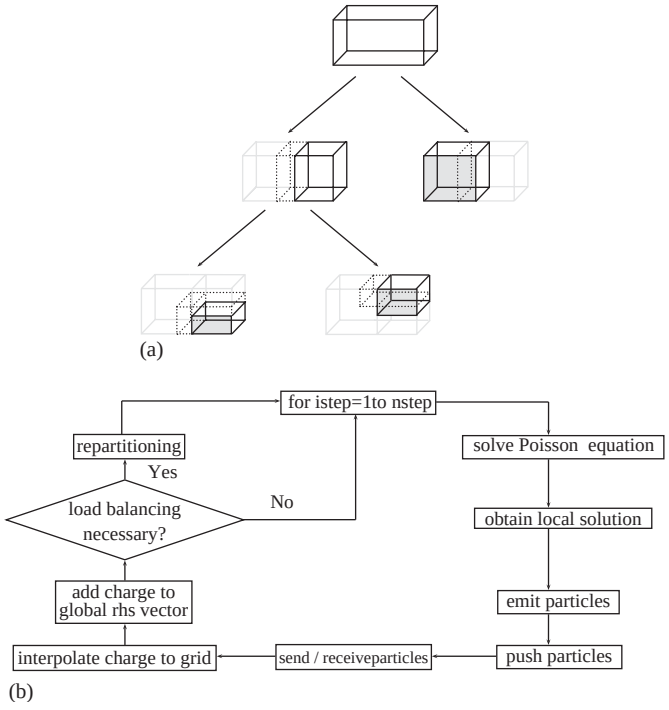



Fig. 1. (a) Hierarchical structure of the domain decomposition for three processes. In a repartitioning step the borders of the domains are adjusted as indicated by the dashed lines, (b) flowchart of the parallel PIC algorithm.

## 3. Simulation results

In a first simulation, multiple electron bunches as occurring in the S-DALINAC accelerator [2] are simulated in the absence of space charge fields. The bunches have a longitudinal extent of $10^{-4}$ m. The distance between two bunches is approximately $10^{-3}$ m. In this configuration a highly nonuniform distribution of the particles is given. In Fig. 2a is shown how the simulation load is efficiently balanced for a two-process-partitioning. In the case of particle tracking the weight function depends only on the number of particles in each cell. Thus, the particles are distributed equally on the processes after a repartition (see Fig. 2b). In Fig. 2c, the parallelization speedup based on the total simulation time is compared for simulations with and without dynamic load balancing, respectively, and for a different number of particles. The speedup curve is very close to the ideal one, if few processes are employed. However, it is less optimal, when the number of processes increases above a certain value. The simulation shows, that the critical number of parallel processes for optimum speedup depends on the number of particles used in the simulation. As repartitioning is expensive in terms of CPU time and interprocess communication, the influence of the parameter $\sigma$ on the parallel performance was studied in the second example of a high perveance electron gun [5]. The geometry of the model and the electron beam is shown in Fig. 3a. The emitted particles are accelerated in the cathode-anode gap of 100 kV voltage and the beam is focused by an immersing magnetic field. Fig. 3b shows the

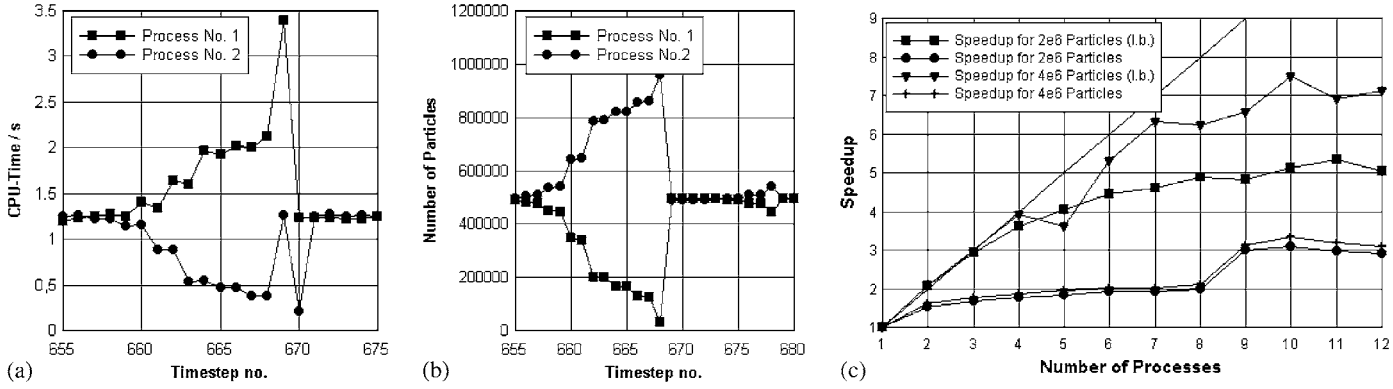



Fig. 2. (a) CPU time needed to complete the PIC loop for a simulation using two processes, (b) decomposition of the particles on the two processes, (c) speedup for different numbers of particles.

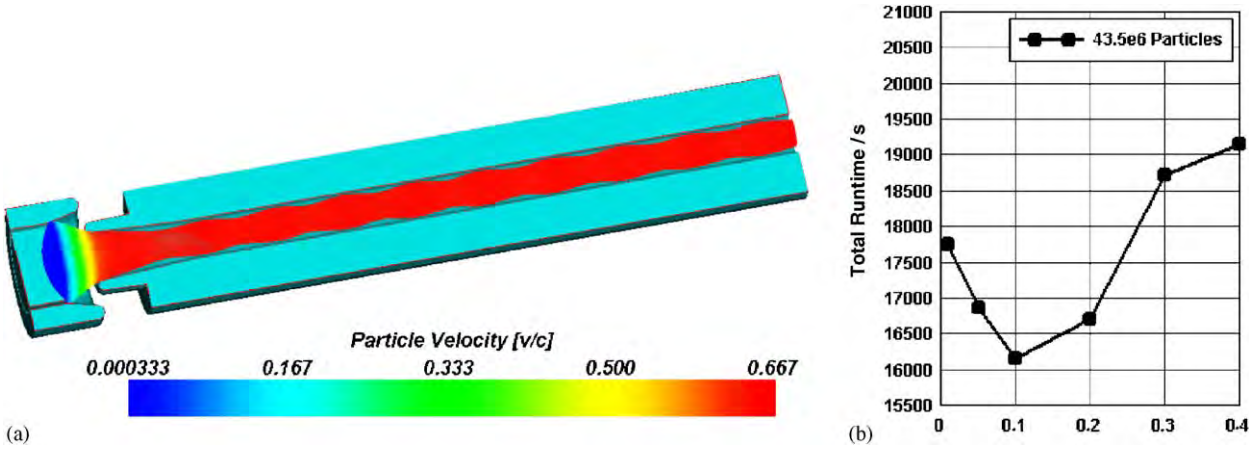



Fig. 3. (a) Geometry of the gun model including the electron beam, (b) dependency of the runtime on the value of $\sigma_{max}$.

total runtime of a simulation involving 43.5 million computational particles for different thresholds $\sigma_{max}$ on 8 processes. An optimal value of $\sigma_{max} \approx 0.1$ can be identified. This confirms the result from the first example, that dynamic load balancing leads to good results, as long as the number of particles per process is large.

## 4. Conclusions

A 3D parallel PIC Code with a dynamic load balancing algorithm was presented and tested in two simulation examples. It was shown, that the algorithm leads to a better parallel performance for particle dominated problems. Care should be taken in selecting the criterion for repartition. Since repartitioning is expensive in terms of CPU time and interprocess communication, performance deterioration may occur, if the repartition procedure is invoked too frequently.

## References

[1] E. Gjonaj, T. Weiland, 3D-modeling of space-charge-limited electron emission. A charge conserving algorithm, Proceedings of the 11th Biennial IEEE Conference on Electromagnetic Field Computation, 2004.
[2] A. Richter, Operational experience at the S-DALINAC, Proceedings of the Fifth European Particle Accelerator Conference, 1996.
[3] C.K. Birdsall, Plasma Physics via Computer Simulation, Institute of Physics Publishing, London, 1991.
[4] T. Weiland, A Discretisation Method for the Solution of Maxwell's Equations for Six-Component Fields, Electronics and Communications (AEÜ), vol. 31, 1977, pp. 116–121.
[5] S. Setzer, Ph.D. Thesis, TU Darmstadt, 2004.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 205–209
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Simulation of RF-focusing in the RFQ matching section by IGUN©

R. Becker*, R.A. Jameson

*Institut für Angewandte Physik der Universität Frankfurt, D-60054 Frankfurt/M, Germany*

Available online 13 December 2005

**Abstract**

In the design and optimisation of the matching between an ion source and a RFQ there is a lacking tool: Either the program for the calculation of ion extraction and matching to the RFQ (e.g. IGUN) has to stop at the entrance flange of the RFQ, because the RF-focusing of the RFQ electrodes (vanes or rods) cannot be taken into account, or the RFQ simulation program, like PARMTEQ, cannot take into account static electric-field, which originate from the matching lens. We therefore have developed a simple theory, how to take into account the RF-focusing of a RF-quadrupole structure as well as the beam spreading by emittance in a simulation with zero emittance by IGUN. This allows very effective use of IGUN to design the matching between an ion source and the RFQ up to the end of the radial matching section of a RFQ. IGUN automatically takes into account the local focusing strength of the RFQ electrodes to adjust the space-charge spreading by virtually adjusting the beam current. Handling of a finite emittance has also been added to IGUN.


*PACS:* 07.05.Tp; 29.17+w; 29.27.Ac

*Keywords:* Numerical simulation; RFQ matching; Ion extraction; Beam dynamics

## 1. Introduction

The direct injection scheme of the RIKEN laser ion source into a RFQ [1], created in the entrance of the RFQ radial matching section a region with overlapping electrostatic and RF-fields. The standard RFQ design program Parmteq in its version of PTEQHI [2] could not handle the plasma meniscus as well as the accelerating static field, while a standard ion extraction program, like IGUN [3–5] could not consider the reduction of beam spreading by starting of RF-focusing in this region or the finite emittance of the real beam (Fig. 1). On the basis of the beam envelope equation we present in this paper a method for IGUN to take into account RF-focusing as well as the enhanced beam spreading by a finite emittance [6].

*Corresponding author.
*E-mail address:* rbecker@physik.uni-frankfurt.de (R. Becker).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.073

## 2. Emittance and RF-focusing simulated by beam spreading in IGUN

The KV-equation, or more generally, the rms envelope equation

$$a'' = \frac{K_p}{a} + \frac{\varepsilon^2}{a^3} - \kappa a \tag{1}$$

describes the variation of the so-called smooth envelope in a RF-focusing channel, for example in the non-modulated entrance part of a RFQ accelerator. Here $a$ is the transverse rms beam size, and $K_p$ is defined as the perveance, representing the beam spreading by space charge. The rms emittance $\varepsilon$ is assumed constant or known a priori and $\kappa$ represents the focusing effect of the RF-field. The program IGUN could handle the beam spreading by space-charge (first term) correctly, but the second and third terms could not be treated. Although beam spreading by a finite emittance could be taken into account in IGUN by

thermal splitting of trajectories, no acceptance can be defined. We therefore modify the space-charge allocation in order to simulate with a zero emittance beam both, the RF-focusing and the finite emittance of an ion beam. How to proceed, can be directly seen from inspection of the KV-equation (Eq. (1)): Adding the last two terms to the first term, will create an "effective" perveance or current to describe beam spreading, emittance spreading and

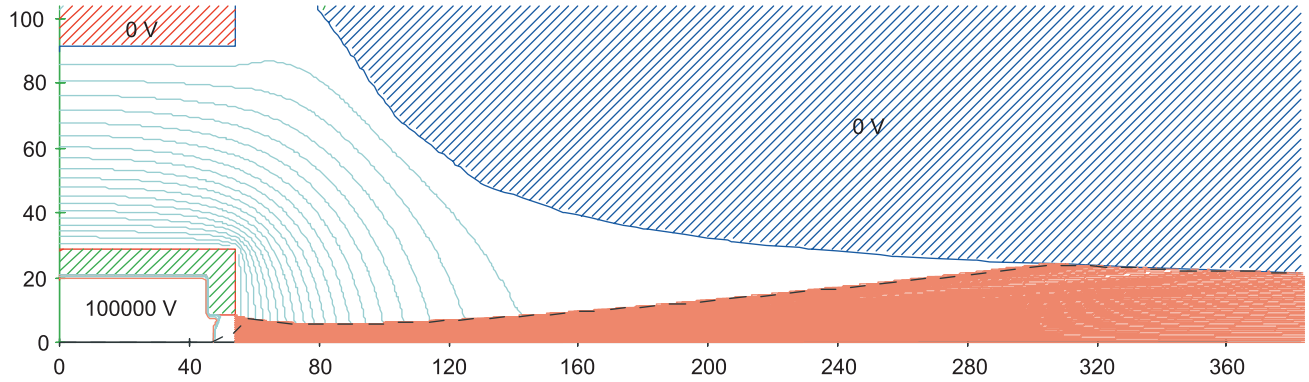



Fig. 1. Direct injection scheme of the RIKEN laser ion source, as an example for the overlap of DC and RF-field in the radial matching section, calculated by IGUN without considering the reduction of beam spreading by RF-focusing.

```
LIS-RFQ at 100 kV
&INPUT1  RLIM= 120, POTN=6, POT=0,100000,0,0,100000,0, BATCH=F,
RFQ=4, APERTF=0.6, ACCEPTF=0.2 &END
0, 0, 0                    20                  2
2, 5, 0                    1, 40, 44           6, 40, 147
2, 5, 11                   5, 40, 75           4, 40, 159
2, 2, 11                   1005, 31, 80        1004, 23.4, 173
2, 2, 13                   5                   14
1002, 7, 9.5               5, 26, 80           4, 11.8, 165.2
5                          1005, 31, 100       4, 9.66, 171.3
2, 12, 0                   5                   4, 8.38, 177.4
1, 40, 0                   5, 40, 105          4, 7.53, 183.6
1, 40, 36                  6, 40, 134          4, 6.92, 189.7
1001, 16.25, 56            1006, 20.25, 154    4, 6.44, 195.8
20                         20                  4, 6.07, 201.9
1, 7.33, 38.1              6, 11.33, 136.1     4, 5.54, 214.1
1001, 8, 40                1006, 12, 138       4, 5, 227
2                          2                   0, 0, 227
1, 7.1, 41.8               6, 10, 138          1, 0, 0
1001, 16.25, 24            1006, 12, 145
&INPUT5 START='PLAS', MASS=  3.0, NS=15,  AV=10,  AVR=1.0, HOLD=20,
AMPSO=0.104,
TE= 50, UI=1.2E3, UP=1.50002E+04, RP=  2, ZP=  13, MAXRAY= 50, STEP=0.2,
TI= 1, NSPLIT=-1, MASS=12,ZEND=147 &END
&BUNDLE ZB1= 12.5, RB2=  2, ZB2= 12.5, EEV= 1.2E+03,CHARGE=4,AMASS=12 &END
```

Fig. 2. An input file for IGUN to calculate the injection system from the RIKEN laser ion source to a RFQ as shown in Fig. 4. The relevant parameters are in line 3 as part of &INPUT1, the geometrical data about the RFQ vane shape in the right column, starting with potential number 4. Naming RFQ = 4 in line 3 is a signal to IGUN, to use the coordinates with potential number 4 in the right column as local values of $A(z)$. For more details about the variables and the interpretation of the geometrical data the reader is referred to the IGUN manual or to the web page www.egun-igun.com

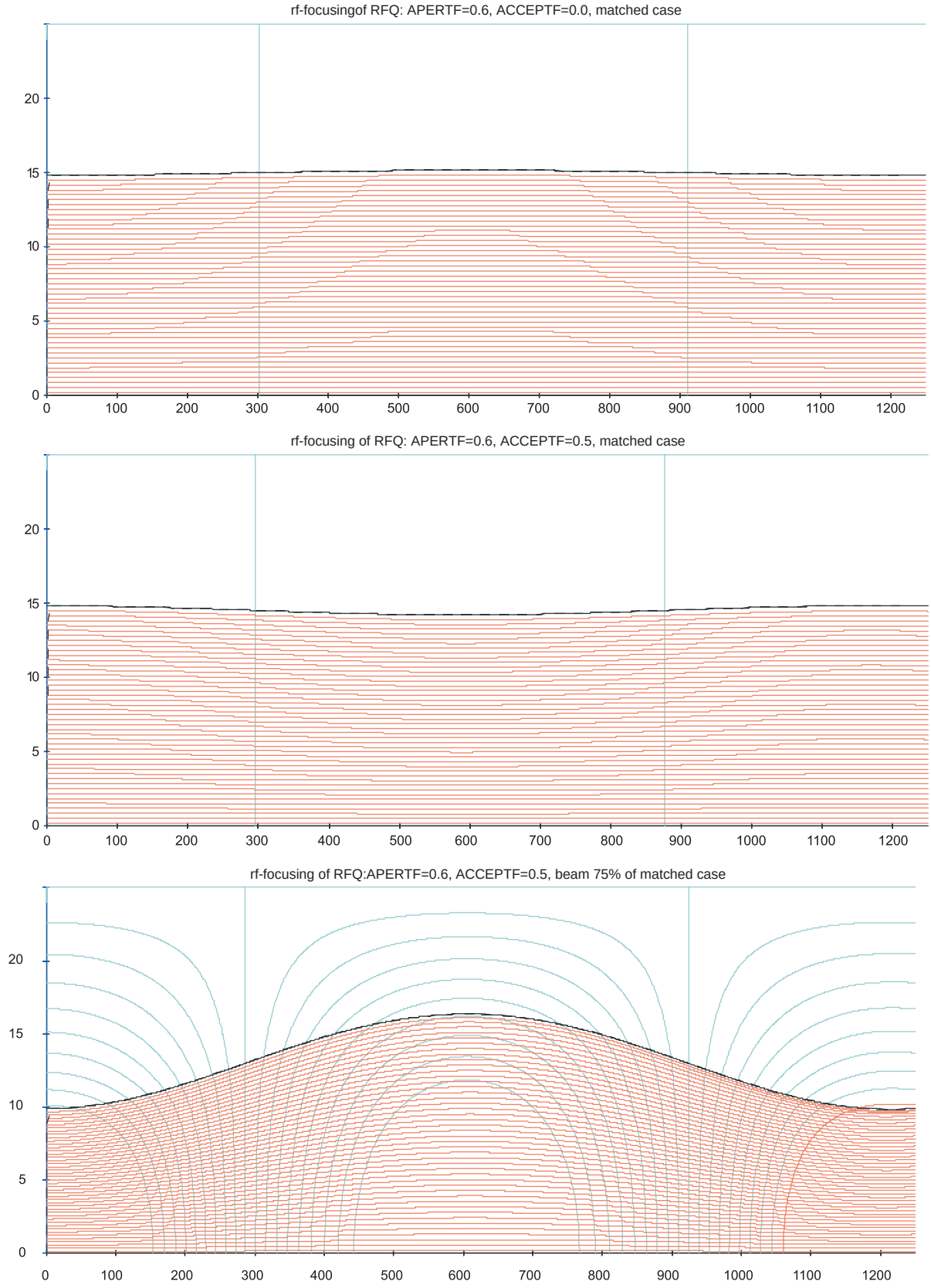



Fig. 3. (Top to bottom) Betatron oscillations in the RFQ as simulated by IGUN. (a) RF-focusing for 60% fill of aperture, no emittance (matched case), (b) RF-focusing for 60% fill of aperture and 50% fill of acceptance by emittance, (c) Mismatch of case (b) by an initial beam radius of 75% of the matched one.

RF-focusing as given for the RFQ by beam spreading of a laminar beam in IGUN.

$$a'' = \frac{K_p}{a}\left\{1 + \frac{\varepsilon^2}{K_p a^2} - \frac{\kappa a^2}{K_p}\right\} = \frac{K_{p,eff}}{a}. \quad (2)$$

We read the factor of the first term as an "effective" perveance and relate it to an effective current by

$$\frac{K_{p,eff}}{K_p} = \frac{I_{p,eff}}{I_p}. \quad (3)$$

The balanced (matched) condition $a'' = 0$ of the envelope equation can be used to express $K_p$ by corresponding values, indicated by subscript "0"

$$K_p = \kappa_0 a_0^2 - \frac{\varepsilon^2}{a_0^2}. \quad (4)$$

In the RFQ the coefficient of the focusing force $\kappa$ changes according to the vane aperture $A$

$$\kappa = \kappa_0 \left(\frac{A_0}{A}\right)^2 \quad (5)$$

where again the subscript "0" indicates the RFQ aperture where the beam is considered matched; for example, a few cells beyond the end of the radial matching section, which is typically 4–6 cells long. Using the definition for the acceptance

$$\alpha = a_0^2\sqrt{\kappa_0} \quad (6)$$

it is finally obtained

$$\frac{I_{p,eff}}{I_p} = 1 + \frac{(a_0/a)^4(\varepsilon/\alpha)^2}{1-(\varepsilon/\alpha)^2} - \frac{(aA_0/a_0A)^2}{1-(\varepsilon/\alpha)^2}. \quad (7)$$

At the match point with $a = a_0$, $A = A_0$ the effective current becomes zero, independent of the emittance to acceptance ratio, reflecting the homogeneous focusing action of an RFQ in smooth approximation with non-modulated electrodes (ignoring the envelope wiggles from the RF-fields).

## 3. Implementation into IGUN

The ratio of Eq. (7) will be used in IGUN to locally and virtually reduce the ion current for the space-charge allocation procedure. For this some variables are known to IGUN by boundary input, e.g. the local aperture $A$, and its special value $A_0$, where matching should occur. IGUN also has control on the beam boundary radii $a$, while $a_0$ can be calculated from the input parameter APERTFILL $= a_0/A_0$. The definition becomes complete by the second input parameter ACCEPTFILL $= \varepsilon/\alpha$, which is the ratio of beam emittance to the RFQ acceptance. (The equations could alternatively be expressed in terms of the "tune depression".). A typical input file for IGUN then may look like Fig. 2.

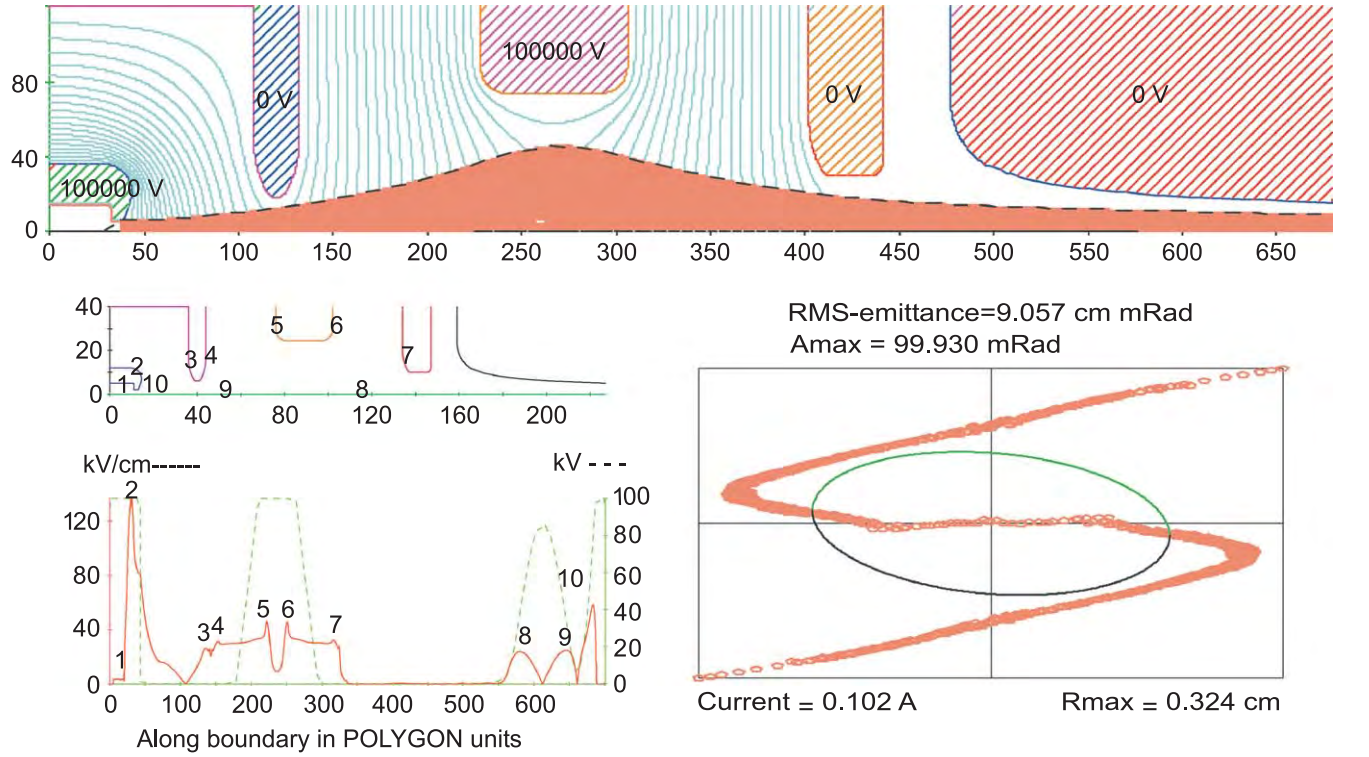



Fig. 4. Matching of the laser ion source of Fig. 1 to the RFQ by means of a decelerating electrostatic lens as ion source potential. The upper panel shows trajectories and equipotential lines, the left lower panel shows surface fields along the boundary, indicating that the critical field maxima at relative negative potential are below 35 kV/cm. In the lower panel to the right the final emittance is plotted, which is dominated by aberrations and exhibits an emittance growth of a factor of 30 with respect to the pure electrostatic ion gun.

## 4. Simulation of betatron oscillations for matched and mismatched beams

For the matched case $a = a_0$, $A = A_0$ the effective current becomes zero, independent of the emittance to acceptance ratio (Fig. 3a,b). For mismatched cases, the beam envelope shows betatron oscillations and a series of negative and positive space charge fields with corresponding equipotential lines for potential wells and hills (Fig. 3c). These potential wells and hills are a footprint of the simulation method: the virtual effective current according to Eq. (7) becomes positive, if the beam radius is smaller than the matched one and negative, if it is larger. In the "real" world, the potential inside the beam will always be positive with respect to the walls.

## 5. Matching calculation

As an example for the application of this new feature of IGUN, we present the result of simulations for the matching of the laser ion source—as shown in Fig. 1—to a RFQ by introduction of a decelerating electrostatic lens. The result of the IGUN simulation is shown in Fig. 4. In the (most likely) case, when the matching conditions are not met properly at the end of the radial matching section, the beam is left in a mismatched situation. As a result, it will perform betatron oscillations, as shown in Fig. 3c, which, however, may be damped by the beginning acceleration in the RFQ.

## 6. Conclusions

For the task of matching a high-intensity ion source to a RFQ a new feature has been added to IGUN, in order to correctly take into account the RF-focusing action of the non-modulated RFQ electrodes in the radial matching section. This allows the use of IGUN from the plasma meniscus to that point inside the RFQ, where axial acceleration starts by modulation of the electrodes. Application of this method will facilitate the end-to-end simulation of accelerating systems, because no transition to a LEBT simulation program will be needed. Additional freedom is available to the design of matching sections, because the end flange of the RFQ tank may be on a high-voltage potential. This has been shown as a very effective way to match high-current beams [7]. The new method has been successfully applied to the problem of matching a high intensity (104 mA $C^{6+}$) ion beam at 50 keV/$u$ energy to a RFQ, providing a decelerating electrostatic lens, which uses the ion source potential. The electrode distances and shapes could be optimised to limit the maximum electric field at relative negative surface parts to below 35 kV/cm.

## References

[1] M. Okamura, et al., Rev. Sci. Instrum. 73 (2002) 761.
[2] R.A. Jameson, A discussion of RFQ linac simulation, Los Alamos National Laboratory Report LA-CP-97-54, September 1997.
[3] R. Becker, W.B. Herrmannsfeldt, Rev. Sci. Instrum. 63 (1992) 2756.
[4] R. Becker, in: Proceedings of EPAC-98, Stockholm, http://accelconf.web.cern.ch/AccelConf/e98/PAPERS/THP44G.PDF
[5] www.egun-igun.com
[6] R. Becker, et al., EPAC-04, paper TUPLT025, http://accelconf.web.cern.ch/Accelconf/e92/PDF/EPAC1992_0816.PDF
[7] R. Becker, in: Proceedings of EPAC 1992, Berlin.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 210–213
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Accuracy of the manufacture of electrodes for a 433 MHz RFQ

A.A. Budtov[a], V.A. Gruzdev[a], V.I. Petrov[a,*], Y.A. Svistunov[a], G.V. Marinin[b]

[a]D.V. Efremov Scientific Research Institute of Electrophysical Apparatus (NIIEFA), Scientific Production Complex of Linear Accelerators and Cyclotrons (NPK LUTS), 196641 St. Petersburg, Russian Federation
[b]Russian Technologies Ltd., 195030 St. Peterburg, Russian Federation

Available online 21 November 2005

**Abstract**

Analysis of the dependence of the accuracy of the interelectrode distance on the accuracy of electrode surface machining for a 433 MHz four-segment radio-frequency quadrupole (RFQ) resonator is reported. The aim of the research was to determine the requirements for measurement methods and machining of the RFQ segments. Analysis of particle capture into acceleration as a function of the electrode modulation amplitude at the RFQ input is discussed.


*Keywords:* Accuracy; Electrodes; RFQ

## 1. Introduction

The accuracy of quadrupole symmetry of the electromagnetic field in a radio-frequency quadrupole (RFQ) structure is mainly determined by the accuracy of the manufacture of the segments. After manufacture and assembly of the four segments (Fig. 5), the electromagnetic field can be corrected using adjustment elements. The higher the manufacturing and assembly accuracy, the less is the necessity to use adjusting elements.

Fig. 1 shows the results of numerical calculations of the electric component of the electromagnetic field in a resonator with exact quadrupole symmetry.

Fig. 2 shows the result for similar calculations for a resonator with a displaced upper electrode; the top of the electrode is displaced to the right by 20 μm, which is $<1\%$ of the distance (2.899 mm) between the tops of neighbouring electrodes.

Analysis shows that realisation of tolerances for interelectrode distances of $\sim$20 μm and less requires a high level of technology, and consequently, methods for measurement and electrode machining corresponding to this level.

## 2. Analysis of tolerances

The accuracy requirements for the manufacture and assembly of RFQ resonators define the accuracy requirements for methods for measurement and machining of resonator segments. The analysis presented allows estimation of the accuracy requirements for these methods.

Fig. 3 shows the cross-section of the resonator. Fig. 6 displays the cross-section at the axes of the resonator. Dimension chains forming interelectrode distances $d = 7.000$ and $e = 2.899$ mm (shown in Fig. 4) were subjected to analysis for the conformity of both dimensions ($d$ and $e$) to a tolerance of 20 μm. It was supposed that the electrode modulation was ideally manufactured using a diamond mill, i.e. the error due to this mill was not taken into account.

A dimension chain determining the dimension of interelectrode distance $d$ is formed by the dimensions $a = 150.000$, $b = 126.500$ and $c = 16.500$ mm (Fig. 3).

Fig. 5 illustrates the dimension scheme for interelectrode distance $d = 7.000 \pm 0.025$ mm.

According to this scheme, the tolerance for dimension $d$ of $\pm 25$ μm is uniformly distributed among the dimensions $a$, $b$ and $c$ (assuming identical accuracy in machining of jointing planes) and hence the tolerance for dimensions $a$, $b$ and $c$ is $\pm 25/3$ μm $= \pm 8$ μm. Thus, the dimensions have values $a = 150.000 \pm 0.008$, $b = 126.500 \pm 0.008$ and

*Corresponding author.
*E-mail address:* npkluts@niiefa.spb.su (V.I. Petrov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.005

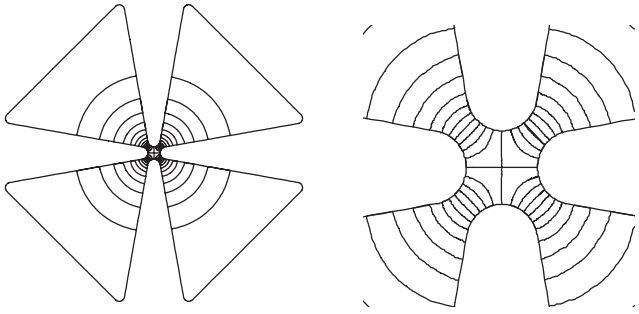


Fig. 1. Electric component of the electromagnetic field in a resonator with exact quadrupole symmetry.

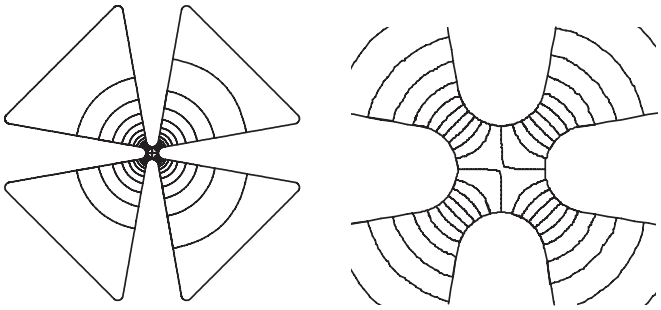


Fig. 2. Electric component of the electromagnetic field in a resonator with a displaced upper electrode.

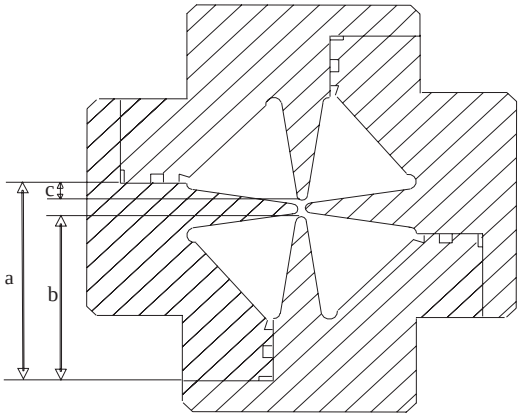



Fig. 3. Resonator cross-section.

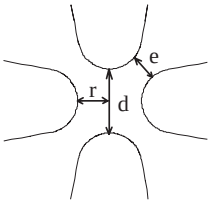



Fig. 4. Interelectrode distances $d$ and $e$ at the tops of the electrodes.

$c = 16.500 \pm 0.008$ mm. The maximum deviation of dimension $d$ takes place when dimension $a$ deflects in antiphase in relation to dimensions $b$ and $c$. For example, $d = a–b–c =$ 149.992–126.508–16.5086.976 mm, i.e., $d = 7.000–0.024$ mm, or $d = a–b–c = 150.008–126.492–16.492 = 7.024$ mm.

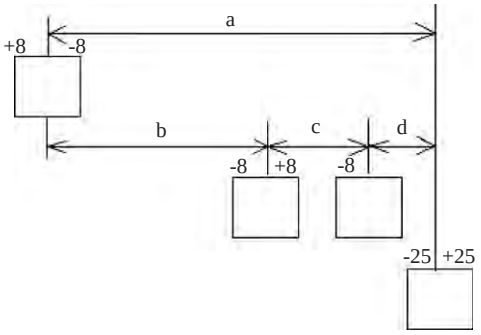



Fig. 5. Dimension scheme for interelectrode distance $d$.

Furthermore, it is necessary to choose a tolerance level for non-flatness of the jointing planes corresponding to the dimension chain considered. The most reliable value for tolerance of non-flatness for jointing planes is one-tenth of the tolerance for the dimensions, i.e. for the case given, $16\,\mu m \times 0.1 = 1.6\,\mu m$, but achievement of such accuracy involves significant costs. A more optimum solution from the point of view of both costs and reliability is a factor of 0.2 of the dimension tolerance, i.e., $16\,\mu m \times 0.2 = 3.2\,\mu m$. Similar arguments are applicable to the measurement error for non-flatness—it is more reliable to have one-tenth of the tolerance value for non-flatness, i.e. $3.2\,\mu m \times 0.1 = 0.32\,\mu m$, but a more realistic variation is $3.2\,\mu m \times 0.2 = 0.64\,\mu m$.

Thus, for the nominal dimension $d = 7.000$ mm and tolerance $\pm 25\,\mu m$, we have tolerance for dimensions $a$, $b$ and $c$ of $16\,\mu m$, non-flatness tolerance of $3.2\,\mu m$, and measurement error for non-flatness of $0.64\,\mu m$.

Fig. 6 shows the cross-section of the RFQ resonator; the dimensions specified form the dimension chain determining the interelectrode distance $e$ in Fig. 4.

Distance $e$ changes to the greatest degree when both adjacent electrodes are simultaneously displaced towards or apart from each other. Thus, the value $\Delta$ (change in interelectrode distance $e$) is connected to $\Lambda$ (displacement of electrodes on the $x$ and $y$ axes; Fig. 6) by $\Delta = \sqrt{2}\Lambda$, i.e., if dimension $e = 2.899$ mm has tolerance $\Delta = \pm 25\,\mu m$, the displacement of each electrode on the $x$ or $y$ axis for the centre of quadrupole symmetry should not exceed $\Lambda = \pm 25\,\mu m/\sqrt{2} \cong \pm 18\,\mu m$; thus, we have $r = 3.5 \pm 0.018$ mm (Fig. 4).

The dimension chain dictating the interelectrode dimension $e$ is shown in Fig. 7. Here $r$ is the distance from the top of the electrode to the centre of quadrupole symmetry (Fig. 4). As the deviation of the electrode from the nominal position is defined by two dimensions, $f$ and $g$, the tolerance for these dimensions will be $\Lambda/2 = \pm 18\,\mu m/2 = \pm 9\,\mu m$, i.e., $f = 130.000 \pm 0.009$ mm and $g = 126.000 \pm 0.009$ mm. Assuming a tolerance for non-flatness of the jointing planes of $18\,\mu m \times 0.2 = 3.6\,\mu m$, we obtain a value for allowable measurement error for non-flatness of $3.6\,\mu m \times 0.2 = 0.72\,\mu m$. A summary of the results is shown in Table 1.

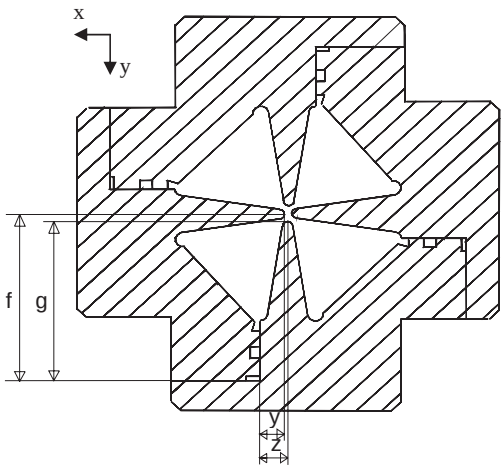



Fig. 6. Resonator cross-section.

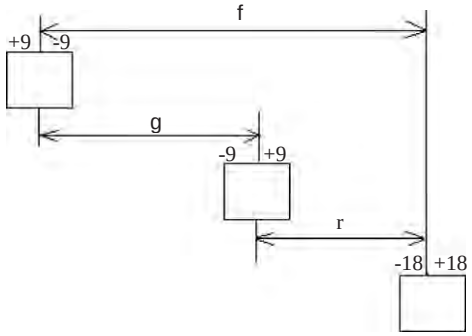



Fig. 7. Dimension chain for interelectrode distance $e$.

Table 1
Summary of calculated results

| | Dimension | |
|---|---|---|
| | $d$ | $e$ |
| Dimension tolerance (μm) | ±25 | ±25 |
| Tolerance for non-flatness of jointing planes (μm) | 3.2 | 3.6 |
| Non-flatness measurement accuracy for jointing planes (μm) | 0.64 | 0.72 |

Practice has shown that such requirements assume temperature monitoring in the room and machine tool capability (elements and detail) to within 1/10 and even 1/100 of 1°. In particular, achieving a measurement error of 1 μm for the NPK LUTS used in the HS328 machine tool required both modification of the machine tool software and improvement of its cooling system to maintain the coolant temperature and elements of the machine tool within a narrower temperature range.

Figs. 8 and 9 show diagrams reflecting the dependence of the tolerance for non-flatness of the jointing planes ε and the required measurement accuracy ζ on tolerance δ for interelectrode distance $d$.

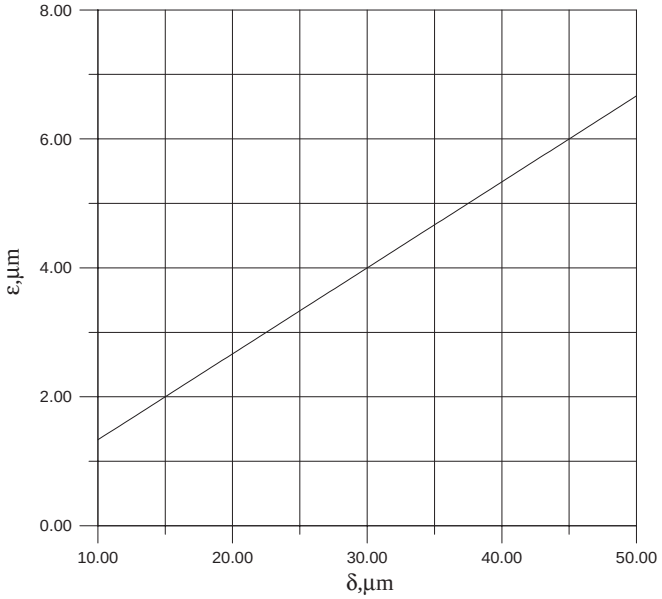



Fig. 8. Dependence of the non-flatness tolerance for jointing planes ε on tolerance δ for interelectrode distance $d$.

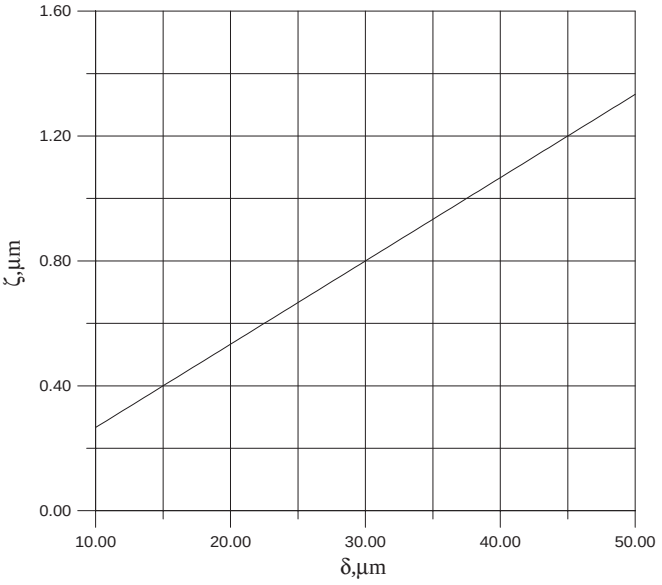



Fig. 9. Dependence of the required measurement accuracy ζ on tolerance δ for interelectrode distance $d$.

## 3. Results of particle dynamics optimisation

A number of measures to improve the HS328 machine tool are now planned to realise an accuracy of 2–3 μm for machining of the jointing planes and electrode modulation.

As stated earlier, numerical calculations for optimisation of the particle dynamics for a 1 MeV deuteron 433 MHz RFQ structure have been carried out. The calculations take into account the opportunity to downsize the electrode modulation amplitude at an accelerating–focusing channel input from 40 to 10 μm. The purpose of this optimisation

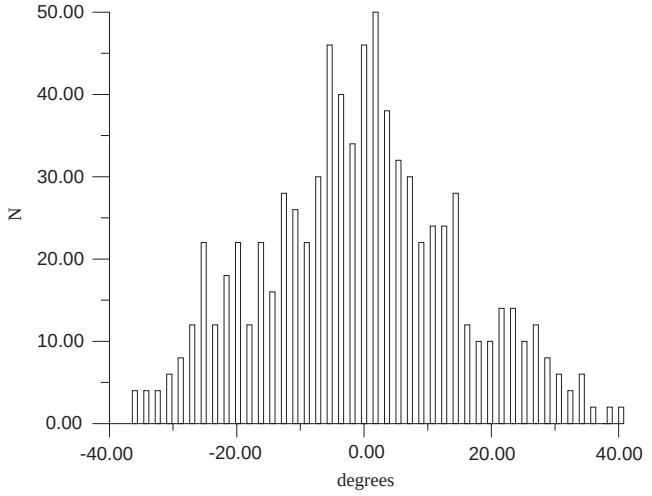



Fig. 10. Phase spectrum of the output beam in the structure before optimisation.

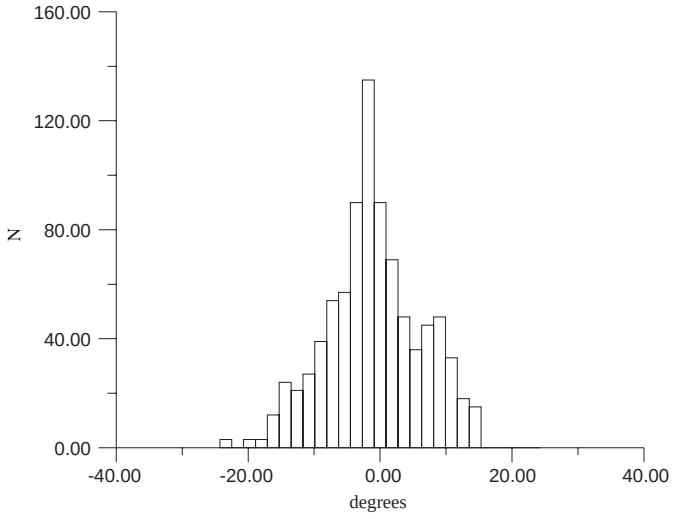



Fig. 11. Phase spectrum of the output beam in a structure with smaller electrode modulation amplitude.

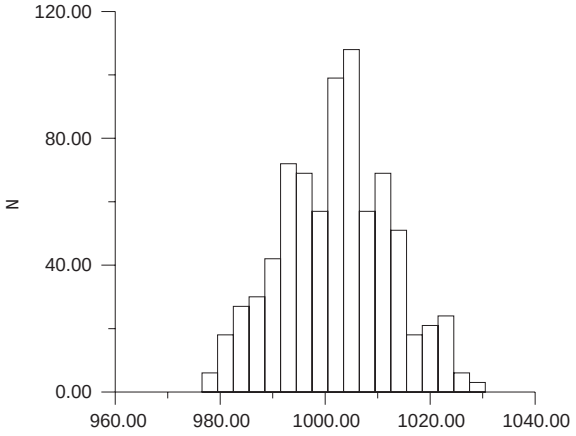



Fig. 12. Power spectrum for the output beam in the structure before optimisation.

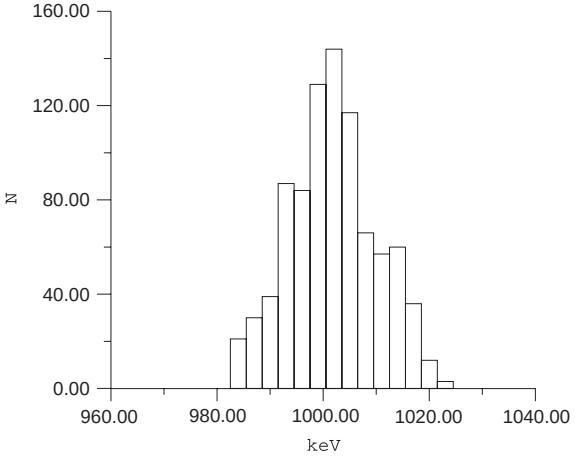



Fig. 13. Power spectrum of the output beam in a structure with smaller electrode modulation amplitude.

was to increase particle capture into acceleration for a 15 mA input beam (while preserving the length of the structure at 2300 mm) with input modulation amplitude of approximately 10 μm.

By optimising the particle dynamics, particle capture into acceleration was increased from 80% to 91% and the phase length of output bunches decreased from 72° to 36°. Fig. 10 shows the phase spectrum for the output beam in the structure before optimisation. Fig. 11 illustrates the phase spectrum in the structure for a smaller electrode modulation amplitude. Figs. 12 and 13 illustrate power spectra for the output beam for these two variations.

## 4. Conclusion

Analysis of the tolerances for a 433 MHz RFQ shows that for tolerance of $\pm 25$ μm for the interelectrode distances, the error in measurements of the machining of jointing planes must be approximately 1 μm. Correspondingly, the tolerance for the non-flatness of jointing planes must be approximately 3 μm. At present, using the HS328 machine tool it was possible to achieve measurement error of 1 μm, while the realisation of 3 μm for non-flatness of the jointing planes will require additional efforts. Resolution of these problems will essentially allow simplification of the procedure for RF tuning.

Available online at www.sciencedirect.com
SCIENCE DIRECT
Nuclear Instruments and Methods in Physics Research A 558 (2006) 214–219
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Investigations of energy penetration for the 2003 design of the target cell at the Hermes experiment

Susan Wipf*

*Deutsches Elektronen Synchrotron, DESY, Hoelderlinstrasse 22, D22607 Hamburg, Germany*

Available online 28 December 2005

**Abstract**

Calculations have been made to investigate the energy penetration through the pumping holes of the beam tube at the interaction point of the HERMES experiment at HERA. A shielding screen made from etched nickel foil, pressed into the aluminium of the target cell, was proposed. Several different models were studied. Calculations for the new design of the spring fingers were also made, and the results for the 2000 design are included for comparison. As there have been no measurements as yet, the 2000 values are the only indication of an upper limit of radio frequency interference which is not tolerable for the detector electronics. The effect of discretisation on the results was also analysed.


*PACS:* 07.05.Tp; 03.50.De; 41.20.–q

*Keywords:* Field computation and analysis; Design of accelerators and components

## 1. Introduction

The 2003 design of the target cell requires six pumping holes on either side of the elliptical beam tube at the interaction point of the HERMES experiment. Fig. 1 shows the present installation. These holes would be stamped out of the sides of the beam tube, 4 mm high, 10 mm long, with a separation of 2.5 mm, see the insert in Fig. 1. The radii of the elliptical beam pipe of the target cell are $4.45 \times 10.5$ mm.

Previous experience with insufficiently shielded gaps in the beam tube at other positions of the target cell, i.e. the spring fingers which connect the target cell to the wakefield suppressor on the one side, at 0.347 m, and the C2 collimator, at −0.24 m, on the other, has shown that high-frequency electro-magnetic fields from the beam can saturate the detector electronics so that data cannot be taken.

A fine screen pressed into the aluminium of the target cell has been proposed for shielding the sensitive detector electronics. The screen would be formed by etching nickel foil, 20 μm thick, leaving ~12 line/cm with 90% transparency. The width of the metal separating the square cells would be 40 μm and the thickness 20 μm. The cells are 0.4 mm square.

Preliminary calculations were first carried out to evaluate possible models for the shielding. Then a simpler model was used to examine the mechanism of coupling through a fine screen. Two other simplified models were used to investigate the effect of discretisation on the results directly and also on the inductance, as a qualitative measure of the coupling. After that, a detailed model was constructed. The calculation used an excitation which allowed a variable mesh in the beam direction. All calculations were carried out with the MAFIA [1] programs.

Calculations for the new design of the spring fingers, using much shorter and narrower gaps are also presented here. This provides a comparison between a model with elliptical beam cross-section and the rectangular approximation, which is used in the other calculations. The previous design had also been calculated and is included

*Corresponding author. Tel.: +49 40 827139, +49 899 82466; fax: +49 899 42466.
*E-mail address:* susan.wipf@desy.de.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.187

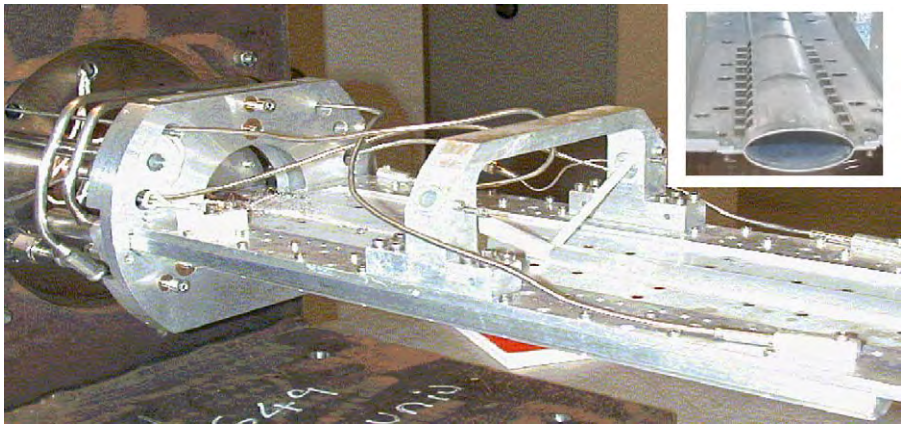


Fig. 1. Hermes 2000 target cell, with beam tube in place (no pumping holes). Inset shows the type of pumping holes used at the target cell.

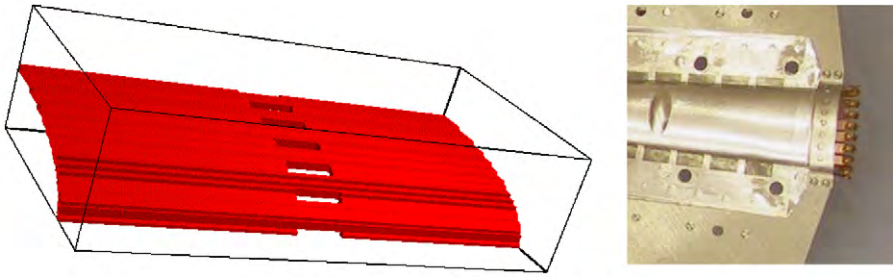


Fig. 2. Elliptical calculation model with a view of new spring fingers.

for comparison, as this is the only available indication of the level of RF power which is not tolerable.

## 2. Spring fingers—comparison with previous calculations

So far there have been no measurements which can be correlated to the amount of RF noise tolerable for the diagnostic equipment. However, in 2000 the level was so high that data could not be taken.

Spring fingers are used in the first place to provide RF shielding and to facilitate the removal of the target cell when necessary, but also as pumping slits. Naturally, pumping power is only achieved at the expense of the shielding. The gaps in the beam tube have been reduced in the 2003 design, reducing the calculated energy penetration by five orders of magnitude, from 183 mW to 1.3 μW.

- 2003 *Design*: fingers 2 mm wide, 5 mm long, with a spacing of 0.8 mm.
- 2000 *Design*: fingers 2 mm wide, 35.5 mm long, with a spacing of 5 mm.

The equipment around the beam tube could not be modelled accurately, thus multiple reflections in that area were neglected. For a good approximation of the actual energy transfer, a half wavelength of empty space round the beam tube was modelled, with sufficiently fine discretisation to allow EM waves to propagate. This means that the model must include at least 10 mm empty space for the case of a 3 mm sigma bunch and then 40 mm for a 12 mm sigma bunch.

It was also shown that using a rectangular instead of an elliptical model for the beam tube underestimated the energy penetration by less than a factor of 2.

Fig. 2 shows the elliptical calculation model and the 2003 design for the spring fingers.

## 3. Preliminary calculations

For the preliminary calculations the screen was modelled with a coarser mesh, $12 \times 2$ wires instead of $23 \times 5$, a rectangular beam tube was used and $\frac{1}{4}$ of the geometry was modelled. *Sheets*, elements with zero thickness available in the MAFIA [1] programs, were used to model the wires of the screen.

The following assumptions were made:

- Shielding is effective: after bunch passage, no coupling between beam tube and outer volume;
- field energy in outer volume is a measure of the coupling;
- energy loss for all bunches adds up incoherently and is absorbed in the outer volume.

Five models were compared. (1) Screen over the holes, surrounded by a cage; (2) cage surrounding open holes;

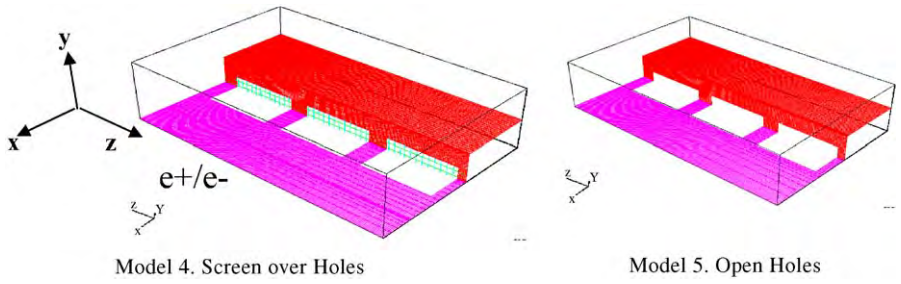



Fig. 3. Geometry of models 4 and 5.

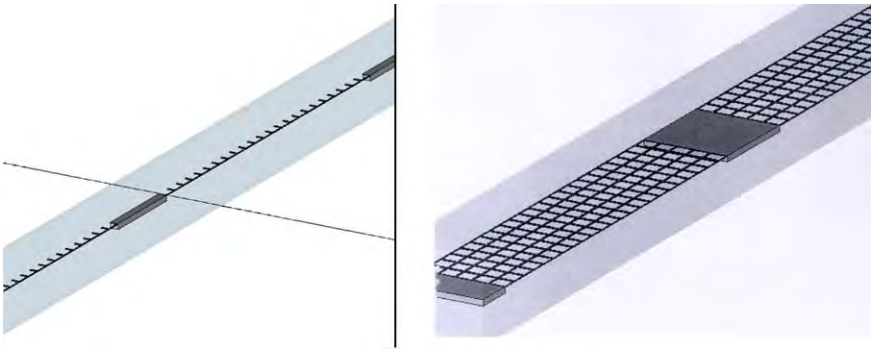


Fig. 4. Left hand side: calculation model. Right hand side: model with boundary conditions.

(3) two horizontal wires; (4) screen over the holes; (5) open holes.

Models 4 and 5, shown in Fig. 3, were chosen for further investigation. Model 3 shields better than Model 4, but the wires cannot be attached firmly. Model 1 provides better shielding but would only be considered if Model 4 proves insufficient, as it is mechanically less stable.

## 4. Simplified models

### 4.1. Transmission line

In order to gain insight into these results, a much simplified calculation model was designed, see Fig. 4; a transmission line consisting of three electric plates with two waveguide ports at each end was used, as shown in Fig. 5. The model represented the six holes in the beam pipe, each 10 mm long with a 2 mm separation bar; the holes in the screen were $400 \times 400\,\mu m$, with $40\,\mu m$ for the width of the wire. Periodic boundary conditions were applied transversally producing an infinitely wide model.

The reflection, ($r$), and transmission, ($t$), together with the coupling, ($c$) and the isolation, ($i$), were calculated (see Fig. 5). The excitation is a Gaussian pulse which travels from the bottom left at the speed of light. As the main scattering effect seemed to come from the edges of the pumping holes and to a lesser extent from the edge of each of the holes in the screen (see Fig. 6), the shape of the screen holes was varied, using rectangular holes two and four times as long. A tapering of the ends of the holes was also tried.

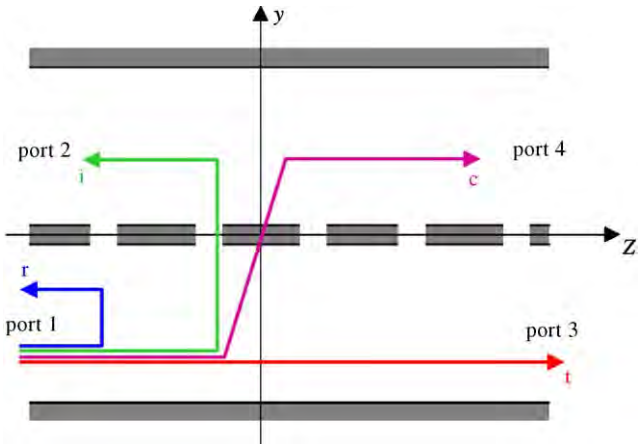



Fig. 5. Four ports are monitored to calculate the transmission ($t$), reflection ($r$), coupling ($c$) and isolation ($i$).

The energy transfer calculated for each port was

for sigma = 3 mm $(r) = 1.5\%$, $(i) = 0.18\%$,
$(t) = 95.2\%$, $(c) = 3.1\%$,

for sigma = 12 mm $(r) = 0.32\%$, $(i) = 0.025\%$,
$(t) = 99.4\%$, $(c) = 0.2\%$.

For a 12 mm bunch the energy penetration was 0.11 mW. Using rectangular holes, with a 2:1 ratio, this could be reduced to 0.051 mW and with 4:1, to 0.038 mW.

### 4.2. Discretisation test—TEM conductors

In so fine a structure as the shielding screen, it was important to establish to what degree the discretisation affected the results obtained. Outside the beam tube there is a large area where waves can propagate. In order to investigate the effect of varying the mesh density within and either side of the screen, a waveguiding structure was introduced; a model with TEM conductor plates, see Fig. 7, guides and reflects the electromagnetic waves, representing a similar physical situation in a smaller volume with better discretisation.

As expected, the actual size of the holes in the screen plays a large role, so that the coarser screen used in Section 3 overestimates the power transmission by two orders of magnitude. However, the fineness of the discretisation within these holes also affects the results, making the screen more transparent. This is not in contradiction to the results in Section 4.1, as there the holes were changed from square to rectangular. It was seen in Model 3 of Section 3 that the wires parallel to the beam direction have a screening effect, while those perpendicular to the beam cause reflections.

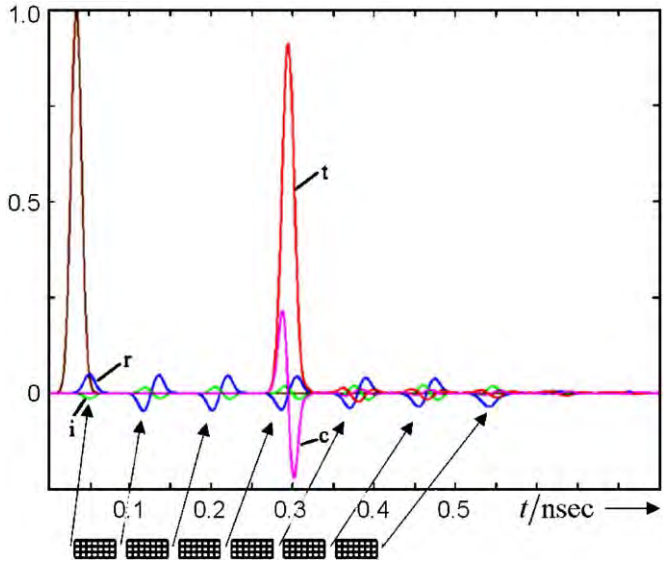



Fig. 6. Energy transfer to the four ports is plotted against time, showing the reflections from the edges of the six holes.

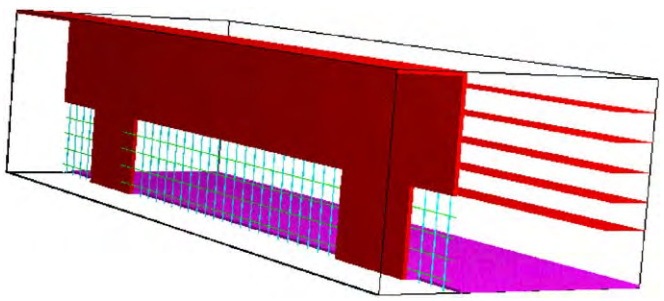


Fig. 7. Model with TEM capacitor plates.

### 4.3. Discretisation test—inductance of wires in screen

The inductance of wires in a screen was investigated with an even simpler model.

The screen consists of many individual wires. Current flowing through the wires causes magnetic fields round each wire. This magnetic field is responsible for the coupling and, as the inductance describes the relation of the current to the field energy, the inductance is a good measure of the coupling strength. Thus, the inductance is used to investigate the effect of discretisation on the results.

A very simple model can be used to consider the effect of varied discretisations on the calculated coupling. A two-plate resonator with periodic structure was chosen, that of a metal post between two metal plates, top and bottom, with periodic boundaries on the other two sides, left and right (see Fig. 8). This represents the resonant conditions for one wire of the screen. The height of the post must be small with respect to the length of the resonator. The frequency shift caused by the presence of the wire is calculated in the frequency domain. The effect of the use of *sheets* on the calculated field penetration through the shielding screen was also investigated. In Fig. 9 the lower curves (yellow and blue) represent the solid post, (almost coincident) and the upper curves show the sheet model with different orientations $L/Z_0$ was normalised to the best discretisation of the solid post, and plotted versus the mesh density. The curves rise steeply initially and then flatten out.

Each wire acts as a very small magnetic filter, thus the thickness of the wire would also affect the coupling. As the calculated fields are constant in each mesh cell, a coarse mesh simulates a thicker wire even when the dimensions of the wire remain constant. *Sheets* increase transmission through the screen by a factor of almost two.

## 5. Full model

In the time domain, when wakefields are not required, it is no longer necessary to maintain an equidistant mesh. A model with solid wires in the screen was generated, shown

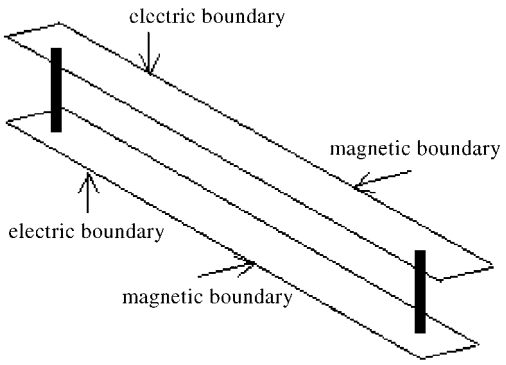



Fig. 8. Simplified model for inductance calculations of wires in a screen as a metal post between two metal plates.

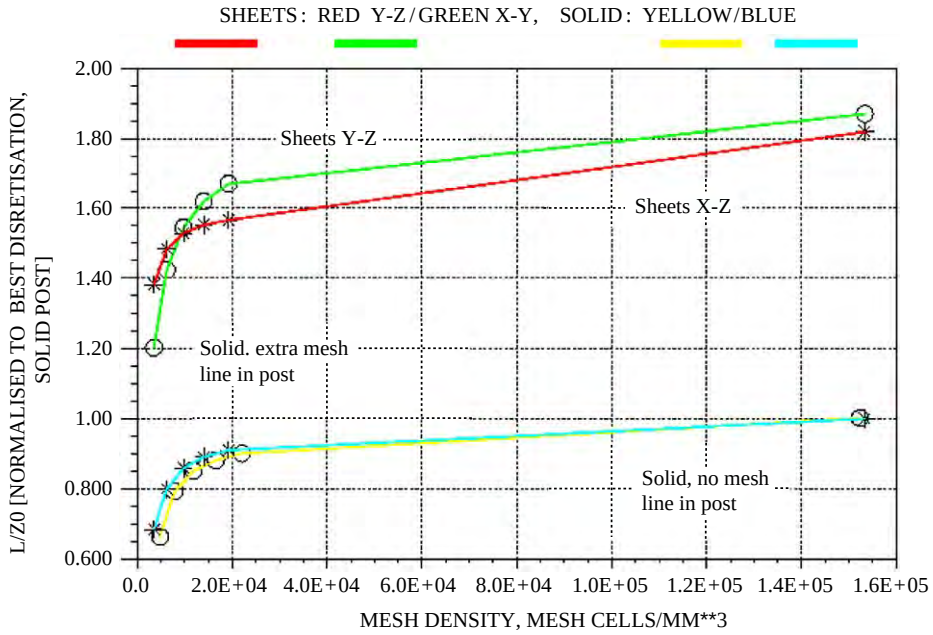



Fig. 9. $L/Z_0$ (normalised to best discretisation of solid post) vs. mesh density. The lower curves (yellow and blue) represent the solid post, (almost coincident). The upper curves show the sheet model with different orientations.

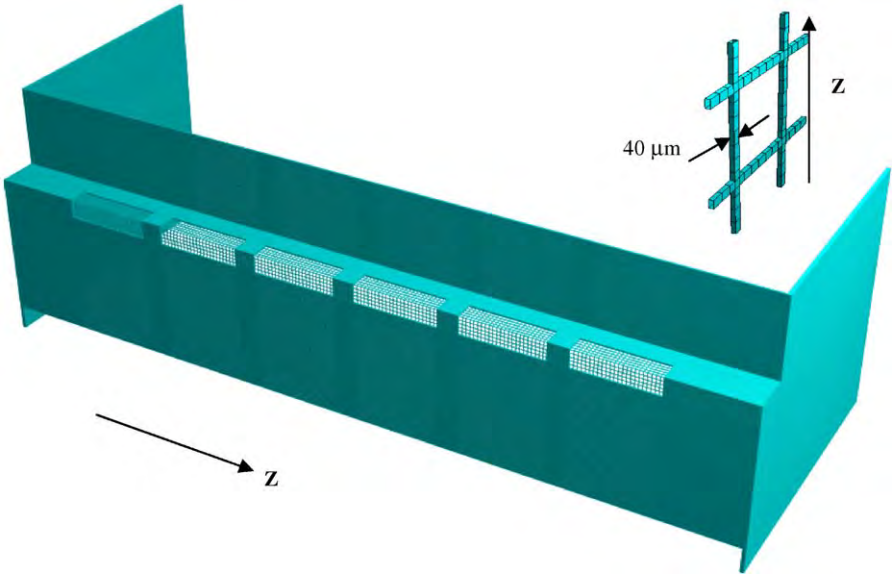



Fig. 10. Full model, a detail of the mesh of the screen is shown in the inset.

in Fig. 10. For comparison, models were also calculated with wires as *sheets* and without screening.

Using *sheets* also allows more energy to reach the outer area and the higher resolution, i.e. more mesh lines between the wires of the screen is also necessary to obtain the correct energy penetration. Table 1 shows the results.

Although the differences between the solid and the sheet model are small when sigma = 12 mm, it is interesting to note that increasing the discretisation makes the screen more transparent to higher frequency components. In Section 4.3, $L/Z_0$ was plotted against the mesh density. The differences found there between the solid and the sheet model are also apparent in these calculations.

## 6. Conclusions

A calculation model was found which gave reliable results within the given restrictions. The factors which affected the results were the use of *sheets*, discretisation, and the coarseness of the screen. The initial results were a factor of 10 too high.

Table 1
Energy penetration through the pumping holes of the full model for two bunch lengths

| Screen model | Resolution | Sigma = 3 mm | | Sigma = 12 mm | |
|---|---|---|---|---|---|
| | | Energy outside[a] ($\times 10^6$ V A s/m$^3$) | Power loss[b] (mW) | Energy outside[a] (($\times 10^3$ V A s/m$^3$) | Power loss[b] ($\mu$W) |
| Solid | Low | 0.3549 | 0.1278 | 0.2693 | $< 1$ |
| Solid | High | 0.3944 | 0.1420 | 0.2951 | $< 1$ |
| Sheets | Low | 0.8490 | 0.3056 | 0.5711 | $< 1$ |
| Sheets | High | 0.9215 | 0.3317 | 0.6136 | $< 1$ |
| Open | | | | | |
| Open holes | Low | 834.4 | 300.4 | 32.13 | 11.57 |

The screen wires were modelled as either sheets or solid, with $6 \times 6$ mesh lines per screen hole (high-resolution), or $4 \times 4$ (low-resolution).
[a]Energy calculated for a bunch with 1 C charge.
[b]Energy scaled to the HERA design current, 60 mA.

The screening reduces the energy penetration by at least two orders of magnitude for a 12 mm beam. This should be sufficient to reduce the radio frequency interference to a tolerable level. Measurements are in progress. To be effective the screening must be firmly attached.

## References

[1] MAFIA, CST GmbH, Bad Nauheimer Str. 19 D-64289, Darmstadt, Germany.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 220–224
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# MMFL $H^+$ and $H^-$ injectors: Computational studies

S.K. Esin, A.V. Feschenko, O.T. Frolov, E.S. Nikulin*, V.P. Yakushev

*Accelerator Complex Division, Institute for Nuclear Research, Russian Academy of Science, 7a, 60th October Anniversary pr., Moscow 117312, Russia*

Available online 1 December 2005

**Abstract**

Results are reported for a computational study of the output beam of the high-current Moscow Meson Factory (MMF) linac $H^+$ and $H^-$ injectors for various ion source operation modes and a variety of accelerating tube optics. Values for the emittance and phase density of the beam at the exit of the accelerating tubes are presented. The possibility of increasing the pulse current and thus the beam brightness with limited emittance growth was investigated. Numerical calculations are compared with experimental data. A computational application for monitoring the beam parameters in television is proposed.


*PACS:* 41.75.Cn; 41.85.−p

*Keywords:* High-current ion source; Proton beam; Negative-ion beam; Beam optics

## 1. Introduction

For a long time, a proton injector [1] has provided the beam for the Moscow Meson Factory high-current linac (MMFL) [2]. In addition, a new $H^-$ injector is under construction.

Modernization of injector optics has been supported by both computational monitoring of the injector output emittance and simulation of accelerator tube (AT) beam transmission matching with a linac low-energy beam transport (LEBT) channel and a radio-frequency quadrupole accelerator (RFQ). The RFQ cavity has an acceptance of $0.35\pi\,\mathrm{cm}\cdot\mathrm{mrad}$. Simulation results for AT beam formation derived using the large-particle method and the Kapchinskij–Vladimirskij microcanonical beam (MCB) model [3–5] practically agree. Because of the great advantage of PC use in terms of the calculation rate and result visualization, the MCB model was later used as a convenient engineering tool. In addition, modernization of the expander–extractor electrode system [6] was accompanied by simulation using the large-particle method.

Emittance measurements, processing and visual representation of results [7] are carried out using LabVIEW software packages.

The accelerating voltage shape and measured beam current at the injector exit are operationally routed into the linac local network, as well as a number of operational injector parameters.

Within the framework of creation of the $H^-$ injector, a computer control system (CCS) project [8] was developed using the National Instruments modules.

Monitoring the cross-sectional structure of beams is also a fundamental requirement for the high-current MMFL. The two-coordinate ionization detector combined with a TV has been applied for non-destructive monitoring of the LEBT beam proton component [9].

## 2. Injector design

The ion source (IS) optics is a set of focusing elements with strong high-order optical effects depending on the aperture. This consideration, combined with the requirements for high-gradient accelerating tubes, makes it attractive to consider simple, linear focusing elements in developing $H^+$ and $H^-$ injectors.

*Corresponding author. Tel.: +7 95 334 0962; fax: +7 95 334 0711.
*E-mail address:* nikulin@inr.ru (E.S. Nikulin).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.006

The MMFL proton injector provides a pulsed beam at the input of the LEBT channel, 2 m downstream of the AT [10] beam formation system.

The MMFL proton injector has been modernized to improve the qualitative performance and to increase the average beam current. This involves reconstructing the AT optics. Research carried out has allowed essential reduction of the emittance and beam current noise. The possibility of increasing the pulse current with limited emittance growth is being studied.

Until summer 2002, the beam pulse current of the proton injector was approximately 100 mA and normalized emittance was not less than $0.24\pi$ cm mrad, depending on the IS operation mode and the AT optical conditions (hereafter, the emittance is given for the beam core, containing 63% of the total beam current).

The necessity for an increase in beam brightness dictated the development of new injector optics. Between September 2002 and April 2003, a greed-free extraction system with conical electrodes was tested. A maximum pulse current of 110 mA and emittance of $0.16\pi$ cm mrad (at 100 mA) were achieved.

At the end of April 2003, a system with a cylinder expander, Pierce geometry and a plate extraction electrode was put into the AT. This allows maximum pulse current of up to 145 mA. At a current of 120 mA, emittance of $0.11\pi$ cm mrad was measured [6].

Designed for proton injector computer models, a system for computational control and diagnostics, as well as the handling of measurement results, has been used in the development of the $H^-$ injector.

The $H^-$ injector has to provide a beam at the input of the linac LEBT channel (hereafter called the injector "target"). The beam radius at the target must not exceed 26 mm, and it is provided only the AT optics for 400 kV, without the additional focusing devices of the LEBT channel. By virtue of the equipment already selected for the injector, the IS output is located approximately 270 cm from the target.

The version of the AT focusing electrodes meeting these requirements is shown in Fig. 1. Beam focusing is achieved by changing the intermediate electrode voltage $U_1$ and the focusing electrode voltage $U_f$. $U_1$ is fixed using the AT voltage divider and can only be changed after removal of a beam. The $U_f$ voltage is supplied from a separate power source and may be controlled operationally. The voltages are determined relative to the AT high-voltage end, which is under potential of 400 kV with respect to ground.

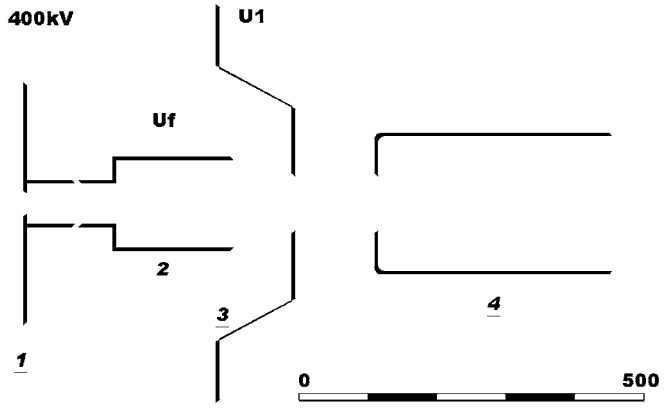



Fig. 1. Electrode scheme: (1) high-voltage AT flange, (2) focusing electrode, (3) intermediate electrode and (4), grounded electrode.

## 3. Simulations

The forming electrodes were developed with the help of the MCB model.

The possibilities for an injector forming system are shown in Fig. 2a. For comparison, the characteristics of a system differing only by a 50-mm-longer focusing electrode are shown in Fig. 2b. The admissible areas for the parameters $U_f$ and $U_1$ are shown when the beam envelope is fitted into the apertures of the forming electrodes and does not exceed the target radius for fixed parameters of the AT input beam. The MCB model allows rapid calculation, obtaining the areas shown in Fig. 2 in 1 h.

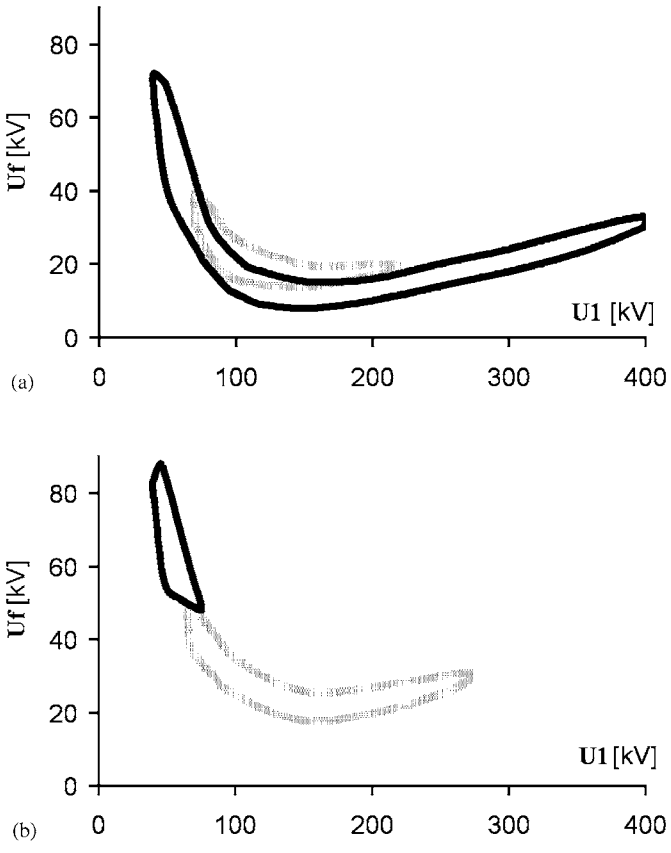



Fig. 2. Areas of admissible values for the focusing and accelerating electrode voltages, calculated using MCB models. (a) Forming system of the $H^-$ injector (Fig. 3); (b) focusing electrode is longer by 50 mm and all other sizes are as for (a). ⋯⋯ Beam with current fluctuations; ▬ beam without fluctuations.

Beams typical for the surface-plasma $H^-$ source, developed at the Institute for Nuclear Physics (Novosibirsk), were taken as the "test" beams [11]. Depending on the IS conditioning and tuning, in particular on hydrogen feed value, both an ordinary beam (with current fluctuations during the impulse) and an optimum beam (when the fluctuations are overwhelmed) can be obtained. For an ordinary beam, a model was used with the following parameters at the AT input: envelope radius, 15 mm; envelope divergence, 15 mrad; and normalized emittance, $2\pi$ mm mrad. The corresponding parameters for an optimum beam are 10 mm, 10 mrad and $0.2\pi$ mm mrad, with ion energy of 20 keV.

It is evident from Fig. 2 that the admissible areas for these beams differ considerably. However, for the $H^-$ injector electrode system selected, there is an essential range of admissible parameters (see Fig. 4) where the areas are close for both beams. $U_1$ voltage in the range 90–120 kV can be easily obtained with the existing AT design. Focusing voltage with regulation in the 20–30 kV range is also suitable for beam control. From the width of the admissible area, it is also possible to conclude as far as beam control achieved is convenient in exploitation. With the same technique, it is possible to calculate other characteristics of a system: beam-forming device (for example, beam current, envelope slope, etc.).

An extraction system with a cylinder expander, Pierce optics and a plate extractor is shown in Fig. 3. Simulation of particle trajectories was carried out with particle velocities perpendicular to the emission surface. The beam at the entrance of AT space is almost parallel. It has a large diameter and may be sufficiently well focused when passing through the AT.

Calculations carried out with a complete AT model that included drift space allowed determination of the transmission conditions for a 200-mA beam to the target aperture. Fig. 4 shows admissible areas for the IS electrodes voltages for different $U_1$ values in the model with Pierce optics.

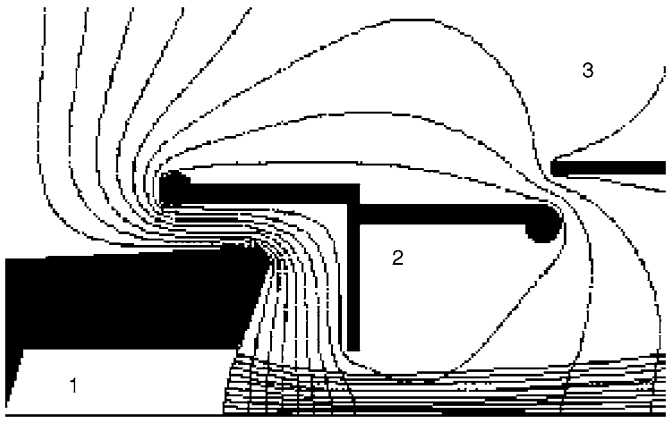



Fig. 3. Extraction system simulation: cylinder expander, Pierce optics, plate extractor, pulse current of 120 mA and extractor voltage of 35 kV. (1) expander, (2) extractor and (3) focusing electrode.

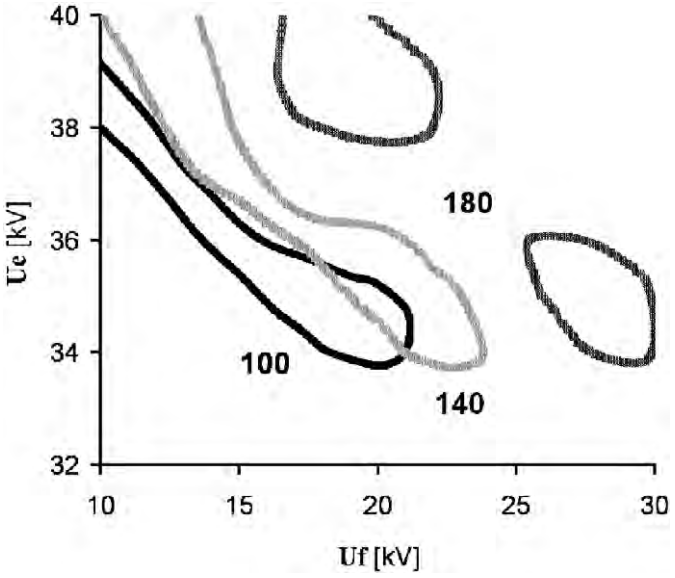



Fig. 4. Admissible voltage values for extracting ($U_e$) and focusing ($U_f$) electrodes for $U_1 = 100$, 140 and 180 kV with beam current of 200 mA.

## 4. Computerized control, data acquisition and processing

At present, the system for measuring HV-modulator parameters and the beam characteristics of the $H^+$ injector has been put into operation. Similar software and hardware for the $H^-$ injector are prepared. The systems are based on LabVIEW software. A workstation for the $H^+$ injector provides control (up to 73 channels) of the 400-kV pulse generator with a pulse repetition rate of up to 100 Hz. Control of ion sources under high potential (45 channels for $H^+$ and 104 for $H^-$) is carried out using a fiber optic cable. The analog channels have a bandwidth of 0–1.6 MHz.

A workstation based on a P-III/800 processor has the potential to exchange data with other linac control workstations using an Ethernet protocol through a 100-Mb/s local network. PCI-1200 and PCI-MIO-16E1 multifunctional input/output modules are built into the workstation for data acquisition and processing for the HV-modulator and auxiliary systems. These modules have four DAC and 24 ADC 12-bit programmable amplification channels with input signals in the range 0.01–10 V and accuracy of $\pm 0.025\%$. There are also 40 TTL input/output channels, and five counter/timer channels with 10-ns precision.

The IS control subsystem located inside the HV electrode is based on an SCXI-1001 microprocessor chassis. The built-in modules (ADC SCXI-1124, DAC SCXI-1140, SCXI-1160 relay and multifunctional SCXI-1200) use 34 analogue channels, 48 discrete channels and six counter/timer channels. In addition, 16 relay channels are used for control of the apparatus, with switching current of up to 2 A at voltage of up to 250 $V_{rms}$. A proportion of the channels is reserved.

The SCXI system stores information and dumps it to the operator workstation through a fiber optic cable. Data are transferred in both directions through the SCXI-2400 connection module using RS232 ports.

A workstation with 256 MB of RAM stores the injector operation history for subsequent analysis. The system periodically dumps accumulated information into an archive.

Beam diagnostics and data processing programs allow control of the qualitative beam parameters [7,12]. The control display of the emittance measurement system is shown in Fig. 5 and the output window of the data processing program for beam emittance measurement is shown in Fig. 6. The following beam parameters are displayed: emittance, average and maximum pulse current, phase density, beam width and position, etc. (Fig. 7).

Among the diagnostics tools, measurement of the beam transverse dimensions and position is the most important. A non-destructive ionization detector is used for this purpose. An optical image of the beam cross-section is recorded from the detector screen by TV camera for monitoring and computer processing. The computer display provides visual control of the beam cross-section, profiles and center-of-mass position (Fig. 8). The frame resolution is $384 \times 288$ pixels with 64 brightness levels. The total number of frames that can be saved in the RAM is 256.

The software provides on-line capability to:

- show background discrimination;
- present 128-point vertical and horizontal beam profiles;

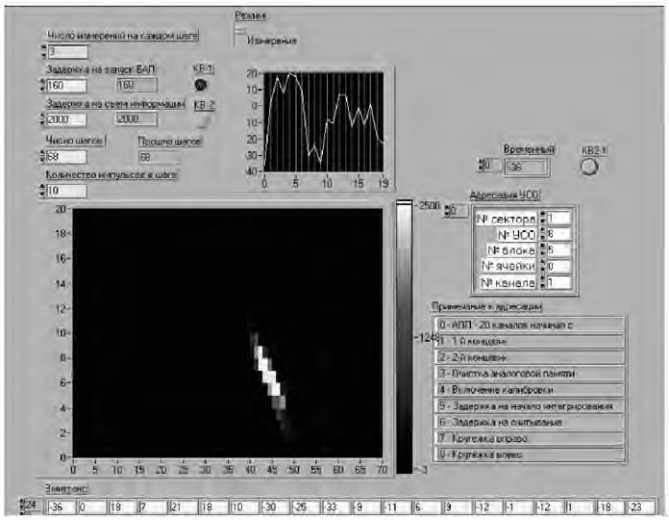


Fig. 5. Control display for the emittance measurement system.

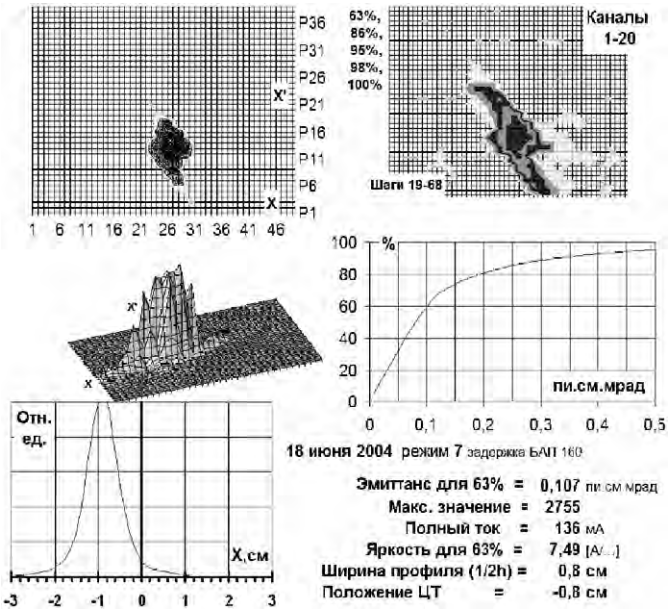



Fig. 6. Output window of the data processing program for beam emittance measurement.

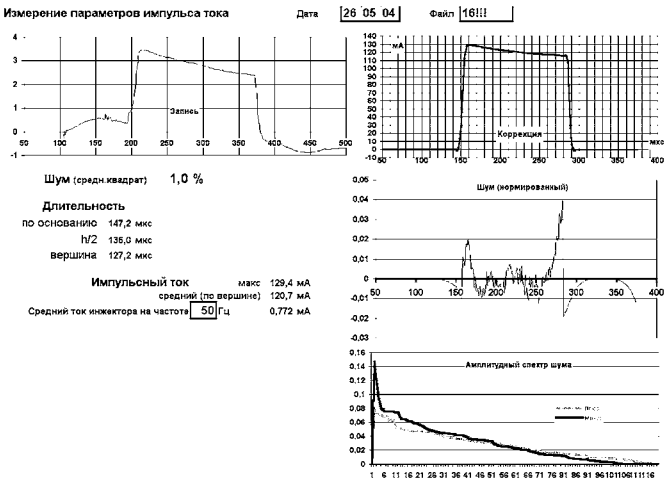



Fig. 7. Output window of the data processing program for measurement of beam current parameters.

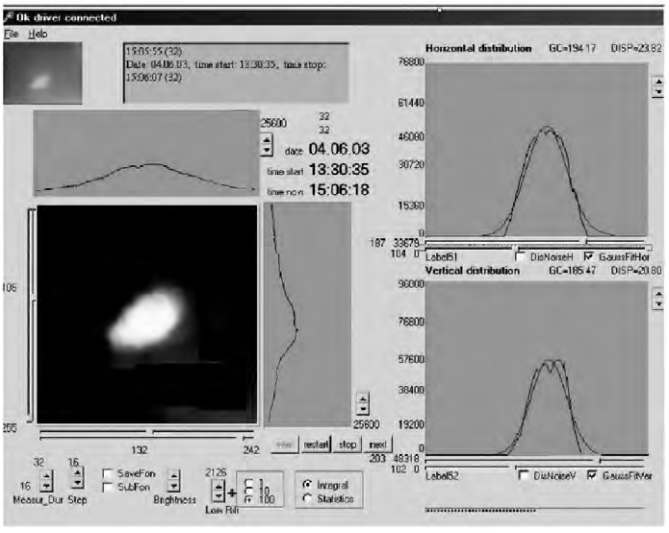



Fig. 8. Characteristic image and profiles of the cross-section of the proton injector beam on the interactive display of the control program.

- present averaged vertical and horizontal profiles;
- accumulate up to 256 frames, resulting in an increase in absolute sensitivity and in signal/noise ratio by more than an order of magnitude;
- show Gaussian distribution fitting of profiles with the least-squares method to determine the beam position; and
- show dispersion of the beam position distribution and average beam position calculated from Gaussian fitting results.

## 5. Conclusions

It is important to minimize ion beam losses and to match emittance with LEBT/RFQ acceptance.

Construction of a linearized injection path from the simplest lenses has successfully demonstrated the possibility of obtaining beam parameters acceptable for either proton or $H^-$ acceleration. Our experimental results provide the basis for checking our simulations of proton and $H^-$ beams.

Future studies are needed to decrease the value of the beam emittance and to include other effects to make the simulations more realistic.

## Acknowledgment

This work was supported by the Russian Foundation for Basic Research under Agreements #01-02-16148 and #04-02-16315.

## References

[1] A.M. Anikeichik, et al., in: Proceedings of the 1988 Linear Accelerator Conference, Newport News, Virginia, CEBAF Report-89-co1, 1989, p. 660.
[2] V.P. Burlakov, et al., in: Proceedings of the 1984 Linear Accelerator Conference, Darmstadt, 1984, p. 9.
[3] I.M. Kapchinskij, V.V. Vladimirskij, in: Proceedings of the International Conference on High-Energy Accelerators and Instrumentation, Geneva, 1959, p. 274.
[4] I.M. Kapchinskij, Dinamika Chastits v Lineinyh Resonansnyh Uskoritelyah, Atomizdat, Moscow, 1966, p. 194 (in Russian).
[5] V.N. Zubetz, et al., Probl. At. Sci. Technol. 34 (3) (1999) 20.
[6] V.I. Derbilov, et al., Probl. At. Sci. Technol. 42 (1) (2004) 13.
[7] V.I. Derbilov, et al., in: Proceedings of the XIVth International Conference on Electrostatic Accelerator Beam Technology, Obninsk, 2001, p. 256.
[8] V.I. Derbilov, et al., Probl. At. Sci. Technol. 39 (5) (2001) 128.
[9] S.K. Esin, et al., Probl. At. Sci. Technol. 42 (1) (2004) 87.
[10] E.G. Komar, Osnovy uskoritelnoij tehniki, Atomizdat, Moscow, 1975, p. 46 (in Russian).
[11] G.E. Dereviankin, et al., Zh. Tekh. Fiz. 48 (1978) 404.
[12] V.I. Derbilov, et al., in: Proceedings of the IXth International BDO Workshop, Saint Petersburg, 2002, p. 91.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 225–226
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Simulation of low energy muon frictional cooling

R. Galea[a,*], A. Caldwell[b], S. Schlenstedt[c], H. Abramowicz[d]

[a]*Nevis Laboratories, Columbia University, Irvington, NY, USA.*
[b]*Max-Planck-Institut für Physik, Munich, Germany*
[c]*DESY, Zeuthen, Germany*
[d]*Tel Aviv University, Tel Aviv, Israel*

Available online 22 November 2005

**Abstract**

Frictional cooling is a proposed method of phase space reduction for a potential muon beam intended for collisions. The basic principle involves compensating for the muon energy loss in media by a constant electric field. The muons are in an energy regime below the ionization peak which for muons in helium is less than 10 keV. Electronic energy loss is treated as a continuous process and all individual nuclear scatters with scattering angles greater than 50 mrad are simulated as discrete processes. Other effects like the Barkas [W.H. Barkas, W. Birnbaum, F.M. Smith, Phys. Rev. 101 (1956) 778.] effect and Muonium formation are also included. The results of our simulations are summarized.


*PACS:* 14.60.Ef; 29.27.Bd; 29.27.Fh; 41.75.Ak

*Keywords:* Frictional cooling; Muon collider; Energy loss

## 1. Introduction

Muon beams are produced occupying a large phase space. The emittance, a quantity quantifying the size of the phase space, must be reduced in order to make luminous collisions possible. The required normalized six-dimensional emittance for several Muon Collider parameter sets is $\varepsilon_{6,N} = 1.7 \times 10^{-10}$ in units of $(\pi\mathrm{m})^3$ [1,2]. This represents six orders of magnitude smaller than typical initial muon beam emittances. This reduction must be accomplished with reasonable efficiency if a Muon Collider is to be a realizable machine.

## 2. Simulation of a Muon Collider based on frictional cooling

The scheme illustrated in Fig. 1 has been simulated. Muons of both signs are produced by scattering an intense proton beam on a target located in a region with a very strong solenoidal magnetic field (20 T, as in the Neutrino Factory study). A drift region with a more moderate magnetic field, of the order of 5 T, allows the bulk of the pions to decay to muons. A combination of MARS [3] and GEANT 3.21 [4] are used to simulate these sections. The muons are then input into the cooling channel, which consists of a cooling cell roughly ~11 m long. The cooling cell contains Helium gas (possibly $H_2$ for $\mu^-$), and an electric field perpendicular to the magnetic field. The electric field direction is reversed periodically, as a function of the position along the axis of the gas cell, to cancel the beam drift. The muons stopped in the cell drift out at a characteristic angle dependent on $\vec{B}$ and $\vec{E}$. They are then extracted through thin windows and reaccelerated.

At the end of the drift region, there is a correlation between the longitudinal momentum of the muons and their arrival time. This allows for a phase rotation, where time varying electric fields are used to increase the number of muons at lower momenta.

## 3. Simulation of physics processes

Frictional cooling cools muon beams to the limit of nuclear scattering. A detailed simulation was therefore performed where all large angle nuclear scatters were simulated. The differential distributions and mean free paths for $\mu$-nucleus scattering were calculated in two

*Corresponding author. Tel.: +1 914 591 2826; fax: +1 914 591 8120.
*E-mail address:* galea@nevis.columbia.edu (R. Galea).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.007

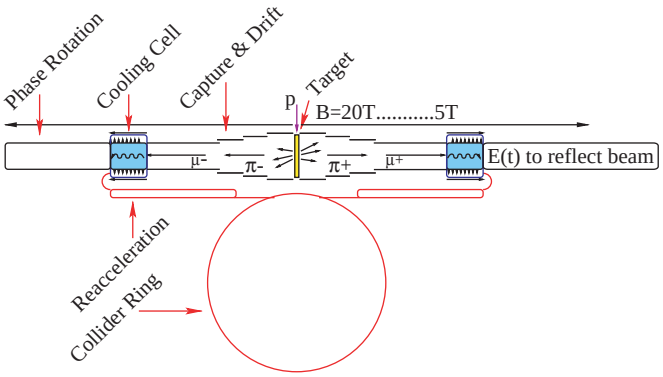



Fig. 1. Overview of a Muon Collider based on frictional cooling. The different sections are not to scale.

different ways depending on the energy regime. A quantum mechanical (Born approximation) calculation was used for the scattering cross section at high kinetic energies ($T_\mu > 2$ keV) and a classical calculation was used at lower kinetic energies. A screened Coulomb potential used in the Born and classical calculations had the form

$$V = \frac{4e^2}{r}\exp\left(\frac{-r}{a}\right),$$

where $a$ is the screening length. For the classical calculation the procedure of Everhart et al. [6] was followed. The differential cross section was calculated by scanning in impact parameter and evaluating the scattering angle at each impact parameter. From the differential cross section, a mean free path for scattering angles greater than a cutoff (0.05 rad) was found, and scatters were then generated according to the differential cross section. This method reproduces the energy loss from nuclear scatters tabulated by NIST [7] for protons. For $\mu^+$, the electronic energy loss is taken from the NIST tables. The suppression for $\mu^-$ (Barkas effect [1]) was parameterized from the results in [8]. The electronic energy loss was treated as a continuous process. The electronic energy losses for $\mu^+$ and $\mu^-$ are shown in Fig. 2. This difference is significant below the ionization peak, which is expected to be due in part to muonium formation.

To simulate the effect of muonium formation in the tracking, an effective charge was used, as given by $\sigma_I/(\sigma_F + \sigma_I)$, where $\sigma_I$ is the cross section for muonium ionization and $\sigma_F$ is the cross section for muonium formation. At energies relevant to frictional cooling of muons, the muonium formation cross section dominates over muonium ionization in all media except for Helium.

Negative muon capture was parameterized from calculations of Cohen [9,10] and included in our simulation. The calculations only extend up to 80 eV. Beyond this, a simple exponential fall off with kinetic energy was assumed.

## 4. Results and conclusions

A Muon Collider complex would be invaluable to the world of high energy physics. In order to make this machine realizable the formidable task of solving the muon cooling problem must be faced. Detailed simulations of such a collider based on the concept of frictional cooling have been performed. They include the simulation and optimization of a collider frontend consisting of

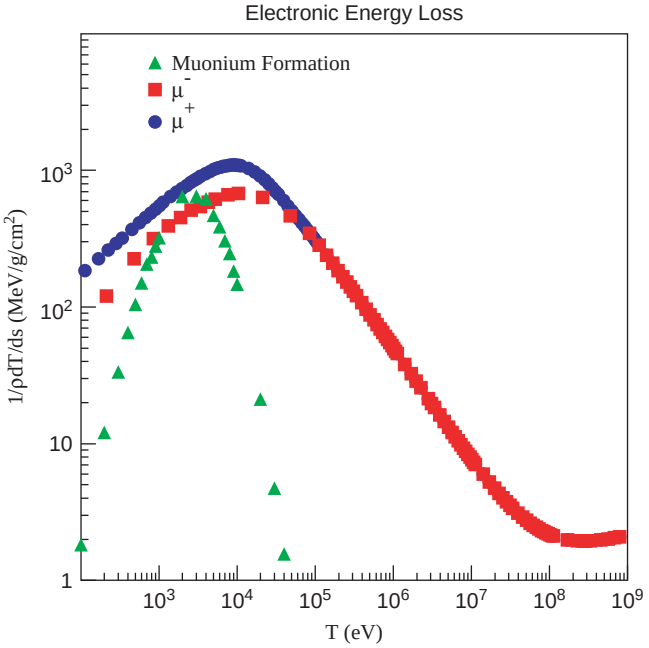



Fig. 2. The electronic energy loss for $\mu^+$ and $\mu^-$. For comparison, the simulated energy loss due to muonium formation (see text) is also shown.

- a 2 GeV proton driver,
- a Copper target 30 cm in length and 0.5 cm thick,
- a drift length of 28 m.

Other physics processes such as muonium formation and muon capture were also addressed as these processes become important to consider in the low energy regime in which frictional cooling functions.

The results of these studies show that final emittances of $(6-8)\times 10^{-11}(\pi\text{m})^3$ can be achieved which offers hope in the pursuit of a successful cooling scheme to make a Muon Collider realizable.

## References

[1] W.H. Barkas, W. Birnbaum, F.M. Smith, Phys. Rev. 101 (1956) 778.
[2] C.M. Ankenbrandt, et al., Phys. Rev. ST Accel. Beams 2 (1999) 081001.
[3] N.V. Mokhov, The MARS Code System User's Guide, version 13 (98), FNAL-FN-628, 1998.
[4] R. Brun, et al., GEANT3, CERN DD/EE/84-1, 1987.
[6] E. Everhart, G. Stone, R.J. Carbone, Phys. Rev. 99 (1955) 1287.
[7] ICRU Report 49, Stopping Powers and Ranges for Protons and Alpha Particles, Issued 15 May, 1993 (see ⟨http://physics.nist.gov/PhysRefData/Star/Text⟩).
[8] M. Agnello, et al., Phys. Rev. Lett. 74 (1995) 371.
[9] J.S. Cohen, Phys. Rev. A 62 (2000) 022512.
[10] J.S. Cohen, J. Phys. B 31 (1998) 833.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 227–234
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Simulation of electron beam dynamics in a high-energy electron cooler

A.V. Ivanov*, V.M. Panasyuk, V.V. Parkhomchuk, V.B. Reva, M.A. Tiunov

*BINP, Acad.Lavrentiev, 11, BudkerINP, Novosibirsk, 630090, Russia*

Available online 29 November 2005

**Abstract**

Electron cooling is now a standard tool for improvement of ion beam parameters in storage rings. In BINP, after successful development of several low-energy electron cooling devices, a project involving a high-energy electron cooler for GSI has been proposed. This cooler has a classical electrostatic scheme with electron energy of up to 8 MeV. Here we present results of numerical simulations of electron beam formation, acceleration and collection for this project. Special attention is paid to a description of the new codes developed in BINP. The electron gun and collector are simulated by the 2D USAM code. This code is modified to calculate collector performance with consideration of secondary emission. The BEAM code is used for simulation of dynamics in the accelerating section. A new 3D electrostatic code, ELEC3D, developed for the simulation of beam dynamics in bends with electrostatic compensation of the centrifugal drift, is described. This code is combined with the existing MAG3D magnetostatic code to provide a universal tool for 3D static calculations.


*PACS:* 29.27.Bd

*Keywords:* Electron cooling; Simulation; Beam dynamics

## 1. HESR electron cooler

The subject of this article is the investigation of the electron beam dynamics in a high-energy electron cooler. Fast electron cooling of antiproton beams in the energy range between 0.8 and 14.5 GeV is a key feature of one of the major objectives of GSI future plans: internal target experiments at the HESR ring [1,2]. Powerful cooling is required for high-resolution experiments investigating the structure of hadrons and the interaction of quarks and gluons in a nuclear medium.

The project involving an 8-MeV electron cooler proposed by BINP is based on a classical electrostatic scheme. The 8-m-high high electrostatic column consists of 80 sections, with potential increasing from the bottom section to the top one. The column has three accelerating tubes, two for the acceleration and deceleration of an electron beam, and one for charging of the column head. For the electrostatic cooler small, a 10-MeV cyclotron of $H^-$ ions is proposed as the charging system. After injection and acceleration, the electron beam is bent in the vertical and horizontal planes and is moved to the cooling section. After the main solenoid, the beam is returned to the electrostatic column, where it is decelerated to recuperate its energy. A longitudinal magnetic field accompanies the electron beam along all of its path.

## 2. Electron gun with variable beam profile

During cooling of intensive ion beams, it was found that there is strong recombination of ions cooled to the low temperature with the electron beam. This recombination leads to considerable losses of the stored beam. Moreover, excessive cooling causes the development of transversal instability that also leads to beam losses. To avoid these effects, the use of hollow electron beams for cooling was proposed [3]. The decrease in electron beam density in the central part, where cooled ions accumulate, leads to a considerable decrease in recombination losses while maintaining the cooling rate of the whole ion beam (because the cooling time decreases rapidly with decreasing oscillation amplitude).

*Corresponding author. Tel.: +7 383 3394554; fax: +7 383 3307163.
*E-mail address:* a.v.ivanov@inp.nsk.su (A.V. Ivanov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.058

An electron gun with variable beam profile was designed and successfully tested in BINP. The gun is shown in Fig. 1. The control electrode is located near the cathode edge, so its potential strictly influences emission from this area. At negative voltage on this electrode, emission is suppressed on the edge of the cathode, so a pencil beam is formed. At positive voltage, a hollow beam is formed. The calculated and measured perveance of the electron gun versus the control electrode potential are shown in Fig. 2.

The gun output current increases with increasing control potential to some value of $U_{control}/U_{anode}$, which depends on the value of $U_{anode}/U_{cathode}$. The formation of a virtual cathode takes place at this point. The control electrode potential causes such a large current that it cannot pass over the anode due to sagging of the potential and part of it is reflected back, whereupon the gun output current remains constant, regardless of further growth of the control electrode potential. At $U_{control}$ values less than the threshold for the virtual cathode, the electron beam is very stable and no fluctuation of space charge was observed.

A thin tungsten wire is used for electron-beam profile measurements. The deposited current was measured during movement of the wire transverse to the beam direction. The beam profiles were recreated using these data. The measured and calculated profiles corresponding to different values of the control electrode potential are shown in Fig. 3.

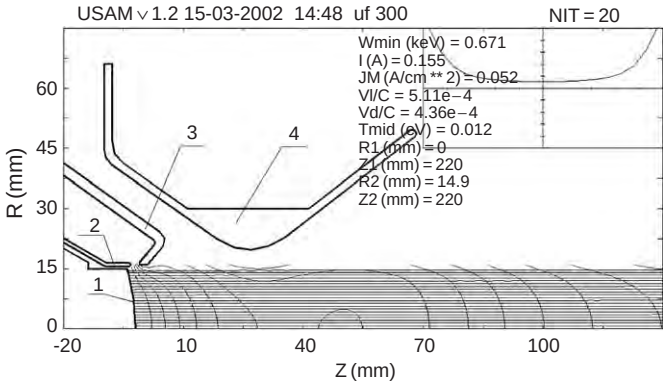



Fig. 1. The geometry of the electron gun with variable beam profile: 1—cathode, 2—forming electrode, 3—control electrode, 4—anode.

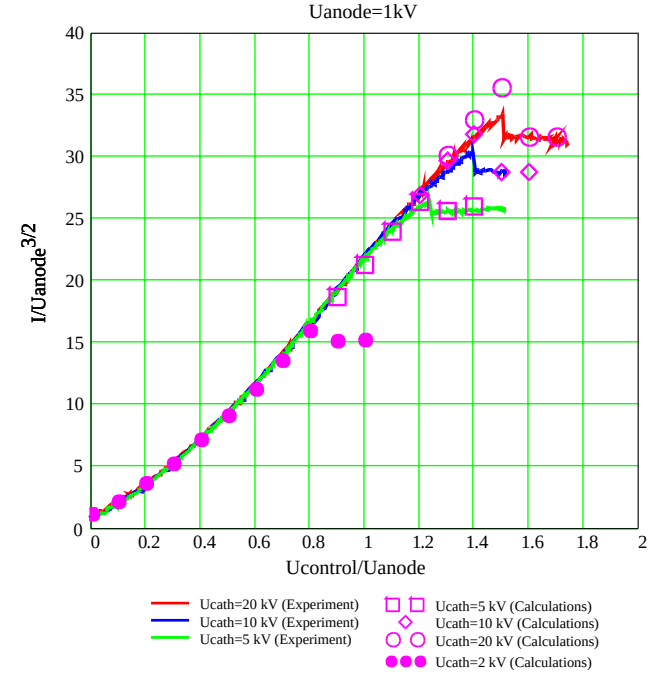



Fig. 2. The perveance of the electron gun as a function of $U_{control}/U_{anode}$.

## 3. Simulation of secondary emission in collector

The main questions that should be considered in designing and simulation of a collector for electron coolers are the prevention of local overheating, achieving higher collector perveance, and consideration of secondary emission. In this article we concentrate on the last question for the collector installed on the EC-300 cooler and proposed for the HESR cooler. This collector is shown in Fig. 4.

To minimize the output current of secondary electrons, a suppressor is installed before the collector. Moreover, the magnetic field in the collector is much lower than in the drift tube, so only secondary electrons with velocities lying

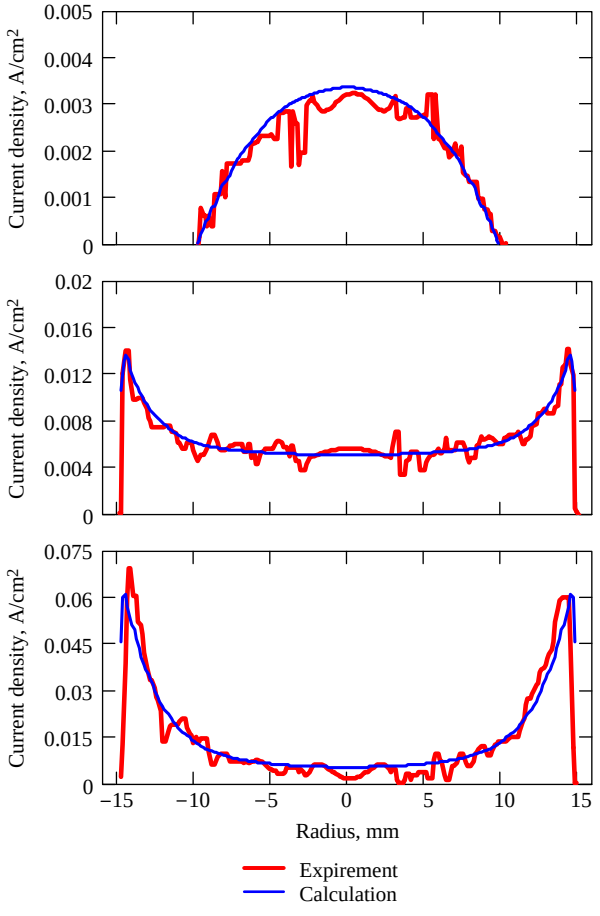



Fig. 3. The calculated and measured beam profiles at $U_{control} = +300$ V, $+100$ V, $-100$ V and $U_{anode} = 500$ V.

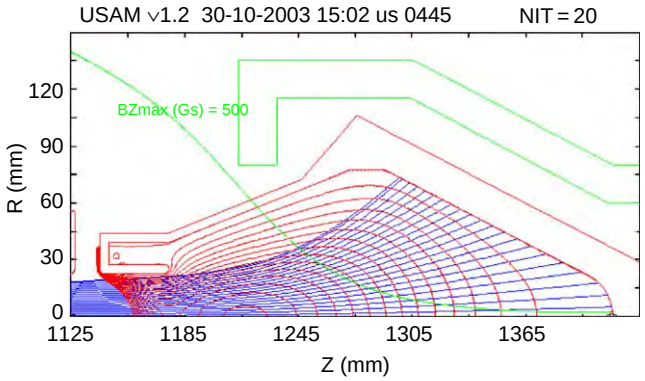



Fig. 4. The collector of EC-300 cooler. Trajectories of particles, equipotentials and magnetic field distribution are shown.

in a small cone

$$\theta = B_{\text{coll}}/B_{\text{tube}} \tag{1}$$

can leave collector, while others are reflected back.

To estimate secondary emission in the collector, the following formulas describing coefficients of emission for re-diffused (2) and elastically reflected (3) electrons are used [4]:

$$\eta(E_0) = \eta_{\text{inf}}(1 - \exp(-(E_0/E_r))) \tag{2}$$

$$r(E_0) = r_{\text{inf}} + (r_0 - r_{\text{inf}})\exp(-(|E_0 - E_e|/W)) \tag{3}$$

where $E_0$ is the primary beam energy.

To describe the dependence of the secondary emission coefficient as a function of angle of incidence formula (4) is used:

$$\eta(E_0, \theta) = \eta(E_0)\exp(K(1 - \cos(\theta))). \tag{4}$$

The energy spectrum of re-diffused electrons is taken as

$$\frac{dn_{\text{rd}}}{dE} = f_{\text{rd}} = N_{\text{rd}}\frac{3}{2}\frac{1}{E_0}\sqrt{\frac{E}{E_0}}.$$

To describe secondary emission from every point of incidence of the primary electrons, a set of secondary electrons is started. The velocities of these secondary electrons are not randomized as in the Monte–Carlo method, but are defined by the primary electron velocity and by numbers of divisions specified by the user. These velocities are determined according to the condition that each secondary particle must carry the same current. Several generations of secondary electrons can be calculated.

Calculated and measured coefficients of secondary emission from the collector are shown as a function of the suppressor potential in Fig. 5. Calculations predict a larger coefficient because the space charge of secondary electrons is not considered. The charge density distribution of the secondary beam is presented in Fig. 6. Due to the diamond-like shape of the collector, the magnetic field on its surface is much higher for larger radii. According to Eq. (1), more secondary electrons from this region are allowed to leave the collector, and the charge density increases on

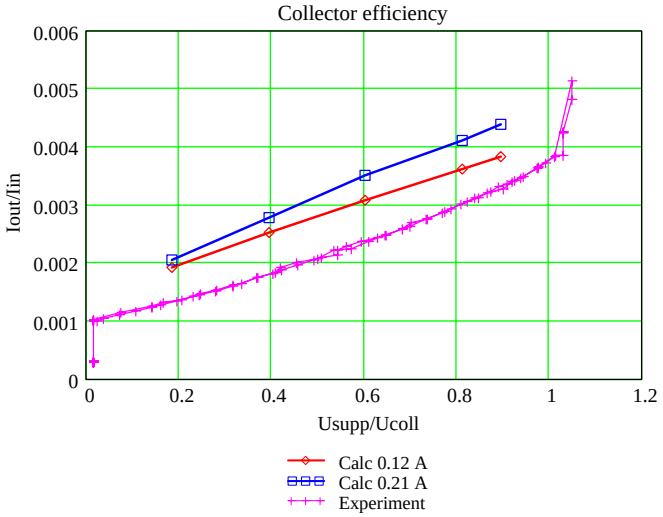



Fig. 5. Collector efficiency as a function of $U_{\text{supp}}/U_{\text{coll}}$.

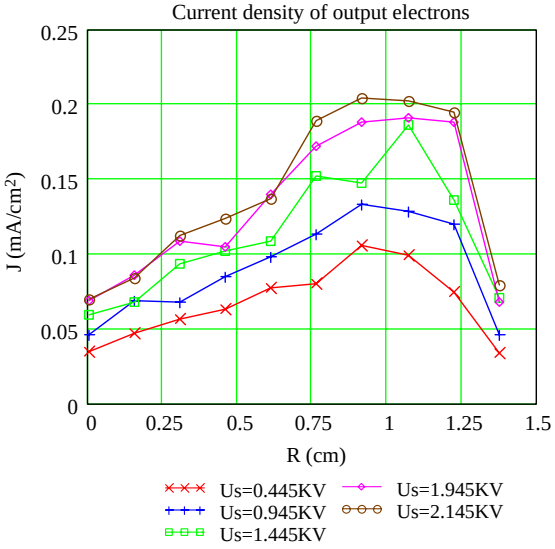



Fig. 6. The secondary beam profile.

the outer beam radius. The energy spectrum of the secondary beam is shown in Fig. 7.

## 4. UltraSAM—2D code for simulation of electron guns

The SAM code package was designed in BINP for simulations of stationary axial-symmetric electron–optical systems and electron guns [5]. The SAM code uses the boundary elements method. The following boundary equations are considered:

$$\int_{L_e+L_d} \sigma(\eta)G(\xi,\eta)\,dl = U_e(\xi) - \int_{S_b} \rho(\eta)G(\xi,\eta)\,dS, \quad \xi \in L_e \tag{5}$$

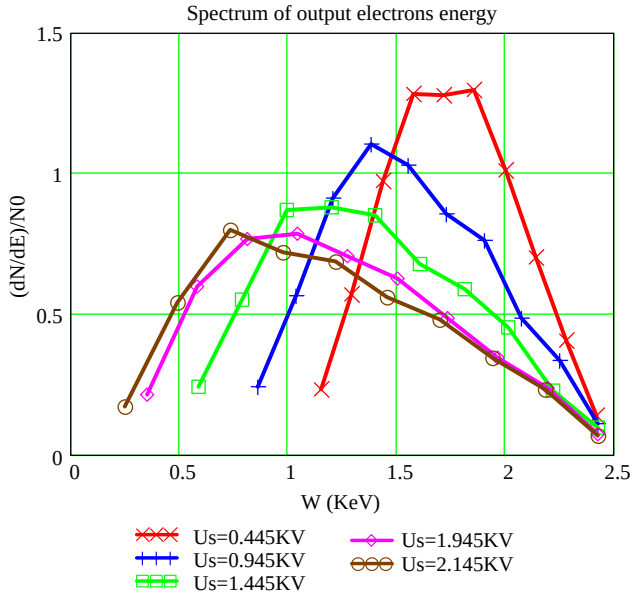



Fig. 7. The energy spectrum of the output beam.

$$2\pi\frac{\varepsilon_1+\varepsilon_2}{\varepsilon_1-\varepsilon_2}\sigma(\xi)-\int_{L_e+L_d}\sigma(\eta)\frac{\partial G}{\partial n_\xi}(\xi,\eta)\,dl = \int_{S_b}\rho(\eta)\frac{\partial G}{\partial n_\xi}(\xi,\eta)\,dS, \quad \xi\in L_d$$

where $L_e$ and $L_d$ are the contours of the electrodes and dielectrics; $\xi=(r_0,z_0)$, $\eta=(r,z)$ are points on these contours; and $S_b$ is the beam cross-section. The kernel of integral Eq. (5)

$$G(\xi,\eta)=\int_0^{2\pi}\frac{r\,d\theta}{\sqrt{r_0^2-2r_0r\,\cos\,\theta+r^2+(z_0-z)^2}} \tag{6}$$

has a simple physical meaning. It is the potential induced at point $(r_0,z_0)$ by an infinitely thin, charged ring with radius $r$ axial coordinate $z$ and unit surface charge density. Integral Eq. (6) can be expressed by the total elliptical integral of the first order.

These integral equations are solved by the collocation method, with spline-interpolation of the solution. Separations of the kernel singularity of the integral equations and of the solution singularity are performed. The contours of the surfaces of the electrodes and dielectrics are defined by segments (lines and arcs). A set of rectangular meshes is used for the space charge distribution. Space charge density is considered to be constant within individual cells. A quasilaminar model of current pipes is used for the description of beam dynamics.

### 4.1. Curved meshes

The most important step in the development of the SAM code and transition to the UltraSAM code is the replacement of rectangular meshes by curved meshes. The contours of curved meshes can be adjusted using the geometry of the electrodes and the beam shape. It is possible to specify several emitters of complex shape. A set of curved meshes is defined for each segment that describes the electron flux from it. These measures greatly increase the precision of a near-cathode area simulation, decrease the numerical aberrations that arise in trajectory analysis due to inaccurate description of the beam self field on its edge, and extend the scope of solvable tasks.

Every cell of the curved mesh is defined by nine points—the cell nodes. The cell is treated as a quadrangular element, which is defined using the finite element method. The cell is transformed into a square situated in logical plane $XY$. In this square the interpolation of every function f is built with the help of basis functions $\psi_i(x,y)$:

$$f(x,y)=\sum_{i=1}^{9}f_i\psi_i(x,y) \tag{7}$$

where $f_i$ are the values of function $f$ at the nodes. The same approach is used for the transformation between real and logical coordinates.

### 4.2. Calculation of potential and field induced by space charge

The potential induced by the beam space charge is equal to the sum of potentials induced by the space charge of individual cells. Using Eq. (7), the potential from one cell can be written as

$$\varphi_{cell}(r_0,z_0)=\sum_{i=1}^{9}\rho_iM_i(r_0,z_0)$$

$$M_i(r_0,z_0)=\iint\psi_i(x,y)G(r_0,z_0;r,z)J(x,y)\,dx\,dy$$

where $\rho_i$ are the values of the space charge density at the nodes, and $J(x,y)$ is the Jacobian for transformation from logical coordinates $(x,y)$ into real $(r,z)$. The coefficients $M_i$ depend only on the cell geometry. The calculation of these coefficients is hindered by the logarithmic singularity of function G at the observation point $(r_0,z_0)$. To maintain the high precision of the potential calculation when this point is situated near the cell or in it, the singularity is separated analytically by the same method as described in Ref. [5].

Calculation of the electric field is achieved in the same way by replacement of $G(r_0,z_0;r,z)$ with $\vec{\nabla}G(r_0,z_0;r,z)$.

### 4.3. Models of emission and beam dynamics

To determine the emitted current density, a special surface placed parallel to a cathode at a small distance $\delta$ is introduced. Knowing the electric field on this surface, the current density can be found using the Child-Langmuir law. Approximation of a plane, cylindrical or spherical diode is used, depending on the shape of the cathode considered and its position. The use of electric field instead

of potential significantly improves the convergence of space charge iterations due to the sharp increase in field near the cathode. The influence of the magnetic field is also considered.

A beam is represented as a set of particle trajectories, each carrying its own part of the full current. By moving through a cell, each particle induces a charge on the nodes of this cell. The charge density at the nodes is calculated from the induced charge as:

$$\rho_i = \frac{Q_i}{V_i}, \quad V_i = \iint 2\pi r(x,y)\psi_i(x,y)J(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

In some systems, heating of the cathode has a significant influence on the beam dynamics. The transversal thermal spread of initial velocities of particles is simulated to consider this effect. From every start point of the trajectories (in the case of a cold cathode), the calculation of several trajectories, each having different additional transverse thermal velocities, is started. These trajectories describe the thermal distribution of the initial velocities.

## 5. Accelerating column

The accelerating column of the proposed cooler consists of many accelerating sections. It is proposed to use 1-m-long accelerating sections, but in the following calculations, standard sections that are two-fold shorter are considered. The presence of gaps between the accelerating electrodes at the points of connection of neighboring sections leads to strong pulsation of the accelerating field. When traveling through the accelerating column, electrons are affected by a radial oscillating force with constant space step created by these pulsations. Moreover, an electron has its own parameter—the space step for Larmour rotation. This step increases with increasing electron energy. When these parameters are comparable, a resonant increase in transverse velocity may occur. Fig. 8 shows that subsequent strikes can increase or suppress the transversal motion, depending on the magnetic field.

To minimize these undesirable pulsations, it is proposed to hit the beam by local variation of the magnetic field in the area where the Larmour rotation step is large enough (see Fig. 9). Calculations show that this local field can be greater than the mean value, or can be smaller. The second case is shown in the figure. To realize this, no special changes in the magnetic system have to be made, apart from the current in one coil must be decreased. Fig. 10 shows how the radius of the pulsations depends on the current and position of this coil. With a change in the total energy, the position and current of the affecting coil must be changed; it is possible to use two or maybe more coils to provide the necessary effect.

## 6. Matching point of magnetic fields

In the electrostatic column of the proposed HESR cooler, the value of the accompanying magnetic field is

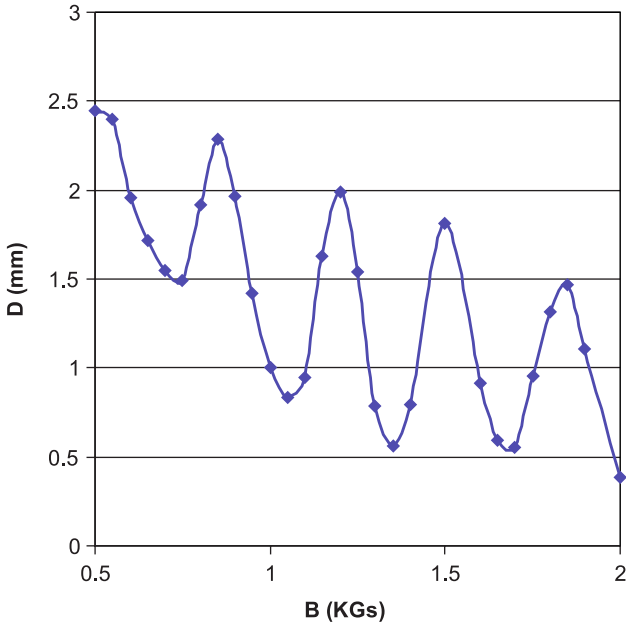



Fig. 8. Diameter of the beam pulsations on the exit of accelerating column as a function of the magnetic field.

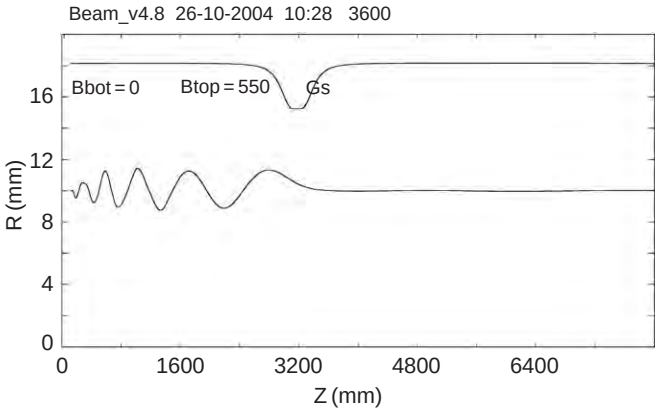



Fig. 9. Minimization of the beam pulsations.

chosen as 500 Gs. In bends and in the cooling solenoid, this field is 10-fold greater, at 5 kGs. Thus, a special matching section must be included to provide a transition of the electron beam between these fields without excitation of the transversal motion. With electron beam energy of 8 MeV, an adiabatic entrance will be too long; shorter field growth leads to strong beam pulsations (see Fig. 11a). To fabricate a compact matching section, another idea was proposed – a quarter-wave transformation section. Between the low and high magnetic fields, there must be a region with field equal to

$$B_{\mathrm{mid}} = \sqrt{B_1 B_2},$$

and the length of this region must be equal to a quarter of the space step for Larmour rotation. In such a configuration of the magnetic field, strikes on the entrance into the section and on the exit compensate each other, and no

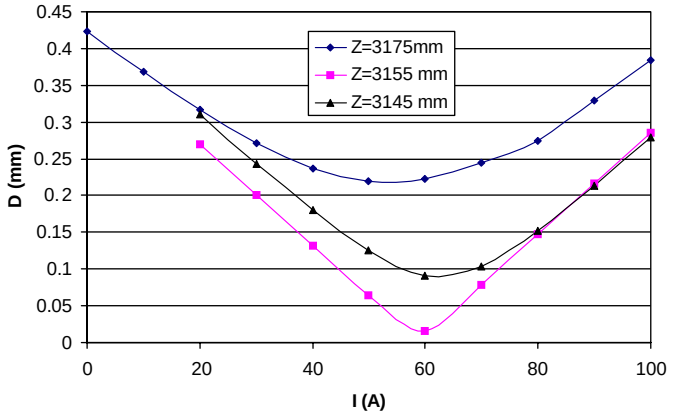



Fig. 10. Dependence of the diameter of the beam pulsations on the position and current of test coil.

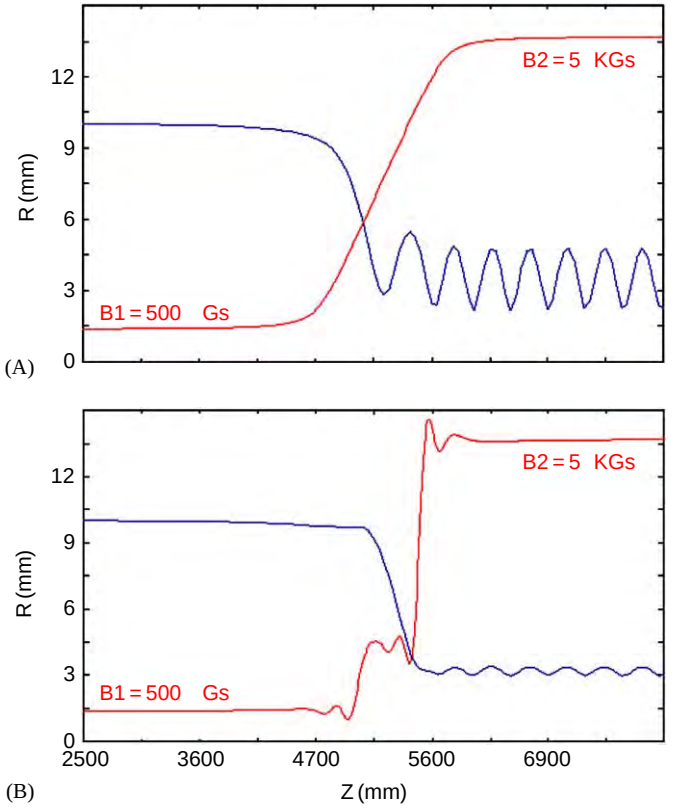



Fig. 11. Entrance into the strong magnetic field. The magnetic field and beam envelope are shown. Smooth growth of field (a) and quarter-wave transformation section (b).

transversal motion occurs. Results of beam dynamics simulation for this section are shown in Fig. 11b.

## 7. Beam code

Beam behavior in the accelerating column and in the matching point of the magnetic fields is simulated by the BEAM code [6]. This 2D code was developed for simulation of dynamics of non-laminar cylindrical beams with high perveance in long electron–optic systems. The initial beam parameters in this code can be defined by the user or can be imported from SAM and UltraSAM codes. In the last case, to describe the non-laminar beam radius of trajectories, the trajectory slope angle and azimuthal speed are represented by power series, with coefficients of these series found by the least-squares method. High-order paraxial approximation is used for the calculation of external electric and magnetic fields. The axial distribution of external fields is calculated using the SAM code. A particle in cell (PIC) method is used to describe the beam transversal electric and magnetic fields. The heat spread of the velocities of beam particles can also be taken into account. This spread is determined by the size and temperature of the cathode and by the beam radius.

## 8. Bends with electrostatic compensation

The motion of an electron in a cooler is not reversible as a rule. This means that electrons reflected from a collector have a small chance of being absorbed by it after that. The magnetic field bending an electron beam in the toroid part of a cooler has an influence in the opposite direction to the reflected electrons. Thus, these electrons fall into the vacuum chamber, inducing an outgassing process and becoming a cause of leakage current. The electrostatic compensation of centrifugal drift can help to achieve reversible dynamics of electrons. Electrons reflected from the collector can return to it and be absorbed. Thus, the leakage current can be strongly suppressed. This effect was observed experimentally on the EC-300 cooler designed and manufactured at BINP for a CSRe ring (IMP, Lanzhou) [7].

To realize electrostatic bending, two plates forming a capacitor are placed in the toroid part. However, this capacitor itself can excite transversal motion of electrons. The curvature of magnetic force lines increases on the edges of toroids, and the electric field itself increases on the edges of the capacitor (Fig. 12). If these two curves are different, beam oscillation is excited. To provide the necessary growth of the electric field, modification of the capacitor edges is proposed (Fig. 13). Calculations show that this modification helps to minimize beam transversal energy. Beam pulsations in a capacitor with non-modified and modified edges are shown in Fig. 14.

## 9. ELEC3D and MAG3D—3D electro- and magnetostatic codes

A new 3D electrostatic code, ELEC3D, was developed in BINP to calculate beam dynamics in bends with electrostatic compensation. This ELEC3D code uses the boundary elements method, which has a number of advantages in comparison with the finite element method. Open regions and extreme aspect ratios do no pose problems for this method, and field solution is perfectly smooth. Moreover, no 3D mesh is necessary in typical cases, and only the surfaces of elements must be described.

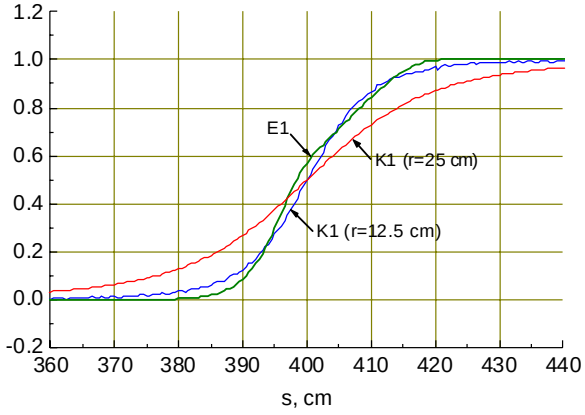



Fig. 12. Matching of the magnetic force line curvature with growth of the electric field.

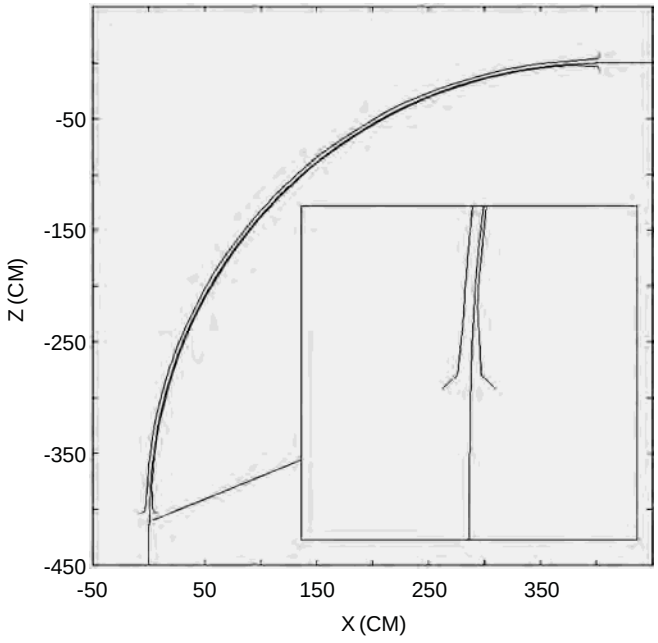



Fig. 13. Modified edges of the capacitor.

In the ELEC3D code, the following boundary equation is considered:

$$\int_{S_e} \frac{\sigma(\vec{r})\,\mathrm{d}S}{|\vec{r}_e - \vec{r}|} = U_e, \quad \vec{r}_e \in S_e \tag{8}$$

where $S_e$ are the surfaces and $U_e$ the potentials of the electrodes. To solve this equation, surfaces of electrodes are divided into cells, and a constant surface charge density is taken within one cell. With this assumption, Eq. (8) can be rewritten as a system of linear equations:

$$\sum_{N\ \text{cells}} C_{ij}\sigma_j = U_i$$

where the coefficients $C_{ij}$ depended only on the geometry of the system. $C_{ij}$ has a simple physical meaning: the potential induced in the center of cell number $i$ by unit surface charge in cell number $j$. One disadvantage of the boundary elements method is that the high non-uniformity of surface

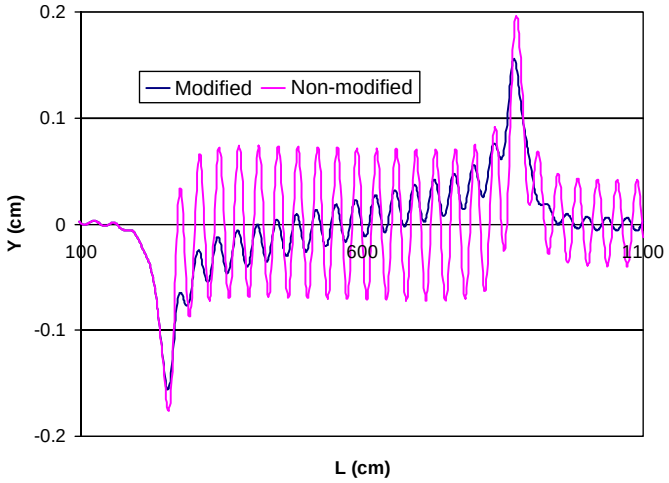



Fig. 14. Beam pulsations in a capacitor with non-modified and modified edges.

charge distribution must be considered by division of surfaces into cells, especially on edges.

Potentials and fields are calculated as the sum of individual potentials and fields induced by cells. The potential of cell number $i$ can be rewritten by transformation of the coordinates as

$$U_i = \sigma_i \int_{S_i} \frac{\mathrm{d}S}{\sqrt{\eta^2 + z^2}} \tag{9}$$

where $S_i$ is the surface of the cell, $\sigma_i$ is its surface charge density, $\vec{\eta}$ is the vector in the plane of the cell, and $z$ is the distance from the point of observation to this plane. To provide precise and fast calculation of this integral, the following analytical method is used. The function under the integral can be represented as a divergence of the function $\vec{\eta}/\sqrt{\eta^2 + z^2} + z$. Using the Gauss theorem, we can rewrite the surface integral as the integral on the sides, which can easily be found analytically. The electric field can be found by differentiation of Eq. (9).

The magnetic field in the electrostatic bends of the HESR electron cooler is calculated using the MAG3D nonlinear magnetostatic code. In the MAG3D code, the magnetic field $\vec{H}$ is considered as a sum of the current field $\vec{H}_c$ and the magnetization field $\vec{H}_m$. $\vec{H}_m$ can be represented as a gradient of scalar function, which is expressed as

$$\Psi(\vec{r}_0) = \frac{1}{4\pi}\int_{V_m} \vec{M}\vec{\nabla}\left(\frac{1}{|\vec{r}_0 - \vec{r}|}\right)\mathrm{d}V$$

where $\vec{M} = \chi(H)\vec{H}$ is the vector of magnetization, $\chi(H) = \mu(H) - 1$, and $V_m$ is the volume of ferromagnetic materials. Thus, the magnetic field $\vec{H}$ can be written as

$$\vec{H} = \vec{H}_c + \frac{1}{4\pi}\int_{V_m} \vec{\nabla}\left[\frac{\chi(\vec{H},\vec{R})}{R^3}\right]\mathrm{d}V. \tag{10}$$

The total volume occupied by ferromagnetic materials is divided in the MAG3D code into small volumes. The

magnetization of each of these volumes is supposed to be constant. With this assumption, Eq. (10) can be rearranged as a system of non-linear equations

$$\vec{H}_i = \vec{H}_{c,i} + \sum_k C_{ik}\chi_k \vec{H}_k$$

where the coefficients $C_{ik}$ are dependent only on the geometry of the system. Coils in the MAG3D code have a rectangular cross-section and consist of straight and arc elements. For straight elements, $\vec{H}_c$ can be found analytically; for arc elements, numerical arc integration should be carried out.

The ELEC3D code is combined with the MAG3D code in one code package to provide a universal tool for 3D static calculations.

## References

[1] W.F. Henning, Proceedings of EPAC, Lucerne, Switzerland, 2004, p. 50.
[2] M. Steck, et al., Proceedings of EPAC, Lucerne, Switzerland, 2004, p. 1969.
[3] A.V. Bubley, V.V. Parkhomchuk, V.B. Reva, Nucl. Instr. and Meth. A 532 (2004) 303.
[4] M. Pivi, M. Furman, Phys. Rev. ST Accel. Beams 6 (2003) 034201.
[5] B. Fomel, M. Tiunov, V. Yakovlev, Preprint BINP 96-11, Novosibirsk, 1996.
[6] M. Tiunov, Preprint BINP 98–78, Novosibirsk, 1998.
[7] E. Behtenev, et al., Proceedings of EPAC, Lucerne, Switzerland, 2004, p. 1419.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 235–239
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Study of transient self-consistent beam dynamics in RF linacs using a particle tracing code

V.V. Mytrochenko*, A. Opanasenko

*NSC KIPT, Akademicheskaya str. 1, Kharkov 61108, Ukraine*

Available online 28 November 2005

**Abstract**

The paper describes a simulation technique for study of unsteady self-consistent dynamics of charged particles in RF linacs that consist of cavities and travelling wave sections. The approach proposed is based on unsteady theories of the excitation of cavities and waveguides by a beam of charged particles and RF feeders. The theory of waveguide excitation is extended to the case of spatially inhomogeneous travelling-wave structures. The SUPERFISH code is used to evaluate the characteristics of the axisymmetric travelling-wave sections. The PARMELA code is applied for simulation of particle motion and to obtain data required for solving the equations for excitation of the RF structures by the beam.


*PACS:* 41.20.Jb; 41.75.Fr

*Keywords:* Linac; Self-consistent beam dynamics; Transients; Travelling wave structures

## 1. Introduction

Acceleration of intense charged beams in RF linacs in the short pulse regime, for which the filling time of standing wave (SW) bunchers and travelling wave (TW) sections cannot be neglected, requires detailed knowledge of transients to diminish their influence on the beam parameters. Although a great deal of information about unsteady acceleration can be obtained using particle-in-cell (PIC) codes, calculation of the beam dynamics in long RF structures using these PIC codes requires rather large computing resources. At the same time it is often necessary to study slow-varying phenomena of a narrow frequency spectrum in the working passband of RF linacs. It has been shown [1] that in this case such well-known codes as PARMELA [2] and SUPERFISH [3], which is usually applied to simulation of steady dynamics, can be used to simulate unsteady, self-consistent particle dynamics in SW cavities. The present work is aimed at developing a simulation technique for unsteady particle dynamics in inhomogeneous TW structures and integrating this technique into a unified algorithm for simulation of self-consistent, unsteady beam dynamics in RF linacs containing both SW and TW structures.

## 2. Method

Extending the theory of excitation of waveguides [4] to inhomogeneous accelerating structures, we suppose that the orthogonal eigen waves $\vec{E}'_{\pm s}$, $\vec{H}'_{\pm s}$ (electrical and magnetic fields, respectively) with the norm

$$N'_s = \frac{c}{4\pi}\int_{S(z)}\left(\vec{E}'_s \times \vec{H}'_{-s} - \vec{E}'_{-s} \times \vec{H}'_s\right)\vec{e}_z \,\mathrm{d}S = \mathrm{const} \qquad (1)$$

can propagate along the longitudinal direction $\pm z$ of an inhomogeneous waveguide. Here $\vec{e}_z$ is the unit vector along the $OZ$ axis and $S(z)$ is the waveguide cross-section at the $z$ coordinate. Then, according to Ref. [4], the Fourier

*Corresponding author.
*E-mail addresses:* mitvic@kipt.kharkov.ua (V.V. Mytrochenko), opanasenko@kipt.kharkov.ua (A. Opanasenko).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.039

harmonics of fields can be presented as follows:

$$\vec{E}_\omega = \sum_s C'_s \vec{E}'_s + C'_{-s}\vec{E}'_{-s} + \frac{4\pi}{\mathrm{i}\omega}\vec{j}_{\omega,z}$$
$$\vec{H}_\omega = \sum_s C'_s \vec{H}'_s + C'_{-s}\vec{H}'_{-s} \quad (2)$$

where $\vec{j}_{\omega,z}$ is the longitudinal component of current density and $C'_{\pm s}(z)$ is a function of $z$ that satisfies

$$\frac{\mathrm{d}C'_{\pm s}}{\mathrm{d}z} = \frac{1}{N'_{\mp s}}\int_{S(z)} \vec{j}_\omega \vec{E}'_{\mp s}\,\mathrm{d}S. \quad (3)$$

If the geometry of a structure varies slowly in the longitudinal direction, the eigen waves $\vec{E}'_{\pm s}, \vec{H}'_{\pm s}$ in cross-section $z$ can be expressed through the eigen waves $\vec{E}_{\pm s} = \mathrm{e}^{\pm \mathrm{i}h_s z}\vec{E}^{(0)}_{\pm s}$, $\vec{H}_{\pm s} = \mathrm{e}^{\pm \mathrm{i}h_s z}\vec{H}^{(0)}_{\pm s}$ that correspond to the homogeneous waveguide with cross-section $S(z)$, as follows:

$$\vec{E}'_{\pm s} = \xi_s(z)\mathrm{e}^{\pm \mathrm{i}\psi_s(z)}\vec{E}^{(0)}_{\pm s}$$
$$\vec{H}'_{\pm s} = \xi_s(z)\mathrm{e}^{\pm \mathrm{i}\psi_s(z)}\vec{H}^{(0)}_{\pm s} \quad (4)$$

where $\mathrm{d}\psi_s/\mathrm{d}z = h_s(z)$, $\mathrm{Re}\{h_s(z)\}$ is a propagation constant, $\mathrm{Im}\{h_s(z)\} = \alpha_s(z)$ is an attenuation constant, and $\xi_s(z)$ is a real function of the $z$ coordinate. It should be noted that Eq. (4) provides adiabatic invariance of the wave power flow.

In a periodic waveguide, functions $\vec{E}^{(0)}_{\pm s}, \vec{H}^{(0)}_{\pm s}$ can be expressed as Floquet's series:

$$\vec{E}^{(0)}_{\pm s} = \sum_{n=-\infty}^{\infty} \vec{E}_{\pm s,n}\mathrm{e}^{\mathrm{i}\frac{2\pi n}{D}z}, \quad \vec{H}^{(0)}_{\pm s} = \sum_{n=-\infty}^{\infty} \vec{H}_{\pm s,n}\mathrm{e}^{\mathrm{i}\frac{2\pi n}{D}z}. \quad (5)$$

Substituting Eq. (4) into Eq. (1), we can obtain a connection between the norms of the waves $(\vec{E}'_{\pm s}, \vec{H}'_{\pm s})$ and $(\vec{E}_{\pm s}, \vec{H}_{\pm s})$ as

$$N'_{\pm s} = \xi_s(z)^2 N_{\pm s}(z) = \mathrm{const}. \quad (6)$$

Differentiating Eq. (6) with respect to $z$ and substituting Eq. (4) into Eqs. (2) and (3), the Fourier harmonics of a field can be found in the following forms:

$$\vec{E}_\omega = \sum_s C_s \vec{E}^{(0)}_s + C_{-s}\vec{E}^{(0)}_{-s} + \frac{4\pi}{\mathrm{i}\omega}\vec{j}_{\omega,z}$$
$$\vec{H}_\omega = \sum_s C_s \vec{H}^{(0)}_s + C_{-s}\vec{H}^{(0)}_{-s} \quad (7)$$

where $C_{\pm s}(z)$ are the new amplitudes of the expansion $C_{\pm s} = \xi_s(z)\mathrm{e}^{\pm \mathrm{i}\psi_s(z)}C'_{\pm s}$ that satisfy

$$\frac{\mathrm{d}C_{\pm s}}{\mathrm{d}z} + \left(\frac{1}{2N_{\pm s}}\frac{\mathrm{d}N_{\pm s}}{\mathrm{d}z} \pm \mathrm{i}h_s\right)C_{\pm s} = \frac{1}{N_{\mp s}}\int_{S(z)} \vec{j}_\omega \vec{E}^{(0)}_{\mp s}\,\mathrm{d}S. \quad (8)$$

In the case of acceleration of a train of bunches, the fields and beam current are specified by narrow frequency spreads around the working frequency $\omega_0$. Thus, accomplishing the inverse Fourier transformation in Eqs. (7) and (8), we can obtain expressions for the fields of forward wave in the lower band in time-dependent form:

$$\vec{E}(t,\vec{r}) = \mathrm{Re}\left\{C_{+0}(t,z)\vec{E}^{(0)}_{+0}(\vec{r}) \times \exp\left(\mathrm{i}\int_0^z h_0(z)\,\mathrm{d}z - \mathrm{i}\omega_0 t\right)\right\},$$
$$\vec{H}(t,\vec{r}) = \mathrm{Re}\left\{C_{+0}(t,z)\vec{H}^{(0)}_{+0}(\vec{r}) \times \exp\left(\mathrm{i}\int_0^z h_0(z)\,\mathrm{d}z - \mathrm{i}\omega_0 t\right)\right\}. \quad (9)$$

The slow-varying amplitude $C_{+0}(t,z)$ obeys

$$\frac{\partial C_{+0}}{\partial z} - \frac{1}{2R_0}\frac{\mathrm{d}R_0}{\mathrm{d}z}C_{+0} + \frac{1}{v_g(z)}\frac{\partial C_{+0}}{\partial t} = \frac{R_0}{2}\mathrm{e}^{-\mathrm{i}\int_0^z h_0(z)\,\mathrm{d}z} \int_{S(z)} \vec{j}_{\omega_0}(t,\vec{r})\vec{E}^{(0)*}_{+0}(\vec{r})\,\mathrm{d}S \quad (10)$$

where $R_0(z) = -4|E^{(0)}_{z,+0}(0)|^2/N_0(z)$ is the serial impedance of the synchronous space harmonic of the accelerated field and $v_g(z)$ is the group velocity. The harmonic of current density is expressed through the Lagrange coordinates $\vec{r}_k$, $\tau_k(t,z)$ of beam particles in the form:

$$\vec{j}_{\omega_0}(t,\vec{r}) = q\frac{\omega_0}{2\pi}\int_{t-\pi/\omega_0}^{t-\pi/\omega_0}\mathrm{d}t \times \mathrm{e}^{\mathrm{i}\omega_0 t}\left\{\sum_{k=1}^{K}\delta[\vec{r}_\perp - \vec{r}_{\perp,k}] \times \delta[t - \tau_k(t,z)]\right\}$$

where $q$ is the charge of the particles. The particle coordinates depend on the field amplitude $C_{+0}(t,z)$, so they are slow-varying functions of time.

It should be noted that the field of the backward wave can be found in the same way.

## 3. Simulation techniques

### 3.1. Travelling wave cells

For numerical solution of Eq. (10) we use the difference equation of the first-order of approximation [5]:

$$\frac{\bar{C}^{(n,m)} - \bar{C}^{(n,m-1)}}{\Delta z} + \frac{\bar{C}^{(n+1,m)} - \bar{C}^{(n,m)}}{v_g^{(m)}\Delta t} = -f^{(m)}I^{(n,m)} \quad (11)$$

where

$$\bar{C}^{(n,m)} = C_{+0}(n\Delta t, m\Delta z)\sqrt{R_0(0)/R_0(m\Delta z)}$$

$$f^{(m)} = \frac{1}{2}\sqrt{R_0(m\Delta z)R_0(0)}\,\mathrm{e}^{\mathrm{i}\int_0^{m\Delta z} h_0(z)\,\mathrm{d}z}$$

$\Delta t$ is the temporal step and $\Delta z$ is the spatial step. The indexes $n$ and $m$ correspond to the discrete time $n\Delta t$ and longitudinal coordinate $m\Delta z$. The Fourier harmonic of a current $I^{(n,m)}$ with mean value $I(n\Delta t)$ over the period $2\pi/\omega_0$ is given in the form:

$$I^{(n,m)} = I(n\Delta t)\sum_k \mathrm{e}^{\mathrm{i}\omega_0\tau_k(n\Delta t, m\Delta z)}. \quad (12)$$

To obtain a self-consistent set of the equations, it is necessary to supplement Eq. (11) by a set of equations of motion of the particles in the fields (Eq. (9)). For solution of the equations of motion we use the PARMELA code, which takes into account the influence of space charge forces, as well as of the focusing system, on the beam dynamics.

By processing the output file of the PARMELA code with particle coordinates, it is possible to calculate the Fourier harmonic of a beam current using Eq. (12) at the end of each TW cell that represents the period of a disk-loaded waveguide (DLW). Obviously, in this case $\Delta z$ must be a multiple of the cell length and $\Delta t$ must be an integer number of periods of an accelerating field. Then $\bar{C}^{(\bar{n},m)}$ determines a field in the $m$th cell for the $n$th time step. Knowing the initial amplitudes in cells $\bar{C}^{(0,m)}$ and the amplitude in the first cell $\bar{C}^{(n,0)}$ as a function of time, Eq. (11) can be solved by evaluating the increments in amplitude for each TW cell at each time step.

To build up field distributions in TW cells, the PARMELA code needs the relative amplitudes of the spatial harmonics. In addition, values of $\alpha_0$, $R_0$ and $v_g$ as functions of $z$ have to be specified. Evaluation of these values can be carried out using the SUPERFISH group of codes. The characteristics of an inhomogeneous DLW are presented as a set of characteristics of a homogenous DLW with iris radii that are equal to each iris radius in the simulated DLW (as follows from Eq. (4)). The parameters of the homogenous DLWs are evaluated using a technique [6] for field patterns in the cavity stacks.

### 3.2. Standing wave cells

Generally, RF linacs, besides TW sections, consist of SW structures. Application of the PARMELA and SUPERFISH codes for simulation of unsteady, self-consistent beam dynamics in SW cavities [1] can be integrated in a common algorithm for full-scale simulation of the RF linac. According to Ref. [1], the slow-varying complex amplitudes $C_r$ of fields

$$\vec{E}(t,\vec{r}) \approx \mathrm{Re}\{C_r(t)\vec{E}_r(\vec{r})\mathrm{e}^{\mathrm{i}\omega t}\}$$

$$\vec{H}(t,\vec{r}) \approx \mathrm{Re}\{C_r(t)\vec{H}_r(\vec{r})\mathrm{e}^{\mathrm{i}\omega t}\}$$

in SW cells (except cells of the input and output couplers) are evaluated from the difference equation:

$$\frac{C_r^{(n+1)}-C_r^{(n)}}{\Delta t}+\left(\mathrm{i}(\omega_0-\omega_r)+\frac{\omega_r(1+\beta)}{2Q_r}\right)C_r^{(n)} = \frac{\omega_r}{Q_rE_0}\sqrt{\frac{Z_{\mathrm{sh}}\beta P_{r0}^{(\mathrm{n})}}{d}}\mathrm{e}^{\mathrm{i}(\varphi_r+\pi/2)} + \frac{Z_{\mathrm{sh}}\omega_r q}{2Q_r\,\mathrm{d}E_0^2}\sum_{k=1}\overline{\vec{v}_k(n\Delta t)\vec{E}_r(\vec{r}_k(n\Delta t))\mathrm{e}^{\mathrm{i}\omega t}}, \quad (13)$$

where $\omega_r$ is the resonant frequency of the cavity, $Q_r$ is the unloaded quality factor, $Z_{\mathrm{sh}}$ is the shunt impedance per unit length, d is the cavity length, $E_0$ is the mean amplitude of the on-axis electric field, $\vec{v}_k$ is the particle velocity, $\beta$ is the coupling factor for a cavity with a feeder, $\varphi_r$ is the phase shift, and $P_{r0}^{(n)}$ is the incident RF power at the given time step. The overbar in Eq. (13) represents time averaging. The field distribution and the values of $\omega_r$, $Z_{\mathrm{sh}}$ and $Q_r$ can be specified with the SUPERFISH codes.

### 3.3. Algorithm features

Let us consider the features of the algorithm designed. If the beam current at a given temporary step is not equal to zero, the running values of the amplitudes and field phases in the cells are written into the input file of the PARMELA code and the simulation of particle motion starts. After the PARMELA code completes the task, its output files carry out-processing to obtain the data necessary to evaluate the increments of the amplitudes. Then the process repeats. Therefore, to keep a physical sense of the results obtained, the step $\Delta t$ should be longer than the time-of-flight of particles through the simulated segment. On the other hand, Eq. (11) gives correct results only at $v_g\Delta t<\Delta z$, so the quantity of TW cells $M$ should be less then $c/v_g$, where $c$ is the velocity of light. Besides, there is artificial tailing of a wavefront along the section if $v_g\Delta t\neq\Delta z$ because of amplitude averaging over the cell length. Therefore, the edges of current and RF power pulses should be at least longer than $M\Delta t$ to diminish errors due to this tailing, which increases with increasing cell numbers. Thus, the typical values of $\Delta t$ and $M$ are 10 periods of RF oscillations and 40 cells. A long linac can be broken into several segments that contain an acceptable number of cells. To inject the beam from one segment into the next, a file with particle coordinates in the six-dimensional phase space is created at each time step. During simulation of the subsequent segment the PARMELA code accepts the file that corresponds to the running time step. It should be noted that the time steps have to be the same throughout simulation of the whole linac. The DLW segments are bounded by half-cells. The field of the last half-cell is the boundary condition used to simulate the next segment.

Testing of the algorithm was carried out in several ways. First, the results of simulation of excitation of pillbox cells and a homogenous DLW by short, ultra-relativistic bunches were compared with the analytical results for a steady-state mode. The simulated field distribution along a DLW containing 40 regular cells agreed within 0.1% with an analytically derived one distribution. The DLW of the KUT linac [7] was chosen to represent an inhomogeneous TW structure. It contains four homogenous segments connected with matching cells, with a total of 35 cells; the phase advance is 120° per cell and $v_g$ drops from 0.024 to 0.009% of $c$. Fig. 1 shows the dependence of the relative group velocity and $\alpha_0$ on distance along the DWL for the KUT linac. An analytical steady-state solution for accelerating field distribution was evaluated from the equation of power diffusion [8]. The test showed that the results

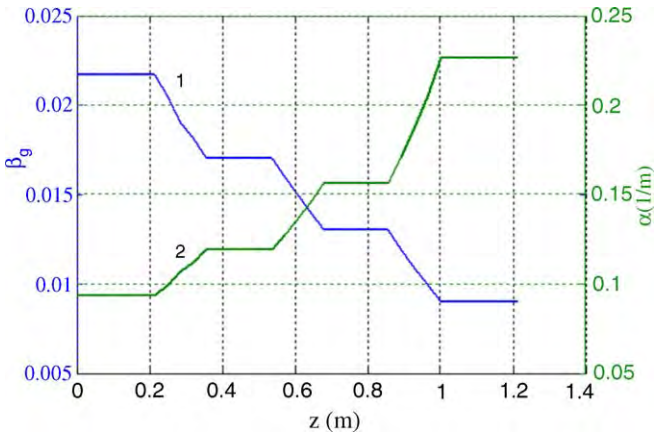



Fig. 1. Dependence of the relative group velocity (curve 1, left axis) and attenuation (curve 2, right axis) on distance along the DWL for the KUT linac.

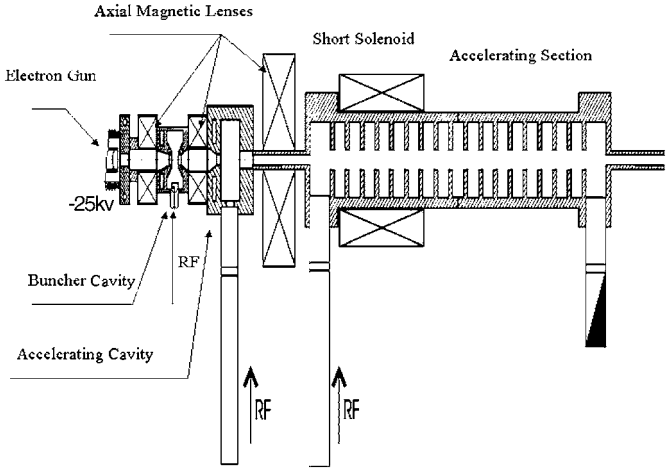



Fig. 2. Linac layout.

Table 1
Specification for the injector cavities

| | Buncher cavity | Accelerating cavity |
|---|---|---|
| $Z_{sh}$ (MΩ/m) | 15.1 | 31.4 |
| $D$ (cm) | 3 | 5 |
| $\beta$ | 10 | 5 |
| Q | 2500 | 10,800 |
| $P_0$ (kW) | 0–0.5 | 1000 |
| $f$ (MHz) | 2797.15 | 2797.15 |

agreed within 3.5%. This is a good result taking into account such a steep change in $v_g$.

## 4. Simulation example

To check the capability of the algorithm designed, full-scale simulation of the KUT linac was carried out. The linac consists of a two-cavity injector and a TW accelerating section, as shown in Fig. 2. The characteristics of the cavities are listed in Table 1. To provide beam transportation, three axial magnetic lenses are installed along the injector and the accelerating section is equipped with a short solenoid. The simulation was carried out at values of the beam current and RF power that are characteristic for KUT operation (accelerated beam current approx. 0.8 A;

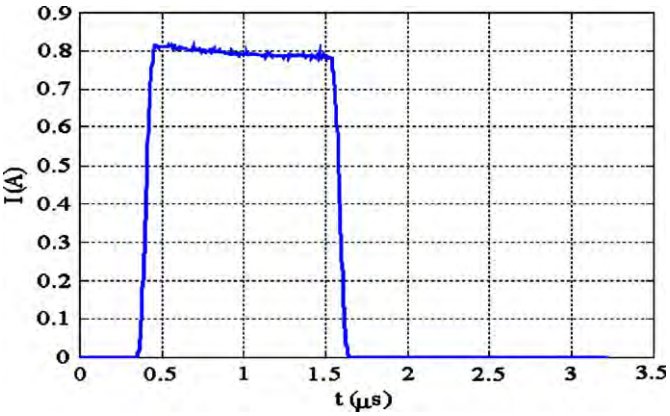



Fig. 3. Waveform of the beam current at a linac exit.

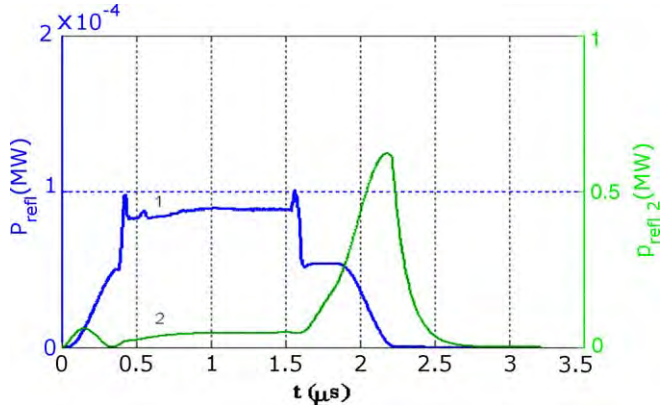



Fig. 4. Waveforms of reflected waves in the feeders of the buncher cavity (1) and the accelerating cavity (2).

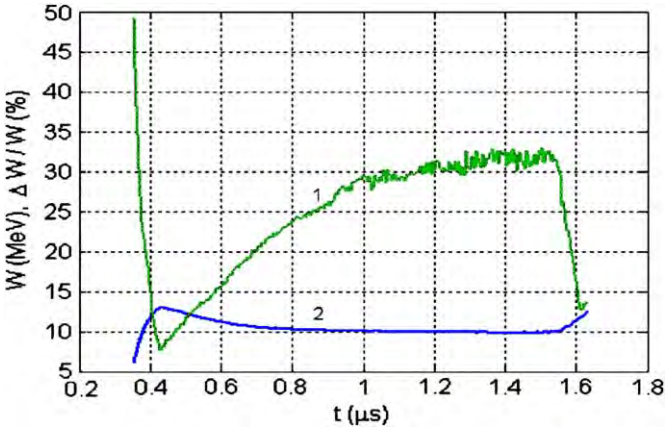



Fig. 5. Time dependence of the energy spread width (1) and the mean beam energy (2).

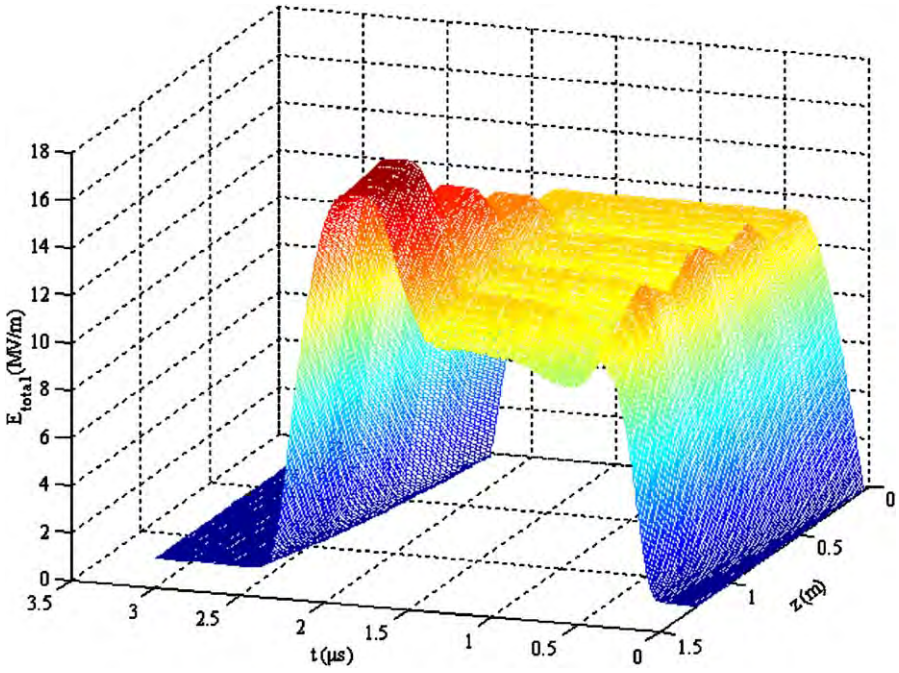



Fig. 6. Time–space distribution of the self-consistent accelerating field.

RF power supply of the buncher cavity, accelerating cavity accelerating section of 70 W, 0.8 MW and 12 MW, respectively).

The duration of a beam current pulse was chosen to be shorter than that of an RF pulse to determine the influence of beam loading. Results of the simulation of self-consistent transient beam dynamics for a certain set of phase shifts between incident waves in the RF feeders are presented in Figs. 3–5. Power waveforms of reflected waves in the feeders of the buncher and accelerating cavities are shown in Fig. 3. Increasing reflection from the bunching cavity at beam injection is evident. The meaning of this effect is given below. Some electrons that are not captured into acceleration can propagate in the backward direction and excite the buncher cavity. Therefore, there is feedback due to the beam. This effect was observed experimentally [9].

Although operation of the injector is stable under the simulation conditions, its adjustment involves fairly complicated work. The time dependence of the energy spread width and the mean beam energy are shown in Fig. 5. It is evident that phasing of the injector and the accelerating section was optimal for the transient region. Such a choice increases the width of the energy spread for the whole pulse. Fig. 6 shows the time–space distribution of the self-consistent field in the DLW excited by the RF source and the electron beam. The accelerated beam substantially decreases the field at the end of the section. On acceleration of the beam with a pulse current of 1.2 A, the field at the end of the section will be approximately zero under the given RF power supply.

## 5. Conclusion

The simulation technique designed for transients in RF linacs allows data on the time-dependent accelerating fields and beam characteristics to be obtained. The testing carried out has shown that the model adequately describes the physical processes involved. The accuracy of simulation of the characteristics of fields and the beam corresponds to the approximations made while developing the technique. The technique proposed can be useful for the design of linacs, as well as in research of beam dynamics.

## References

[1] V.V. Mytrochenko, V.A. Kushnir, A.N. Opanasenko, Simulations of transient phenomena in thermionic RF guns, in: Proceedings of EPAC '02, Paris, 2002, p. 1669.
[2] L.M. Young, PARMELA. LA-UR-96-1835, Los Alamos, 1996.
[3] J.H. Billen, L.M. Young, Poisson Superfish LA-UR-96-1834, Los Alamos, 1996.
[4] L.A. Vainshtein, V.A. Solntsev, Lectures on Microwave Electronics, Soviet Radio, Moscow, 1973 (in Russian).
[5] N.I. Aizatsky, Zh. Tech. Phys. 57 (1987) 1532 (in Russian).
[6] G.A. Loew, R.H. Miller, R.A. Early, K.L. Bane, IEEE Trans. Nucl. Sci. NS 26 (1979) 3701.
[7] N.I. Aizatsky, Yu.I. Akchurin, V.A. Gurin, et al., KUT industrial technological accelerator, in: Proceedings of the 14th Workshop on Accelerators of Charged Particles, vol. 4, Protvino, 1994, p. 259.
[8] O.A. Valdner, A.V. Shalnov, A.N. Didenko, Accelerating Wave Guides, Atomizdat, Moscow, 1973.
[9] N.I. Aizatsky, V.I. Beloglasov, V.N. Boriskin, et al., Experimental studies on an injector complex with RF bunching and pre-accelerator system, in: Proceedings of the International Workshop on $e^+$ and $e^-$ Sources and Pre-Accelerators for Linear Colliders, Schwerin, 1994, p. 409.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 240–246
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Beam dynamics in super-conducting linear accelerator: Problems and solutions

Yu. Senichev*, A. Bogdanov, R. Maier, N. Vasyukhin

*IKP, Forshugszentrum Julich, Germany*

Available online 28 November 2005

**Abstract**

The linac based on SC cavities has special features. Due to specific requirements the SC cavity is desirable to have a constant geometry of the accelerating cells with limited family number of cavities. All cavities are divided into modules, and each module is housed into one cryostat. First of all, such geometry of cavity leads to a non-synchronism. Secondly, the inter-cryostat drift space parametrically perturbs the longitudinal motion. In this article, we study the non-linear resonant effects due to the inter-cryostat drift space, using the separatrix formalism for a super-conducting linear accelerator [Yu. Senichev, A. Bogdanov, R. Maier, Phys. Rev. ST AB 6 (2003) 124001]. Methods to avoid or to compensate the resonant effect are also presented. We consider 3D beam dynamics together with space charge effects. The final lattice meets to all physical requirements.


*PACS:* 29.17.+w; 29.27.−a; 45.20.Jj

*Keywords:* Linear accelerators; Charged-particle beams; Hamiltonian; Nonlinear optics

## 1. Introduction

Super-conducting (SC) RF technology allows getting the high gradient acceleration. Especially it is effective in case of low current, since the power required for the beam acceleration is much smaller than the RF power needed for the RF field creation in a normal-conducting cavity. Additional argument for a SC cavity is a possibility to accelerate particles with different charge-mass ratio due to independent feeding of cavities. However, due to lot of reasons a super-conducting cavity does not have flexible tuning as a normal-conducting cavity, and geometry of SC cavity has to be simplified as much as possible. In particular, it is desirable to have SC cavity with constant geometry of accelerating cells and changing from one family of cavities to another. It means the phase velocity changes step by step from family to family as well. The particles are sliding down or up relatively of RF wave in dependence on ratio between the particles and the wave velocities. Thus, the particles are almost never in synchronism with the equivalent traveling wave, and they have not instantaneous stability. Nevertheless, abandoning from synchronism, we acquire a freedom in choice of RF phase shift between relatively short SC cavities, which can provide the stable quasi-synchronous motion in whole accelerator.

For cryogenics all SC cavities are divided into modules, and each module is housed into one cryostat. However, some equipment, in particular, diagnostics, vacuum pumps and focusing elements desirable to place outside of cryostat under normal conditions. It requires an additional drift space between cryostats. The number of SC cavities in one module is determined by required length between focusing elements. In the same time drift spaces can be considered as a parametric perturbation of the longitudinal motion. Evident advantages of SC cavities initiate its application starting with the lowest possible energy. Moving down with energy the longitudinal frequency grows and it can be comparable with the repetition frequency of drift spaces. In this case, the resonance condition for particle in longitudinal plane can be realised. Since the drift space length can be comparable with the length of module, the resonance width can be significant and it could

*Corresponding author.
*E-mail address:* y.senichev@fz-juelich.de (Yu. Senichev).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.055

dramatically affect on beam stability. In this paper we study the resonant effects due to the inter-cryostat drift space. Developed theory let us follow the beam dynamics under the influence of different order parametric resonances, as well as estimate resonance width and find the most dangerous resonances. Methods to avoid or to compensate the resonant affect are also presented.

At present SC linear accelerator is considered as candidate for intensive beam with high duty cycle. We investigated the space charge effect in SC linear accelerator based on module lattice with long distance between focusing elements and strong defocusing RF field.

Due to absence of synchronous particle the SC linear accelerator has specificity in RF amplitude and phase tuning. In this article, we consider the modified delta-T procedure for SC linear accelerator.

All practical applications are considered on the example of COSY SC linac [2].

## 2. SC accelerators with stepped RF phase structure

We already mentioned that in SC linear accelerator the cavities are joint in families, and all $n$ cavities of one family have the same structure phase velocity $\beta_{str} = \text{const}$ for $i \in 1-n$. Thus, particles move in a cavity, where the structure phase velocity $\beta_{str}$ is constant. Therefore, they oscillate around $\varphi_s = 0^0$ (for "sin" wave). In the considered case the particle velocity deviation from the structure phase velocity $\Delta\beta = \beta_{str} - \beta$ can exceeds the velocity spread of stationary separatrix with a synchronous level $\beta = \beta_{str}$. Fig. 1 explains the mechanism of acceleration. It is based on an RF shift for each cavity. The particles have to return back each time after passing through a cavity to get an average phase $\overline{\phi_s}$ over all cavities. By a proper choice of the RF phase shift $\Delta\varphi_{RF}$ between cavities one can create a

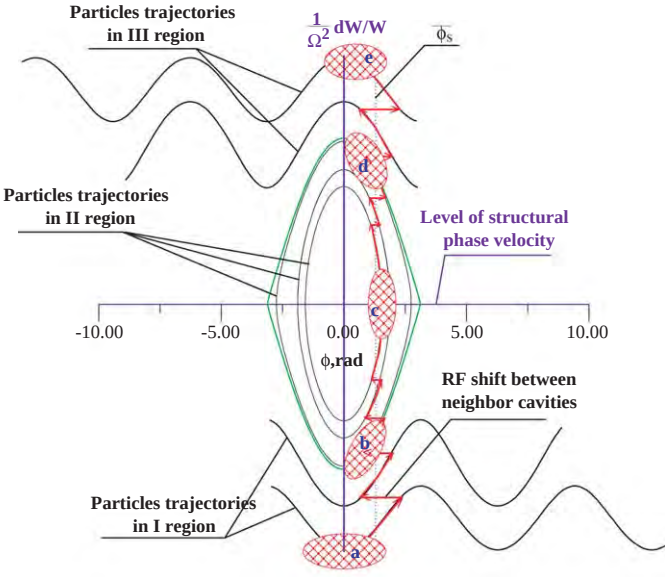



Fig. 1. Longitudinal motion in stepped RF phase structure.

quasi-synchronous motion and in total a stable motion in the whole accelerator. The quasi-synchronous particle oscillating in a cavity around $\varphi = 0^0$ is forced by the inter-cavity RF shift to oscillate around $\overline{\phi_s}$. In one cavity we can write the equations:

$$\frac{d\phi}{d\varsigma} = \frac{\beta_{str}}{\beta} - 1$$
$$\frac{d\beta}{d\varsigma} = \frac{eE_{ac}\lambda\beta_{str}}{2\pi m_0 c^2 \gamma^3 \beta}\sin\phi \quad (1)$$

where $d\varsigma = 2\pi(dz/\beta_{str}\lambda)$ is new normalized longitudinal coordinate, $E_{ac}$ is accelerating field and $\lambda$ is wave length. The separatrix in case of the stepped RF phase is created in the following way [1] . From Eq. (1) one derives the phase oscillation equation $d^2\phi/d\varsigma^2 + \Omega^2\sin\phi = 0$, where $\Omega^2 = A_E \cdot \beta_{str}^2/\beta^3$ is determined by means of parameter $A_E = eE_{ac}\lambda/2\pi m_0 c^2\gamma^3$.

Obviously, if one does not undertake some action with phase $\phi$, the particles will accelerate around phase $\phi = 0$, and acceleration will be absent, since $\Delta W \propto \overline{\sin\phi_s}$. To correct this situation, one adds the external phase shift $\varphi_{str}(\varsigma)$ to the phase $\phi$, and the corrected phase is:

$$\phi(\varsigma) = \int_0^{\varsigma}\frac{\beta_{str}\,d\xi}{\beta} - \int_0^{\varsigma} d\xi + \varphi_{str}(\varsigma) \quad (2)$$

and the first equation of system (1) takes the form:

$$\frac{d\phi}{d\varsigma} = \frac{\beta_{str}}{\beta} - 1 + \frac{d\varphi_{str}(\varsigma)}{d\varsigma}. \quad (3)$$

Then the phase oscillation equation is:

$$\frac{d^2\phi}{d\varsigma^2} + \Omega^2\sin\phi - \frac{d^2\varphi_{str}}{d\varsigma^2} = 0. \quad (4)$$

Thus, the second derivative of $\varphi_{str}(\varsigma)$ defines the acceleration rate in the stepped RF phase structure.

However, the RF phase of each cavity is fixed and changes from cavity to cavity step by step. The step value is proportional to periodicity $T$ of the cavities. It is a stepwise function with an average value $\overline{\varphi}_{str}$ coinciding with an RF shift of the ideal case. Introducing a "triangular" function $\tilde{\varphi}_{norm} = \sum_{m=1}^{\infty}(1/\pi m)\sin m\nu_{ph}\varsigma$, the real $\varphi_{str}(\varsigma)$ could be submitted through the sum:

$$\varphi_{str}(\varsigma) = \overline{\varphi}_{str}(\varsigma) + \frac{d\overline{\varphi}_{str}}{d\varsigma}\cdot 2\pi\cdot T\cdot\tilde{\varphi}_{norm}(\varsigma). \quad (5)$$

As distinct from the stepped phase velocity structure in the considered structure the amplitude of triangular function is proportional to the derivative $(d\overline{\varphi}_{str}/d\varsigma)\cdot 2\pi\cdot T$:

$$\tilde{\varphi}_{str} = \frac{d\overline{\varphi}_{str}}{d\varsigma}\cdot 2\pi\cdot T\cdot\sum_{m=1}^{\infty}\frac{1}{\pi m}\sin m\nu_{ph}\varsigma \quad (6)$$

where $\nu_{ph} = 1/T$. Substituting Eqs. (5), (6) in Eq. (4), one gets

$$\frac{d^2\phi}{d\varsigma^2} + \Omega^2 \sin\phi - \Omega^2 \sin\overline{\phi_s} = 2\frac{d^3\overline{\varphi}_{str}}{d\varsigma^3}\sum_{m=1}^{\infty}\frac{\sin m\nu_{ph}\varsigma}{m\nu_{ph}} + 4\frac{d^2\overline{\varphi}_{str}}{d\varsigma^2}\sum_{m=1}^{\infty}\cos m\nu_{ph}\varsigma - 2\frac{d\overline{\varphi}_{str}}{d\varsigma}\sum_{m=1}^{\infty} m\nu_{ph}\sin m\nu_{ph}\varsigma. \quad (7)$$

To compare the contribution of the terms on the right side of Eq. (7) all derivatives are written in more explicit form using the current velocity $\beta$:

$$\frac{d\overline{\varphi}_{str}}{d\varsigma} = 1 - \frac{\beta_{str}}{\beta(\varsigma)}; \quad \frac{d^2\overline{\varphi}_{str}}{d\varsigma^2} = \Omega^2(\varsigma)\sin\overline{\phi_s}$$

and

$$\frac{d^3\overline{\varphi}_{str}}{d\varsigma^3} = 3\frac{\beta(\varsigma)}{\beta_{str}}\Omega^4\sin^2\overline{\phi_s}. \quad (8)$$

In the vicinity of $\overline{\phi_s}$ one can write for (7)

$$\phi(\varsigma) = A\cos(\Omega\cos^{1/2}\overline{\phi_s}\cdot\varsigma + \chi) + \overline{\phi_s} - 2\frac{d\overline{\varphi}_{str}}{d\varsigma}\sum_{m=1}^{\infty}\frac{m\nu_{ph}\cdot\sin m\nu_{ph}\varsigma}{\Omega^2\cos\overline{\phi_s} - m^2\nu_{ph}^2}. \quad (9)$$

As in the case of the stepped phase velocity one defines the particle oscillating only with external frequency as a quasi-synchronous particle:

$$\phi_{qs} = \overline{\phi_s} - 2\frac{d\overline{\varphi}_{str}}{d\varsigma}\sum_{m=1}^{\infty}\frac{m\nu_{ph}\cdot\sin m\nu_{ph}\varsigma}{\Omega^2\cos\overline{\phi_s} - m^2\nu_{ph}^2}. \quad (10)$$

Let us pass on to a coordinate system moving together with the quasi-synchronous particle $\psi = \phi - \overline{\phi_s} - f(\varsigma)$, where $f(\varsigma)$ is:

$$f(\varsigma) = -2\frac{d\overline{\varphi}_{str}}{d\varsigma}\sum_{m=1}^{\infty}\frac{m\nu_{ph}\cdot\sin m\nu_{ph}\varsigma}{\Omega^2\cos\overline{\phi_s} - m^2\nu_{ph}^2}. \quad (11)$$

For $\nu_\beta \gg \Omega\cos^{1/2}\overline{\phi_s}$ one retains the first harmonic only $f(\varsigma) = \varphi_a\sin\nu_{ph}\varsigma$ with $\varphi_a = (2/\nu_{ph})\cdot(d\overline{\varphi}_{str}/d\varsigma)$.

Using new variables $\{\Delta\beta = \beta_{qs} - \beta; \psi = \phi - \overline{\phi_s} - f(\varsigma)\}$ one gets the equation system [1]:

$$\frac{d\psi}{d\varsigma} = p_\psi$$
$$\frac{dp_\psi}{d\varsigma} = -\Omega_0^2[\sin(\psi + \overline{\phi_s}) - \sin\overline{\phi_s}] \quad (12)$$

where $p_\psi = (\beta_{str}/\beta_{qs}^2)\cdot(\beta - \beta_{qs})$ is the relative velocity spread and $\Omega_0^2 = (\beta_{str}^2/\beta_{qs}^3)A_E J_0(\varphi_a)$. Expanding the right-side of the second equation (12) into a series about $\psi$, using new impulse and new time variable $p = p_\psi/\Omega_0\cos^{1/2}\overline{\phi_s}$; $\tau = \Omega_0\cos^{1/2}\overline{\phi_s}\cdot\varsigma$, and then passing to "action-angle" $(I,\theta)$ coordinates $I = \frac{p^2}{2} + \frac{\psi^2}{2}$; $\psi = \sqrt{2I}\cdot\cos\theta$; and $p = \sqrt{2I}\cdot\sin\theta$, we write Hamiltonian in the new coordinates:

$$H(I,\theta) = I - \frac{2^{3/2}}{3!}\tan\overline{\phi_s}\cdot I^{3/2}\cos^3\theta - \frac{2^2}{4!}\cdot I^2\cos^4\theta + \frac{2^{5/2}}{5!}\tan\overline{\phi_s}\cdot I^{5/2}\cos^5\theta + \cdots \quad (13)$$

## 3. Longitudinal motion parametrization by inter-cryostat drift space

Now we consider the motion with drift space between cryo-modules, when the particle passes through the system "module + drift" and the field $E_{ac}$ can be submitted through Fourier series:

$$E_{ac} = E_0\frac{L_{mod}}{L_T}\cdot\left(1 + 2\sum_k\frac{\sin(k\pi L_{mod}/L_T)}{k\pi L_{mod}/L_T}\cos k\nu\varsigma\right) \quad (14)$$

where $\nu = 1/T_p$ is the perturbation frequency and the period of perturbation measured in number of $\beta_{str}\lambda$, that is $T_p = L_T/\beta_{str}\lambda$ and $L_T$—space length of the perturbation period.

Taking into account the field representation in an accelerator with drifts and supposing that one of the resonant condition $\nu \approx n\Omega$ is fulfilled, we retain only slow-oscillating terms $\cos(n\theta - (\nu/\Omega)\tau)$. Hence, the phase $\vartheta = \theta - (\nu/n\Omega)\tau$ changes very slowly, and $\vartheta$ can be considered as parameter. Using the generating function $F(\overline{I},\vartheta,\tau) = -\overline{I}\cdot(\vartheta + (\nu/n\Omega)\cdot\tau)$, we pass to the resonant Hamiltonian $\bar{H}_{r,n}(\bar{I},n\vartheta) = \overline{H}(\bar{I},n\vartheta) - (\nu/n\Omega)\cdot\bar{I}$ with variables $\overline{I},n\vartheta$. Here we introduce the parameter $\Delta_n = 1 - (\nu/n\Omega)$, showing how far the eigen frequency of small oscillation from $n$-integer resonance, or by another words, it is the frequency detuning from resonance, when $\overline{I} \ll 1$ and $\dot{\theta} \approx 1$. In new variables the resonant Hamiltonian is for the first-integer resonance $\dot{\theta} \approx \nu/\Omega$

$$\overline{H}_{r,1}(\overline{I},\vartheta) = \Delta_1\cdot\overline{I} - \frac{1}{16}\overline{I}^2 - \frac{2^{1/2}}{8}e_1\tan\overline{\phi_s}\cdot\overline{I}^{3/2}\cos\vartheta + \frac{2^{1/2}}{96}e_1\tan\phi_s\cdot I^{5/2}\cos\vartheta \quad (15)$$

for the second-integer resonance $2\dot{\theta} \approx \nu/\Omega$

$$\overline{H}_{r,2}(\overline{I},2\vartheta) = \Delta_2\cdot\overline{I} - \frac{1}{16}\overline{I}^2 + \frac{1}{4}e_1\cdot\overline{I}\cos 2\vartheta - \frac{1}{24}e_1\cdot\overline{I}^2\cos 2\vartheta \quad (16)$$

for the third-integer resonance $3\dot{\theta} \approx \nu/\Omega$

$$\overline{H}_{r,3}(\overline{I},3\vartheta) = \Delta_3\cdot\overline{I} - \frac{1}{16}\overline{I}^2 - \frac{2^{1/2}}{24}e_1\tan\overline{\phi_s}\cdot\overline{I}^{3/2}\cos 3\vartheta + \frac{2^{1/2}}{192}e_1\tan\overline{\phi_s}\cdot I^{5/2}\cos 3\vartheta \quad (17)$$

and so on.

The resonant Hamiltonian (15–17) describes the motion in the coordinates system rotating with angular frequency $\omega_n = \nu/n\Omega$ in normalized time $\tau = \Omega \cdot \varsigma$. One can see the non-linear approach gives many additional resonances in comparison with linear case. As example Fig. 2 shows the curves of $\Omega(W)$ and $\frac{\nu}{n}(W)$ together with $n$-integer resonance boundaries.

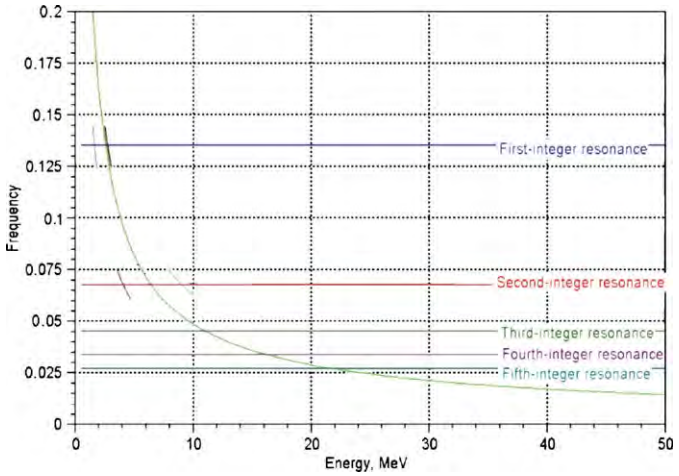



Fig. 2. Non-linear resonances in COSY injector.

## 4. Direct and inverse bifurcations under non-linear resonance crossing

For low energy accelerator the second-integer resonance is the first, which one affects the beam with initial energy, in particular, in COSY 2.5 MeV (see Fig. 2). It has the quadrupole symmetry topology. Therefore, for the more visual interpretation let us pass in Cartesian coordinates

$$H_{r,2}(\bar{\psi},\bar{p}) = \frac{\bar{\psi}^2}{2}\left(\Delta_2 + \frac{1}{4}e_{1,2}\right) + \frac{\bar{p}^2}{2}\left(\Delta_2 - \frac{1}{4}e_{1,2}\right) - \frac{\bar{\psi}^4}{64}\left(1 + \frac{2}{3}e_{1,4}\right) - \frac{\bar{p}^4}{64}\left(1 - \frac{2}{3}e_{1,4}\right) - \frac{1}{32}\bar{\psi}^2\bar{p}^2 \quad (18)$$

where $\bar{\psi},\bar{p}$ describe the averaged motion. Here we introduce new parameters $e_{1,2}$ and $e_{1,4}$. They both equal to $e_1$, but the second index indicates on their origin. The first of them $e_{1,2}$ is due to the quadrupole term in Hamiltonian, and the second $e_{1,4}$ is due to the octupole term. Fig. 3 shows the separatrix of second integer resonance together with the isolated singular points and distances from the origin for detuning $|\Delta_2| < \frac{1}{4}e_{1,2}$.

During acceleration the longitudinal frequency $\Omega$ of particle motion goes down (Fig. 2), and detuning

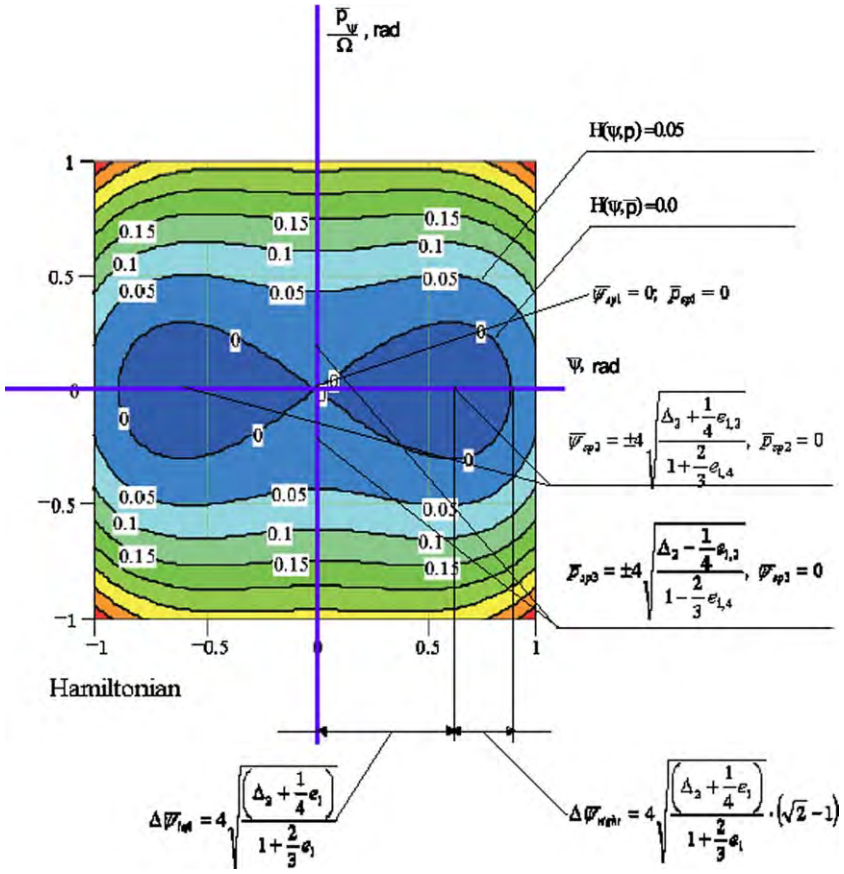



Fig. 3. Separatrix of second integer resonance.

parameter $\Delta_2 = 1 - (\nu/2\Omega)$ changes its value from positive to negative. Depending on trajectories behavior in the vicinity of the origin coordinates we will subdivide the detuning values on the resonant $|\Delta_2| < e_{1,2}/4$ and non-resonant $|\Delta_2| > e_{1,2}/4$ regions.

Since in the non-resonant region Hamiltonian (18) can be submitted as $(\bar{\psi}^2/2)(\Delta_2 + e_{1,2}/4) + (\bar{p}^2/2)(\Delta_2 - (e_{1,2}/4)) = C_i$, where both coefficients of $\psi^2$ and $p^2$ have either positive ($\Delta_2 > \frac{1}{4}e_{1,2}$), or negative ($\Delta_2 < -\frac{1}{4}e_{1,2}$) signs simultaneously, the phase trajectories at small $\psi$ have an elliptical form with the center in the origin of coordinates $\bar{\psi} = 0$, $\bar{p} = 0$.

The oscillation frequency of particles with a small deviation is $\dot{\vartheta} \approx (\Delta_2^2 - \frac{1}{16}e_{1,2}^2)^{1/2}$. Far from the resonance $|\Delta_2| \gg e_{1,2}/4$ the ratio of $\bar{\psi}_m$ to $\bar{p}_m$ axes is $\bar{\psi}_m/\bar{p}_m = \sqrt{\frac{\Delta_2 - (e_{1,2}/4)}{\Delta_2 + (e_{1,2}/4)}} \to 1$ and $\dot{\vartheta} \to \Delta_2$. It means that outside the resonance, when detuning $|\Delta_2| \gg e_{1,2}/4$, particles have the circular phase trajectories, and the total frequency in the original phase space $\Delta_2 + (\nu/2\Omega)$ equals to unit. Hence, in accordance with definition of the detuning parameter $\Delta_2 = 1 - (\nu/2\Omega)$ the trajectories are unperturbed. For the non-resonant cases $\Delta_2 > e_{1,2}/4$ and $\Delta_2 < -e_{1,2}/4$ the phase trajectories have elliptical form in the central part.

Now let us consider the resonant case, when the coordinates origin losses stability $|\Delta_2| \leqslant \frac{1}{4}e_1$. In this moment both coefficients of $\psi^2$ and $p^2$ in Hamiltonian have different signs, and how we mentioned already, the center of oscillation degenerates into the saddle. The phase trajectories have the form of hyperbola in the central region $\psi^2 = ((\frac{1}{4}e_{1,2} - \Delta_2)/(\Delta_2 + \frac{1}{4}e_{1,2})p^2)$. In the periphery due to the non-linear third term of Hamiltonian they are closing with two islands formation. From Eq. (18) we can find the equation of small oscillation inside these islands:

$$\Delta\ddot{\bar{\psi}} + 2\left(\frac{e_{1,2}^2}{16} - \Delta_2^2\right) \cdot \Delta\bar{\psi} = 0. \tag{19}$$

Thus, in the resonance $\Delta_2^2 < e_{1,2}^2/16$ there are two regions: outside islands particles oscillate with frequency as before $\dot{\vartheta} \approx \sqrt{\Delta_2^2 - \frac{1}{16}e_{1,2}^2}$, and inside with frequency

$$\dot{\vartheta} = \sqrt{2\left(\frac{e_{1,2}^2}{16} - \Delta_2^2\right)}.$$

The transformation process of the elliptical motion with one stable centre ($\Delta_2 > \frac{1}{4}e_{1,2}$) into the hyperbolical motion with two stable centres ($\Delta_2 < \frac{1}{4}e_{1,2}$) is called the direct bifurcation. And the inverse transformation of the two centers motion into the one center motion is called the inverse bifurcation. During the resonance crossing particles experience both bifurcations. It is possible, when neighbour particles begin to move with the different frequencies. After resonance the frequencies become equal. Bifurcation causes a significant growth of the effective emittance. We have considered bifurcation under second order resonance crossing. The same phenomena can be observed under higher order resonance.

Moving away from the second integer resonance, we approach to the third integer resonance. The force of the third integer resonance is much smaller than the second integer resonance for the central part of phase space, and it can be stronger for the larger deviation from centre. Many authors investigated the resonances crossing. In particular, in Ref. [3] the effect of passing through the isolated resonance is considered. In our case the particles not leaving the previous resonance they are coming in the next one. It is interesting to see the integrated effect of resonance crossing, when the detuning changes dynamically during the acceleration process with Lorenz damping. The phase portrait strongly depends on how we come through resonances. Hamiltonian formalism does not allow doing that, and we used the numerical simulation. Fig. 4 shows the option nearest to real case of COSY, when we start from the detuning from the second order resonance $\Delta_2 = 0.3$ and finish by the detuning from the third order resonance $\Delta_3 = -0.13$.

There are different methods how to decrease influence of resonance. It is obviously in order to avoid the resonance crossing we should change the frequency perturbation $\nu$, for instance, by decreasing cavities number in one cryo-module. But it is not always convenient from technical point of view.

Another method is the recovering of phase shape of bunch by two harmonics debuncher [4]. This debuncher consists of two cavities: cavity with normal frequency and cavity with doubled frequency. As soon as the phase shift is small enough we can say that both harmonics act simultaneously. This system can be represented as system with two harmonic simultaneously: $(d^2\varphi/dt^2) + \Omega^2(\sin(\varphi) + \alpha\sin(2\varphi)) = 0$, or

$$\frac{d^2\varphi}{dt^2} + \Omega^2 \sin(\varphi)(1 + 2\alpha\cos(\varphi)) = 0.$$

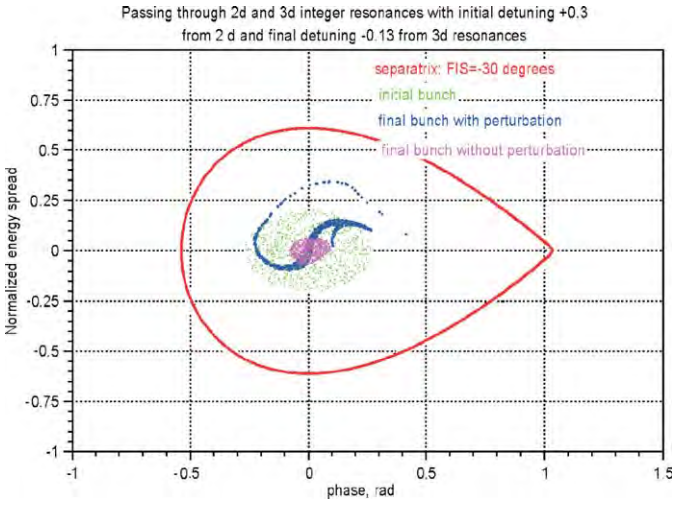



Fig. 4. Dynamical change of detuning from $\Delta_2 = 0.3$ to $\Delta_3 = -0.13$ ($e_1 = 1.2, \alpha = 0.14$).

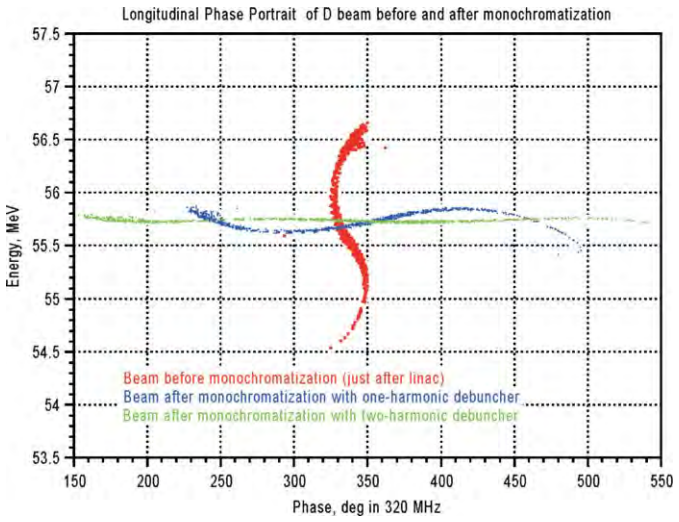



Fig. 5. Phase portrait on linac exit with and without debuncher.

Now we construct the Hamiltonian $H = (P^2/2) - \cos(\varphi) - (\alpha/2)\cos(2\varphi)$ in normalized time $\tau = \Omega t$ and normalized momentum $P = \frac{1}{\Omega}(\mathrm{d}\varphi/\mathrm{d}t) = \mathrm{d}\varphi/\mathrm{d}\tau$. We can see that the form of Hamiltonian seems like resonant Hamiltonian of 2nd order for $\alpha \approx -0.5$, what means it acts oppositely to linear accelerator itself. Therefore, after debuncher the phase portrait becomes more straight (see Fig. 5).

## 5. Tune-up procedure of SC linear accelerator

Development of the tune-up procedure for a linear accelerator is the next important stage after the design is complete. Conventional $\Delta T$ procedure developed by Crandall for a tuning of a normal-conducting (NC) linear accelerator allows setting up of the accelerating field amplitude and phase in cavities with design phase velocity [5]. In contrast the quasi-synchronous phase velocity in SC linear accelerator is determined by an amplitude and RF phase of cavities. And the phasing itself must be established in frame of the tune-up procedure. The lack of coincidence between the experimental and design values of the equivalent phase velocity results in the uncertainty of the final energy. Moreover, a SC cavity is short and has small phase advance of the longitudinal motion, which leads to an insensibility of particles motion to variation of the electric field inside the cavity.

The main task for the phasing of the cavities is to ensure a quasi-synchronous motion of the particles. In order to realize quasi-synchronous motion in the real accelerating structure the phasing must be based on the measured data of the beam. We scan the initial phase of a particle and search for the particle with zero energy gain. Than the phase of the sought particle differs from the measured by $-(90^\circ + \overline{\varphi_s})$. After all the cavities have been phased the equivalent phase velocity of the structure is adjusted.

Because of small phase advance the accuracy of time flight procedure is greatly insufficient to feel changing of the field amplitude and phase inside the cavity. On the other hand it means the particle motion remains stable with respect to the amplitude and phase errors. This stability allows joining several cavities in one module without losing the quality of the beam. And the further tuning deals with such a module. The phase advance equal to the half of one longitudinal oscillation gives the maximum sensitivity. Hence the optimal number of cavities in a module must provide around $\pi$ of the phase advance.

We have investigated how variation of field amplitude and phase along one module with discontinuity of one cavity affects on the equivalent phase velocity. One of the ways to change the equivalent phase velocity of some particular module is to vary an average field level in the whole module, after which it is necessary to re-phase the cavities inside the module. In order to control the behavior of the equivalent phase velocity and to compare it with the design value we analyze the $\Delta T$ plane. After we adjusted the equivalent phase velocity of the real accelerating structure we use the conventional $\Delta T$-procedure to finalize tuning. In details $\Delta T$-procedure for SC linear accelerator is described in Ref. [6].

The modified tune-up gives one of the most attractive features of SC linacs. It is the possibility of the smooth adjustment of the final energy. The phasing of the cavities, which allows changing the equivalent phase velocity, provides a basis for such a possibility. This fact puts the SC linear accelerator in one series of accelerators-candidates for medical therapy.

## 6. Features of high intense beam dynamics in SC linear accelerator

From mentioned above we see that SC linear accelerator has a lot of advantages in comparison with room temperature accelerators. In case of intense beam with high duty cycle and high accelerating gradient the SC linear accelerator is the only candidate. But together with benefits there are new problems for high intense beam. In particular, the focusing elements of SC linear accelerator usually are located between cryo-modules, what makes the focusing period longer and the space charge effect stronger. Simultaneously, together with strong RF defocusing factor the high accelerating gradient gives the phase advance in longitudinal plane comparable with the phase advance in radial plane. Due to all these factors the SC linear tune shift can reach the value up to 50%. Nevertheless, compensating the linear part of space charge force, we have the strong influence of non-linear space charge part, which affects on hallo particles and causes the losses. Since the limitation for high current accelerators usually is coming from losses, approximately 5 W/m, the non-linear space charge acquires significant role. The aperture of SC cavity can be done larger than quadrupole, therefore the main part of losses usually is located in quadrupoles. In the same time maximum quadrupole aperture is restricted by

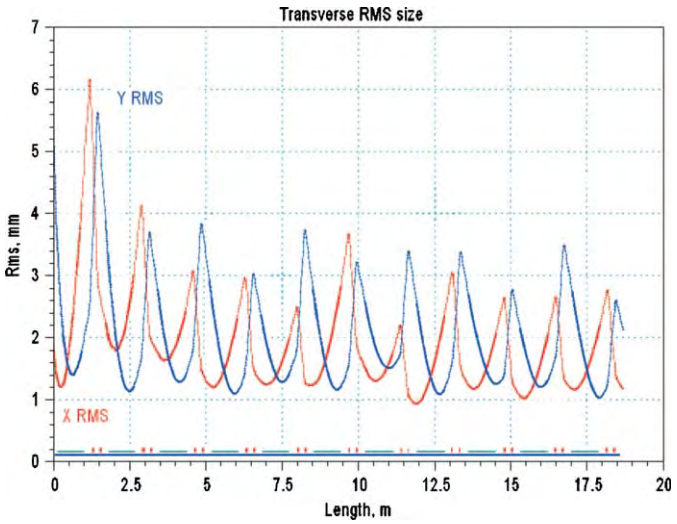



Fig. 6. Phase portrait on linac exit with and without debuncher.

magnetic saturation on pole, and for 2 cm radius aperture the gradient is limited by 50 T/m.

As example we consider COSY injector design. Focusing period consist of four SC half wave cavities [2] in cryomodule and quadrupole doublet. Numerical calculation shows it is possible in this structure to accelerate 30 mA of average current with losses of order 5 W/m and duty factor 100%. The beam is accelerated from energy 2.5 MeV up to 50 MeV in two kinds of resonators with operation frequency 160 MHz on first 5 periods and 320 MHz on rest 6 ones. Transverse RMS are shown at Fig. 6.

Since the matched acceptance depends on RF phase the total transverse RMS emittance growth by ~30%. Obviously, for shorter focusing period the effect of space charge is weaker. In particular, the linear accelerator with 2 resonators in cryo-module and twice shorter focusing period has maximum current 160 mA. In detail the space charge effect is investigated in Ref. [7].

## 7. Conclusion

In SC linear accelerator is no synchronism between particle and phase wave. To create a quasi-synchronous motion the stepped RF phasing of cavities is used. In this article we have used the developed earlier separatrix formalism for super-conducting linear accelerators in absence of synchronism.

We have investigated the inter-cryostat drift space influence on the longitudinal motion in a super-conducting linear accelerator. It is known that in the non-resonant case the drift decreases the separatrix momentum spread proportionally to $\Delta p/p \propto \sqrt{L_{\mathrm{cav}}/(L_{\mathrm{cav}}+L_{\mathrm{drift}})}$ and keeps the same phase length. However, in case of the resonance $\nu = n\Omega$ the influence of drift space plays significant role. The shape of both separatrix and bunch can change dramatically. Using Hamiltonian formalism, we determined the resonance boundary and qualitatively described the behavior of particles versus values of the perturbation and detuning from resonance.

We have studied analytically and numerically the direct and inverse bifurcation phenomena. The last one allowed us understanding of bunch shape distortion. Comparing the results of this model with numerical simulation in 3D fields of MAFIA, we could explain the bunch behavior in the COSY SC linear accelerator.

We have developed two possible methods of avoidance of the strong resonance influence. The first obvious solution is a resonance bypass. The first cavities are housed in one cryo-module two by two. After the resonance passing we return to the option four cavities in one cryo-module. The second method is two harmonics buncher-rotator. Its action is so adjusted in order to unbend S shape bunch after a linear accelerator. For the COSY SC linac we made a choice for second method, since it gives additional beam monochromatization, which one is important in case of injection in a ring.

The modified $\Delta T$-procedure has been developed for SC linear accelerator. It has the next main features. Firstly, we tune up several cavities at once. Secondly, before we can use the conventional $\Delta T$ procedure, it is essential to adjust the equivalent phase velocity of the real accelerating structure. This adjustment can be done by variation of the field level inside the module being tuned up. And only after that we are able to apply convenient $\Delta T$-procedure for the final tuning.

High intense linear accelerators based on SC structures gives the advantage like lower power supply at higher accelerating gradient. Simultaneously, SC linear accelerator has the strong RF defocusing factor and longer focusing period in comparison with conventional NC accelerator. We have investigated how critically 6D mismatching affects on losses and determined criteria for SC accelerator parameters choice.

## References

[1] Yu. Senichev, A. Bogdanov, R. Maier, Phys. Rev. ST AB 6 (2003) 124001.
[2] R. Toelle, et al., A Superconducting Injector LINAC for COSY, EPAC 2002.
[3] G. Guignard, Effect des champs magnetiques perturbatears d'un synchrotron, CERN 70-24, 1970 (Chapter 4).
[4] N. Vasyukhin, R. Maier, Y. Senichev, R. Stassen, R. Toelle, Energy monochromatization system for injection H and D beams in synchrotron, ICANS-XYI, 2003, pp. 977–981.
[5] K.R. Crandall, The delta-$T$ tune-up procedure for the LAMPF 805 MHz linac, LANL Report LA-6374-MS, June 1976.
[6] A. Bogdanov, R. Maier, Yu. Senichev, Delta-T procedure for SC linear accelerator, EPAC 2004, pp. 1249–1251.
[7] N. Vasyukhin, R. Maier, Yu. Senichev, Space charge problem in low energy SC linear accelerator, EPAC 2004, pp. 1999–2001.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 247–248
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Low-energy beam line at KVI

Dragan Toprek[a,*], Iwo Formanoy[b], Sytze Brandenburg[b]

[a]VINCA Institute of Nuclear Sciences, Laboratory for Nuclear and Plasma Physics 011, P.O. Box 522, 11001 Belgrade, Serbia and Montenegro
[b]Kernfysisch Versneller Instituut (KVI), Zernikelaan 25, 9747 AA, Groningen, The Netherlands

Available online 21 November 2005

**Abstract**

This paper describes the redesign of the low-energy beam line at KVI. Redesigned and properties of the optical elements of the transport beam line is done by using the code COSY INFINITY in the third-order of approximation. The effects of fringe fields of the optical elements are also taken into account in the third order of accuracy. Examples of beam transport calculations and emittance measurements are shown in the case of $^3He^+$. The beam emittance is estimated by varying quadrupole method. By using COSY code it is concluded that the settings of some optical elements in the present design of the beam line should changed to increase the beam transmission. In this way the beam transmission is increased 2–3 times. To achieve higher beam transmission the redesign of the beam line is necessary.


*PACS:* 29.27.Eg; 29.27.Fh; 41.75.Ak

*Keywords:* Beam; Transmission; Fringe field; Emittance; Quadrupole

## 1. Introduction

The beam current of $^3He^+$ beam just after the analyzing magnet M72 (see Fig. 1 in the Ref. [1]) is about 20 μA but after the matching section (matching section consists of five electrostatic quadrupoles; from EQ8 to EQ12 in Fig. 1) is only 2 μA. To improve the transmission of the beam through the injection beam line, the optical properties of the transport beam line and the beam emittance should be known. The beam emittance is measured at the three different positions along the beam transport line and in the both planes, horizontal and vertical, by varying the strength of an upstream quadrupole. The first horizontal/vertical profile grid (marked by letters HH1/HV1 in the Fig. 1) is located 7.17 m/7.27 m downstream from the ion source, the second profile grid (HH2/HV2) is located 10.20 m/10.30 m and the third one (HH4/HV4) is located 15.40 m/15.41 m downstream from the ion source.

Redesign of the transport beam line (within the existing building) and its optical properties is done by using the code COSY INFINITY [2–5] in the third order of approximation, taking into account in the third order of accuracy the effects of the fringe fields also.

## 2. COSY calculations for the present beam line design

The optical properties of the low energy beam line are studied for the $^3He^+$ beam. It is considered the case of $\varepsilon_x = \varepsilon_y = 50\pi$ mm mrad for the beam emittances and 0.5% of the half of the momentum spread.

By using COSY code it has come to conclusion that the strength of the EQ5 and EQ7 electrostatic quadrupoles should increase to increase the beam transmission 2–3 times.

After the M72 magnet and after a long drift space (210 cm) is a matching section which consists of five electrostatic quadrupoles. After the matching section the beam is bend 90° up by an electrostatic deflector, into the cyclotron. The beam envelopes, calculated by the code COSY INFINITY, are presented in Fig. 1a (horizontal plane) and Fig. 1b (vertical plane). The double vertical lines are the positions of the profile grids in the horizontal (first line) and the vertical (second line) plane.

*Corresponding author.
*E-mail address:* dragan.toprek@accel.de (D. Toprek).

0168-9002/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.008

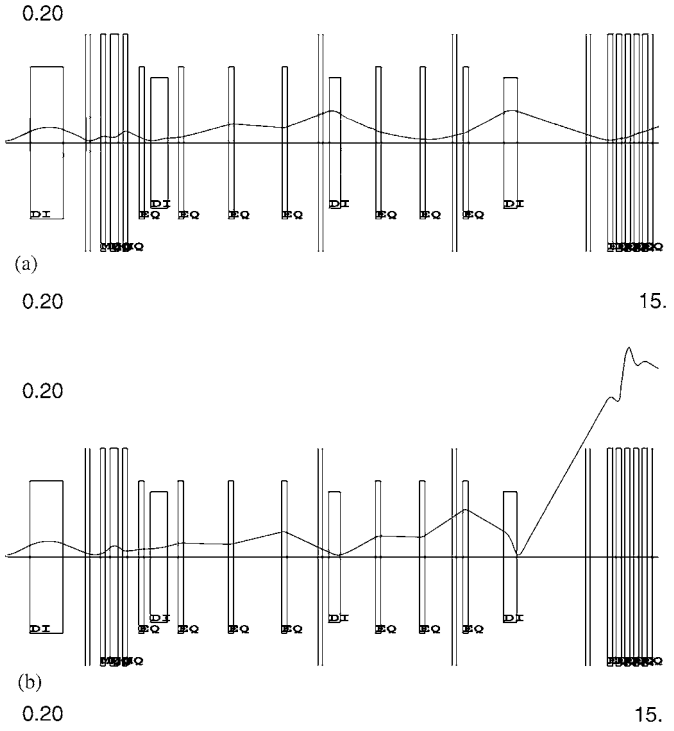



Fig. 1. Envelopes of the beam for the present beam line design calculated by the code COSY INFINITY in the horizontal (case *a*) and the vertical (case *b*) plane. The double vertical lines are the positions of the profile grids in the horizontal (first line) and the vertical (second line) plane.

From the Fig. 1b we can see that the beam is over-focused by the M72 bending magnet. Due to shortage of a vertical focused quadrupole in the long drift space after the M72 magnet the beam blew up in the vertical plane. The bending magnet M72 and the drift space between the M72 and the matching section are the main places of the beam loosing.

Before we present the new design of the low-energy beam line we will turn our attention to the matching section. The matching section consists of five electrostatic quadrupoles. Half of the aperture and the length of each of them are 5 and 12 cm, respectively. The distance between each of them is only 8 cm. The consequence of this geometry is that the fringe field effects in this region are very strong. One of the ways to decrease the fringe field effects in the matching

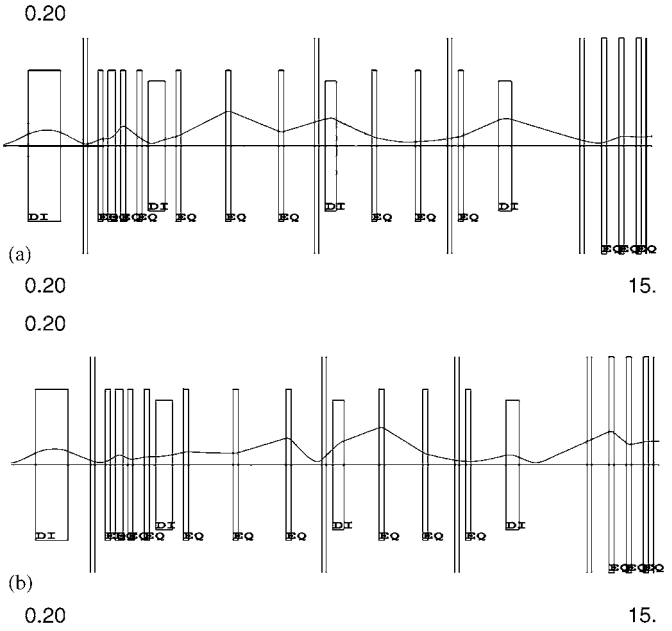



Fig. 2. Envelopes of the beam for the redesign beam line calculated by the code COSY INFINITY in the horizontal (case *a*) and the vertical (case *b*) plane.

section is to remove the electrostatic quadrupoles EQ9 and EQ11 (see Fig. 1 in the Ref. [1]). In the new design of the beam line the matching section is made by three electrostatic quadrupoles (EQ8, EQ10 and EQ12; see Fig. 1).

The envelope of the new (redesign) transport beam line is presented in Fig. 2a (horizontal plane) and Fig. 2b (vertical plane).

## References

[1] Paper P-071 in this Proceedings.
[2] M. Berz, Nucl. Instr. and Meth. A 298 (1990) 473.
[3] M. Berz, Proc. International Workshop on Nonlinear Problems in Accelerator Physics, Berlin, 1992.
[4] M. Berz, Proceedings of computation and accelelerator physics conference, 1993, AIP Conf. Proc. 297 (1994) 267.
[5] M. Berz, Physics of particle accelerators, 1990, AIP Conf. Proc. 249 (1991) 456.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 249–252
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Experimental characterization and numerical simulations of the electron source at PITZ

K. Abrahamyan[a], J. Bähr[a], J.P. Carneiro[b], K. Flöttmann[b], J.H. Han[a], M.v. Hartrott[c], M. Krasilnikov[a,*], D. Lipka[a], V. Miltchev[a], A. Oppelt[a], B. Petrossyan[a], D. Pose[a], D. Richter[c], S. Schreiber[b], L. Staykov[a], F. Stephan[a]

[a]*DESY, Zeuthen, Germany*
[b]*DESY, Hamburg, Germany*
[c]*BESSY GmbH, Berlin, Germany*

Available online 28 November 2005

**Abstract**

The Photo Injector Test facility at DESY Zeuthen (PITZ) works on the optimization of electron sources for Free Electron Lasers (FELs). The main motivation and challenge of PITZ is the production of electron beams with high charge and small normalized transverse emittance. Recent results of electron beam studies at PITZ are presented in this paper. The experimental characterization of the electron source is discussed together with beam dynamics simulations.


*PACS:* 41.75.Ht; 41.85.Ar

*Keywords:* Photo injector; Emittance

## 1. Introduction

The Photo Injector Test Facility at DESY Zeuthen (PITZ) has been built in order to test and optimize electron sources for Free Electron Lasers (FELs) and future linear colliders. PITZ is a facility that has the purpose to produce intense electron beams with minimum transverse emittance and short bunch length as required for FEL operation. The experimental set-up consists of a 1.5 cell L-band RF gun with a $Cs_2Te$ photocathode, a solenoid system for space charge compensation, a photocathode laser capable to generate long pulse trains with variable temporal and spatial pulse shape, and an extensive diagnostics section. The slit masks technique is used for the electron beam emittance measurements, the emittance measurements system (EMSY) including single and multislit masks is located at $z = 1.62$ m from the cathode. A schematic layout of PITZ is shown in Fig. 1.

The first stage of the PITZ project has been successfully completed at the end of 2003. A smooth commissioning procedure of the RF gun cavity yielded an operation with up to 900 μs long RF pulses at 10 Hz repetition rate and an accelerating gradient at the cathode of about 42 MV/m [1]. That corresponds to a maximum average power of 27 kW in the gun cavity with 0.9% duty cycle. This long RF pulse operation fulfills the TTF2 requirements. The startup conditions of TTF2 on normalized projected beam emittance [2] have been fulfilled. By optimization of the photocathode laser properties together with the RF field and solenoid parameters a minimum normalized projected emittance of $1.7\,\pi$ mm mrad has been achieved for a 1 nC electron beam at a mean longitudinal momentum of 4.7 MeV/$c$.

## 2. Photocathode laser

One of the key issues of a photo injector is the photocathode laser. The laser temporal and spatial profiles have a significant impact on the electron beam performance. In

*Corresponding author.
*E-mail address:* mikhail.krasilnikov@desy.de (M. Krasilnikov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.048

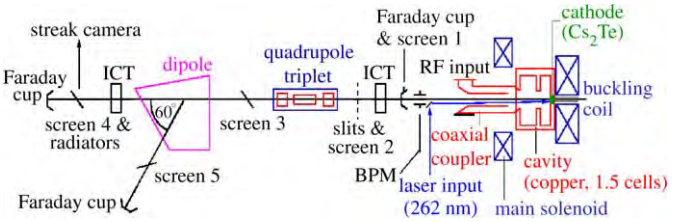



Fig. 1. Layout of PITZ.

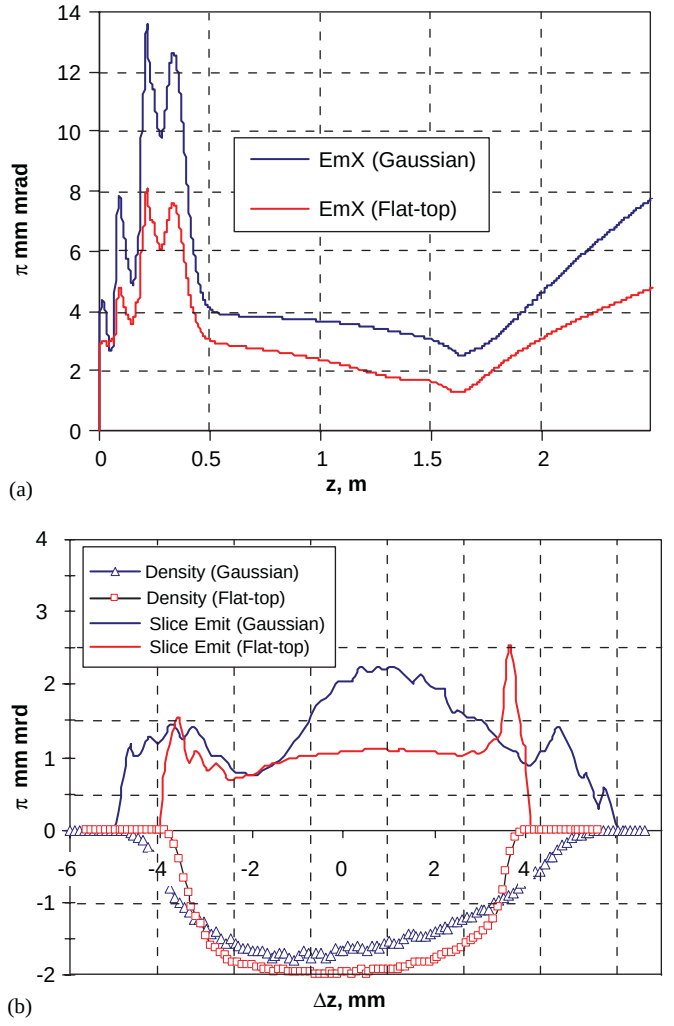



Fig. 2. (a) Projected normalized beam emittance along the beam line for two longitudinal profiles: Gaussian ($\sigma_t \approx 5.2$ ps, transverse profile: homogeneous, $\sigma_x = \sigma_y = 0.5$ mm) and flat-top (20 ps FWHM, 2 ps rise/fall time, transverse profile: homogeneous, $\sigma_x = \sigma_y = 0.75$ mm). (b) Slice emittance for both longitudinal laser profiles at the EMSY position ($z = 1.62$ m from the photocathode). Charge densities are plotted in arbitrary units on the negative axis.

order to reduce the space charge influence on the transverse beam emittance the longitudinal laser profile at PITZ has been modified from a Gaussian ($\sigma_t \approx 7$ ps) to a flat-top distribution (18–23 ps FWHM). ASTRA [3] beam dynamics simulations for both laser profiles show that the optimum emittance (minimum emittance at EMSY location $z = 1.62$ m) for the flat-top profile (20 ps FWHM, 2 ps rise/fall time) is a factor of ~2 smaller than for the best Gaussian profile ($\sigma_t \approx 5.2$ ps). The simulated normalized beam emittance as a function of longitudinal position along the beam line is shown in Fig. 2a. The slice beam emittance inside the electron bunch at $z = 1.62$ m is shown in Fig. 2b for both longitudinal laser profiles together with the corresponding charge density profiles.

## 3. Charge production study

Faraday Cups and integrating current transformers (ICT) have been used to measure the electron beam charge. A basic measurement is the so-called phase scan: the accelerated charge downstream of the gun is measured as a function of the launch phase, the relative phase of the laser pulses with respect to the RF. Phase scan studies (measurements and simulations) for the normal operating

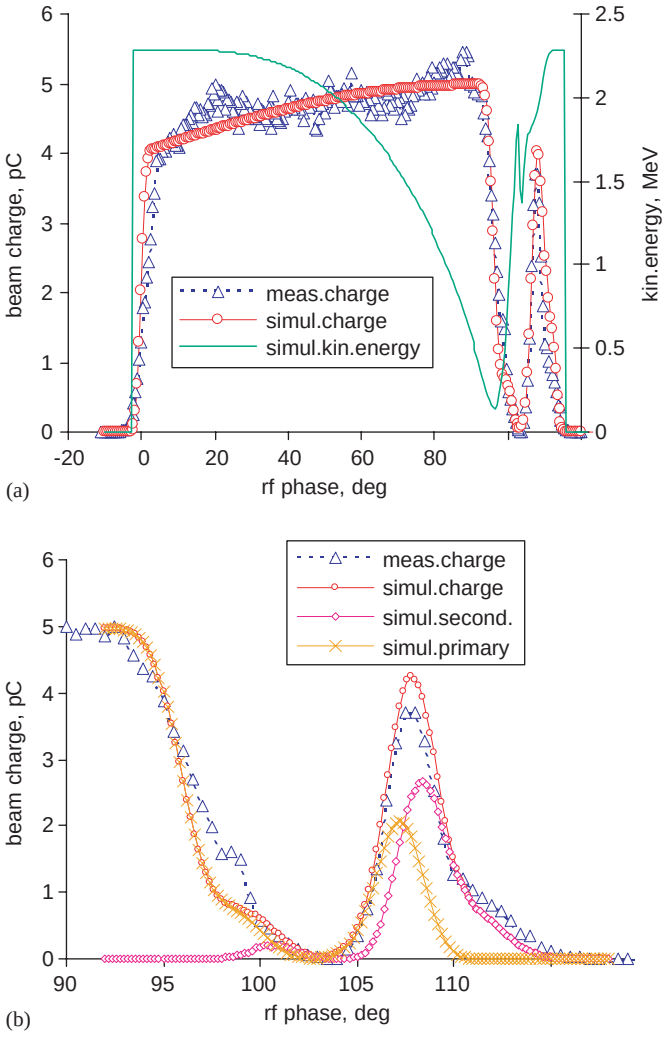



Fig. 3. (a) Measured and simulated phase scan (beam charge vs. RF phase). The simulated mean kinetic energy is plotted on the right axis. (b) Detailed phase scan for the "bump" phases. The simulated contributions of primary and secondary electrons (simulated) are shown in addition to their sum (simulated charge).

conditions (1 nC bunch charge and 42 MV/m accelerating field at the cathode) have been presented in Refs. [1,4]. For a detailed study of photo injection processes phase scans have been done also for low bunch charges (~4–5 pC) and moderate RF field on the photocathode (~21 MV/m). Corresponding simulations have been performed as well. Solenoids were off for this study. A typical phase scan for this case is shown in Fig. 3. In order to simulate the phase scan the Schottky effect—the dependence of the charge emission on the actual electric field (RF and space charge)—has been modeled by using the formula $Q[\mathrm{pC}] \propto (0.4 + 0.045 \cdot E_{\mathrm{RF+SC}}[\mathrm{MV/m}])$. Another charge production mechanism can be illustrated by a detailed study of the RF phase range ~100–115°. This corresponds to a "bump" in the phase scan. For these RF phases the beam acceleration takes place in a non-smooth way: besides the particles accelerated downstream there is an amount of particles reflected back to the cathode. Under some circumstances reflected particles can produce secondary electrons. Simulations show (Fig. 3b) that secondary electrons can explain at least partially these phase scan peculiarities. Furthermore, momentum measurements and simulations [5] show that the beam momentum distribution contains a secondary electron part.

The discrepancy between measurements and simulations can be explained by a non-perfect beam and beamline alignment during measurements ($Q^{\mathrm{measured}} < Q^{\mathrm{simulated}}$ due to electrons lost on the aperture), as well as by restrictions of the primary and secondary emission modeling ($Q^{\mathrm{measured}} > Q^{\mathrm{simulated}}$). Another probable reason of the discrepancy is the uncertainty in the RF field balance of the gun cavity (ratio $E^{\mathrm{Cathode}}/E^{\mathrm{FullCell}}$) that has been used in the simulations.

## 4. Emittance study at PITZ

Measurements of the transverse emittance were performed using a single-slit scan technique [1]. Beamlets from three slit positions were taken into account for the emittance calculation. Beamlet profiles were observed at screen 3, 1010 mm downstream of the single-slit mask (see Fig. 1).

Besides the photocathode laser parameter optimization other RF gun parameters have been optimized in order to reach a minimum normalized beam emittance, namely RF phase, main and bucking solenoid current. The emittance has been measured for different RF phases and main solenoid currents $\varepsilon_{xy}(\Phi_{\mathrm{RF}}, I_{\mathrm{main}})$ [6]. The photocathode laser with flat-top longitudinal profile (20 ps FWHM and ~5 ps rise/fall time) and transverse rms ($X/Y$) sizes 0.51/0.63 mm has been tuned in order to keep the beam charge at 1 nC. The applied RF gradient on the cathode corresponded to ~42 MV/m. During the two-dimensional ($\Phi_{\mathrm{RF}}, I_{\mathrm{main}}$) parameter scan the bucking solenoid was off, so the magnetic field at the cathode is supposed to be small but not zero. For the best point ($\Phi_{\mathrm{RF}} = -5°$, $I_{\mathrm{main}} = 305$ A) the bucking solenoid current was fine tuned in order to compensate the remnant longitudinal magnetic field on the cathode. The measured horizontal and vertical emittance as a function of the bucking solenoid current is shown in Fig. 4a. The asymmetry in the laser transverse profile leads to the observed beam shape and emittance asymmetry. Another asymmetry factor can be the impact of the vacuum mirror, used to position the laser beam onto the photocathode [7].

The normalized beam emittance has been simulated using ASTRA for injector parameters close to the ones observed during the emittance measurements, except the transverse laser profile which has been assumed rotationally symmetric (to apply a 2D $r-z$ space charge routine with equivalent rms sizes). The results of the simulations are shown in Fig. 4b. Simulated slice parameters of the optimum beam emittance ($I_{\mathrm{buck}} = 20$ A) are plotted in Fig. 5a,b. for $z = 1.62$ m (EMSY location).

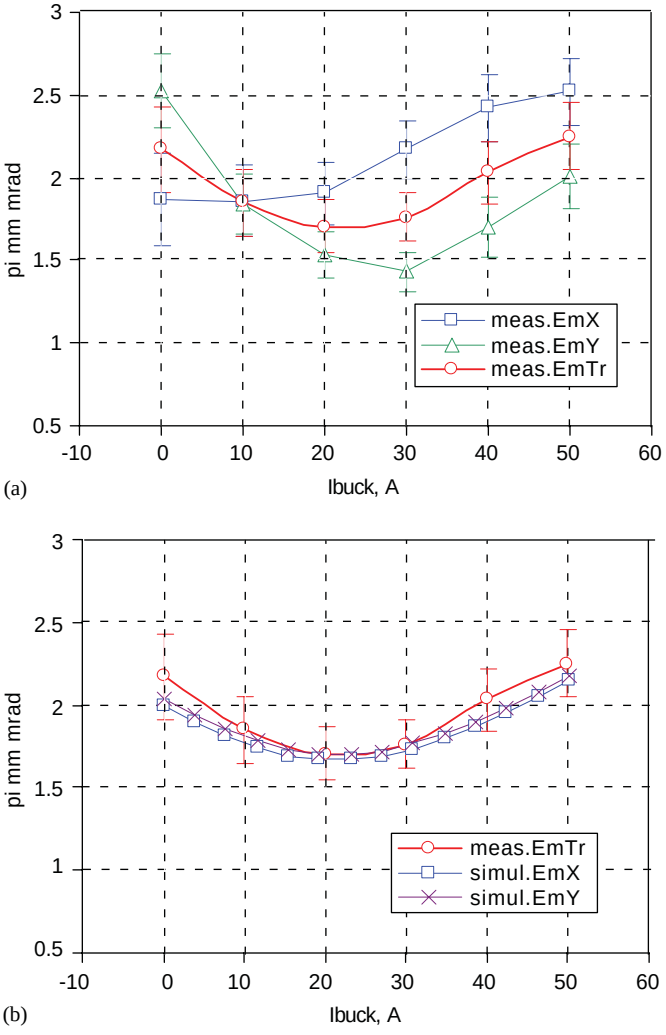
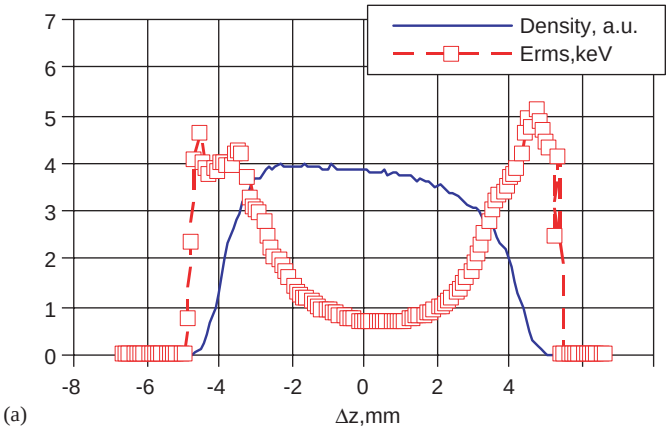
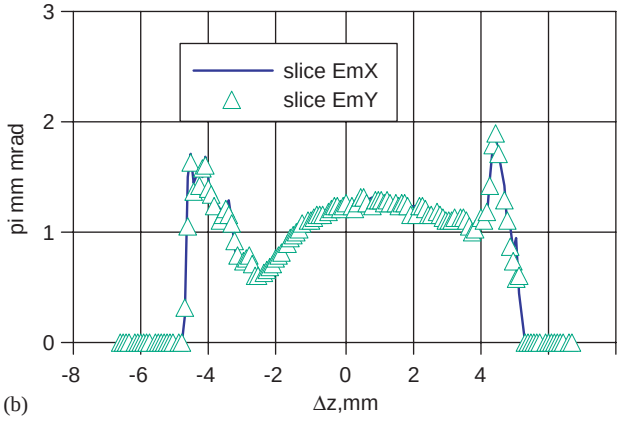



Fig. 4. (a) Measured normalized beam emittance as a function of the bucking solenoid current. (b) Measured transverse emittance ($\varepsilon_{\mathrm{tr}} = \sqrt{\varepsilon_x \cdot \varepsilon_y}$) and simulated horizontal and vertical emittance vs. bucking solenoid current.




(a)




(b)

Fig. 5. Slice parameters for the optimum beam at $z = 1.62$ m: (a) Charge density and slice energy spread. (b) Slice horizontal and vertical emittance. The projected transverse normalized beam emittance is $1.7\,\pi$ mm mrad.

## 5. Conclusions

The optimization of the electron source at the Photo Injector Test facility at DESY Zeuthen (PITZ) yielded a minimum normalized projected emittance of $1.5\,\pi$ mm mrad in one plane and of $1.7\,\pi$ mm mrad as a geometrical average of the horizontal and vertical emittances. This result has been obtained by extensive variation of the photo injector parameters within a wide range. A flat-top longitudinal photocathode laser profile has been used. Detailed experimental and numerical studies of the photo injector have been performed. As next step, a large extension of the facility by including a booster cavity for emittance conservation studies is foreseen as well as further improvements of the photocathode laser properties.

## References

[1] M. Krasilnikov, et al., Characterization of the electron source at the photo injector test facility at DESY Zeuthen, FEL 2003, Tsukuba, September 2003.
[2] SASE FEL at the TESLA Facility, Phase2, TESLA FEL Reports 2002-01.
[3] ASTRA User Manual, http://www.desy.de/~mpyflo/Astra_documentation/.
[4] M.v. Hartrott, et al., Experimental characterization of the electron source at the photo injector test facility at DESY Zeuthen, PAC 2003, Portland, May 2003.
[5] J.H. Han, M. Krasilnikov, K. Flöttmann, Secondary electron emission in a photocathode rf gun, Phys. Rev. ST Accel. Beams 8 (2005) 033501.
[6] M. Krasilnikov et al., Optimizing the PITZ electron source for the VUV-FE, in: Proceedings of the Ninth European Particle Accelerator Conference (EPAC2004), Lucerne, 2004, pp. 360–362.
[7] S. Setzer, et al., Inuence of beam tube obstacles on the emittance of the PITZ photoinjector, in: Proceedings of the Ninth European Particle Accelerator Conference (EPAC2004), Lucerne, 2004, pp. 1984–1986.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 253–255
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Field emission gun for X-ray tubes

A.S. Baturin*, A.I. Trufanov, N.N. Chadaev, E.P. Sheshin

*Moscow Institute of Physics and Technology, Institutskii per. 9, Dolgoprudny, Moscow Region 141700, Russia*

Available online 28 November 2005

**Abstract**

The attempt to replace hot cathode with field emission cathode in a commercial X-ray tube have been made. We are continuing investigation of emission properties of the carbon fibers. Cathodes for X-ray tube were made from bundle of carbon fibers. The prototype provides 0.1 mA with focus spot about ∅2 mm under continued 40 kV mode. Each tube contain cathode made of single fibers bundle. The matrix cathode to increase total tube current up to 10 mA/cm$^2$ for macro cathode structure is also considered.


*PACS:* 79.70.+q

*Keywords:* X-ray; Field emission; Carbon fiber

The main direction of portable X-ray equipment evolution is the downsizing, reliability improvement and power consumption decreasing. It requires improvement of X-ray tube as a key element of X-ray equipment. So it is necessary to enhance efficiency and downsize of X-ray tube. The performance of X-ray tube basically depends on electron gun construction. The requirements for the electron gun could be determined in the following way: the dimensions about several millimeters, high efficiency (more than 90%) with instant turning on and low power consumptions.

Mainly the commercial X-ray tubes have the electron gun with hot cathode. Still these tubes do not fulfill the requirements for portable X-ray equipment.

The electron gun with field emission cathode (FEC) is more promising for application in portable devices. It is possible to significantly decrease the electron gun dimensions due to the high current density of the field emission (FE). The FE is very effective phenomenon because it does not consume the additional energy in comparison with the thermal electron emission. FEC is instantly ready for operation.

Nowadays the field emission properties of various materials are investigated. Nevertheless majority of cathode materials require the vacuum condition better than $10^{-7}$ Torr for a high emission current stability and long lifetime [1].

The technology used for producing of X-ray tubes with hot cathode provides residual gas pressure about $10^{-6}$ Torr. The improvement of vacuum will essentially increase the production cost. Therefore it is reasonable to develop FECs, which are able to operate in vacuum $10^{-6}$ Torr. The carbon fibers are the most suitable material for this purpose [2]. The FECs made of carbon fibers (CF) posses high emission current stability and lifetime about 10 000 h in vacuum $10^{-6}$ Torr.

Due to the wide (about 120°) and not uniform angular distribution of emitted electrons [3,4] there are some difficulties in developing of electron gun with CF cathode.

In our laboratory we are considering various methods of cathodes producing from carbon fibers. The encapsulating bundle of fibers into glass capillary is the most promising methods [2]. However there are some fibers features, which makes difficulties in application of fibers bundle as FEC. The bundle of fibers is a multi tips system. The peripheral fibers of the bundle can deviate from main bundle under electrostatic load and change the emission characteristic of the cathode. The special forming process of cathodes by corona discharges under atmosphere condition was used. At work [5] the method of fibers sharpening is described. We used similar tools to make special profile of our

*Corresponding author. Tel.: +7 095 408 5944; fax: +7 095 409 9543.
*E-mail address:* baturin@lafeet.mipt.ru (A.S. Baturin).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.033

cathode and to minimize the influence of peripheral fibers (see Fig. 1). After forming process emission is more stable.

Cathode long test in vacuum chamber showed that the cathode material was evaporated during cathode operating. The surface near the cathode was being covered by film of cathode evaporation product. In some cases a leakage between cathode and modulator electrodes were observed. There are special vacuum gaps in the design of cathodes-modulator unit for X-ray tube to avoid leakages.

For our experiments we chose commercial X-ray tube with hot cathode of indirect heating (manufacturer Svetlana-X-ray Company, Russia—www.svetlana-x-ray.ru). The initial parameters of tube are: anode voltage—50 kV, anode current—0.3 mA, focus spot—∅ 3–4 mm. Fig. 2 shows the stem of the tube with field emission cathode-modulator unit.

The pilot lot of tubes were assembled and subjected standard pumping down process with training on a vacuum station.

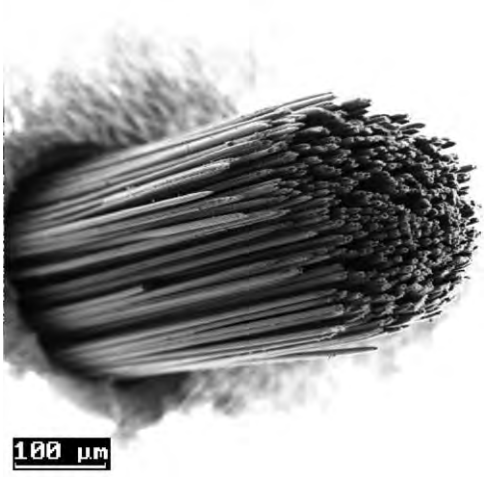



Fig. 1. Bundle of carbon fibers after corona discharge forming.

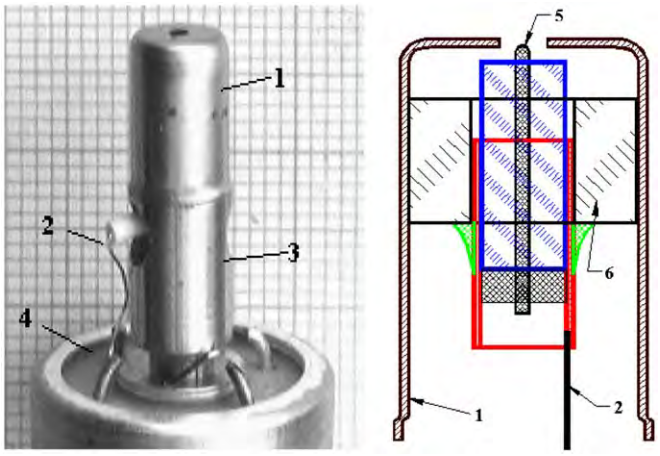



Fig. 2. X-ray tube stem with field emission cathode-modulator unit: 1—modulator, 2—cathode wire, 3—muff, 4—stem of tube, 5—carbon fibers, 6—glass insulator.

There was some problem with current leakage between pins of the tube stem. The control voltage for our cathode is up to 1.2 kV, but the stem pins provide only 1 kV isolation. We have to coat stem with special high voltage compound to increase stem insulatation of the tubes.

HV power supplies (+40 kV for anode and 2.5 kV for control voltage) of our own design were used to test tubes.

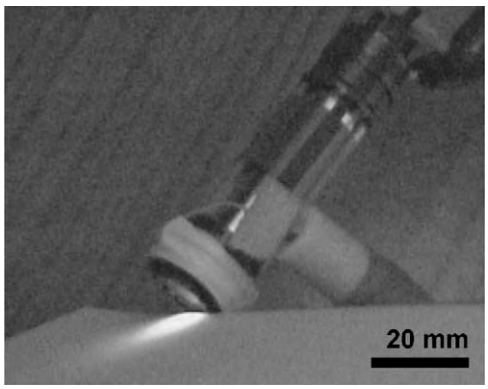



Fig. 3. X-ray tube with FEC is turned on in continuous mode with 40 kV and 0.1 mA on the anode.

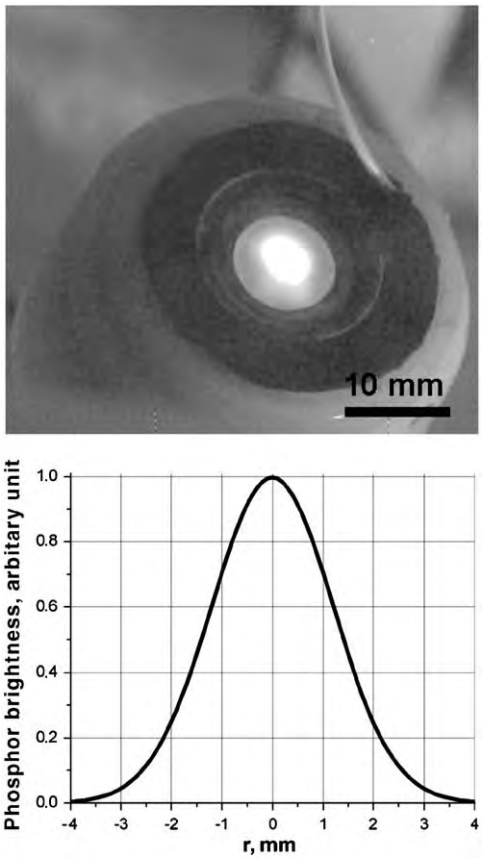



Fig. 4. The phosphor is covered on the tube anode to evaluate size and shape of the focal spot under various anode voltages and currents. The figure represents the tube under +40 kV and 0.1 mA on the anode.

The anode side and cathode side of tubes were coated with HV compound to test tubes in normal atmosphere. The intensifying X-ray screen was utilized to observe X-ray (see Fig. 3).

The beryllium exit window of tube was covered with phosphor to observe the shape and size of focal spot under various anode voltages and currents (see Fig. 4). The photos were made with mirror to prevent a camera from X-ray.

Some tubes were tested during 50 h without any degradation tendency in the cathodes emission characteristics. At present time we are planning to develop special measuring bench for long test of X-ray tubes with field emission cathodes.

The cathode described above can provide the tube current up to 0.25 mA. However the more convenience way is to utilize a matrix cathode made of bundles of carbon fibers. The packing density of such cathode is 100 bundles per $cm^2$ with total current 10 $mA/cm^2$ from macro cathode structure. This is comparable with nanotubes cathodes [6] for X-ray tubes. The expected lifetime of such structure is several thousand of hours.

The authors wish to thank the company "Svetala–Rentgen" for technical support of the experiments.

## References

[1] H. Sugie, M. Tanemura, et al., Appl. Phys. Lett. 78 (17) (2001) 2578.
[2] A. Baturin, A. Trufanov, et al., J. Vac. Sci. Tech. B 21 (1) (2003) 354.
[3] M. Mousa, Appl. Surf. Sci. 94/95 (1996) 129.
[4] A. Bobkov, et al., J. Vac. Sci. Tech. B 19 (1) (2001) 32.
[5] F. Baker, et al., J. Phys. D: Appl. Phys. 7 (1974).
[6] G. Yue, et al., Appl. Phys. Lett. 81 (2) (2002) 355.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 256–259
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# The influence of the ion bombardment on the current stability of field-emission cathodes of carbon nanotubes

Vitaly S. Bormashov*, Andrey S. Baturin, Konstantin N. Nikolskiy, Roman G. Tchesov, Evgeniy P. Sheshin

*Moscow Institute of Physics and Technology, Department of Vacuum Electronics, Institutskiy per. 9, 141700 Dolgoprudniy, Moscow Region, Russia*

Available online 1 December 2005

**Abstract**

The significant current degradation of the carbon nanotubes cathode was found in the presented work. This current instability came from the adsorption desorption processes on the cathode surface. During the long-time tests the periodical changing of the cathode electron work function was shown. The proposed model allowed to find the dependence of the electron work function changing with the time. This model has a good agreement with the experimentally obtained results.


*PACS:* 79.70.+q; 73.63.Fg; 79.20.Rf

*Keywords:* Field emission; Carbon nanotubes; Degradation model; Sorption processes

## 1. Introduction

One of the most promising applications of the carbon nanotubes (CNT) is the field electron emission [1]. The main advantage of the CNT in comparison with other carbon materials [2] is the significant value of the field amplification factor. One of the most important characteristics of the field emission cathodes and the devices based upon them is the long-term emission current stability. However, the inconsistent results were obtained during the long-time field emission tests of the CNT cathodes [3]. The drastic decrease of the emission current at once after the cathode is switching on occurs so the stable operation of the cathode during long time was obtained [3].

The saturation phenomenon was found in the work [4]. This effect consists of the changing of the current voltage characteristic Fowler–Nordheim behavior during the applied voltage increase. To clarify this phenomenon different models were suggested [5,6]. However, the more completed study of the effect was made in the work [7]. The authors associate this phenomenon with the sorption process, which occurs on the surface of the carbon nanotube. The results of different works show the influence of the sorption process on the long-term current stability [8,9].

At the experiment described below an interesting effect was observed during the investigation of the long-term current stability. Two different time periods were found characterized by different degradation ratio (the current voltage characteristic of the cathode shifts to the higher voltage region). In the presented work, the model was proposed to describe the observed effect by the mean of the sorption process and ion bombardment.

## 2. Field-emission cathode production

The screen-printing technique was applied to produce the field-emission cathodes. The carbon powder obtained in the arc discharge was mixed with the organic binder. The carbon power contained more than 80% of the multiwall carbon nanotubes. The got paste was drown through the mask to the glass substrate covered by the ITO layer. The area of the cathode was 0.125 $cm^2$. To remove the organic binder the cathode was baked at 450 °C for 10 min. The surface of the produced cathode is shown in Fig. 1.

*Corresponding author. Fax: +7(495)4099543.
*E-mail address:* vitaly@lafeet.mipt.ru (V.S. Bormashov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.012

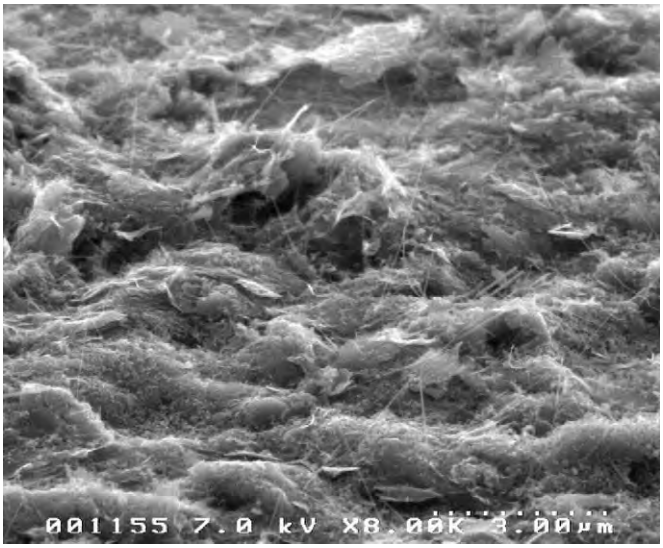



Fig. 1. The SEM image of field-emission cathode surface.

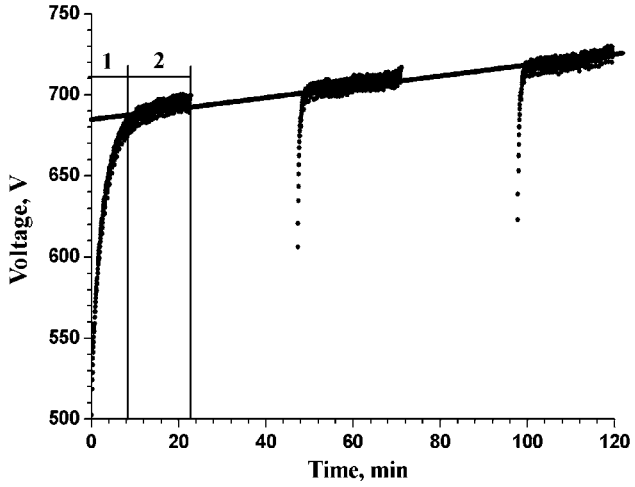



Fig. 2. Voltage versus time curve of the carbon nanotubes FEC. j-const. $j = 0.5\,mA/cm^2$.

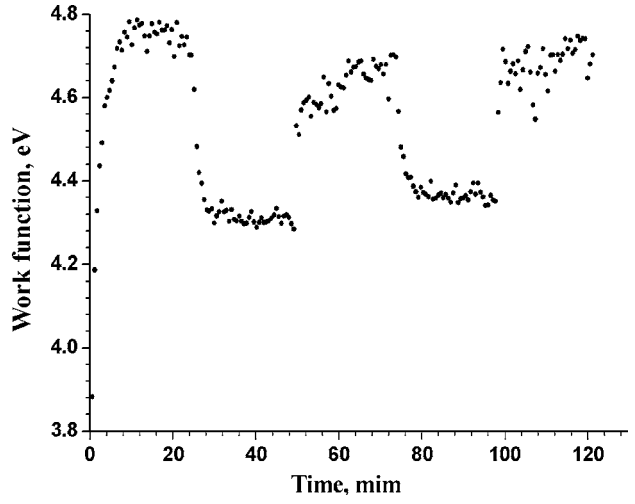



Fig. 3. Electron work function change during long-time field-emission tests.

## 3. Field emission tests

The diode tester was made of the field-emission cathode. The distance between anode and cathode was fixed by the spacer and this distance was 100 μm. Then the device was pumped and sealed off till the technical vacuum condition ($\sim 10^{-6}$ Torr). The long-term stability of the emission current was investigated at the current stabilization mode. The applied voltage versus time characteristics was measured. During the long-term tests the interesting effect was observed. It was the decreasing of the applied voltage after the cathodes switching off and following switching on (see Fig. 2). Furthermore the voltage–time characteristic has different time periods (Fig. 2, time period 1 and 2) characterized by a different degradation ratio.

The first time period is characterized by the significant changing of the current voltage characteristic. The second time period has slower changing of the characteristic. Such significant changing is attributed to the electron work function of the cathode excursion, which occurs due to the sputtering of the chemical adsorbed molecules by ion bombardment of the cathode surface. As it is known from different works [8,9] some types of the molecules (for example $H_2O$) adsorbing to the surface of the carbon nanotubes decrease the electron work function.

During the long-time test the current voltage characteristics were measured every 1 min. It is possible to estimate a field amplification factor of the cathode by making an assumption that the electron work function is equal to 4.7 eV at the end of the second time period (see Fig. 2). Then the values of the electron work function at the different times of the cathode operation were calculated (see Fig. 3). There is a periodical changing of the electron work function in the presented figure. It is attributed to formation of the adsorbed layer of the molecules during non-operating time period and to the destruction of the layer by ion bombardment during the cathode operation. The obtained results show that the ion bombardment greatly influences the stability of the emission current of the cathode based on carbon nanotube.

## 4. The model of the adsorbed layer formation

Let us consider the time period, which is characterized by the absence of the ion bombardment. For the average flow of the molecules per unit area in a unit of time the following expression can be written:

$$V_a\left[\frac{1}{cm^2 c}\right] : \frac{2.3 \times 10^{22} P\,(\mathrm{Torr})}{\sqrt{MT[\mathrm{K}]}}. \quad (1)$$

The covering efficiency of the surface Θ can be introduced as a ratio of the number of molecules $N$

covering the surface to the number of molecules $N_0$ forming the monolayer.

Using the accommodation coefficient $K_{ak}$, which has a decreasing value during the increasing covering efficiency [10], it is possible to write the expression to characterize the ratio increasing of the covering efficiency:

$$v_+ = K_{ak}(\Theta)\frac{V_a}{N_0}. \quad (2)$$

Let us consider two energy states of the adsorbed molecules. These states correspond to the physical and chemical adsorptions. The physical adsorption state has energy $U_m$. For a molecule to make a transition to the chemical adsorption state, it must get the energizing $Q_a$. Using the typical times of lifetime of the adsorbed molecules in these energy states and the typical time transition from the one energy state to another it is possible to write the expressions (3,4) characterizing the ratio changing of the covering efficiency.

$$\dot{\Theta}_1 = K_{ak}\frac{V_a}{N_0} - \frac{\Theta_1}{\tau_0}\left[\exp\left(-\frac{U_m}{kT}\right) + \exp\left(-\frac{Q_a}{kT}\right)\right] \quad (3)$$

$$\dot{\Theta}_2 = \frac{\Theta_1}{\tau_0}\exp\left(-\frac{Q_a}{kT}\right) - \frac{\Theta_2}{\tau_0}\exp\left(-\frac{U_a}{kT}\right) \quad (4)$$

where $\dot{\Theta}_1$ is the ratio changing of the covering efficiency by the physical adsorbed molecules, $\tau_0$ the period of the thermal oscillation ($\tau_0 = h/kT$), $\dot{\Theta}_2$ the ratio changing of the covering efficiency by the chemical adsorbed molecules, and $U_a$ the energy in the chemical adsorbed state.

To estimate the number of ions, which are generated in the anode–cathode spacing in a unit of time, it is possible to apply the following expression [11]:

$$N_i = \frac{I_e}{e}\frac{pL}{kT}\sigma \quad (5)$$

where $I_e$ is the electron emission current, $k$ the Boltzmann constant, $L$ the anode–cathode distance, $T$ the temperature, $e$ the electron charge, $p$ the pressure, and $\sigma$ the ionization cross-section.

Let us introduce the effective cathode area of the ion bombardment $S_e$. In general this value is not equal to the total area of the cathode. The ion current can be given by the expression

$$j_i = \frac{N_i}{S_e} = \frac{I_e}{e}\frac{pL}{kT}\frac{\sigma}{S_e}. \quad (6)$$

We can introduce the spattering coefficients of the chemical and physical adsorbed molecules $D_1$ and $D_2$, correspondently. Expressions (3) and (4) can be written in the following forms when the ion bombardment occurs:

$$\dot{\Theta}_1 = K_{ak}\frac{V_a}{N_0} - \frac{\Theta_1}{\tau_0}\left[\exp\left(-\frac{U_m}{kT}\right) + \exp\left(-\frac{Q_a}{kT}\right)\right] - j_i\frac{D_1}{N_0}\Theta_1 \quad (7)$$

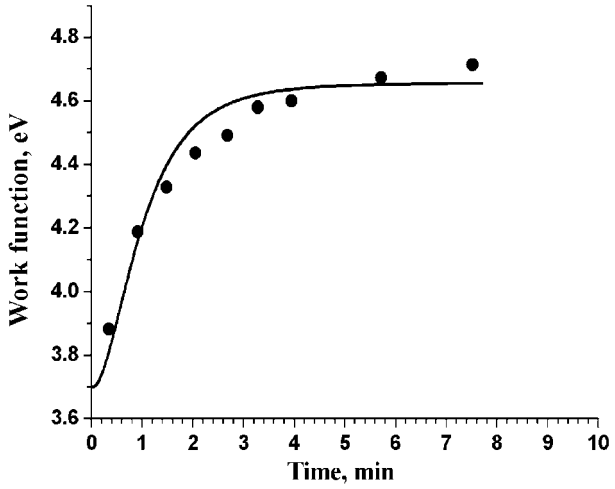



Fig. 4. The electron work function changing during the field-emission cathode operation, the solid line—the simulation results; the black points—the value of the electron work function measured by the mean of the FN-charts.

$$\dot{\Theta}_2 = \frac{\Theta_1}{\tau_0}\exp\left(-\frac{Q_a}{kT}\right) - \frac{\Theta_2}{\tau_0}\exp\left(-\frac{U_a}{kT}\right) - j_i\frac{D_2}{N_0}\Theta_2. \quad (8)$$

Using the numerous works [12], which describe the results of the investigation of the adsorption process of the active metal on the field emission cathodes surface, one can expect that the first layer of the adsorbed atoms result in the significant changing of the electron work function. The electron work function changing depending on the covering efficiency of the surface by the chemical adsorbed molecules can be given by the empirically obtained formula:

$$\varphi(\Theta) = (\varphi_0 - \varphi_{min})\exp\left(-\frac{\Theta}{\Theta_{opt}}\right) + \varphi_{min} \quad (9)$$

where $\varphi_0 = 4.7\,\text{eV}$ is the electron work function of the carbon nanotubes, $\varphi_{min}$ the minimum experimentally obtained value of electron work function of the carbon nanotubes.

Using the proposed model the numerical simulation of the electron work function changing by ion bombardment was made. Having set a value of the effective spattering coefficients, the energy of the chemical adsorbed state of the molecules [10] the dependence of the electron work function upon time was calculated (see Fig. 4).

Fig. 4 shows that the proposed model has a good agreement with the experimental data, which were obtained during the field emission tests of the cathode of carbon nanotubes.

## 5. Conclusions

During the experiment it was found that the ion bombardment influences the long-term current stability of the field-emission cathodes of carbon nanotubes. The significant shift of the current voltage characteristic to the higher voltage region occurred during the cathode

operation under technical vacuum conditions. This behavior of the characteristic took place due to the distraction of the layer formed by adsorbed molecules, which leads to the electron work function increase. The physical model was proposed to describe the adsorption desorption processes of the residual gases on the cathode surface. This model takes into account the ion bombardment destructing the adsorbed layer. The simulation results have a good agreement with the experimental data. Using the proposed model it is possible to optimize the operation regimes of the carbon nanotubes field-emission cathode. For example, one can select the most optimal composition of the residual gases, which allow to decrease the electron work function. Choice of the density of electron emission current will not allow to destruct the adsorbed layer significantly. The considered model makes it possible to predict the long-term behavior of the field-emission cathode.

## References

[1] J.M. Bonard, J.P. Salvetat, T. Stockli, L. Forro, A. Chatelain, Appl. Phys. A 69 (1997) 245.
[2] E.P. Sheshin, Ultramicroscopy 79 (1999) 101.
[3] J.M. Bonard, et al., Solid-State Electr. 45 (2001) 893.
[4] P.G. Collins, A. Zettl, Phys. Rev. B 55 (1997) 9391.
[5] J.M. Bonard, F. Maier, T. Stockli, A. Chatelain, et al., Ultramicroscopy 73 (1998) 7.
[6] X. Xu, G.R. Brandes, Appl. Phys. Lett. 74 (1999) 2549.
[7] K.A. Dean, B.R. Chalamala, Appl. Phys. Lett. 76 (2000) 375.
[8] A. Wadhawan, R.E. Stallcup, J.M. Perez, Appl. Phys. Lett. 78 (2001) 108.
[9] A. Akwani, A. Wadhawan, R.E. Stallcup, K.F. Stephens, J.M. Perez, Appl. Phys. Lett. 79 (2001) 1867.
[10] M. Grujicic, G. Cao, B. Gersten, Appl. Surf. Sci. 206 (2003) 167.
[11] M.S. Mousa, T.F. Kelly, E.P. Sheshin, et al., Abstract Book 47th IFES, 2001, EP 20.
[12] E.W. Muller, J. Appl. Phys. 26 (1955) 732.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 260–264
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# High energy electron bunch generation by using a plasma separator

S. Miyazaki*, K. Miyauchi, K. Sakai, T. Kikuchi, S. Kawata

*Department of Material Science and Engineering, Utsunomiya University, Utsunomiya 321-8585, Japan*

Available online 28 November 2005

**Abstract**

In this paper, we propose an electron acceleration by using an intense short pulse laser and a thin slab plasma separator. When an intense short pulse laser illuminates electrons in vacuum, the electrons are accelerated by the ponderomotive force at the front of laser pulse and the electrons accelerated loose their energies at the tail of laser pulse. This is one of serious problems in the ponderomotive electron acceleration. In order to suppress the energy loss, we propose an overdense plasma separator to extract the electrons before entering the deceleration region of laser pulse. In our electron acceleration mechanism, only the laser is reflected by the overdense plasma separator and the electrons pass through the plasma separator. Consequently, the electrons can obtain a net energy after the interaction with the laser pulse.


*PACS:* 41.75.Jv; 29.27Fh

*Keywords:* Electron acceleration; Pondermotive force; Plasma separator

## 1. Introduction

In recent years, many particle acceleration schemes have been proposed by an invention of the Chirped Pulse Amplification (CPA) technique [1,2], such as electron acceleration in vacuum [3–13], high energy ion production [14–20] and so on. The ponderomotive electron acceleration is one of the typical mechanisms in an electron acceleration in vacuum, and it has been investigated well in previous researches [7–11]. However, the ponderomotive electron acceleration has some problems. One is an electron scattering in the transverse direction by the transverse ponderomotive force. A $TEM_{00}$ mode laser, which is generally employed, has an intensity peak at the laser axis. Therefore, the transverse ponderomotive force scatters electrons in transverse. The transverse electron scattering is suppressed successfully by using a $TEM_{10} + TEM_{01}$ mode laser, which has been investigated in our previous research [11].

Another problem is a phase difference between the electrons and the laser. When an intense short pulse laser illuminates the electrons in vacuum, the electrons are accelerated by the ponderomotive force at the front of laser pulse. However, the electrons accelerated loose their energies at the laser pulse tail. Even if the laser is strongly focused, the energy loss appears. Most simple idea of a suppression of the electron deceleration is a separation of the electrons from the laser pulse before entering the laser deceleration region [12,13]. In this paper, we focused on the electron separation from the laser deceleration region by using an overdense plasma separator. The electrons are accelerated by the ponderomotive force in the head of laser pulse, and then the laser pulse and electrons hit the overdense plasma. The laser is reflected by the plasma and only the electrons pass through the plasma. After the laser reflected, the laser does not influence the electrons and a high energy electron bunch is generated. In our calculations, we employ 2.5-dimensional particle-in-cell (PIC) simulations, and the relativistic equation of motion is used.

In Section 2, the calculation model and parameters are presented. Section 3 shows the simulation results of the

*Corresponding author.
*E-mail address:* dt030106@cc.utsunomiya-u.ac.jp (S. Miyazaki).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.013

electrons acceleration by using the plasma separator. The conclusions are described in Section 4.

## 2. Simulation model

Fig. 1 shows a schematic view of 2.5-dimensional PIC simulations. The laser propagates in the $x$ direction and is polarized in the $y$ direction. A focused $TEM_{00}$ mode laser is adopted in our calculations. The longitudinal and transverse electric fields induced at the left boundary of the simulation box are written by

$$E_x = \frac{i}{k}\frac{\partial E_y}{\partial y}, \tag{1}$$

$$E_y = E_0 \frac{w_0}{w(x)} \exp\left\{-\left[\frac{(y-y_c)}{w(x)}\right]^2 - \left(\frac{\eta}{\tau/2}\right)^2\right\} \times \sin\left\{\left[k\eta - \phi(x) + \frac{ky^2}{2R(x)}\right]\right\}. \tag{2}$$

Here $k$ is the laser wave number, $E_0$ is the amplitude of electric field of laser, $w_0$ is the radius of laser, $w(x) = w_0\sqrt{1+[(x-x_f)/x_R]^2}$, $\eta = x - x_f - ct$, $x_f$ is the laser focal point, $x_R$ is the Rayleigh length, $y_c$ is the center of the laser in the transverse direction, $\tau$ is the pulse length of laser, $\phi(x) = \tan^{-1}[(x-x_f)/x_R]$, and $R(x) = x[1+[x_R/(x-x_f)]^2]$. In our simulations, $y_c = 20\lambda$, where $\lambda$ is the laser wavelength. The laser enters from the left boundary ($x = 0$), and then the laser propagation is solved by the Maxwell equation. To solve the motion of particles the relativistic equation of motion $d\boldsymbol{P}/dt = q(\boldsymbol{E} + \boldsymbol{V} \times \boldsymbol{B})$ is employed, where $\boldsymbol{P}$ is the particle momentum, $q$ is the particle charge, and $\boldsymbol{E}$ and $\boldsymbol{B}$ are the electric fields and magnetic fields, respectively.

In our study, the laser wavelength $\lambda = 1.053\,\mu m$, $w_0 = 3\lambda$, the laser intensity is $2.67 \times 10^{19}\,W/cm^2$ ($a_0 \sim 4.65$), the pulse duration $\tau = 9\lambda$ ($\sim 30\,fs$), the focal point of laser $x_f = 15\lambda$, the initial electrons velocity in the $x$ direction is $0.1c$, and the initial density of electrons injected is $1.0 \times 10^{14}\,cm^{-3}$. Here $a_0$ is the dimensionless parameter of the laser field, $e$ and $m_e$ are the electron charge and mass, $\omega$ is the laser angular frequency, and $c$ is the speed of light in vacuum. The laser and the electrons injected propagate in the $x$ direction, and the center of the laser and electron bunch coincide at the focal point $x_f$. The electron bunch is located in $6\lambda \leqslant x \leqslant 18\lambda$ initially, and the transverse size is $9\lambda$. The separator plasma consists of a proton and electron. At the initial time, the separator plasma charge is neutral, and the number density $n_p = 3n_c$. Here $n_c$ is the critical density. The thickness of plasma is $2\lambda$, and is located in $35\lambda \leqslant x \leqslant 37\lambda$. The electron bunch and plasma are in the Maxwell distributions with the temperature of the electron bunch $T_e = 1\,eV$ and the plasma temperature $T_p = 100\,eV$ initially.

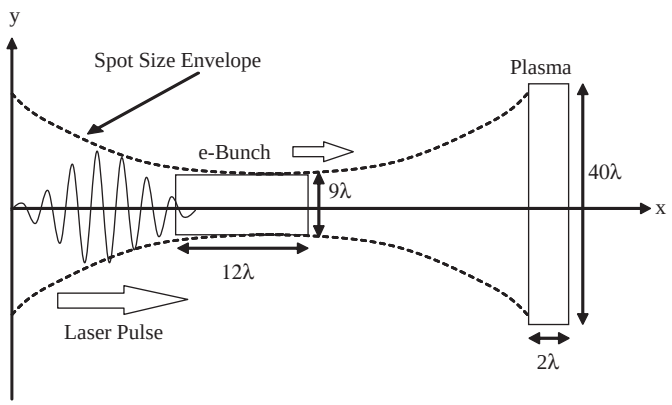



Fig. 1. A schematic view of the simulation model.

The computational domain is given by $R = \{(x,y)|0<x<80\lambda, 0<y<40\lambda\}$. The mesh size in the $x$ and $y$ directions $\Delta x = \Delta y = 0.05$, and the number of super particle in one mesh is 16.

## 3. Simulation results

In this section, first we show simulation results of the ponderomotive electron acceleration in the case without the plasma separator. Fig. 2 presents time developments of the distributions of electron kinetic energy in the $x-y$ plane at the times of $t = 112, 169$, and $253\,fs$. The center of laser and electron bunch injected reach to the focal point at $t = 82.7\,fs$. The electrons, which are located near the central axis, are accelerated strongly by the ponderomotive force, and their kinetic energies reach about 9.27 MeV at $t = 169\,fs$. However, the electrons are decelerated and scattered in the transverse direction at the tail of laser pulse. Consequently, the electron bunch accelerated is scattered in transverse. Fig. 3 shows a history of the relativistic factor $\gamma$ of typical electron in the $x$ axis. The electrons are trapped by the laser field, and after the passage of the electrons through the laser focal point, the electrons are scattered gradually in the transverse direction with the expansion of laser in transverse as shown in Fig. 2. At the same time the electron energy $\gamma$ also decreases at the laser tail deceleration. The final energy of typical electron is about $\gamma_f = 10.5$ (see Fig. 3), and it agrees well with the estimation of final electron energy expressed by $\gamma_f \sim a_0^2/2 = 10.9$ [9]. Such the acceleration is the typical ponderomotive electron one.

To extract the high energy electrons before entering the laser deceleration region, the plasma separator is employed in this paper. Distributions of the kinetic energy of electrons are presented in Fig. 4. The parameters of the laser and the initial electron bunch are same as those in the case without the plasma separator. Before the electrons and laser hit the plasma, the distribution of electron energy is same with that in Fig. 2a. When the electrons and the laser reach to the plasma separator, only the laser is reflected by the plasma separator, and the electrons accelerated pass through the plasma separator. Consequently, the electron bunch accelerated keeps their energy, and the transverse divergence of electrons accelerated is suppressed. Fig. 5

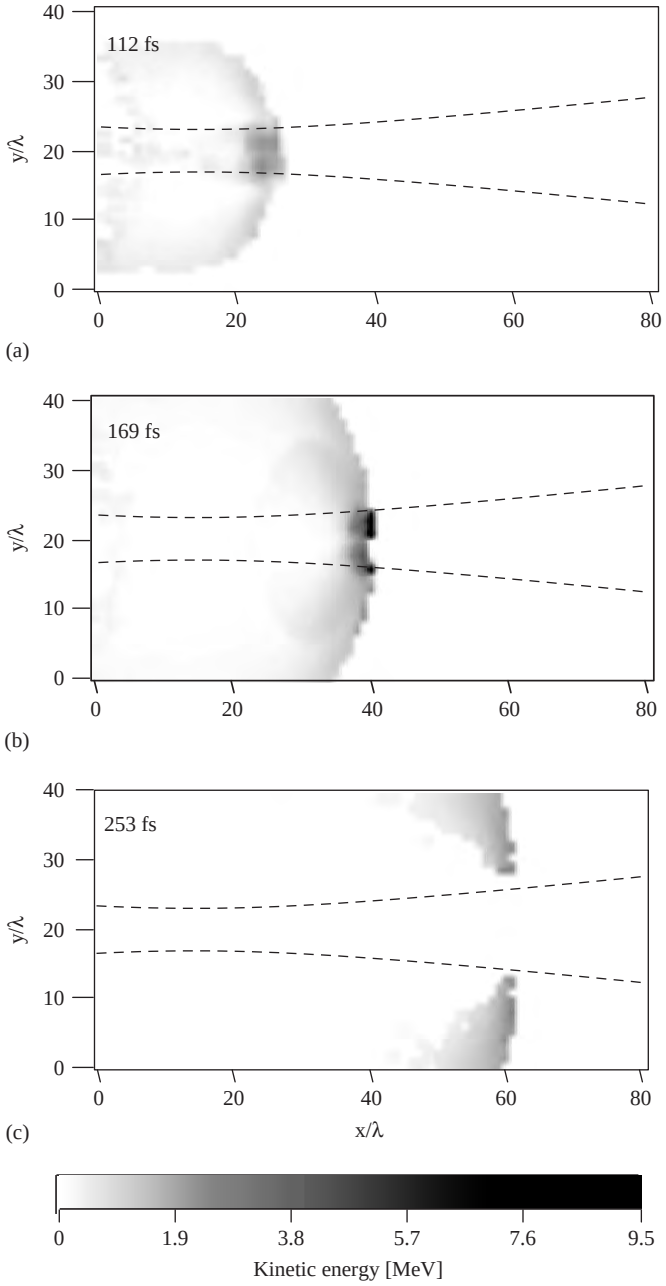



Fig. 2. Distributions of electron kinetic energy in the case without the plasma separator at (a) $t = 112$ fs, (b) 169 fs, and (c) 253 fs. The dashed line shows the spot size envelope of the laser.

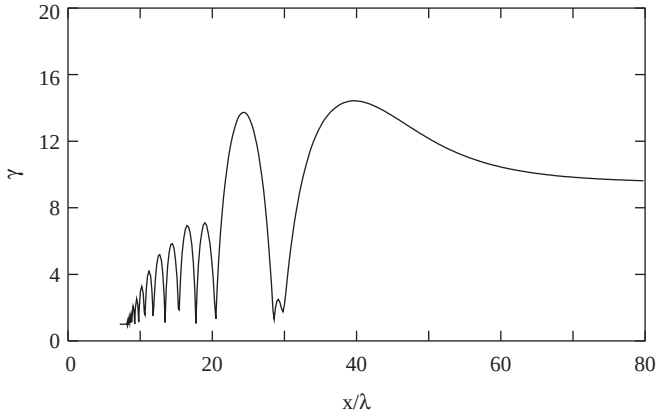



Fig. 3. A history of the relativistic factor $\gamma$ of the electron accelerated versus $x$-axis in the case without the plasma separator.

presents an energy $\gamma$ history of electron accelerated. In the above description, the electrons can obtain a net energy after the passage of the electrons through the plasma separator. The final energy $\gamma$ of the electron accelerated reaches about $\gamma = 17.7$. The electron slightly loses the energy at the left-side surface of the plasma separator as shown in Fig. 5. This is caused by the laser reflected. After the laser is reflected, the electron accelerated is not influenced by the transverse ponderomotive force of the laser field, and consequently the high energy electron bunch locates near the central axis as shown in Figs. 4b and c. The averaged energy of the electron bunch accelerated is about $\gamma = 3.92$, and the rms emittances in the transverse and longitudinal directions are about 0.527 mm mrad and 0.697 mm mrad, respectively. Figs. 6a and b show the averaged kinetic energy and the rms emittances of the high energy electron bunch with the different laser intensities and the pulse lengths. In these cases, the total laser energy is kept constant when the pulse length changes. From Fig. 6, the averaged electron energy becomes large and the transverse and longitudinal rms emittances decrease with the increase in the laser intensity. Because the ponderomotive force is proportional to the laser intensity, at the low intensity region the electrons are not accelerated sufficiently by the laser, and are scattered in the transverse direction. At the laser intensity of $1.60\times 10^{19}$ W/cm$^2$, the averaged energy of the electron bunch is about $\gamma = 2.27$.

Fig. 7 presents an influence of the location of plasma separator to the electron bunch accelerated. Fig. 7a shows the averaged kinetic energy of electron bunch accelerated, and the transverse and longitudinal rms emittances are presented in Fig. 7b. By setting the plasma separator leftward from the optimal position, the electrons are separated by the plasma separator before the sufficient acceleration. Therefore, the electron energy is slightly decreased. When the plasma separator is placed right from the optimal position, the low energy electrons are scattered in transverse before entering the plasma separator, and only the high energy electrons pass through the plasma separator. Consequently, the electron bunch with the low divergence is produced as shown in Fig. 7b. However, the averaged energy is slightly decreases because the electrons accelerated enter to the laser deceleration region. From Fig. 7, the difference of the electron energy and the rms emittances are relatively insensitive to the position of the plasma separator.

The maximum growth rate of the two-stream instability and the filamentation instability between the electron bunch accelerated and the separator electrons

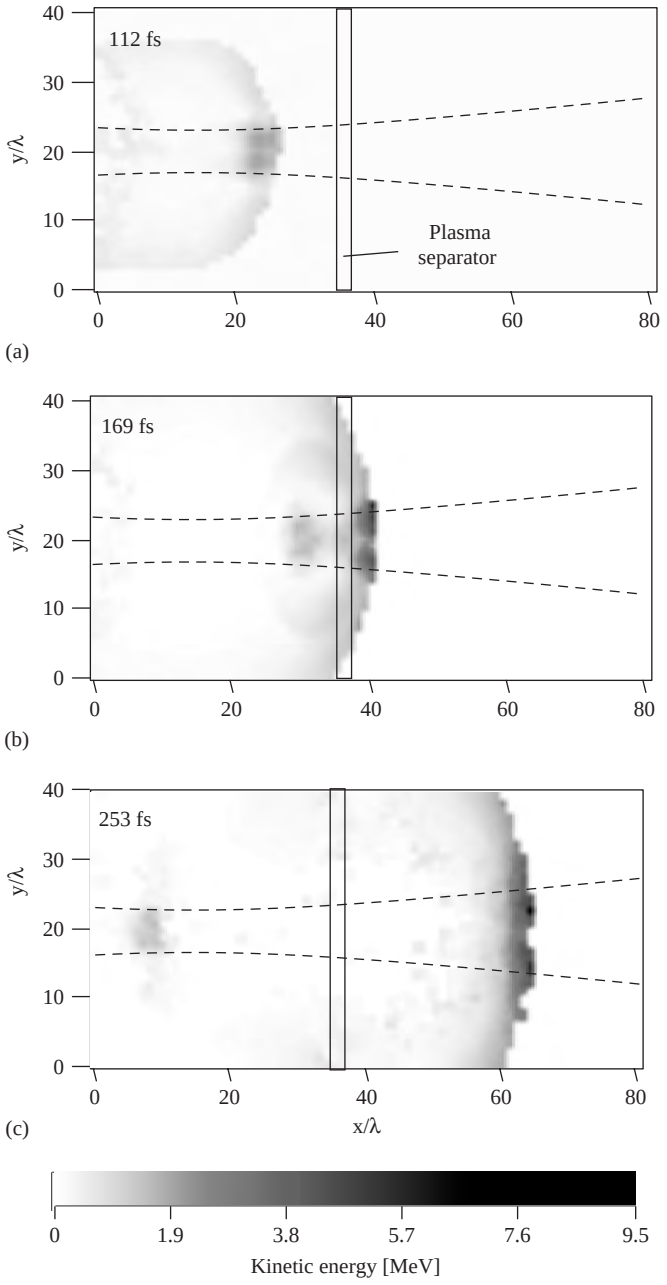



Fig. 4. Distributions of the electron kinetic energy in the case with the plasma separator at the times of (a) $t = 112$ fs, (b) 169 fs, and (c) 253 fs.

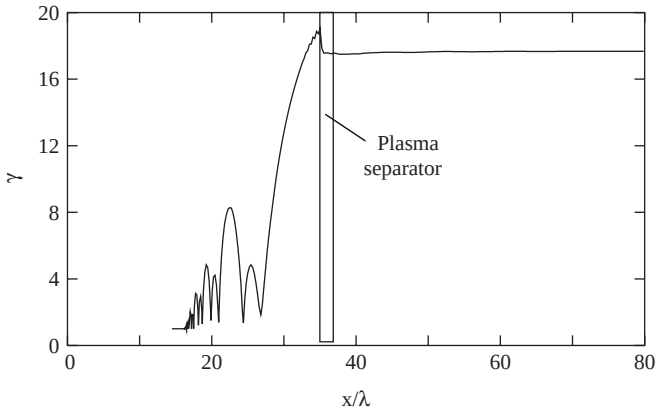



Fig. 5. An energy $\gamma$ change of the electron accelerated as a function of $x$ in the case with the plasma separator.

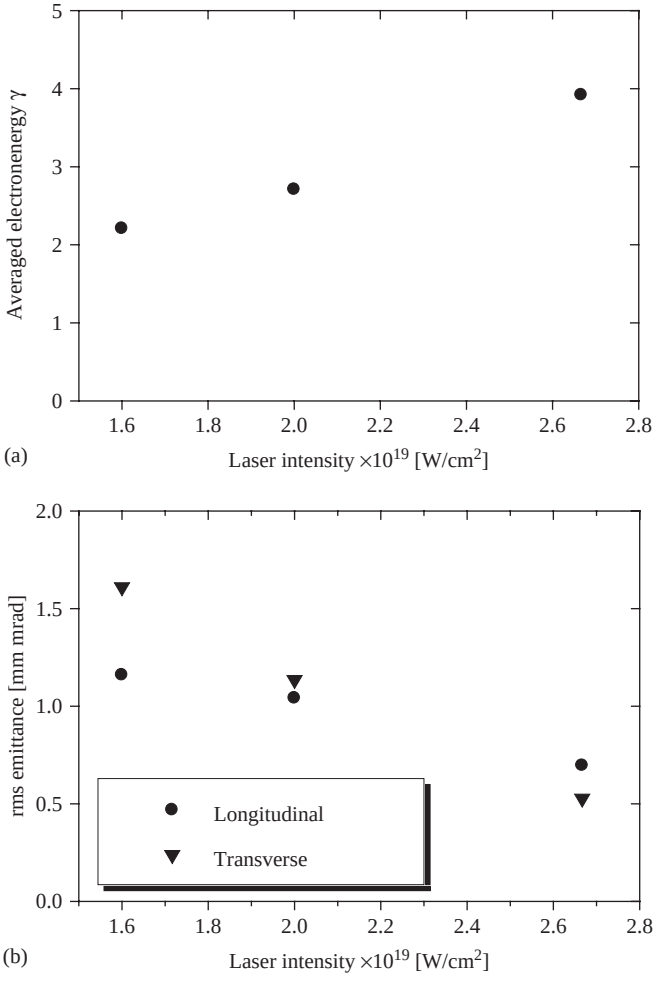



Fig. 6. Influences of the laser intensity to the electrons accelerated. (a) The averaged energy of electrons accelerated, and (b) the transverse and longitudinal rms emittances. Here the total laser energy is kept constant, when the laser intensity changes. The other parameters are same as those in the case without the plasma separator.

are estimated below. The two-stream instability is written by [21]

$$\delta_{\max} = -\frac{\nu_b}{2} + \sqrt{\frac{\pi}{2}}\frac{\omega_b^2}{\omega_e}\frac{V_b^2}{u_b^2}\exp\left(-\frac{1}{2}\right). \quad (3)$$

Here $\nu_b$ is the collision frequency between the beam electron and the plasma electron, $\omega_b$ is the electron bunch plasma frequency, $\omega_e$ is the electron frequency of the plasma separator, $V_b$ is the electron bunch velocity, and $u_b$ is the thermal velocity of the electron bunch. In our case the passage time of the plasma separator is about $\tau_{ele} = 7.28$ fs. The maximum growth rate of the two-stream instability is estimated by $\delta_{\max} \sim 1.98 \times 10^{13}$ 1/s, that is, stable. The filamentation instability is

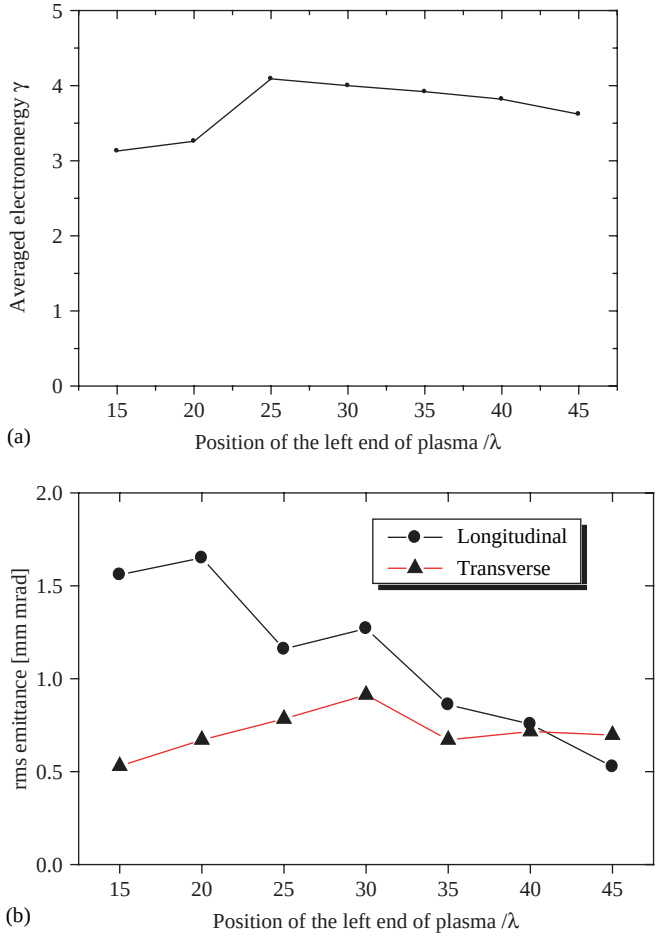



Fig. 7. Relations between the plasma position and the electron bunch accelerated. The left end of plasma is changed by $15\lambda$, $20\lambda$, $25\lambda$, $30\lambda$, $35\lambda$, $40\lambda$, and $45\lambda$. The other parameters are same as those in Fig. 4. The averaged energy of electron bunch accelerated is shown in (a), and (b) presents the rms emittances in the transverse and longitudinal directions.

estimated by [22]

$$\delta_{\max} \sim v_e \beta / 3, \tag{4}$$

at $k = (2w_b^2/9v_e^2)^{1/4}(\omega_e/c)$. Here $\beta = \omega_b^2/\omega_e^2 + k^2c^2/\omega_e^2$, and $v_e$ is the collision frequency of the electron–electron in plasma. This equation is derived under the conditions of $|(\omega_c + iv_e)/(ku_e)| > 1$, $|\omega_c/(ku_e)| > 1$, and $v_e > \delta$, where $\omega_c$ is the complex wave frequency, and $u_e$ is the thermal velocity of plasma electron. These conditions are fulfilled in our case. The maximum growth rate of the filamentation instability is about $\delta_{\max}\tau_{ele} \sim 9.90 \times 10^{-4} \ll 1$. If the thickness of the plasma separator is not too thick, it serves a stable plasma separator. The electron energy loss by the collisions [23] is also estimated. The parameters we employed the collisional energy loss is 1.31 eV. Therefore the energy loss by the collision is negligible.

## 4. Conclusions

In summary, we demonstrated the successful acceleration and separation of the electrons by using the slab plasma separator. In this paper, the plasma separator separates the electrons accelerated before entering the laser deceleration regions, and the transverse divergence induced by the transverse ponderomotive force of $TEM_{00}$ mode laser is also suppressed. For the laser intensity of $2.67 \times 10^{19}$ W/cm$^2$ ($a_0 \sim 4.65$), the averaged energy of the electron bunch accelerated reaches about $\gamma = 3.92$, and the transverse and longitudinal rms emittances are about 0.527 and 0.697 mm mrad, respectively. In this paper, the relation between the position of the plasma separator and the electron bunch accelerated is also investigated. The influence of the location of the plasma separator to the electron acceleration is relatively weak. The plasma separator proposed in this paper could be simply used in the electron acceleration in vacuum.

## Acknowledgements

The authors would like to extend their thanks to Profs. K. Mima, K. Tachibana, K. Nakajima, S. Kurokawa, N. Yugami, J. Limpouch and A. A. Andreev for their valuable discussions and suggestions on this subject. This work was partly supported by JSPS (Japan Society for the Promotion of Science) and MEXT.

## References

[1] D. Strickland, G. Mourou, Opt. Commun. 56 (1985) 219.
[2] G. Mourou, C.P.J. Barty, et al., Phys. Today 51 (1998) 22.
[3] S. Kawata, T. Maruyama, et al., Phys. Rev. Lett. 66 (1991) 2072.
[4] Q. Kong, S. Miyazaki, et al., Phys. Rev. E 69 (2004) 056502.
[5] P.X. Wang, Y.K. Ho, et al., Phys. Lett. 78 (2001) 2253.
[6] G. Malka, J.L. Miquel, Phys. Rev. Lett. 77 (1996) 75.
[7] G. Malka, E. Lefebvre, et al., Phys. Rev. Lett. 78 (1997) 3314.
[8] J.L. Chaloupka, D.D. Meyerhofer, Phys. Rev. Lett. 83 (1999) 4538.
[9] G.V. Stupakov, M.S. Zolotorev, Phys. Rev. Lett. 86 (2001) 5274.
[10] S. Miyazaki, Q. Kong, et al., J. Phys. D 36 (2003) 2878.
[11] Q. Kong, S. Miyazaki, et al., Phys. Rev. E 69 (2004) 056502.
[12] V. Vshivkov, N. Naumova, et al., Nucl. Instr. and Meth. A 410 (1998) 493.
[13] W. Yu, V. Bychenko, et al., Phys. Rev. Lett. 85 (2000) 570.
[14] S.V. Bulanov, F. Califano, et al., Rev. Plasma Phys. 22 (2001) 227.
[15] T. Esirkepov, M. Borghesi, et al., Phys. Rev. Lett. 92 (2004) 175003.
[16] M. Allen, Y. Sentoku, et al., Phys. Plasma 10 (2003) 3283.
[17] M. Borghesi, A.J. Mackinnon, et al., Phys. Rev. Lett. 92 (2004) 055003.
[18] T. Nakamura, S. Kawata, Phys. Rev. E 67 (2003) 026403.
[19] K. Matsukado, T. Esirkepov, et al., Phys. Rev. Lett. 91 (2003) 215001.
[20] T.E. Cowan, J. Fuchs, et al., Phys. Rev. Lett. 92 (2004) 204801.
[21] M.N. Rosenbluth, et al., Handbook of Plasma Physics, Volume 1, Basic Plasma Physics I, North-Holland, Amsterdam, Netherlands, 1983, p. 540.
[22] T. Okada, K. Niu, J. Phys. Soc. Jpn. 50 (1981) 3845.
[23] J.D. Huda, NRL PLASMA FORMULARY, Naval Research Laboratory, Washington, 1998, p. 31.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 265–270
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Suppression of high-energy proton beam divergence in laser–foil interaction

S. Miyazaki[a,*], R. Sonobe[a], T. Kikuchi[a], S. Kawata[a], A.A. Andreev[b]

[a]*Department of Material Science and Engineering, Utsunomiya University, Utsunomiya 321-8585, Japan*
[b]*Scientific Center, S. I. Vavilov State Optical Institute, St. Petersburg, Russia*

Available online 22 November 2005

**Abstract**

A suppression of transverse proton divergence by a localization of high-energy electron bunch generated by an interaction between an intense laser and a foil target is investigated in this paper. When an intense laser illuminates a foil target, electrons are accelerated longitudinally by the ponderomotive force and produce a strong magnetic field. The electrons are confined by the magnetic field and an electric field is generated locally. Protons are extracted and mainly accelerated longitudinally by the local electric field. Consequently, the transverse divergence of high-energy protons is successfully suppressed.


*PACS:* 52.38Kd; 52.65Rr; 41.75.Jv

*Keywords:* Laser–foil interaction; Electron localization; High-energy proton

## 1. Introduction

In recent years, we have seen remarkable progress in the laser intensity improvement [1,2], and various researches have been made on laser–matter interaction, for example, electron acceleration in vacuum [3–7], high-energy ion bunch generation [8–17] and so on. When an intense laser illuminates a thin slab foil, electrons are accelerated by the ponderomotive force of the laser and oscillate around the foil target. The accelerated electrons produce an electric field near both the sides of the target (see for example, Ref. [14]) and accelerate ions. Such a high-energy ion production has been well-investigated experimentally and theoretically. In the future, an improvement in a quality of accelerated ions becomes very important, and researches in the progress of the beam quality are just stated [16,17]. In this paper, we focused on suppression mechanism of the high-energy proton divergence in a laser–foil interaction by using 2.5-dimensional particle-in-cell (PIC) simulations.

In the laser–foil interactions, the spread of the protons is strongly related to the behavior of the electrons. Recently, it has been proposed in Ref. [16] to use the underdense plasma for protons acceleration. In the mechanism proposed in Ref. [16] underdense plasma electrons are accelerated by the ponderomotive force, and the accelerated electrons produce a strong magnetic field. Such a magnetic field confines the electrons and produces an electric field in a long time, and the proton energy is higher than in the case of overdense plasma in the same parameter range.

In our study, we employ the ultra-intense laser and a solid target to suppress the transverse divergence of the accelerated protons. When the laser intensity is high enough, even in the case of the solid target the electrons are accelerated strongly in the direction of laser propagation, and a strong magnetic field, which has a transverse confinement effect on the electrons, is generated near the target. The electrons are localized in transverse and longitudinal directions near the opposite side of the irradiation surface of laser, and a static electric potential is formed locally. The localization of static electric potential produces a transverse electric field toward the

*Corresponding author.
*E-mail address:* dt030106@cc.utsunomiya-u.ac.jp (S. Miyazaki).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.014

laser axis. The protons are extracted and accelerated by the transverse and longitudinal electric fields. Therefore, even in the slab solid target one can expect the suppression of fast protons divergence.

In order to investigate the transverse suppression effect of the protons divergence, four different combinations of the laser intensity and pulse duration with a same laser energy are calculated in this paper.

In our suppression mechanism, the behavior of the electrons influences strongly the protons beam divergence. In the calculation results presented in this paper, the maximum proton energy is almost same for all the parameter sets. However, even if the laser energy is same, a transverse spread of the electrons accelerated by the ponderomotive force becomes large with a reduction of the laser intensity, and consequently the protons also diverge. This is caused by the difference in the distribution of the magnetic field generated by the fast electrons accelerated by the laser. The simulation model and parameters are presented in Section 2. In Section 3, the simulation results of the suppression processes of the transverse protons divergence are presented. In Section 4, concludes the paper.

## 2. Simulation model

In our study, we employ 2.5-dimensional particle-in-cell (PIC) simulations. Fig. 1 shows the schematic view of simulation model. The foil target consists of a $H^+$ layer of $1.5\lambda$ thickness with an additional $1.0\lambda$ linearly changing density gradient and the peak density is a solid density. Here $\lambda$ is the laser wavelength. At initial time, the temperatures of protons and electrons are in the Maxwell distribution with 1 keV. The mass ratio of the proton and electron $m_i/m_e$ is 1836. The laser propagates in the $x$ direction and is linearly polarized in the $y$ direction with the Gaussian profile in the transverse and longitudinal directions. The laser wavelength $\lambda = 1.053\,\mu m$ and the laser diameter $r_s = 4\lambda$. The center of laser in the transverse direction $y_c = 11\lambda$. The simulations are carried out with the four different combinations for the laser intensity $I$ and pulse duration $\tau$ as shown in Fig. 1. In this paper, the target side illuminated by the laser is called the laser side and other side is called the rear side.

The computational domain is given by $R = \{(x, y)|0 < x < 50\lambda, 0 < y < 22\lambda\}$. At the initial time ($t = 0$), the foil target is located in $23\lambda < x < 24.5\lambda$. The mesh width in the $x$ and $y$ directions $\Delta x = \Delta y = 0.05\lambda$, the computational time step $\Delta t = 0.04\Delta x$, and the total number of super particle is equal to $3.0 \times 10^5$ for the electrons and protons, respectively.

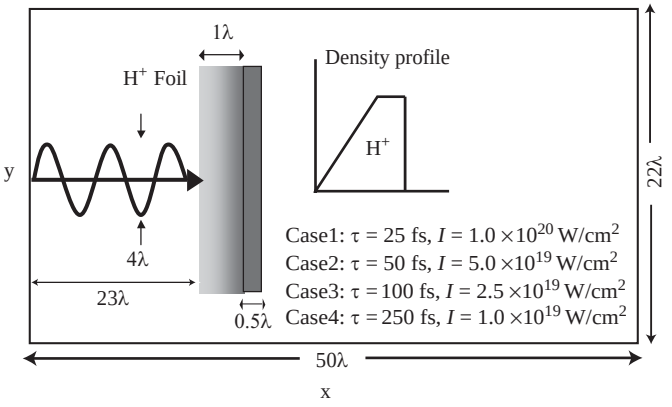



Fig. 1. A schematic view of the simulation model. The foil target consists of protons and electrons and the charge of foil target is neutral initially. The thickness of foil target is $1.5\lambda$ with an additional $1\lambda$ linear density slope.

## 3. Simulation results

The simulation results for the suppression of the transverse proton spread are shown in this section. First we investigate the magnetic field distributions in the cases of the parameter sets of Cases 2 and 4.

Fig. 2(a) shows the distribution of high kinetic energy ($>1$ MeV) electrons in the $x-y$ plane at the $t = 140$ fs for

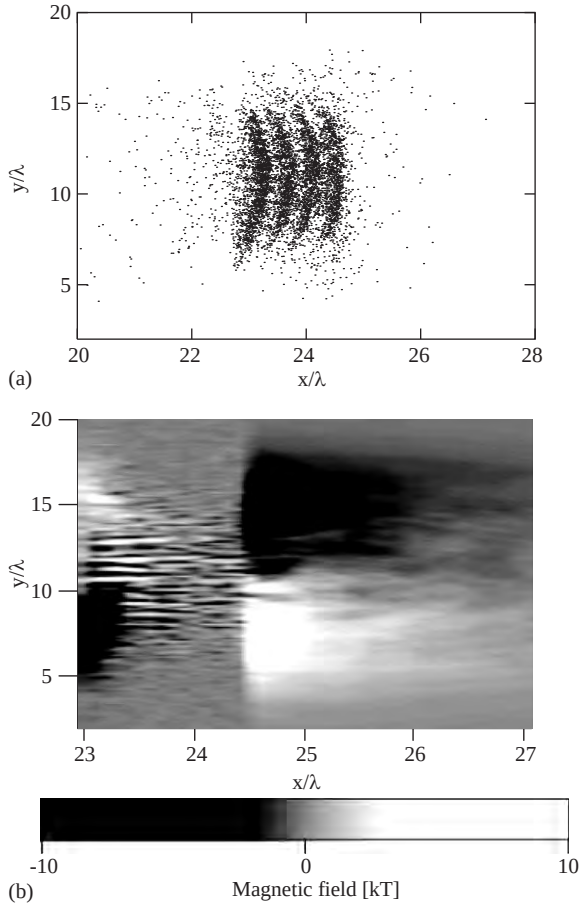



Fig. 2. (a) The kinetic energy distribution of electrons accelerated ($>1$ MeV), and (b) the magnetic field $B_z$ distribution in the $x-y$ plane at the time of $t = 140$ fs for Case 2. The intensity peak of laser reaches the left side of foil at time $t = 130$ fs.

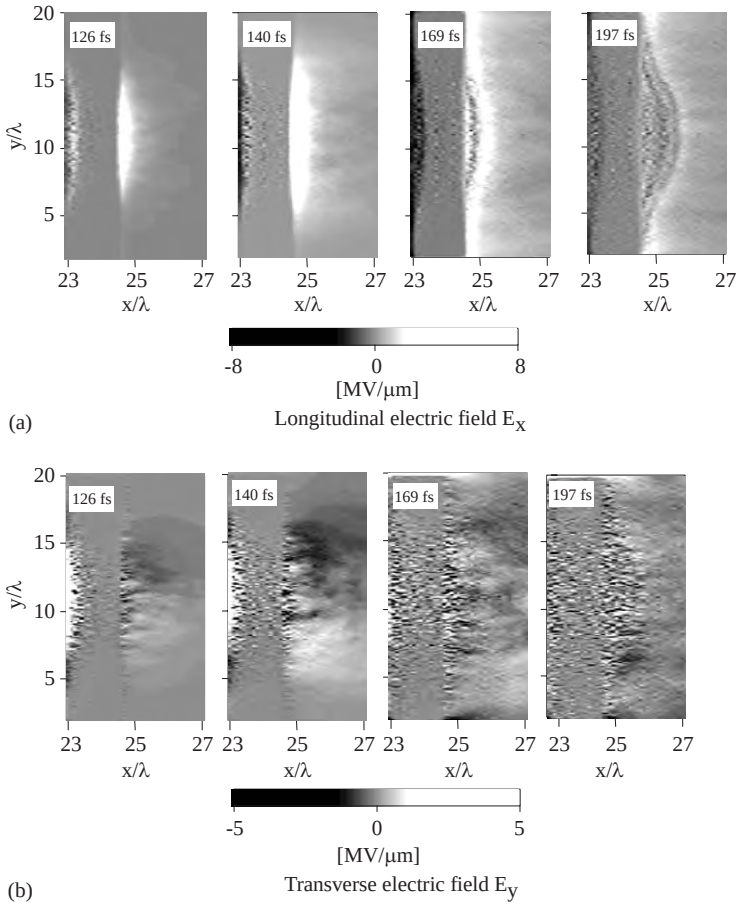



Fig. 3. (a) The snapshots of the longitudinal and (b) the transverse electric fields in the $x-y$ plane at the $t = 126, 140, 169$ and 197 fs for Case 2.

Case 2, and the magnetic field $B_z$ distribution in the $x-y$ plane at the time of $t = 140$ fs is presented in Fig. 2(b). In Case 2, the peak of the laser pulse reaches the left end of foil surface in 130 fs. Because the laser intensity is high ($I = 5 \times 10^{19}$ W/cm$^2$) and the pulse duration is short ($\tau = 50$ fs), the electrons are rapidly accelerated by the ponderomotive force in the longitudinal direction and the high-energy electron bunches are produced clearly as shown in Fig. 2(a). The maximum kinetic energy of electrons is about 6.89 MeV. The distances of the electron bunch accelerated correspond to the half laser wavelength $\lambda/2$. As seen from Fig. 2(b), such electron bunches generate strong magnetic field $B_z$, and the electrons are confined by this magnetic field. The maximum amplitude of the magnetic field $B_z$ reaches about $10^4$ T, and a strong magnetic field is formed in wide area.

Fig. 3 presents the distributions of the longitudinal and transverse electric fields $E_x$ and $E_y$ in the $x-y$ plane at the times of $t = 126, 140, 169$ and 197 fs. The magnetic field confines the electrons strongly at the rear side, and a static electric potential is formed locally near the right side of target surface. At this time, most protons are not emitted from the foil target because of the difference in mass ratio of electron and proton. Consequently, a local electric field $E_x$ is formed by the localization of the electrons at the rear side, and the transverse electric field $E_y$ is also formed toward the center in transverse as shown in Fig. 3(b). Although the amplitude of the longitudinal electric field $E_x$ is slightly higher than the transverse electric field $E_y$, both $E_x$ and $E_y$ are about few MV/μm. After that, the protons are mainly accelerated longitudinally by the longitudinal electric field $E_x$. The potential formed at the front of the foil target is negated by the protons and the transverse electric field $E_y$ also vanishes immediately. However, the transverse divergence of the protons is suppressed by this transverse electric field $E_y$ generated by the localization of electrons.

Fig. 4 shows the high kinetic energy (>1 MeV) electron distribution and the magnetic field distribution $B_z$ in the $x-y$ plane at $t = 337$ fs for Case 4. Here the laser intensity peak arrives to the left side of the target surface at the time of 330 fs. In Case 4, because of the low intensity ($I = 1 \times 10^{19}$ W/cm$^2$) and the long pulse ($\tau = 250$ fs) laser, the electrons are spread gradually in the transverse and longitudinal directions, receiving the laser energy. The maximum kinetic energy of electron is about 6.00 MeV, and it is nearly same for Case 2. However, the fast electron bunches are not formed clearly in the foil by the electron

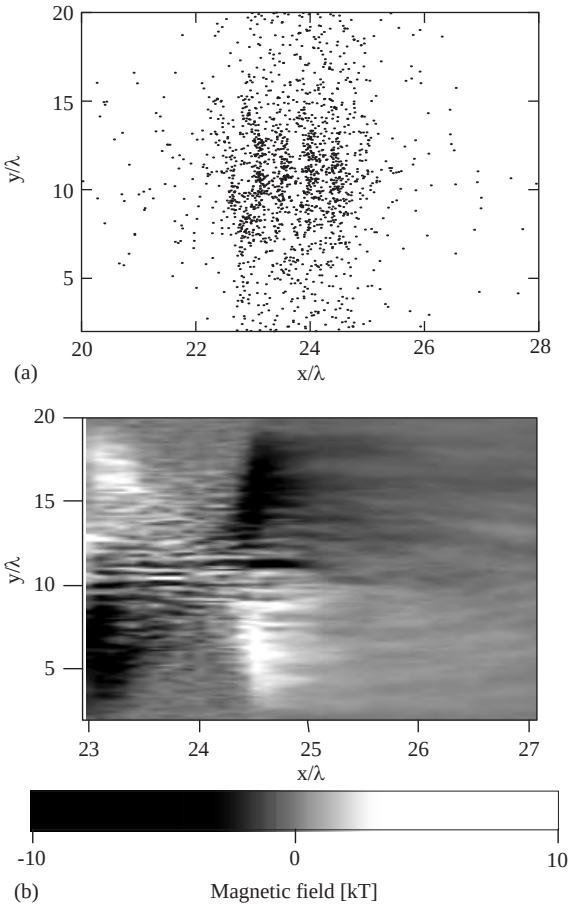



Fig. 4. The distributions of (a) the high kinetic energy (>1.0 MeV) electrons and (b) the magnetic field $B_z$ distribution in the $x-y$ plane at time $t = 337$ fs for Case 4. The intensity peak of laser arrives at the left side of foil at $t = 330$ fs.

divergence as shown in Fig. 4(a). The maximum amplitude of the magnetic field $B_z$ is about $6.71 \times 10^3$ T, and the areas of strong magnetic field are narrow compared to that in Case 2. Therefore, the electron confinement effect of the magnetic field $B_z$ is also weak and the electrons spread slowly in the transverse and longitudinal directions like a thermal expansion from the target.

Fig. 5 shows the longitudinal and transverse electric fields, $E_x$ and $E_y$, distributions in the $x-y$ plane at $t =$ 295, 337, 379 and 421 fs in the parameter sets of Case 4. The electrons are slowly emitted from the target and the localization of the static electric potential is not generated. Therefore, the transverse electric field $E_y$ generated by the localization of electrons, which is seen in Case 2, is not generated, and the protons are accelerated by the thin layer of electric field formed by electron expansion. Consequently, the protons spread in the transverse direction with electron divergence.

The beam intensity distributions of accelerated protons and the phase space $V_x - V_y$ distributions of protons accelerated for Cases 2 and 4 are shown in Fig. 6. The protons are mainly accelerated in the $x$ direction and produce the high intensity regions at the rear side in both the cases. The maximum intensity of protons accelerated for Cases 2 and 4 are $2.68 \times 10^{14}$ W/cm$^2$ and $2.94 \times 10^{14}$ W/cm$^2$, respectively, and the particle number of the proton bunch accelerated for Cases 2 and 4 are about $2.09 \times 10^{12}$/μm and $2.45 \times 10^{12}$/μm, respectively. As a result, the number of protons accelerated is nearly same for Cases 2 and 4. However, the remarkable difference appears in the spread of the transverse velocity of the protons accelerated. Although the accelerated protons are gathered around $V_y = 0$ in Case 2, the protons of Case 4 are diffused. This is caused by the difference in the distributions of the transverse electric field $E_y$ generated by the electron localization as shown in Figs. 3(b) and 5(b).

Fig. 7(a) shows the transverse and longitudinal rms emittances for Cases 1–4. The longitudinal rms emittance is nearly same in all cases. However, the transverse rms emittance becomes large as the laser pulse duration becomes long. The maximum kinetic energy of proton for all cases is presented in Fig. 7(b). Since the laser energy is fixed, the maximum kinetic energy of proton is almost same.

## 4. Conclusions

In this paper, we investigated the suppression effect of the transverse proton divergence in the laser–foil interaction by using 2.5-dimensional PIC simulations. We kept the laser energy constant and calculated different laser intensities and pulse durations. In the short pulse and high intensity laser, the electrons are accelerated longitudinally to the order of MeV by the laser and produce a localized static electric potential quickly at the rear side. At the same time the strong magnetic field confines the electrons in transverse. As a result, the transverse electric field appears toward the direction of the laser axis, and the spread of the transverse direction of the protons is suppressed. For the low intensity and the long pulse laser, the electrons expand in the transverse and longitudinal directions and accelerate the protons gradually. At this case, the suppression effect of the transverse electric field does not appear and the protons expand transversely. At the laser intensity of $I = 5 \times 10^{19}$ W/cm$^2$ and the pulse duration of $\tau = 50$ fs, the transverse and longitudinal rms emittances are about 0.119 mm mrad and 0.417 mm mrad, respectively, the maximum kinetic energy of proton reaches about 8.12 MeV, and the averaged kinetic energy is about 3.35 MeV. By using the long pulse duration and low intensity laser, although the energy of the particle hardly changed compared with the high intensity and short pulse laser, the transverse rms emittance becomes large.

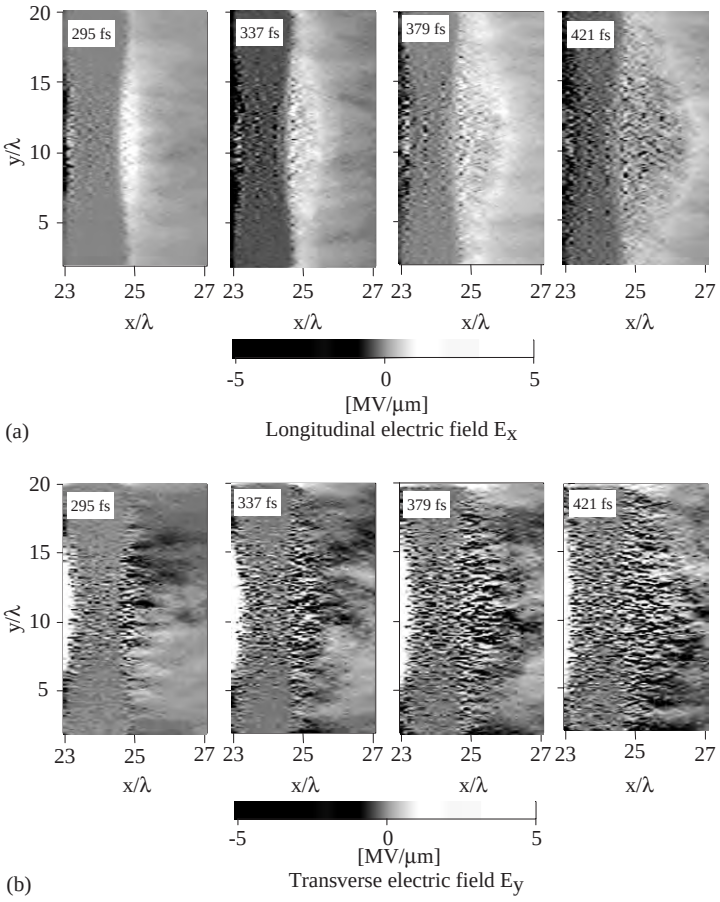



Fig. 5. The distributions of (a) the longitudinal and (b) the transverse electric fields at $t = 295, 337, 379$ and 421 fs in Case 4.

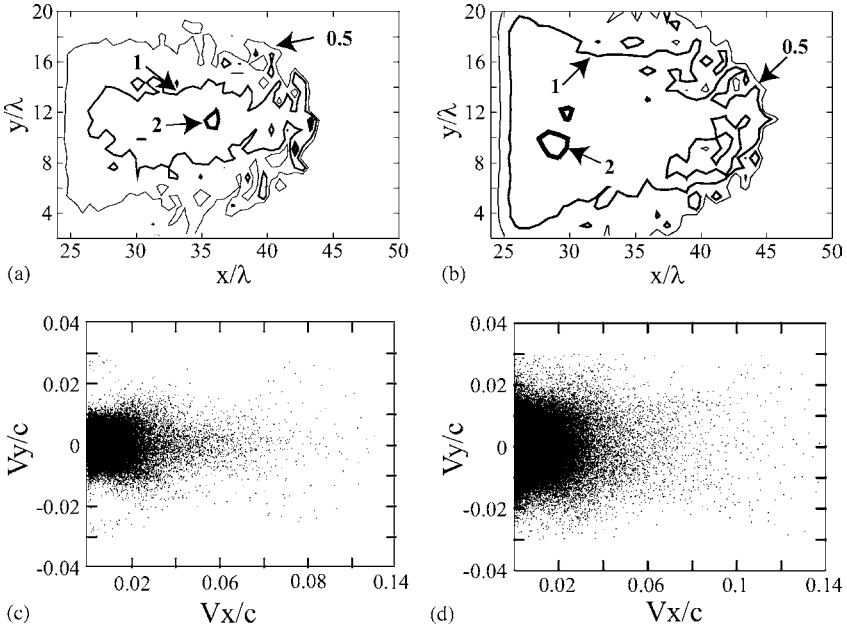



Fig. 6. The intensity distribution of accelerated protons for (a) Case 2 and (b) Case 4 after interaction with the laser. The phase space $V_x - V_y$ distributions for (c) Case 2 and (d) Case 4. Here, the values of intensity are normalized by $10^{14}\,\mathrm{W/cm^2}$.

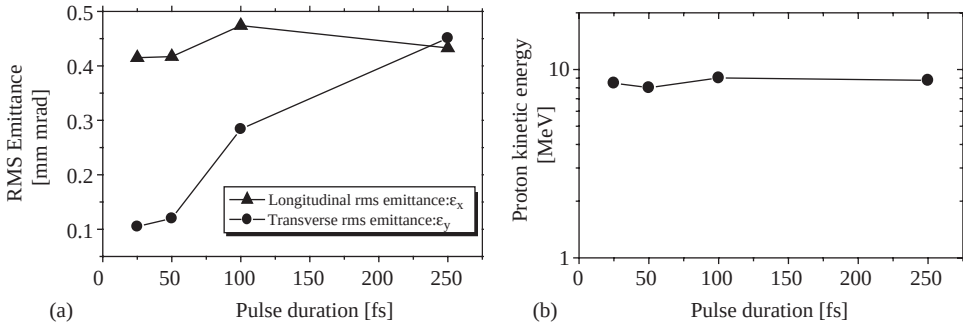



Fig. 7. (a) The transverse and longitudinal rms emittances for Cases 1–4 after the acceleration. (b) The maximum proton kinetic energy for all the parameter sets.

## Acknowledgements

The authors would like to extend their thanks to Prof. J. Limpouch, Prof. K. Mima, Prof. K. Tachibana, Prof. S. V. Bulanov, Prof. K. Nakajima, Prof. S. Kurokawa and Prof. N. Yugami for their valuable discussions and suggestions on this subject. This work was partly supported by JSPS (Japan Society for the Promotion of Science) and MEXT.

## References

[1] D. Strickland, G. Mourou, Opt. Commun. 56 (1985) 219.
[2] G. Mourou, C.P.J. Barty, et al., Phys. Today 51 (1998) 22.
[3] S. Kawata, T. Maruyama, et al., Phys. Rev. Lett. 66 (1991) 2072.
[4] Q. Kong, S. Miyazaki, et al., Phys. Rev. E 69 (2004) 056502.
[5] P.X. Wang, Y.K. Ho, et al., Phys. Lett. 78 (2001) 2253.
[6] G. Malka, J.L. Miquel, Phys. Rev. Lett. 77 (1996) 75.
[7] G. Malka, E. Lefebvre, et al., Phys. Rev. Lett. 78 (1997) 3314.
[8] S.V. Bulanov, F. Califano, et al., Rev. Plasma Phys. 22 (2001) 227.
[9] S.V. Bulanov, T.Zh. Esirkepov, et al., Phys. Lett. A 299 (2002) 240.
[10] S.C. Wilks, A.B. Langdon, et al., Phys. Plasmas 8 (2001) 542.
[11] T. Esirkepov, M. Borghesi, et al., Phys. Rev. Lett. 92 (2004) 175003.
[12] M. Allen, Y. Sentoku, et al., Phys. Plasma 10 (2003) 3283.
[13] M. Borghesi, A.J. Mackinnon, et al., Phys. Rev. Lett. 92 (2004) 055003.
[14] T. Nakamura, S. Kawata, Phys. Rev. E 67 (2003) 026403.
[15] P.K. Patel, A.J. Mackinnon, et al., Phys. Rev. Lett. 91 (2003) 125004.
[16] K. Matsukado, T. Esirkepov, et al., Phys. Rev. Lett. 91 (2003) 215001.
[17] T.E. Cowan, J. Fuchs, et al., Phys. Rev. Lett. 92 (2004) 204801.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 271–273
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Multi-tip field emission-based electron sources

Vladimir M. Zhukov

*St. Petersburg State University, Universitetskij pr., 35, St. Petersburg 198504, Russia*

Available online 22 November 2005

**Abstract**

The principal parts of accelerators are electron sources, which operate in some cases in field electron emission mode and can be developed as multi-tip field emission arrays (FEAs). Using an FEA with thermally smoothed tips, field emission current of 20.5 A was obtained in the pulse mode. The highest field emission current obtained using a non-thermally smoothed FEA was 175 A in the pulse mode.


*PACS:* 85.45.Db; 85.45.Bz; 79.70.+q; 41.75.Fr; 68.37.Vj

*Keywords:* Field emission; Cathodes; Pulse current

## 1. Introduction

The principal elements of accelerator devices are charged-particle (electrons and ions) sources.

In many cases, electron sources (cathodes) operate in the field electron emission (FEE) mode [1], which is mostly obtainable for tip emitters with small tip curvature radius.

A traditional method to obtain sufficiently high FEE current at relatively low voltage is the development of multi-tip field emission arrays (FEAs) with high identity order of the individual emitters.

Multi-tip systems were first described and approved by Dyke and co-workers [2]. In their work, preprocessed tips were fixed onto FEA bodies by point welding. Disadvantages of this technique are the inconsistency of tip spacing, differences in height and parallelism, low mechanical strength, and discontinuities in heat and electrical contacts between the tips and the conducting base. These factors negatively influence the simultaneous sharpening and formation of emitters.

Golov proposed [3] that field-emission cathode arrays could be produced from refractory metal foils (in particular, tungsten) by spark cutting. Using this technique, the array body and preprocessed tips are a single entity, and high uniformity of the bar geometry is ensured.

## 2. Experimental

The arrays used in this work (Fig. 1) were prepared using a modified technique.

Separate emitter tips were sharpened using a typical electrochemical etch in NaOH solution.

Further processing of the multi-tip cathodes may be carried out directly in the electron-accelerator vacuum chamber by thermal or thermo-field treatment (“build-up”). Using this procedure, the heated tip is exposed to an electron decelerating voltage (“+” onto the tip).

Using multi-tip field-emission arrays with this final processing step for thermal smoothing of tips at high temperature in single pulse mode, FEE current of 20.5 A was obtained (at a voltage pulse duration of $\tau = 2\,\mu s$, maximum voltage $U = 167\,kV$, and $N = 12$ emitters) [4].

In this system a separate tip was “blunted” up to conditions when the tip radius was 3.3 μm. From this tip, FEE current of 4.12 A (at a voltage pulse amplitude of 153 kV and pulse duration of 2 μs) was obtained.

Multi-tip thermo-field-processed systems that are not thermally smoothed are quite promising. This is due to the following reason.

It has already been noted that exposure of a highly heated emitter tip to a strong electric field (of non-emissive direction for electrons) results in changes in the tip shape—so-called “build-up”.

*E-mail address:* dvzhuk@nevsky.net.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.021

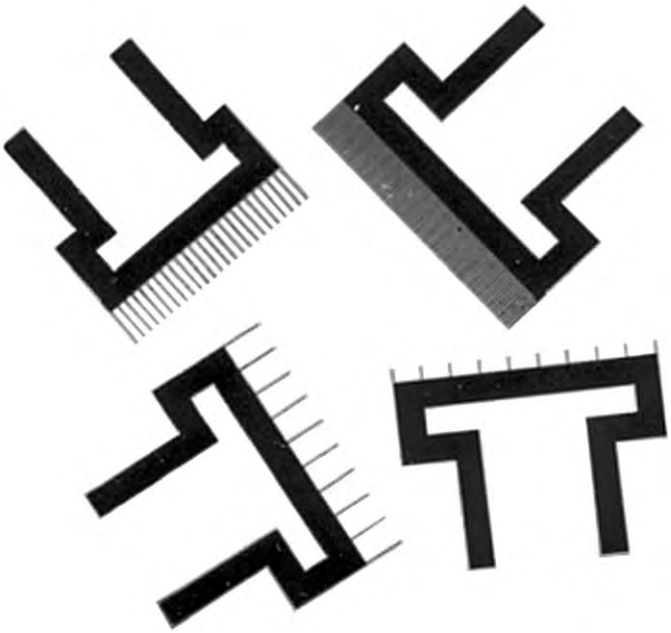


Fig. 1. Linear field-emission cathode arrays.

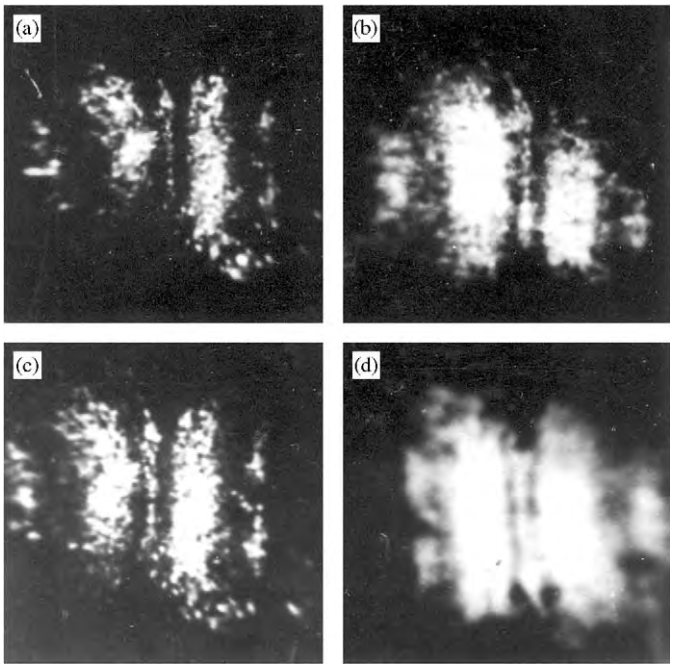



Fig. 3. Emission images of FEAs. Stationary mode: (a) before pulse FEE current extraction; and (b) after explosive rupture of some emission centres. Pulse mode: (c) at 50 A of FEE current; and (d) at 75 A of FEE current.

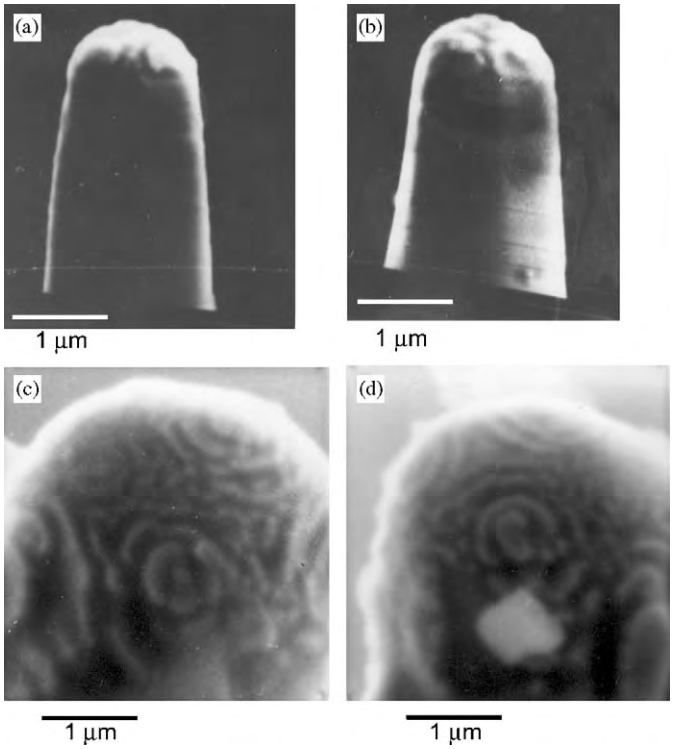



Fig. 2. Emitter tips at different ''build-up'' stages: (a,b) early stages after relatively short thermo-field treatment; and (c,d) after a long treatment time.

In the ''build-up'' process the emitter consecutively runs some quasistable phases with corresponding changes in emission images (which may be obtained using Müller's projector, ''–'' onto the tip) and the voltage needed for fixed FEE current extraction [5]. During this process, steps are generated on the emitter tip, and emission centres corresponding to the corners of these steps become visible in the emission images.

In Fig. 2 photographs of emitter tips at different stages of ''build-up'' are presented, taken using a scanning microscope.

It follows from the emission images, shadows and scanning micrographs that the characteristic scale parameters of these centres (such as the effective ''radius'', and tip ''cone angle'') are of an order that can provide high FEE current densities and magnitudes for ''built-up'' emitters.

One of the systems investigated [6,7] consisted of two fixed, parallel Π-shaped arrays (see Fig. 1) carrying 20 tips on each array. After thermo-field processing, many emission centres appeared on the surface of each separate tip (which was proved by the emission images), which were rather uniformly distributed over the surface (Fig. 3).

## 3. Theoretical background

It is well known [1] that the dependence of FEE current $I$ on voltage $U$ (between the cathode and anode) has an exponential shape.

The current–voltage characteristics of the FEA described are presented in Fig. 4. The solid line corresponds to the theoretical curve $I = C_1 U^2 \exp(-C_2/U)$, where $C_1 = 1.88\,\mathrm{A/\ kV^2}$, $C_2 = 184.5\,\mathrm{kV}$ for the geometry of the

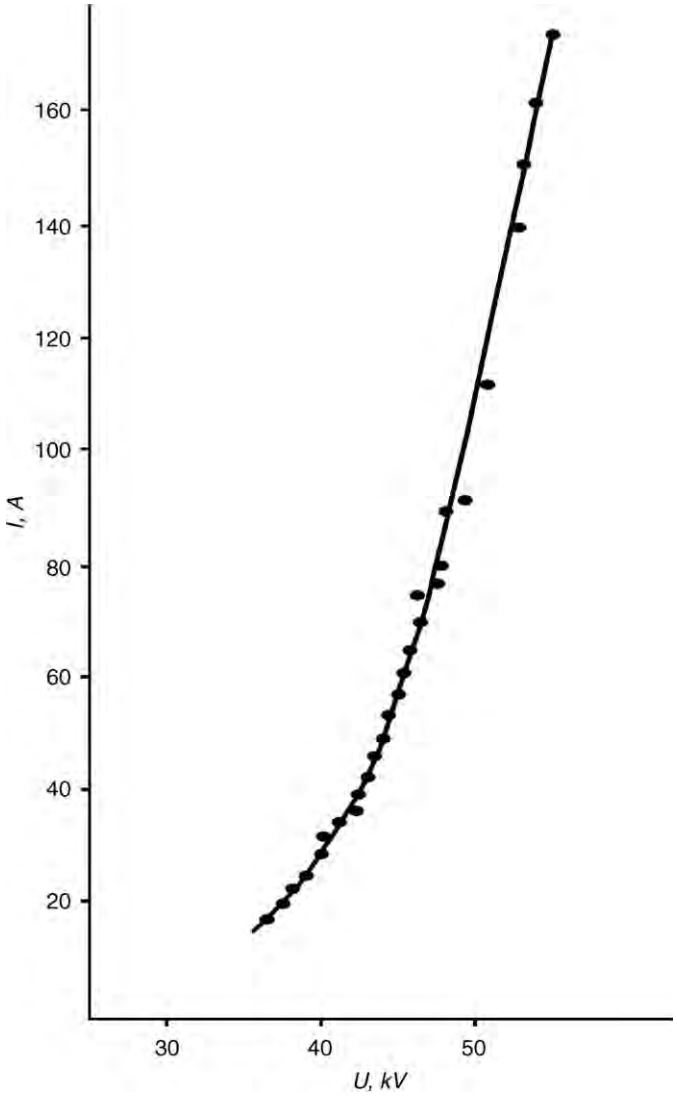



Fig. 4. Current–voltage characteristics of the FEA. Peak FEE current 175 A.

multi-tip cathode and the plane anode. The good coincidence between the theoretical and experimental curves may be explained by the contribution of the adjoining emission-centre parts of the tips to the FEE current, despite blunting of the individual centres in the emission process.

## 4. Conclusion

The highest FEE current obtained was 175 A at a voltage square-pulse duration of 20 ns and a voltage amplitude of 53.5 kV [7]. From a single emitter at $\tau = 20$ ns and $U = 44$ kV, FEE current of 5.5 A was obtained [6].

These values have not been exceeded by other researchers.

## References

[1] R.H. Fowler, L. Nordheim, Proc. R. Soc. A 119 (1928) 173.
[2] W.P. Dyke, Sci. Am. 210 (1964) 108.
[3] A.A. Golov, et al., USSR Author's Certificate No. 342241, Bulletin of Inventions No. 19 (1972) (in Russian).
[4] B.I. Grishanov, V.M. Zhukov, et al., Radiotekh. Elektron. 23 (1978) 575 (in Russian).
[5] V.M. Zhukov, S.A. Polezhaev, Zh. Tech. Phys. 57 (1987) 1133 (in Russian).
[6] V.M. Zhukov, S.A. Polezhaev, Radiotekh. Elektron. 33 (1988) 2153 (in Russian).
[7] V.M. Zhukov, Abstracts of the Third International Workshop on Beam Dynamics & Optimization, St. Petersburg, 1996, p. 39.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 274–276
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Flexible implementation of the Ensemble Model with arbitrary order of moments ☆

W. Ackermann*, T. Weiland

*Technische Universität Darmstadt, Institut für Theorie Elektromagnetischer Felder (TEMF), Schloßgartenstraße 8, D 64289 Darmstadt, Deutschland, Germany*

Available online 1 December 2005

**Abstract**

The *Ensemble Model* takes advantage of an approach to express the phase space particle distribution function in terms of the first, second and higher order moments instead of considering individual particles. Based on a new flexible implementation, an arbitrary number of orders can be processed and automatically converted into proper update equations for the simulation program *V-Code*. In this paper the influence of the introduction of higher order moments on the beam dynamics simulation is investigated. The achievable accuracy and the numerical efforts are compared with the ones obtained from the lower order calculations.


*PACS:* 05.20.Gg; 29.27.Bd; 41.75.–i

*Keywords:* Beam dynamics simulation; Ensemble model; Moment approach

## 1. Ensemble Model

To characterize the particle density distribution function $f$ in the space of coordinates and normalized momenta, the *Ensemble Model* utilizes an approximation of the multi-dimensional FOURIER transformation of the particle distribution instead of the density function itself [3]. The phase space coordinates $\{x, y, z, p_x, p_y, p_z\}$ are composed of $\vec{r} = \{x, y, z\}$ and $\vec{p} = m\vec{v}/m_0c = \{p_x, p_y, p_z\}$ where all vector components are given in cartesian coordinates. Since the principal behaviour of the *Ensemble Model* can still be traced using a simplified description, it is convenient to limit the derivation of the method to the 2D case. All necessary relations can then be obtained for the real model by expanding the corresponding fundamental equations to the 6D space.

In the first instance it is advisable to consider the 2D-FOURIER transformation

$$F(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \times \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{(jux)^k}{k!}\frac{(jvy)^l}{l!}\,\mathrm{d}x\,\mathrm{d}y \tag{1}$$

with the common kernel exponential function expanded as a power series. Due to the fact that the infinite expression is uniformly convergent, the order of summation and integration can be interchanged. The FOURIER transformed function is therefore given by

$$F(u,v) = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \frac{j^{k+l}}{k!\,l!} M_{kl} u^k v^l$$
$$M_{kl} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^l f(x,y)\,\mathrm{d}x\,\mathrm{d}y \tag{2}$$

where the coefficients $M_{kl} = \langle x^k y^l \rangle$ are called the moments of the distribution function. In practice the series in Eqs. (1) and (2) are not expanded to infinity but are rather truncated at a given maximum order. The approximated FOURIER transformed function is then used in a computer code to describe the characteristics of a real particle beam. In most cases it is sufficient to concentrate on the different

☆ This work was supported by DESY Hamburg.
*Corresponding author.
*E-mail addresses:* ackermann@temf.tu-darmstadt.de (W. Ackermann), thomas.weiland@temf.tu-darmstadt.de (T. Weiland).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.066

kinds of moments without the necessity to go back to the actual particle distribution. This is due to the fact that the moments instantaneously describe such physical identities like mean values

$$\bar{\mu} = \langle \mu \rangle, \quad \mu \in \{x, y, z, p_x, p_y, p_z\} \tag{3}$$

or variances

$$\sigma_\mu = \sqrt{\langle \mu^2 \rangle - \langle \mu \rangle^2}, \quad \mu \in \{x, y, z, p_x, p_y, p_z\}. \tag{4}$$

In statistics all moments are given in two forms: in addition to the already introduced raw moments, centralized moments of the form

$$\bar{M}_{kl} \coloneqq \langle (\mu - \langle \mu \rangle)^k (\nu - \langle \nu \rangle)^l \rangle, \quad \mu, \nu \in \{x, y, z, p_x, p_y, p_z\} \tag{5}$$

are also in use. Any raw moment of the distribution function depends on the real location of the particles and they are typically used to describe the average positions whereas the centralized moments are commonly translatory invariant and therefore employed to determine the shape of the particle distribution function regardless of the actual position.

The amount of involved moments is linked ultimately to the employed order and can be determined by the binomial coefficients

$$\binom{6 + k + l - 1}{k + l}$$

in 6D space. To solve the initial value problem on the basis of the ensemble model, a set of consistent initial conditions has to be applied. The time evolution of either raw or central moments can then be derived by integrating the first order differential equations

$$\frac{\partial \langle \mu \rangle}{c \partial t} = \frac{\partial}{c \partial t} \int \mu f \, \mathrm{d}\vec{r} \, \mathrm{d}\vec{p} = \int \frac{\partial \mu f}{c \partial t} \, \mathrm{d}\vec{r} \, \mathrm{d}\vec{p} = \int \left( f \frac{\partial \mu}{c \partial t} + \mu \frac{\partial f}{c \partial t} \right) \mathrm{d}\vec{r} \, \mathrm{d}\vec{p} \tag{6}$$

where the notations $\langle \mu \rangle = \int \mu f \, \mathrm{d}\vec{r} \, \mathrm{d}\vec{p}$ with $\mu \in \{x, y, z, p_x, p_y, p_z, x^2, \ldots\}$ for the raw and with $\mu \in \{(x - \langle x \rangle)^2, (x - \langle x \rangle)(y - \langle y \rangle), \ldots\}$ for the central moments are used. Application of a partial integration together with a proper ordering of the different integrals allows then to rewrite the fundamental differential equations in the compact form

$$\frac{\partial \langle \mu \rangle}{c \partial t} = \langle \mathrm{grad}_{\langle \vec{r} \rangle}(\mu) \rangle \left\langle \frac{\vec{p}}{\gamma} \right\rangle + \langle \mathrm{grad}_{\langle \vec{p} \rangle}(\mu) \rangle \left\langle \frac{\vec{F}}{m_0 c^2} \right\rangle + \left\langle \mathrm{grad}_{\vec{r}}(\mu) \frac{\vec{p}}{\gamma} \right\rangle + \left\langle \mathrm{grad}_{\vec{p}}(\mu) \frac{\vec{F}}{m_0 c^2} \right\rangle \tag{7}$$

which, however, cannot be used directly for a systematic implementation. In order to proceed, it is necessary to express the right-hand side in such a way that it consists of

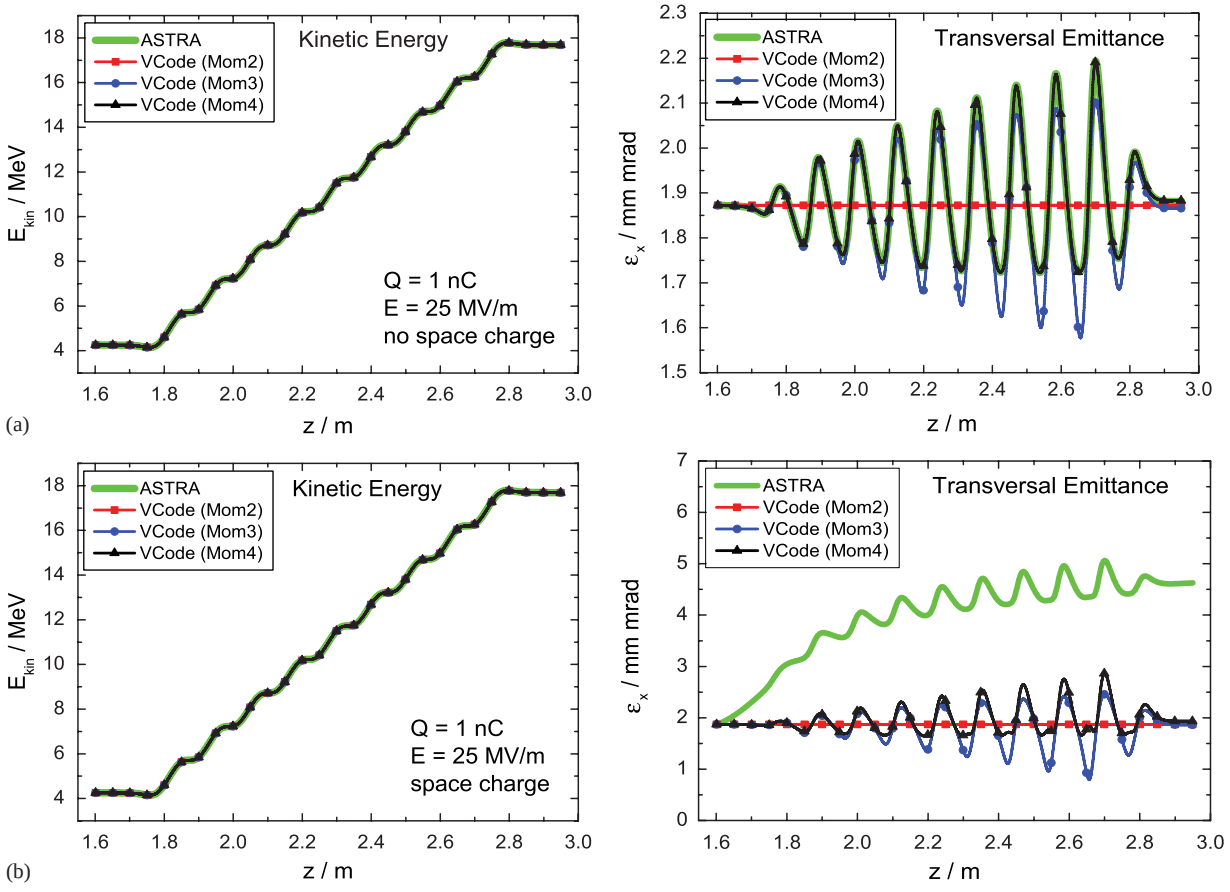



Fig. 1. Calculated kinetic energy and transversal emittance for the cavity part: (a) space charge forces are omitted; (b) space charge forces are included.

the time-dependent ensemble parameters only, i.e. the underlying moments. This is achieved by providing a series expansion of $1/\gamma$ as well as by a series expansion of the applied forces $\vec{F}$ in a given operating point and utilizing a truncation which is strongly related to the regarded order of moments. If the energy spread is small enough compared to the mean energy of the whole particle ensemble, it is reasonable to use just a linear approximation of $1/\gamma$; in all other cases higher order approximations are required. The series expansion for the forces can be formally performed by dividing them into the internal space charge motivated forces and those due to external fields.

## 2. Simulation

The described formalism for the *Ensemble Model* has been implemented into the beam dynamics simulation program *V-Code* and is already in use [1,2]. To study the influence of different orders of moments, a representative beam line layout of the PITZ gun has been chosen. For test simulations, a TESLA 9-cell-cavity has been placed subsequent to a gun. After the gun simulation, the calculations for the cavity part of the model have been performed. The operating conditions were adjusted in such a way that at the end of the gun a bunched beam with a minimum transversal emittance was obtained. The cavities mechanical dimensions range from $z = 1.6$ to $z = 3.0\,\mathrm{m}$ and an electrical field with a strength of $E = 25\,\mathrm{MV/m}$ at $f = 1.3\,\mathrm{GHz}$ was excited.

## 3. Conclusion

The principal behaviour of the moment based approach can be traced from the simulated results presented in Fig. 1. The diagrams for kinetic energy and transversal emittance show different tracking capabilities when either external forces or space charge forces are considered. The simulations demonstrate that the bunch dimensions as well as the energy are properly reproduceable even with lower order moments. Due to the implemented linear space charge model, the evolution of the sensitive emittance can be correctly performed with increasing order of moments in the presence of only external forces.

## References

[1] A. Novokhatski, T. Weiland, The model of ensembles for the beam dynamics simulation, Proceedings PAC1999, New York, 1999.
[2] M. Krassilnikov, A. Novokhatski, B. Schillinger, S. Setzer, T. Weiland, V-code beam dynamics simulation, Proceedings ICAP2000, Darmstadt, 2000.
[3] P.J. Channell, The moment approach to charged particle beam dynamics, IEEE Trans. Nucl. Sci. NS-30 (1983) 2607.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 277–279
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Dynamics of charged particles in optical traps

A.I. Dzergatch*, V.A. Kuzmin, S.V. Vinogradov

*Moscow Radiotechnical Institute RAS, 132 Warshawskoe sh., 117519 Moscow, Russia*

Available online 22 November 2005

**Abstract**

Analysis and numerical modeling show the possibility of the confinement of semi-relativistic electron motion in large (dimensions $\gg$ wavelength) cylindrical optical traps for low-density ($n_p/n_{crit}\sim 0.01$) plasmoids. The traps are based on circular polarized cylindrical waves $TE_{mn}(r,\varphi)$. Perspectives of other forms and volume variation of traps are briefly discussed.


*PACS:* 4262; 5275

*Keywords:* Optical traps; Underdense plasma

## 1. Analysis

High-frequency traps or RF potential wells [1] are based on the alternating gradient (AG) principle (ponderomotive force, light pressure). A simple system has been described [2,3] as a collective accelerator with small traps (dimensions $\ll$ wavelength $\lambda$). The present investigation of the semi-relativistic electron dynamics in higher types of vacuum strong (e.g., $\sim$100 GV/m) cylindrical waves in oversized (dimensions $\gg$ wavelength $\lambda$) volumes shows possibilities of synthesis, acceleration and compression of large (dimensions of many $\lambda$) optical traps for underdense plasmoids ($n_p/n_{crit}\sim 0.01$). These waves may be filtered out from an ultra-short pulse of a powerful laser by means of a surrounding hologram.

In the case of cylindrical waves, a 2-D trap is based on a circularly polarized wave $TE_{mn}(r,\varphi)$ [4]:

$$E_z = 0; \quad E_r = \hat{E}x^{-1}J_n(x)\sin\psi;$$
$$E_\varphi = -n^{-1}\hat{E}J_n'(x)\cos\psi;$$
$$\psi = \omega t - k_z z - n\varphi; \quad k^2 = k_z^2 + k_r^2 = \omega^2/c^2 \tag{1}$$

where $E$ is the field amplitude and $k_r = A_{mn}/R$ is the ratio of the $m$th root derivative of a Bessel function to the waveguide radius. The formulas for the magnetic field are omitted. Large $R \gg \lambda$ gives high $\hat{E}\sim 0.2mc^2/e\lambda \approx 0.1MV/\lambda$ in the focal zone. It ensures good AG focusing of electrons near $x \approx n$, which can trap the plasma ions.

Substitution of the field [Eq. (1)] into the equation of motion

$$\mathrm{d}(m\vec{v})/\mathrm{d}t = \mathrm{e}(\vec{E} + \vec{v}\times\vec{B}) \tag{2}$$

gives a system of connected linear equations with periodical coefficients for small oscillations $\Delta r(t), \Delta\varphi(t), \Delta z(t)$ of an electron around an arbitrary point $r_0, \varphi_0, z_0$ near the point $x \approx n$. If the velocities are also small, the equations reduce to separated equations with periodic coefficients; then the $z$-motion is inertial, and the transversal motion is described by two non-uniform Mathieu equations:

$$\begin{aligned}(m/e)\Delta\ddot{r} &= E_{r0} + E_{r0}'\Delta r = Ex_0^{-1}J_n(x_0)\sin\psi_0 \\ &\quad + E\partial/\partial x[x^{-1}J_n(x)]_{x=x0}\sin\psi_0\Delta x, \\ (m/e)\Delta\ddot{\varphi} &= -n^{-1}EJ_n'(x_0)\cos\psi_0\Delta\varphi.\end{aligned} \tag{3}$$

The radial motion $\Delta r(t)$ consists of forced harmonic oscillation of frequency $\omega$ around the point $r = r_0$, and a relatively slow free betatron or "Mathieu" oscillation around the same point; both oscillations depend on the initial conditions $r_0, \dot{r}_0$. The azimuthal motion $\Delta\varphi(t)$ is a free "Mathieu" oscillation around $\varphi_0$. Its amplitude is defined by the initial conditions. Thus, we have a large 2-D trap (hose) with a diameter of $\sim 2n\lambda$ for electrons. The

*Corresponding author.
*E-mail address:* adzergatch@mtu-net.ru (A.I. Dzergatch).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.023

Coulomb force of the electrons can hold positive ions of the plasmoid, as in the case in Refs. [2,3].

## 2. Numerical modeling

Numerical modeling was carried out using the precise Eqs. (1) and (2). Modeling of the case $k_z = 0$, $n = 14$ has confirmed the above analysis for small and large oscillations. Preliminary studies were carried out for the cases of two almost parallel crossed plane waves and of a Gaussian beam; the new results are shown in Figs. 1–4.

Fig. 1 shows the computed transversal (in coordinates $x$, $y$) motion of an electron during 12384 periods of the HF field (1). An electron with energy of 2.5 keV ($\beta\gamma = 0.1$) begins its motion from the wave axis, moves in the radial direction and, after reflections from the area of relatively high field near the trap boundary $r = R_0 = 2.5\lambda$, returns back to the center. The focusing tube field acts like a mirror, returning the particle to the center after each reflection. This was observed for $10^6$ periods, practically without change of the pulse near the axis. The difference with respect to a mirror is that after each two reflections the pulse deviates counterclockwise at 0.78° in the direction of the wave field rotation. Fig. 1 shows 115 radial oscillations during 12384 periods of the HF field. During this time the direction of the oscillations turns by 90°. The trajectories fill the first and third quadrants (only the first is shown).

Fig. 2 shows the phase plane for one such radial oscillation, which lasts for 107 periods of HF. At radius values of up to half of the trap radius, the particle moves uniformly. Then the field sharply increases, causing an increase in the amplitude of the radial oscillations and their complication near the caustic boundary. The oscillations of the radial pulse are first accompanied by a decrease in its mean value and then by a change in sign of the radial pulse. The oscillations are obviously non-linear. For an increase in the initial pulse $\beta\gamma$ from 0.04 to 0.1, the period of radial motion of the electron only increases from 93.7 to 107 HF periods, despite the fact that in the central region with a weak field, particles with higher energy pass 2.5-fold faster.

Fig. 3 shows (in coordinates $x$, $y$ during 752 HF periods) the trajectory of the electron starting with the same pulse along a straight line shifted by $0.5\lambda$ from the axis, i.e. it has a pulse momentum relative to the axis. Its trajectory is also straight inside the central area and there are almost mirror reflections from the strong field region at caustic boundary.

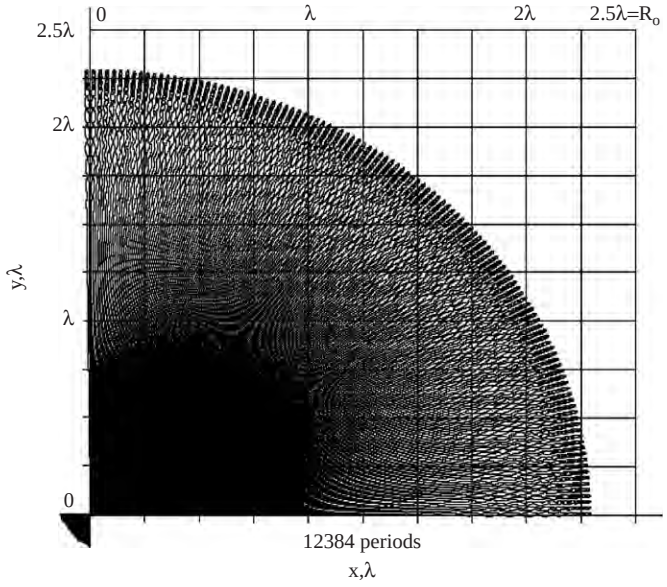



Fig. 1. 12384 periods.

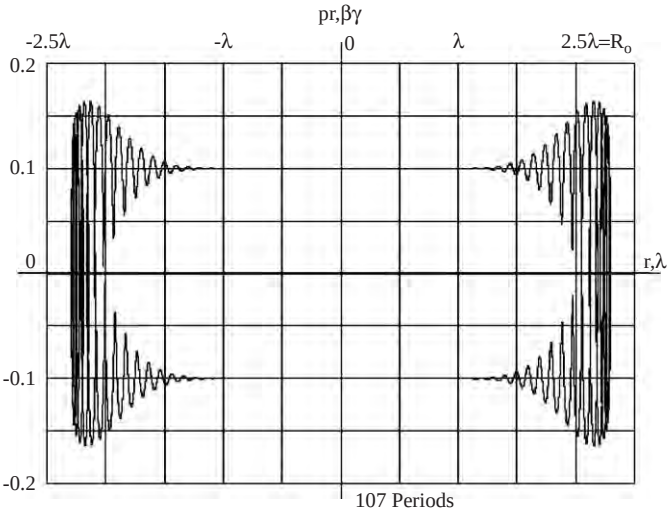



Fig. 2. 107 periods.

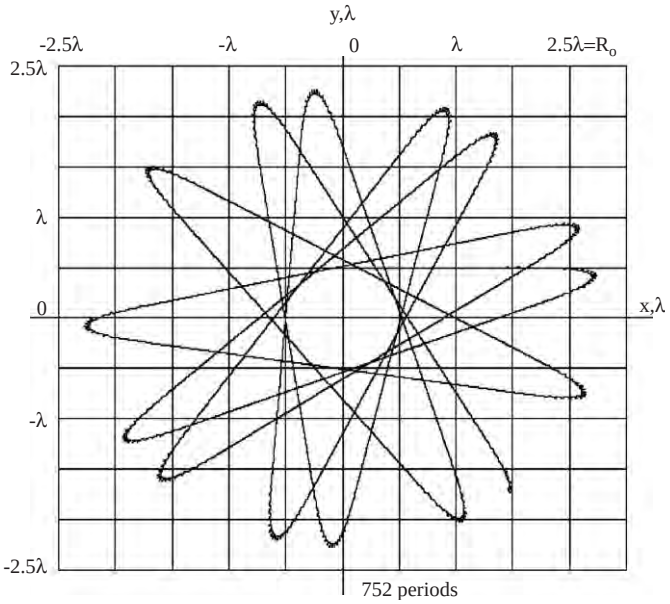



Fig. 3. 752 periods.

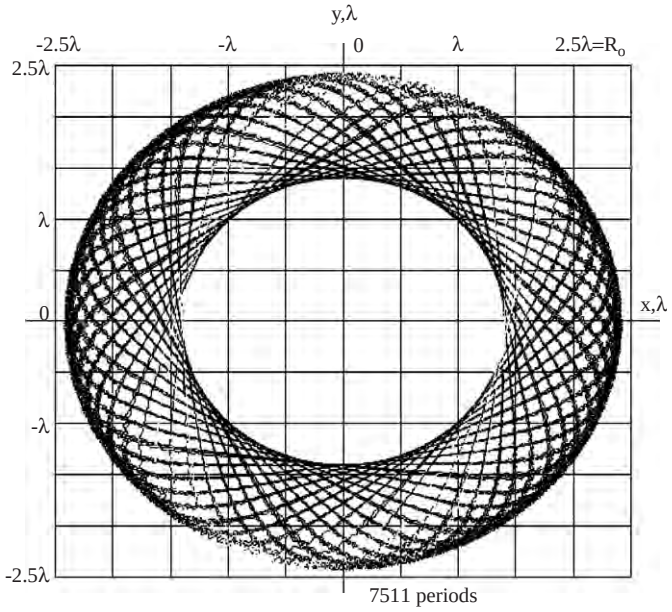



Fig. 4. 7511 periods.

The trajectory is similar to a star break line with smoothed edges. The values of the pulse and pulse momentum in the central area are constant, as well as the minimal and maximal radii from the axis. This takes place in all $10^6$ periods of calculation. There is a small difference from the trajectory of the electron with the same value, but with pulse momentum in the opposite direction. The angle variations for reflections differ to some degree, i.e. of the same order as pulse direction deviations for reflections in Fig. 1.

Fig. 4 shows the trajectory of an electron starting with the same pulse along a straight line shifted at $1.4\lambda$ from the axis during 7511 HF periods. It differs from the previous case by a 3-fold larger pulse momentum and minimal radius and a slightly larger maximum radius. In both cases, the rings denoting possible radius values are filled with trajectories. All cases demonstrate the reliable trapping of particles with transverse energy of up to 2.5 keV inside a tube wave (1) without notable variation of the parameters during $10^6$ HF periods.

## 3. Discussion and conclusion

Relatively low values of index $n$ (up to 14) in the modeling were defined by standard programs. A transition from the present idealized case of cylindrical traps to finite traps can be made by bending the cylinder into a tor of small curvature, or by using spherical circular polarized waves. Compression of traps may be based on the use of overlapped intervals of Bessel functions near their first maximum. The supposed process includes a series of steps from one multipole with $n$ pairs of poles to another with $n$–1 pairs of poles at the corresponding frequencies $\omega_n$ and $\omega_{n-1}$.

Thus, the analysis and numerical modeling of non-interacting electron dynamics in a cylindrical trap show that the motion of electrons is confined by the rotating wave; the particles move inertially within a field-free cylinder, and are reflected by the AG of the field near the inner field caustic.

Taking into account the results of Refs. [1–3], we might expect that plasmoids would be held during several ten thousand periods of the field in the case of weakly interacting plasma particles, $n_p \sim 0.01\, m\varepsilon_0\omega^2/e^2$. Collective centripetal acceleration of underdense plasmoids in the varying volume traps can be useful in various systems for collective acceleration and colliding beams. Realization of the traps is possible in a general case on the base of volume holograms.

## Acknowledgments

The authors thank A.A. Vassilyev and A.M. Sessler for help and discussions.

## References

[1] I.R. Gekker, The Interaction of Strong Electromagnetic Fields with Plasmas (Transl. from Russian), Clarendon Press, Oxford, 1982.
[2] A.I. Dzergatch, in: Proceedings of the EPAC 2000, Vienna, pp. 883–885.
[3] A.I. Dzergatch, V.S. Kabanov, V.A. Kuzmin, S.V. Vinogradov, in: Proceedings of the EPAC 2000, Vienna, pp. 880–882.
[4] W.R. Smythe, Static and Dynamic Electricity, McGraw-Hill, New York, 1950.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 280–281
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# The buncher optimization for the biperiodic accelerating structure with the high-frequency focusing

A.I. Fadin

*Moscow Engineering Physics Institute (State University), Kashirskoe Shosse 31, Moscow, Russian Federation*

Available online 22 November 2005

**Abstract**

The bunching part optimization results of an on-axis-coupled biperiodic accelerating structure for electron linac with high-frequency focusing are presented. System is intended for operation in the continuous regime at operating frequency of 2856 MHz and input power 5.5 MW. The basic development challenge for such installations on average input currents is the effective beam transfer through the structure. Some variants of the bunching sections distinguished by number of bunching cells were considered. The optimum capture ratio and an acceptable spectrum are provided by structure with five bunching cells. Optimization was carried out by means of dynamics simulation code PARMELA and a package of applied programs for the axial symmetric structures calculation SUPERFISH. Taking into account, space-charge limitation, the maximum capture ratio is 55%.


*PACS:* 29.25.Bx

*Keywords:* Buncher; Linac

## 1. Introduction

In this paper [1] four variants of the accelerating structures consisting of a buncher with varied number of cells and regular section of 20 cells were surveyed. Feed power and average effective shunt impedance was 5.5 MW and 70 MOhm/m. From the point of view of capture efficiency were of interest the third and fourth variants at which bunchers comprised of three and four cells, accordingly. In the given work, the problem consist in optimization of a bunching section of accelerating structure, with the goal to increase the greatest possible capturing at a beam injection current of 0.5 A.

## 2. Calculation results

Initially the task was solved without taking into account a space-charge effect. The beam with energy $W_{in} = 0.040$ MeV, radius $R_{in} = 0.059$ cm and emittance $\varepsilon = 0.002$ pi rad cm was injected. In a crossover, on distance $L_w = 0.8$ cm from a plane of a cell wall, it had radius $R_w = 0.05$ cm. By varying number and geometry of bunching cells, the fifth variant of structure in which the buncher will consist of five cells has been made. The first cell has the small relative phase velocity 0.4 close to a relative phase velocity of injected electrons, performing klystron buncher function. In Table 1, cells relative phase velocity $\beta_w$ and normalized to a rest energy field $\alpha_m$ are listed. Number of a variant corresponds to number of bunching cells in the accelerating structure. The maximum capture ratio obtained at modeling a beam dynamics, was 69.7% i.e., 18% higher than for three- and four-cell bunchers.

By the following step of beam dynamic it was modeled in view of a beam space charge. In Table 2 the capture ratio coefficient $K$, target current $I_{out}$, average energy $W_{av}$ and relative power spread $\delta W$ are presented. Phase and power distribution of the beam particles obtained on the accelerator end are represented in Fig. 1.

*E-mail address:* alexey_fadin@land.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.024

Table 1
Linac parameters

| Variant | | $N_{cell}$ | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| 1 | $\beta_w$ | 0.64 | 0.99 | 0.99 | 0.99 | 0.99 |
| | $\alpha_m$ | 2.88 | 2.88 | 2.88 | 2.88 | 2.88 |
| 2 | $\beta_w$ | 0.66 | 0.76 | 0.99 | 0.99 | 0.99 |
| | $\alpha_m$ | 1.23 | 2.88 | 2.88 | 2.88 | 2.88 |
| 3 | $\beta_w$ | 0.66 | 0.68 | 0.84 | 0.99 | 0.99 |
| | $\alpha_m$ | 0.82 | 2.67 | 2.88 | 2.88 | 2.88 |
| 4 | $\beta_w$ | 0.66 | 0.74 | 0.86 | 0.98 | 0.99 |
| | $\alpha_m$ | 0.92 | 2.67 | 2.88 | 2.88 | 2.88 |
| 5 | $\beta_w$ | 0.40 | 0.66 | 0.80 | 0.86 | 0.90 |
| | $\alpha_m$ | 0.31 | 1.13 | 1.64 | 2.05 | 2.47 |

Table 2
Beam dynamics calculation results

| Variant | $K$ | $I_{out}$ (A) | $W_{av}$ (MeV) | $\delta W$ (%) |
|---|---|---|---|---|
| 1 | 0.38 | 0.19 | 10.78 | 22.1 |
| 2 | 0.46 | 0.23 | 11.20 | 20.5 |
| 3 | 0.48 | 0.24 | 12.05 | 13.9 |
| 4 | 0.48 | 0.24 | 12.61 | 15.0 |
| 5 | 0.55 | 0.27 | 11.81 | 15.4 |

## 3. Conclusion

Capture in five-cell buncher is 13.5% higher in comparison with two other alternatives at comparable average energy and energy spread. Thus, by insignificant increase length of a buncher, the target current of a beam enlarged by 13%.

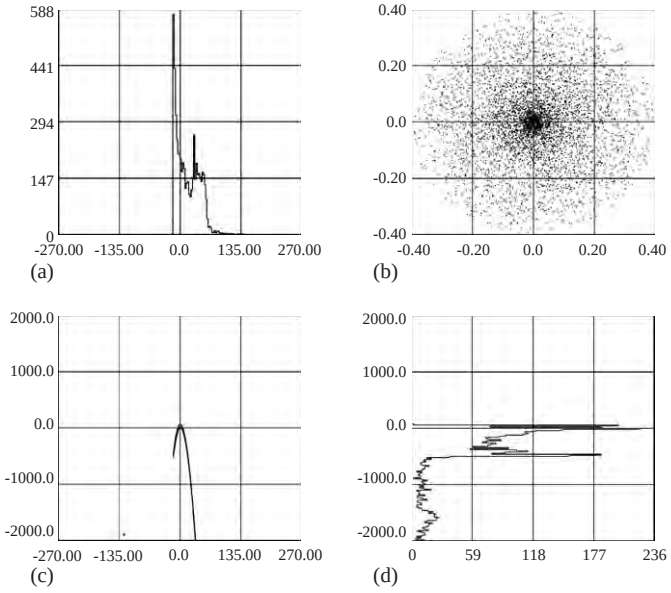



Fig. 1. Phase and energy distribution of particles for variant number 5: (a) phase spectrum, (b) $y$ vs. $x$ cross-section, (c) phase energy space, (d) energy spectrum.

## References

[1] A.I. Fadin, The choice of the buncher for the accelerator on biperiodic structure with a high-frequency focusing, Scientific session of MEPHI, 2003. Conf. Scientific—educational center CRDF, Basic researches of a substance in extreme states, The collection of proceedings, 2003, pp. 59–60.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 282–291
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Staging acceleration and cooling in a Neutrino Factory

C. Johnstone[a,*], M. Berz[b], K. Makino[b]

[a]*Fermi National Accelerator Laboratory, P.O. Box 500, Batavia IL 60510, USA*
[b]*Department of Physics and Astronomy, Michigan State University, East Lansing MI 48824, USA*

Available online 1 December 2005

**Abstract**

All schemes to produce intense sources of high-energy muons—Neutrino factories, beta beams, Colliders—require collection, RF capture, and transport of particle beams with unprecedented emittances, both longitudinally and transversely. These large initial emittances must be reduced or "cooled" both in size and in energy spread before the muons can be efficiently accelerated to multi-GeV energies. The acceleration stage becomes critical in formulating and optimizing muon beams; individual stages are strongly interlinked and not independent as is the case in most conventional acceleration systems. Most importantly, the degree of cooling, or cooling channel, depends on the choice of acceleration. In the current US baseline scenario, the cooling required for acceleration is about a factor of 10 in transverse emittance per plane. Longitudinal cooling is also required. In the proposed Japanese scenario, using an alternative acceleration scheme, no cooling is presumed. This work discusses two basic, but different approaches to a Neutrino Factory and how the optimal strategy depends on beam parameters and method of acceleration.
Published by Elsevier B.V.

*PACS:* 41.75–i; 41.85Ja; 07.77Ka

*Keywords:* Neutrino Factory; Muon acceleration; FFAGs

## 1. Introduction

The important stages in the US scheme for a Muon Collider [1] and Neutrino Factory [2–4] are outlined in Fig. 1 (left and right, respectively). Muons are created via the decay of pions, and pions are produced by directing an intense beam of protons onto a production target. The initial stage of a muon facility is considered to be a proton driver capable of delivering an ultra-short (3 ns long), high-intensity ($10^{14}$ p/pulse) beam. Collection, capture and bunching of pions and muons following the production target are the next major systems. First, a 50 m long channel allows the majority of pions with momentum less than $\sim$1 GeV/$c$ to decay into muons; the muons are then captured and bunched into 200 MHz RF buckets. Just after production, the captured particle distribution exhibits an enormous rms momentum spread of $\pm 55\%$. The tremendous energy spread of the muons is reduced through phase rotation in an induction linac or, more recently, in consecutive RF stations with varying frequencies. The combined bunching and phase rotation process produces a train of approximately a hundred 200 MHz bunches with a reduced $\delta p/p$ and a final rms value of about 10 cm in bunch length. (Both the induction linac and the RF-based bunching and rotation schemes produce similar final momentum spreads and bunch lengths so the criterion in choice of technique is merely cost.) The transverse emittance successfully captured is approximately 16 cm rad (full, normalized) at a momentum of 200 MeV/$c$ (mainly due to the strong, large-aperture solenoid surrounding the production target). The effective range in momentum, however, of captured particles extends from approximately 25 MeV/$c$ up to a cut off near 1 GeV/$c$. Since these large emittances cannot be efficiently accelerated, a "cooling stage" for emittance reduction precedes acceleration. To be effective, a cooling channel must be able to accept not only large transverse emittances (implying large-apertures common to both magnetic and RF components), but also a

*Corresponding author. Tel.: +1 630 840 3794; fax: +1 630 840 6311.
*E-mail addresses:* cjj@fnal.gov (C. Johnstone), berz@msu.edu (M. Berz).

0168-9002/$ - see front matter Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.078

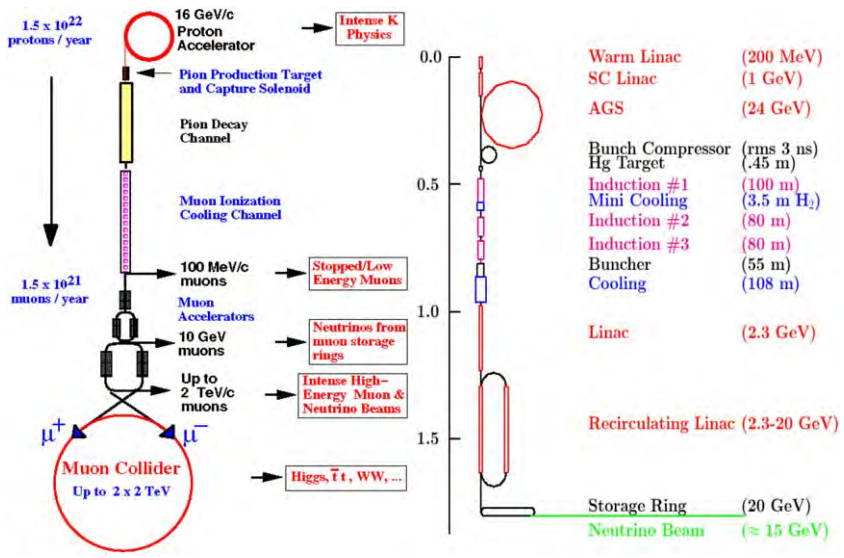



Fig. 1. Schematics of a Muon Collider [1] (left) and a Neutrino Factory [2] (right).

large (full) momentum spread of at least 40% $\delta p/p$ (generally quoted as $\pm 20\%$ $\delta p/p$ about a reference energy). The acceleration and storage rings—although nonconventional due to the large admittance and rapid cycle requirements that are imposed by large emittances and short muon lifetimes—represent the final stages of these facilities.

In all of the scenarios developed for Neutrino Factories and Muon Colliders, the captured transverse and longitudinal beam emittances are unprecedented. In comparison with high-energy hadron facilities, the transverse emittance is a factor of 1000 larger in each plane and the longitudinal emittance is 20–100 × larger even after bunching and phase rotation. Acceleration and collision of intense muon beams becomes impractical without a significant reduction, or cooling, of incipient emittances—transversely by a factor of 2.5–10 per plane for a Neutrino Factory [2–4] and at least a factor of 1000 for a Muon Collider [1]. (In the Neutrino Factory the required emittance reduction is tailored to the conditions for acceleration and in the Collider for the storage or collision ring.) Longitudinally, the degree of cooling differs drastically depending on the acceleration method, which is discussed here. The challenge in the design of these facilities, in particular the Neutrino Factory, lies primarily in accelerating the large beam emittances, a task which is further complicated by the short muon lifetime, or timescale on which these facilities must operate.

## 2. The acceleration stage

In a Neutrino Factory, the ability of, or limits to, accelerating large-emittance beams determines the specifications which upstream systems must meet, particularly the cooling. The downstream storage rings and experiments are presently not the limiting constraint. Acceleration proves, then, not only a difficult stage to develop, it becomes the pivotal one in the path to this facility. To further complicate issues, acceleration must occur rapidly because of potentially heavy losses from decay [1]. Linear accelerators are the optimal choice in this respect, but, above a few GeV, they become prohibitively expensive. Conventional synchrotrons cannot be used because normal conducting magnets cannot readily cycle in the ramping times [1] required by muon decay, nor do they support ultra-large beam emittances. In the past, the US baseline relied on recirculating linacs (RLAs) with separate, fixed-field arcs for each acceleration turn. Separate arcs allow control over the pathlength as a function of energy, allowing traversal times to be matched to the RF phase requirements for stable acceleration. Alternative approaches have focused on adapting the Fixed-Field Alternating Gradient (FFAG) accelerator first developed and tested at MURA [6], primarily because of its inherently large longitudinal acceptance. The Japanese approach (KEK) [7], for example, supports a radial-sector FFAG accelerator, but only in the context of a single-muon bunch and low frequency, broadband RF. Recent breakthroughs and improvements have demonstrated a new design for a FFAG accelerator [8,9] that can support stable, multi-turn acceleration for a high-frequency bunch train—the US scenario.

Transverse cooling (the upstream stage is termed precooling) is preferable in any approach after collection and capture to avoid enhanced component apertures, power levels, and, hence, cost of acceleration systems. However, it is the choice of accelerator that determines the longitudinal emittance that can be effectively accelerated and, therefore, also the degree of, or even the need for, longitudinal cooling. At a very fundamental level the overall design and staging of a muon facility depends on the method of acceleration chosen, and, for RLAs vs.

FFAGs, the longitudinal acceptance of the two machines embodies the most significant difference between the two scenarios. Since the RLA approach has been discussed extensively in past feasibility studies, and the intent of this paper is to focus on larger-acceptance accelerators, the RLA system requirements will be discussed only briefly. Instead this paper focuses on the FFAG approach, addressing the relaxed cooling requirements and reformulation of upstream stages relative to the RLA. Additional references to FFAG acceleration are included in Ref. [10] for a more complete overview.

The acceleration stage is composed of two subsystems: a linear pre-accelerator and a recirculating-beam accelerator (in this case either the RLA or the FFAG machine). The preaccelerator will not be discussed in any detail here, but certain assumptions in its parameters are required to extrapolate to the next stage of acceleration.

### 2.1. Pre-accelerator

A linear accelerator is optimal for the first acceleration stage in order to bring the low-energy muon beam delivered by the cooling stage to an energy at which it becomes feasible to inject into a re-circulating-beam accelerator. It is advantageous to initiate acceleration after moderate cooling (the precooling stage) in order to mitigate aperture-related costs with components and power. Without precooling, the $16\pi$ cm rad emittance (full, normalized) which is captured gives rise to meter-scale apertures at 200 MeV/$c$, since peak beta functions can assumed to be $\sim$3 m (characteristic of 1–2 T solenoidal confinement [2] or quadrupole confinement in a short, 2 m, 90° FODO cell [11]). Even with minimal cooling (defined here as a factor of 2.5 per plane in transverse emittance), linac apertures decrease by 40%, which is a substantial reduction when applied to a meter.

## 3. Recirculating accelerators

Efficient injection into a recirculating accelerator precludes an extended transverse beam size; one-half to meter beam sizes are definitely problematic. The beam size at injection into the recirculating acceleration stage depends most strongly on the performance of the cooling stage in combination with the capability or acceleration installed in the pre-accelerator. Without the pre-accelerator, and even assuming the maximum cooling factor of 10, the beam size remains large: $\pm$0.15 m (full) for a beta function of 3 m. Furthermore, recirculating accelerator designs have peak beta functions near 10 m, which implies apertures must reach at least 0.5 m to accommodate injected beam. Clearly adiabatic cooling by accelerating to 2.5 GeV (implying a further reduction in geometric emittance relative to 200 MeV/$c$ by a factor of 12) reduces the beam size substantially, by factor of 3.5. In order to achieve injection and more reasonable component apertures in the second acceleration system, a 2.5 GeV pre-accelerator is always assumed in this work. When combined with modest transverse cooling, the maximum beam size in the second accelerator becomes less than, or approximately, $\pm$8–16 cm (full) at 2.5 GeV, where the smaller number assumes a factor of 10 and the larger number a factor of 2.5 in cooling per plane. Beam sizes in the injection straight are lower than the maximum by about 25%, so injection remains challenging.

The transverse acceptance of the two proposed recirculators is described in the following sections. The emittance reduction factors are quoted relative to normalized units for convenience; they were calculated based on a 2.5-GeV pre-accelerator (which sets the injection energy) along with lattice design and practical assumptions for component apertures in the RLA and the FFAG cases, respectively. It is further assumed that a pre-accelerator can be designed that is capable of linearly accelerating both large transverse and large longitudinal emittances, with the latter particularly relevant to the FFAG scenario. (Linear acceleration implies that the effective normalized transverse emittance and the absolute momentum spread do not dilute significantly.)

### 3.1. Recirculating linear accelerator (RLA)

In an RLA, the beam is injected into a linac, accelerated, and returned by separate, fixed-field arcs on each acceleration turn, thereby achieving multiple passes of acceleration through the same linac. At the exit of each linac, the beam is sorted by energy and directed into a separate arc for transport on each acceleration pass. At the end of each arc, the trajectories from all arcs are recombined for acceleration in the opposing linac.

In spite of separate transport channels, however, the need to match to and maintain the RF bunching imposes the ultimate constraint on momentum acceptance within an arc and, indirectly, on its transverse acceptance. The implied condition on the longitudinal motion is that the value for M56 in each arc remain approximately constant over the accepted momentum range. For this purpose, strong families of sextupoles are used in the arcs, thereby achieving and fixing a maximum momentum acceptance in $\delta p/p$ of $\pm$5–10% for an RLA for a Neutrino Factory. Since sextupoles produce geometric aberrations, they also limit the useful dynamic aperture although with only a few turns, this is not a strong effect. Simply increasing the dispersion, thereby decreasing the strength of chromatic correction, does not appear to be effective in increasing either the longitudinal or the transverse acceptance of the RLA beyond this value due to aperture limitations and difficulty with high-order dispersion suppression in the linac sections.

Another, major difficulty in the design of the recirculating linacs lies in directing a beam with both a large transverse emittance and a large momentum spread into separate arcs on each acceleration pass. Clearly, to separate cleanly in a passive magnetic system (the only option for

the microsecond circulation times), the energy width must be less than the energy difference between consecutive acceleration passes. A large transverse beam size implies a further increase in the distance required for achieving clean separation, promoting an unavoidable conflict between longitudinal and transverse acceptance. Transversely, the full acceptance achieved so far in the recirculating linac approach lies between $1.5\pi$ and $2\pi$ cm rad for the momentum spread quoted. The ability to separate this beam passively into independent component channels then sets the minimum acceleration requirement per turn and, therefore, the maximum number of recirculation turns achievable. Given these transverse emittances and quoted beam sizes ($\pm 8$ cm), the momentum spread that can be practically accelerated to extraction energy appears to be $\pm 1\%$ for each 400 MeV of acceleration per turn until the $\pm 10\%$ limit set by chromatic correction is reached. Correspondingly, only 4 acceleration turns have been reported for RLAs in order to accommodate large momentum spreads: $\sim\pm 5\%$ $\delta p/p$ for a 3–11 GeV RLA and $\sim\pm 10\%$ for a 3–20 GeV RLA [2]. The corresponding momentum spreads, 250 and 500 MeV/$c$, can be compared to the useful momentum width of the muon beam after bunching and phase rotation, which appears to be ~600–700 MeV/$c$. (Much below 100 MeV/$c$ beam is lost due to decay and above 700 MeV/$c$, the number of muons captured decreases rapidly.)

In conclusion, since the full initial normalized beam emittance, as defined by the present bunching scheme, is about $16\pi$ cm rad, this incipient emittance must be reduced or cooled by a factor of 8–10 transversely and at least 1.4–3 longitudinally before acceleration can be accomplished in an RLA designed for a Neutrino Factory.

### 3.2. FFAG acceleration option

Alternatively, a circular accelerator system can be devised with magnetic fields that remain constant during acceleration by adopting an alternating gradient focusing lattice. The arcs of such machines, composed of large aperture magnets, can be designed to accommodate the large-energy range in acceleration. The beam centroid orbit is not fixed as in a ramped machine, but rather moves across the magnet aperture during acceleration. Lattices have been developed which can contain an energy change of at least a factor of four, although current lattices have converged on a factor of 2 as most feasible technically. In a FFAG accelerator orbit length unavoidably changes with energy; this change can be substantial and can result in a significant phase-slip relative to the RF waveform (unless low-frequency RF is employed). The phase slip accumulates on a per turn basis and eventually prevents acceleration to the extraction energy. This effect limits the number of turns that can be supported under conditions of rapid acceleration when the RF phase cannot be adjusted on a corresponding timescale. Recent improvements in lattice design, however, have resulted in significant enhancement of the number of acceleration turns that can be supported over the RLA, 10–15 as compared with 4–5 turns for the same energy gain. A larger number of turns ($>20$) is not desirable in muon applications because of decay, especially at the lower energies. Consequently, a dramatic reduction in RF voltage is not gained using the FFAG, but there is significant cost reduction over its RLA counterpart.

There are several classifications of FFAGs which refer to the lattice and momentum dependence of the optics. Since the concern here is with machine admittance, only this general property will be advanced for the different machines. The so-called *scaling* FFAG accelerators, such as the radial or spiral sector, display an almost unlimited momentum acceptance, but transverse acceptance remains somewhat restricted. Another approach to FFAGs, referred to as *nonscaling*, employs only linear magnetic elements [12] (quadrupoles and dipoles). Although the transverse optics changes slowly with energy, this latest type of FFAG demonstrates both strong momentum acceptance and unlimited dynamic aperture (limited only by the physical restrictions of the components). However, conditions of rapid acceleration are required to avoid beam loss from betatron resonances, a condition which also applies to muon acceleration.

The *scaling* FFAG designs have successfully achieved dynamic apertures of a few centimeter-radians (full), thereby mitigating the transverse cooling requirements by about a factor of 2 relative to the RLA and yielding an overall transverse reduction factor of ~5. *Nonscaling* FFAG designs exist with component apertures of 30–40 cm (~20 cm) horizontally (vertically) that accept a value of $6.4\pi$ cm rad for the full normalized emittance at the 2.5 GeV injection energy. The demonstrated emittance reduction factor is modest: only 2.5 in cooling prior to acceleration. Another advantage in both the *scaling* and *nonscaling* cases is total elimination of longitudinal cooling. In recent *nonscaling* FFAG designs, component apertures are now comparable to the RLAs, both horizontally and vertically. This is mainly due to the fact that the large dispersion in the RLA (~2–3 m required to separate beam with $\pm 5$–$10\%$ $\delta p/p$ into different arcs) is comparable to the shifts of the beam centroid during acceleration in the FFAG, even in the presence of lower transverse cooling where beam sizes increase by 40–100%. (Beta functions are similar in the RLA and the FFAG.) To fully realize the reduced transverse cooling factor, however, the pre-accelerator must be capable of linearly accelerating both large transverse and large longitudinal emittances. (Again, linear acceleration implies that normalized emittances remain reasonably unchanged by the acceleration.)

At injection the momentum acceptance of the FFAG is, practically, about $\pm 14\%$, or $\pm 350$ MeV/$c$. This limit is set by optimizing the FFAG design parameters which places the injection energy at a cell phase advance of $0.7\pi$. The upper limit to phase advance for injection is about $0.8\pi$ in order to keep the injection optics stable and away from the $\pi$ stability limit (the optics of the FFAG are FODO-like).

This upper limit on phase advance effectively sets the lowest momentum accepted for injection. Assuming an even spread about the central energy, gives the $\pm 14\%$ in $\delta p/p$. Even with a lattice optimized for acceleration and cost, the FFAG scheme does not appear to require any longitudinal cooling.

A simplistic view of cooling for a Neutrino Factory is illustrated in Fig. 2 including reference to different acceleration schemes. The technique of ionization cooling permits reduction of transverse emittances (4D phase space), or beam sizes, to levels acceptable for injection into accelerators with large momentum acceptance, or into ring coolers [4]. Ring coolers are multi-turn cooling channels designed to further reduce the transverse plus longitudinal emittance (6D phase space) to the smaller values required by a Neutrino Factory based on a RLA accelerator or by a Muon Collider.

### 3.3. Acceleration summary

Elimination of all longitudinal cooling, involving further muon decay, inevitable transmission losses, injection and extraction difficulties, and general R&D issues associated with both the optics and component design of advanced cooling channels, makes a persuasive argument to adopt the FFAG as the acceleration stage. Simple adiabatic damping during acceleration (if the dynamics are adequately conserving) is sufficient for the final transmitted emittances to meet the requirements of the storage ring and experiment. The present FFAG designs appear to satisfy this criterion, and, therefore, it appears sufficient to consider only a simple cooling stage for an FFAG: a straight or linear channel for transverse cooling (bends are required only for emittance exchange or longitudinal cooling). Such a linear channel must bring the initial, precooled emittance of $16\pi$ cm rad down to $3.2\pi$ cm rad for the *scaling* FFAG and $6.4\pi$ cm rad for the *nonscaling* FFAG accelerator. It should be noted, that a linear transverse precooler stage is relevant not only for the FFAG scenario, it is needed in RLAs and Muon Colliders as well with ring coolers representing a later stage of cooling. (Ring coolers also require "precooled" beam since their present design cannot accept large transverse emittances.)

## 4. The cooling stage

### 4.1. Ionization cooling

Emittance reduction, or ionization cooling, occurs because the muon beam loses momentum in all directions when traversing a target, or so-called absorber, and this energy loss can be replaced solely in the longitudinal direction by re-acceleration in an RF cavity; thus decreasing the beam's divergence for a given transverse dimension. The designs of stable optical configurations for cooling channels are particularly challenging because the straightforward cooling dynamics described above compete with the stochastic processes in the absorber; predominately multiple, or Coulomb, scattering which re-heat the beam. A net cooling effect can be achieved only if the cooling terms surpass the reheating ones, a state achieved through proper optics design in a cooling channel. The equation which follows represents the emittance increase in an absorber due to multiple scattering and, in the presence of cooling (reacceleration by an RF cavity), the minimum emittance [1] achievable for a specific channel design. From an optics standpoint, it is clear from equation 1 that the lower the average beta across the target, the proportionally lower the emittance increase from re-heating (multiple scattering), and therefore, the lower the achievable equilibrium emittance.

$$e_{\mathrm{N,min}} = \frac{\beta_{\perp}(14\,\mathrm{MeV})^2}{(2\beta m_{\mu} L_R\,\mathrm{d}E/\mathrm{d}S)} \tag{1}$$

where $\beta_{\perp}$ is the transverse beta function at the absorber, $\beta$ the relativistic velocity, $m_{\mu}$ the mass of the muon, $L_{\mathrm{R}}$ the radiation length of the absorber material, and $\mathrm{d}E/\mathrm{d}s$ the energy lost per meter in the absorber.

Therefore, the performance of a cooling channel depends not only on the beta functions at the absorber and on their constancy across a large momentum range,

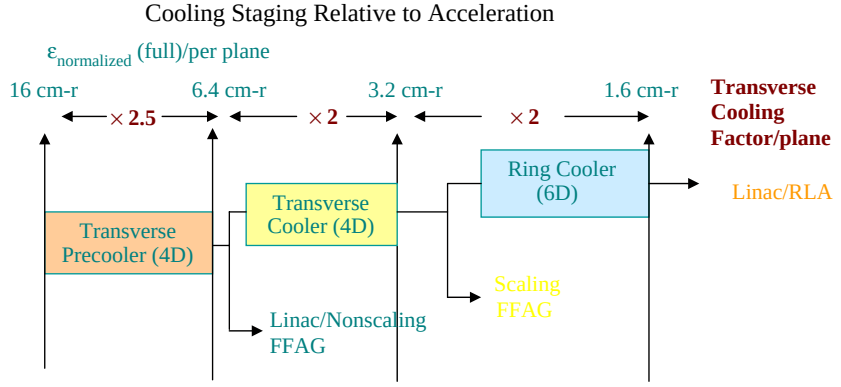



Fig. 2. Rough schematic of staged cooling relative to the acceleration model for a Neutrino Factory.

but, as important, also on the relative value of the starting emittance to the equilibrium emittance; the ratio needs only to be larger than ~1.5 for effective cooling. This observation indicates that the extremely low beta functions required in the latter stages of cooling are not a prerequisite during the early stages and may actually be less optimal from both a technical and nonlinear standpoint. Relaxing the low-beta conditions at the absorber impacts tremendously the design, stability and strength of the elements used in the upstream cooling. This observation will later be used to optimize the design of the cooling channels and adapt to the mode of acceleration chosen.

### 4.2. Cooling channel specifications

Assuming that a full transverse acceptance of $1.5\pi$ to $2\pi$ cm rad, as required for an RLA, corresponds to $2.5\sigma$ of a Gaussian beam profile, the rms normalized emittance demanded from transverse cooling is $2.4\pi$ to $3.2\pi$ mm rad. This degree of cooling (according to Eq. (1)) corresponds to a cooling channel with an average beta function of $<0.5$ m at the absorber. In the FFAGs, the relaxed transverse conditions allow betas at the absorber to increase to $\leqslant 1$ and $\leqslant 2.5$ m for *scaling* and *nonscaling* machines, respectively.

In momentum the practical range for ionization cooling extends roughly from 100 MeV/$c$ to a cutoff between 800 and 900 GeV/$c$. Application of all or a fraction of this range in the specification and design of a cooling channel depends on the longitudinal emittance capability of the acceleration system. The $\pm 5\%$ momentum acceptance of an RLA, for example, translates into a 250 MeV/$c$ total width at an injection energy of 2.5 GeV, implying the optimal cooling range should potentially encompass 150–400 MeV/$c$. Early cooling channels [2,3] based on large-aperture solenoids with a 155–245 MeV/$c$ cooling acceptance clearly were not optimally matched even to the requirements of the RLA.

In the case of FFAG acceleration one argues for a cooling channel which can accept and cool over as much of this momentum range as possible. The $\pm 14\%$ momentum spread accepted by current FFAG designs translates easily into an absolute momentum range of 700 MeV/$c$ at the 2.5 GeV injection energy. Assuming no or slight increase in this spread, then all of the effective momentum spread of beam from the bunching/phase rotator stage can be utilized. In light of the relaxed transverse specifications and no prerequisites for longitudinal cooling, a simple, quadrupole-based channel [11] is found to be well-suited to FFAG acceleration. The concepts for this channel will be described in the following sections. (Such a channel could also serve as a precooling stage for RLA acceleration.) Since the baseline has been solenoidal, not quadrupole cooling, first a brief discussion of solenoidal cooling will be presented along with comparative arguments for quadrupole cooling.

### 4.3. Solenoidal cooling channel

The transverse cooling stages described in the feasibility studies have been based on extremely large aperture solenoids with strong, sometimes superconducting, fields (1.25–5 T), one baseline example being the sFOFO channel. This channel is well documented and will not be revisited in any detail [2,3]. The sFOFO channel is capable of achieving a value for beta at the absorber of about 0.4 m across a momentum range of 150–250 GeV/$c$. Substituting in the above equation produces an equilibrium emittance value of $1.7\pi$ mm mr (normalized, rms). The channel can be expected practically to deliver a muon beam with an rms normalized emittance of $2.5\pi$ mm mr (1.5 above equilibrium), clearly sufficient to drive an RLA acceleration stage.

For this value of low-beta, however, the momentum range is constricted to lying between the momentum limits of 155 and 250 GeV/$c$, with these two limits representing strong channel resonances, $2\pi$ and $\pi$, respectively [3]. Beam is essentially captured up to a momentum of close to a GeV, but in this channel the majority cannot be cooled or even transmitted. (In fact, the minimum in the re-heating terms which contribute to Eq. (1) actually occurs at a momentum of 400 MeV/$c$.) These and other concerns listed below exist with the solenoidal channels.

- discrete liquid hydrogen absorbers and RF cavities are housed physically within the solenoidal aperture increasing component aperture, cost, field nonlinearities, and power requirements;
- strong longitudinal to transverse correlations develop in solenoids;
- nonlinear dynamics increase dramatically with solenoid aperture–simple or low-order models are found to be inadequate[ ];
- strong, especially superconducting field strengths promote increased sensitivity of beam parameters at such low energies (0.1–1 GeV);
- the momentum reach (155–250 MeV/$c$) of the sFOFO is limited relative to beam from the bunching and phase rotation stage (later designs have improved acceptance [5]);
- operational problems have been discovered experimentally in the R&D of RF cavities operated in a strong solenoidal field.

These beam control issues and technical difficulties provide incentive to develop a channel based on more conventional magnetic elements: large-aperture quadrupoles rather than solenoids. Some of the technical motivations for a quadrupole-based channel are as follows:

- liquid hydrogen absorbers and RF cavities extracted from magnetic apertures;
- recent developments indicate the entire channel could be pressured with gaseous hydrogen (windowless absorber)

including the RF cavity which halves the length of the channel and increases dramatically the cavity gradient [13];
- magnetic fields are reduced to nonsuperconducting levels, with more stable beam dynamics, and less sensitivity to field imperfections and nonlinearities;
- quadrupoles are strong-focusing rather than "soft-focusing" elements with reduced nonlinear contributions from nonparaxial terms or kinematical effects.

### 4.4. Linear quadrupole precooler

When beta functions at the absorber exceed or approach 1 m, the focusing strength is dramatically reduced and the absorber no longer has to be located at the lowest or a very-low beta point, allowing more flexibility in the choice of optical structure and focusing elements. This observation represents the basis for designing a competitive cooling channel based on normal-conducting quadrupoles in a simple lens, or FODO-cell configuration. Its application is solely as an upstream stage of cooling and, being a linear channel with no bends, serves to reduce the large transverse beam size in preparation for acceleration in an FFAG or for injection into ring coolers.

#### 4.4.1. Optical structure

With a sufficiently relaxed beta at the absorber, one can consider a short, alternating quadrupole lens structure. The advantages of a short FODO cell structure over a doublet or triplet quadrupole telescope are primarily in the acceptance and stability of optical parameters over a tremendous chromatic range. The dynamical range in telescope structures is about $\pm 5\%$ $\delta p/p$, beyond which there is no closed-orbit solution for off-axis beam. The limited momentum acceptance of the triplet/doublet quadrupole channels restrict their implementation to after longitudinal, or momentum, cooling has occurred and are not considered further here. However, in standard (implying repetitive) FODO-cell optics, the minimum beta in one plane is located at the maximum beta in the other. A minimum beta or beam size cannot be established simultaneously in both planes, and, therefore, the absorber cannot be located at the lowest beta point in this type of channel. The smallest beta for both planes combined is found halfway between the quadrupoles, at the "crossing point" in $\beta_x$ and $\beta_y$. Due to this limitation, the valid application of a FODO-based cooling channel is just after capture and phase rotation.

#### 4.4.2. Transverse cooling

For a short FODO cell, the average beta in both planes is equal and lies between 1 and 2 m for normal conducting quadrupoles and short (~0.5 m) spacing between them. The value of the beta functions at the crossing point is unusually stable over a large momentum range: from $-20\%$ to almost $+100\%$ if the phase advance is adjusted properly. The optics rationale for its design and stability will be discussed after a presenting the physical parameters chosen for this channel.

The physical parameters chosen for this channel shown in Fig. 3 were initially chosen to be comparable, or competitive with the sFOFO channel [2]. The aperture of the quadrupole was chosen somewhat conservatively—its length is constrained to be equal to its aperture—in order that the quadrupole field profile and therefore the optics are not fringe-field dominated.

The average beta achieved at the absorber in this channel is 1.6 m at 200 MeV/$c$ (this is the defined central momentum of the sFOFO channel [2]). This absorber beta yields a design equilibrium emittance (rms, normalized) of $6.8\pi$ mm rad, or a practical rms final beam emittance of $10.2\pi$ mm mrad (factor of 1.5 above equilibrium). Assuming a $2.5\sigma$ Gaussian, the full final transverse beam emittance is then $6.4\pi$ cm rad, or a factor of 2.5 below the $16\pi$ cm rad emittance arriving from the upstream bunching stage.

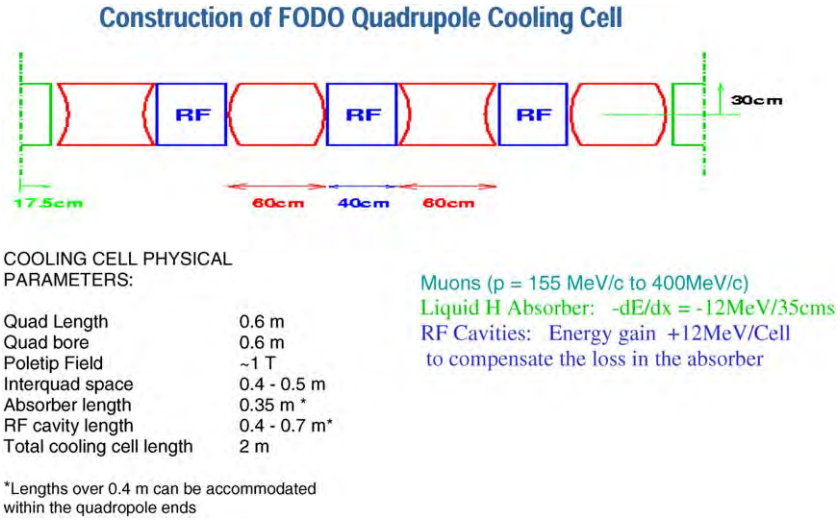



Fig. 3. Physical parameters and layout of the FODO-based quadrupole cooling channel.

*4.4.3. Momentum acceptance*

With the extreme demands placed on momentum performance, it is instructive to examine the FODO cell under the precepts of thin-lens conditions. First, it is useful to choose a reference momentum, $p_0$, and study the phase advance as a function of momentum relative to this reference in order to evaluate performance limits. For such a study, it becomes practical to assign a working point, or initial cell phase advance, to this reference momentum and one which is centrally located between stability limits: 0° and 180°. Clearly 90° is an obvious choice, hopefully optimizing the momentum reach of the channel. This choice of phase advance was applied to a $p_0$ of 200 MeV/$c$, a value chosen to be comparative with current cooling channel designs. The phase advance dependence, $\varphi$, on momentum can now be obtained in the thin-lens approximation (see reference)

$$\sin\frac{\varphi}{2}=\frac{(1/\sqrt{2})p_0}{p}\ \text{with a clear lower limit of stability,}\quad p=\frac{1}{\sqrt{2}}p_0. \tag{2}$$

Differentiating gives

$$\frac{1}{2}\cos\frac{\varphi}{2}\,\mathrm{d}\varphi=-\frac{p_0}{\sqrt{2}p^2}\mathrm{d}p\quad\text{or}\quad\frac{\mathrm{d}\varphi}{\mathrm{d}p}=\frac{\sqrt{2}p_0}{p^2\sqrt{1-p_0^2/2p^2}}. \tag{3}$$

Notice that for a $p_0$ of 200 MeV/$c$, the above analysis (Eq. (2)) gives a lower momentum cutoff for the channel of ~140 MeV/$c$ and, at large $p$, the phase advance varies more and more slowly, as $1/p^2$. The results of this analysis are graphed in Fig. 4 clearly demonstrating the large play in momentum of the simple-lens FODO cell. When compared with calculations, an almost constant factor of 0.8 was needed to translate the changes in phase advance from the thin-lens model to ones accurate for the channel as designed.

The slow variation in phase advance does not set restrictions on the length of the cell, but the variation of the peak beta function with momentum does. Using the definitions above, the peak beta function[ ] for a FODO cell is given by

$$\beta_{\max}=L\frac{\kappa(\kappa+1)}{(\kappa^2-1)^{1/2}},$$

$$\frac{\mathrm{d}\beta_{\max}}{\mathrm{d}p}=L\frac{(\kappa^2-\kappa-1)}{(\kappa-1)^{3/2}(\kappa+1)^{1/2}}\frac{\mathrm{d}\kappa}{\mathrm{d}p}=0\ \text{for a minimum.} \tag{4}$$

The variable $\kappa$ is defined by the following thin-lens equation.

$$\sin\frac{\varphi}{2}=1/\varsigma\quad\text{where }\varsigma=f/L\ \text{(thin lens).} \tag{5}$$

Here $\varphi$ is the phase advance of the FODO cell, $f$ is the focal length of $\frac{1}{2}$ of a full quadrupole, and $L$ is the length of a half-cell from quadrupole center to center (see, for example references listed in Ref. [14]).

In the above Eq. (4), $(\kappa^2-\kappa-1)$ can only be set to 0 locally (at ~76°), but this does not guarantee stability in the beta function over a large range in momentum. The only approach that minimizes $\mathrm{d}\beta_{\max}/\mathrm{d}p$ over a broad spectrum is to let $L$ approach 0. No drift between quadrupoles is optimal, but the choice of a short drift of ~0.5 m (which corresponds here to a half-cell length of 1 m) intentionally slows the variation of the maximum beam size with energy and at the same time insures a more feasible technical channel design. (Here absorbers and RF cavities are not installed inside magnet apertures.) The variation of the maximum beta with momentum for this design is shown below.

When the momentum dependence of the average beta at the absorber was studied, the change was found to increase slowly with energy and slowly relative to the increase in normalized admittance of the channel with energy (which is due to slowly varying peak beta values as a function of momentum). This increase in normalized emittance indicates cooling takes place over a large momentum range despite the increase in $\beta_\perp$ at the absorber.

Both the sFOFO and quadrupole cooling channels were fully modeled and tracked to high-order using the code COSY [15]. The simulation included:

- full nonlinear terms;
- with full solenoidal [16] and quadrupole fringe fields [17–19] (including different models);

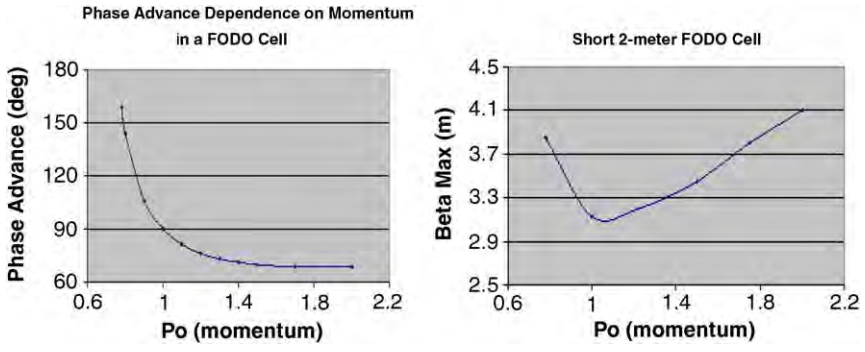



Fig. 4. On the left is the phase advance of a FODO cell plotted with respect to an arbitrary momentum, $p_0$, whose phase advance has been set to 90°. On the right is the variation of the peak beta function relative to $p_0$ for a half-cell length of 1 m.

- multiple scattering (absorbers + windows);
- energy loss including straggling and spin;
- $dE/dx$ as a function of energy;
- 200 MHz sinusoidal RF;
- full quadrupole fringe fields (both Enge representation[ ] and actual measured fields[ ]).

For the tracking, particles were launched in 2 cm steps along both axes and along the diagonal starting at the center of the absorber with and without the cooling (the hydrogen absorber). With cooling turned on, the transmission losses in the quadrupole channel exclusive of muon decay were almost negligible—less than 1% over the momentum range accepted by the sFOFO channel (155–245 MeV/$c$). (This transmission corresponds to an rms bunch length of 7.5 cm, a $\sigma_E$ of 12 MeV, $\varphi_s = 60^\circ$, $\Delta\varphi_s = \pm 54^\circ$, which corresponds to the 200 MHz RF bucket being about half filled.) Since the cooling rapidly reduces the emittance, the dynamic aperture is almost not relevant because a beam that fills the entire quadrupole aperture is cooled and is not lost even in the presence of fringe fields.

The predicted cooling behavior was observed; if a Gaussian distribution is launched which fills the quadrupole aperture, then the final rms of the distribution was found to be near 6.8 mm rad (normalized) for this specific channel (Fig. 5). The longitudinal losses of the quadrupole channel appear to be less than the solenoidal channel, the reason being the absence of longitudinal transverse correlations that plague solenoids.

As mentioned it is important to calibrate the expanded momentum reach of the quadrupole cooling channel. As noted above, the geometrical acceptance is almost constant, therefore, the normalized acceptance is increasing (since the relativistic velocity is not changing significantly). Hence, the absorber beta can be allowed to track the increase in normalized emittance acceptance. When benchmarked against the full cooling simulation performed at a $p_0$ of 200 MeV/$c$, the channel cools beyond 400 MeV/$c$. The momentum reach of the cooling in the quadrupole channel appears to be significantly larger than in the sFOFO, which extends from 155 to 245 MeV/$c$. Current effort is underway to launch a realistic beam from the bunching/phase rotation stage with varying momentum cuts to determine accurately the extent of the momentum cooling achieved in this channel.

### 4.5. Cooling summary

It is clear that a simple quadrupole cooling channel can be considered as the only cooling stage necessary for FFAG acceleration in a Neutrino Factory. At 200 MeV/$c$ and a 60 cm diameter bore, it first appears that the aperture is insufficient to accommodate the full $16\pi$ cm rad emittance delivered by the upstream bunching and phase rotation system (assuming ~3 m peak beta function in the quadrupole). However, an elliptical beam pipe can be installed in the quadrupole which extends beyond the poletip diameter by 50% and still maintain good field quality if the poletips are properly designed and separated (F. Mills, private communication). This would accommodate full beam using FODO channel optics. It is also interesting to note that the 200 MeV/$c$ central beam momentum was chosen due to the already strong solenoidal fields employed in the sFOFO channel and any increase in the central momentum implied an increase in field strength which is limited in this particular channel design. There is no reason not to accommodate a larger central momentum, and one which is more optimally matched to the production spectrum which extends from ~100 MeV/$c$ up to ~1 GeV/$c$. Moreover, the minimum of the sum of the re-heating terms (multiple scattering plus

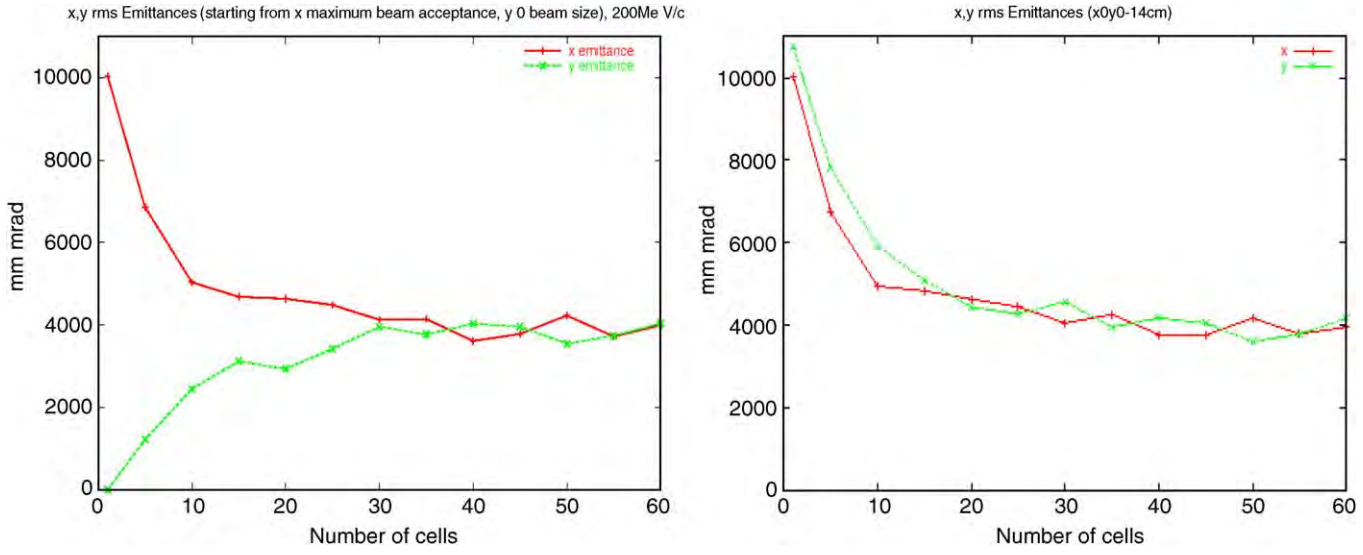



Fig. 5. Emittance reduction as a function of number of cooling cells. Left plot shows this starting emittance in $x$, no emittance in $y$ with corresponding emittance growth up to the equilibrium emittance. Plot on right are particles launched in $x$,$y$ Gaussian distributions, but along the diagonal.

straggling) actually occurs at 400 MeV/$c$. For the quadrupole channel design, the poletip field is not near a technical limit and can be increased. However, an increase in quadrupole strength is not required as this channel cools efficiently and well beyond the momentum reach of the sFOFO channel, from 150 to $>400$ MeV/$c$, so its central momentum is actually closer to 300 MeV/$c$. Further there are gains to be made by reducing losses from muon decay by propagating higher momenta, on average, through the channel.

## 5. Conclusions

In this paper, optimizing the stages in a Neutrino Factory has been presented. Clearly the staging and optimization are critically dependent on the choice and format of accelerator. It has been demonstrated that possibly the simplest, lowest-cost scenario is a *nonscaling* FFAG machine coupled to a linear (no bending) transverse cooling channel constructed from the simplest quadrupole lens system, a FODO cell. Transverse cooling demands are reduced by a factor of 4 and no longitudinal cooling is required relative to the RLA option. Detailed simulations further show that a quadrupole-based channel cools efficiently and much beyond the momentum range of a sFOFO [2] cooling channel with similar magnetic apertures. Current effort is underway to characterize the exact energy extent of the cooling. Applying different—both assumed and measured fringe fields—to represent the quadrupole elements fully has been an integral part of the simulations and ensures feasibility in quadrupole design and performance. One important observation is that such a channel cools effectively over a large variation in the fringe-field profile. Extensive design and simulation work are currently in progress on the proposed *nonscaling* FFAG and results are also encouraging. This cooling/acceleration scheme potentially represents the baseline scenario for the next US feasibility study.

## References


[1] The $\mu^+$ $\mu^-$ Collaboration, $\mu^+\mu^-$ collider: a feasibility study, BNL-52503, Fermi-Conf-96-/092, LBNL-38946, July, 1996.

[2] S. Ozaki, R. Palmer, M. Zisman, J. Gallardo (Eds.), Feasibility study II of a muon-based neutrino source, BNL-52623, June, 2001, available at http://www.cap.bnl.gov/mumu/studii/FS2-report.html.

[3] N. Holtkamp, D. Finley (Eds.), A feasibility study of a neutrino source based on a muon storage ring, Fermilab-Pub-00/108-E, 2000.

[4] Muon Collider/Neutrino Factory Collaboration, Phys. Rev. St. Accel. Beams 6 (2003) 081001.

[5] The Neutrino Factory and Muon Collider Collaboration, Neutrino factory and beta beams and development, Study 2-a, Lab Reports, BNL-72369-2004, FNAL-TM-2259, July 30, 2004, available at http://www.cap.bnl/mumu/study2-a and http://www.interactions.org/neutrinostudy.

[6] K. Symon, MURA-KRS-6, (MURA-43)(1954);
K. Symon, et al., Phys. Rev. 103 (1956) 1837.

[7] S. Machida, et al. Beam optics design of an FFAG synchrotron, (MOP1B20), Y. Sato, et al, Development of a FFAG proton synchrotron, (MOP1B21), EPAC 2000, submitted.

[8] C. Johnstone, S. Koscielniak, Nucl. Instr. and Meth. A 519 (2004) 472.

[9] S. Koscielniak, C. Johnstone, Nucl. Instr. and Meth. A 523 (2004) 25.

[10] M. Craddock, Cern Courier, V44, No. 6 (2004). FFAG Workshop, Vancouver, CA, April 15–24 2004, http://www.triumph.ca/ffag2004;
S. Koscielniak, M.K. Craddock, "simple analytic formulae for properties of nonscaling FFAG lattices", Proceedings of the Ninth European Particle Accelerator Conference, July 2004, Lucerne Switzerland (EPAC 2004);
S. Koscielniak, Comparison of COSY DA Maps with analytic formulae for orbit functions of a non-scaling FFAG accelerator, Proceedings of the Ninth European Particle Accelerator Conference, July 2004, Lucerne Switzerland (EPAC 2004);
C. Johnstone, S. Koscielniak, Optimizing non-scaling FFAG lattices for rapid acceleration, Proceedings of the Ninth European Particle Accelerator Conference, July 2004, Lucerne Switzerland (EPAC 2004).

[11] C. Johnstone, et al., Nucl. Instrum. and Meth. A 519 (2004) 472.

[12] C. Johnstone, A. Garren, in: Proceedings of the 1999 Particle Accelerator Conference, New York, NY, March 29–April 2, 1999 pp. 3068.

[13] R. Johnson, High pressure, high gradient RF cavities for muon beam cooling, in: Proceedings of the Linac2004, 17 August 2004, Lubeck Germany, available at http://www.muonsinc.com/TU203.pdf.

[14] E.D. Courant, H.S. Snyder, Ann. Phys. 3 (1958). Also, H. Bruck, "Accélérateurs de Perticules, Press Universitaires de France, Paris, 1966; English translation: Los Alamo Report LA-TR-72-10.

[15] M. Berz, K. Makino, COSY INFINITY Version 8.1—user's guide and reference manual. Technical Report MSUHEP-20704, Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, 2001, see also http://cosy.pa.msu.edu.

[16] K. Makino, M. Berz, D. Errede, C. Johnstone, Nucl. Intr. and Meth. A 519 (2004) 162.

[17] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999;
Also, G. Hoffstatter, M. Berz, Phys. Rev. E 54(4) (1996) 5664;
G.H. Hoffstatter, Rigorous bounds on survival times in circular accelerators and efficient computation of fringe-field transfer maps, Ph.D. thesis, Michigan State University, East Lansing, Michigan, USA, 1994.

[18] B. Erdelyi, Simplectic Approximation of Hamiltonian Flows and Accurate simulation of fringe field effects, Ph.D. thesis, Michigan State University, East Lansing, Michigan, USA, 2001, http://bt.pa.msu.edu/cgi-bin/display.pl?name=erdelyiphd.

[19] S. Kowalski, H. Enge, RAYTRACE, Technical Report, MIT, Cambridge MA, 1985.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 292–294
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Discrete optimization problem in beam dynamics

E.D. Kotina

*St. Petersburg State University, Universitetskij pr. 35, Petrodvoretz, St. Petersburg, Russia*

Available online 22 November 2005

**Abstract**

Discrete systems of equations have been often used when considering charged particle dynamics in accelerating and focusing structures. In this paper, a discrete optimization approach is considered for a problem in beam dynamics. The conventional approach for the design of various controlled systems involves the initial calculation of programmed motion (e.g. motion of a synchronous particle) and subsequent examination of perturbed motion with the use of equations for deviations. This approach, however, does not always lead to satisfactory results. Thus, when analyzing perturbed motions significantly dependent on programmed motions, it can happen that the dynamic characteristics of the perturbed motions obtained are not satisfactory from one point of view or another. A mathematical model that allows simultaneous optimization of programmed motion and an ensemble of perturbed motions in discrete systems is suggested.


*PACS:* 02.30.Xx; 02.30.Yy; 29.27.−a

*Keywords:* Controlled systems; Optimization; Functional; Discrete system

## 1. Introduction

The problem considered is a non-standard problem of the theory of optimal control in discrete systems. The problem of control of a particle trajectory and ensembles of trajectories (beams of trajectories) has been considered under various quality criteria. Along the trajectories of the system, we consider functions characterizing the dynamics of programmed motion and the dynamics of perturbed motion. The mathematical model suggested in this work also allows calculation of the density of particle distributions in phase space. A method for solving this problem is proposed. Analytical expressions for functional variations that help in constructing various directed methods of simultaneous optimization are suggested. The mathematical apparatus given can be effectively used in the optimization of the dynamics of charged particles in linear accelerators.

## 2. Statement of the problem

Let us consider the following system of discrete equations:

$$x(k+1) = f(k, x(k), u(k)), \tag{1}$$

$$y(k+1) = F(k, x(k), y(k), u(k)) \tag{2}$$

for $k = 0,\ldots,N-1$, where $x(k)$ is the $n$-dimensional phase vector defining programmed motion, $y(k)$ is the $m$-dimensional phase vector defining perturbed motion, $u(k)$ is the $r$-dimensional control vector, and $f(k, x(k), u(k))$ is the $n$-dimensional vector function defining the process dynamics at each step. For all $k \in \{0,1,\ldots,N\}$, the vector function $f(k) = f(k, x(k), u(k))$ is assumed to be definite and continuous on $\Omega_x \times U(k)$ in all its arguments $(x(k), u(k))$, along with partial derivatives with respect to these variables. $F(k) = F(k, x(k)y(k), u(k))$ is the $m$-dimensional vector function; for all $k \in \{0,1,\ldots,N\}$, it is assumed to be definite and continuous on $\Omega_x \times \Omega_y \times U(k)$ in all its arguments $(x(k)y(k), u(k))$, along with the partial derivatives with respect to these variables and second partial derivatives. $\Omega_x$ is the domain in $R^n$, $\Omega_y$ is the domain in $R^m$,

*E-mail address:* ekotina@compmath.spbu.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.025

and $U(k)$, $k = 0,1,\ldots,N-1$, is a compact set in $R^r$. In this case, we consider the Jacobian function

$$J(k) = J(k,\, x(k),\, y(k),\, u(k)) = \left|\frac{\partial F(k)}{\partial y(k)}\right|$$

to be non-zero for all changes of $k$, $x(k)$, $y(k)$ and $u(k)$.

Eq. (1) describes the dynamics of programmed motion. Eq. (2) is the equation for deviations, which describes perturbed motions.

We assume that $x(0) = x_0$ is fixed and the initial state of system (2) is described by set $M_0$—a compact set of non-zero measures in $R^m$. We call the sequence of vectors $\{u(0),\, u(1),\ldots,\, u(N-1)\}$ the control of the system described by Eqs. (1) and (2) and denote it by $u$ for brevity. We call the corresponding sequence of vectors $\{x(0),\, x(1),\ldots,\, x(N)\}$ the trajectory of programmed motion and denote it by $x = x(x_0,u)$. We denote the phase state of a programmed particle for the $k$th step by $x(k) = x(k,\ x_0,\ u)$. Similarly, we call the sequence of vectors $\{y(0),\, y(1),\ldots,\, y(N)\}$ the trajectory of perturbed motion and denote it by $y = y(x,\, y_0,\, u)$. We denote the phase state of the particle for the $k$th step by $y(k) = y(k,\, x(k),\, y_0,\, u(k))$.

The set of trajectories $y(x,\, y_0,\, u)$ corresponding to the initial state $x_0$, the control $u$ and different states $y_0 \in M_0$ are referred to as an ensemble of trajectories, or the beam of trajectories. The phase state of the beam at the $k$th step is also called the cross-section of the beam of trajectories and is denoted by $M_{k,u}$, i.e.,

$$M_{k,u} = \{y(k) : y(k) = y(k,\, y_0,\, x(k),\, u(k)),\, y_0 \in M_0\},$$

and the controls satisfying conditions $u(k) \in U(k)$, $k = 0,1,\ldots,N-1$ are admissible.

Let $x_0$ be the initial $u(k) \in U(k)$, $k = 0,1,\ldots,N-1$, of a synchronous particle and $M_0$ be the set of initial phase states of charged particles with density distribution $\rho_0(y_0) = \rho(0,\ y_0)$. We would like to determine how the distribution density function along the beam trajectories transforms. Let us fix an instant $k+1$ and the point $\bar{y}_{k+1} \in M_{k+1,\,u}$. Let $\bar{y}(k+1)$ be an image of the point $\bar{y}(k)$ in view of Eq. (2). We denote by $G(y(k))$ the set of points $y^i(k) \in M_{k,\,u}$, such that the trajectories of the system emanating at the instant $k$ from $y^i(k)$ at the step $k+1$ fall within a certain $r$-neighborhood $S_r(\bar{y}(k+1))$ of the point $\bar{y}(k+1)$.

By the distribution density of trajectories from Eq. (2) at the point $\bar{y}(k+1)$ on the $k+1$th step, we mean the limit

$$\rho\,(k+1,\ \bar{y}_{k+1}) = \lim_{r\to 0}\left\{\frac{1}{\mathrm{mes}\,(S_r(\bar{y}(k+1)))} \times \int_{G(y(k))} \rho(k,\ y_k)\,\mathrm{d}y_k\right\} \tag{3}$$

where

$$\mathrm{mes}\,(S_r(\bar{y}\,(k+1))) = \int_{S_r(\bar{y}(k+1))} \mathrm{d}y\,(k+1). \tag{4}$$

From the one-to-one correspondence of the sets $G(y(k))$ and $S_r(\bar{y}(k+1))$, the integral in Eq. (3) transforms to the form

$$\int_{S_r(\bar{y}(k+1))} \rho(k, y_k) J_k^{-1}\,\mathrm{d}y(k+1)$$

where $y_k = y(k)$ and

$$J_k^{-1} = J^{-1}(k,\, x(k),\, y(k),\, u(k)) = \det\left(\frac{\partial y(k)}{\partial y(k+1)}\right) = \det{}^{-1}\left(\frac{\partial F(k,\, x(k),\, y(k),\, u(k))}{\partial y(k)}\right).$$

In view of this and the form $\mathrm{mes}\,(S_r(\bar{y}(k+1)))$, we obtain the following equation for $\rho(k,y_k)$:

$$\rho(k+1,\ y(k+1)) = J_k^{-1}\rho(k,\ y_k), \quad \rho(0,\ y_0) = \rho_0(y_0). \tag{5}$$

The function $\rho(k) = \rho(k,y_k)$ denotes the distribution density functional for the $k$th step.

We introduce the following function:

$$I(u) = \sum_{k=1}^{N-1} \int_{M_{k,u}} \varphi_k(x_k,\, y_k,\, \rho(k,y_k),\, u_k)\,\mathrm{d}y_k + \int_{M_{N,u}} g(y_{N,}\, \rho(N,y_N))\,\mathrm{d}y_N \tag{6}$$

where $x_k = x(k)$, $\varphi_k$ and $g$ are continuously differentiable functions.

The functional (6) allows the simultaneous estimation of programmed and perturbed motions, as well as their simultaneous optimization and calculation of the density of particle distribution in phase space.

## 3. Functional variation

Let us rewrite Eqs. (1) and (2) and an equation for the distribution density along trajectories of the system (2):

$$\begin{aligned} x(k+1) &= f(k,\, x(k),\, u(k)),\\ y(k+1) &= F(k,\, x(k),\, y(k),\, u(k)),\\ \rho(k+1) &= J^{-1}(k,\, x(k),\, y(k),\, u(k))\rho(k) \end{aligned} \tag{7}$$

for $k = 0,\ldots,N$–1.

Let us denote variations of trajectories of system (7) as $\delta x(k)$, $\delta y(k)$ and $\delta\rho(k)$, with admissible variation of control $\Delta u$ and a given $u$.

Now, we define the corresponding equations for variations:

$$\delta x(k+1) = \frac{\partial f(k)}{\partial x(k)}\,\delta x(k) + \frac{\partial f(k)}{\partial u(k)}\,\Delta u(k), \tag{8}$$

$$\delta y(k+1) = \frac{\partial F(k)}{\partial x(k)}\,\delta x(k) + \frac{\partial F(k)}{\partial y(k)}\,\delta y(k) + \frac{\partial F(k)}{\partial u(k)}\,\Delta u(k), \tag{9}$$

$$\delta\rho(k+1) = \rho(k)\frac{\partial J^{-1}(k)}{\partial x(k)}\,\delta x(k) + \rho(k)\frac{\partial J^{-1}(k)}{\partial y(k)}\,\delta y(k) + J^{-1}(k)\delta\rho(k) + \rho(k)\frac{\partial J^{-1}(k)}{\partial u(k)}\,\Delta u(k). \quad (10)$$

We also have the following equation [4]:

$$\mathrm{div}_y\,\delta y(k+1) = \mathrm{div}_y\,\delta y(k) + J^{-1}(k)\left(\frac{\partial J(k)}{\partial y(k)}\,\delta y(k) + \frac{\partial J(k)}{\partial x(k)}\,\delta x(k) + \frac{\partial J(k)}{\partial u(k)}\,\Delta u(k)\right) \quad (11)$$

where $\mathrm{div}_y\,\delta y(k) = \sum_{i=1}^{m}\frac{\partial\delta y_i(k)}{\partial y_i(k)}$.

Taking into account Eqs. (8)–(11), the initial values of variations $\delta x(0)=0$, $\delta y(0)=0$, $\delta\rho(0)=0$, $\mathrm{div}_y\,\delta y(0)=0$, and using methods of investigation for functionals of type (6) [1], variation of functional (6) (for admissible variation of control $\Delta u$) can be represented in the following form:

$$\delta I = \sum_{k=0}^{N-1}\int_{M_{k,u}}\left(J(k)p^{\mathrm{T}}(k+1)\frac{\partial F(k)}{\partial u(k)} + J(k)\gamma^{\mathrm{T}}(k+1)\frac{\partial f(k)}{\partial u(k)} + J_k\xi(k+1)\rho(k)\frac{\partial J^{-1}(k)}{\partial u(k)} + q(k+1)\frac{\partial J(k)}{\partial u(k)} + \frac{\partial\varphi(k)}{\partial u(k)}\right)\mathrm{d}y_k\,\Delta u(k) \quad (12)$$

where $p(k)$, $\gamma(k)$, $\xi(k)$ and $q(k)$ are the following auxiliary functions:

$$p^{\mathrm{T}}(N) = \left(\frac{\partial g(y_N,\rho_N)}{\partial y(N)}\right),$$

$$\xi(N) = \left(\frac{\partial g(y_N,\rho_N)}{\partial\rho(N)}\right),$$

$$q(N) = g(y_N,\rho_N), \quad \gamma(N) = 0,$$

$$q(k) = J(k)q(k+1) + \varphi(k), \quad \xi(k) = \xi(k+1) + \frac{\partial\varphi(k)}{\partial\rho(k)},$$

$$p^{\mathrm{T}}(k) = J(k)p^{\mathrm{T}}(k+1)\frac{\partial F(k)}{\partial y(k)} + J(k)\,\xi(k+1)\rho(k)\frac{\partial J^{-1}(k)}{\partial y(k)} + q(k+1)\frac{\partial J(k)}{\partial y(k)} + \frac{\partial\varphi(k)}{\partial y(k)}$$

$$\gamma^{\mathrm{T}}(k) = J(k)p^{\mathrm{T}}(k+1)\frac{\partial F(k)}{\partial x(k)} + J(k)\gamma^{\mathrm{T}}(k+1)\frac{\partial f(k)}{\partial x(k)} + J(k)\,\xi(k+1)\rho(k)\frac{\partial J^{-1}(k)}{\partial x(k)} + q(k+1)\frac{\partial J(k)}{\partial x(k)} + \frac{\partial\varphi(k)}{\partial x(k)}$$

for $k = 1,\ldots,N-1$.

Eq. (12) for functional variation allows the construction of various methods of optimization of the function in Eq. (6).

## 4. Conclusion

Simultaneous optimization of programmed and perturbed motions under various quality criteria was considered in previous works [2–4] and the results were applied to the optimization of beam dynamics in linear-tube accelerators [3]. The analytical representation obtained in this paper for variation of the functional examined allows simultaneous optimization and calculation of the distribution density of trajectories in the beam.

## Acknowledgement

This work was supported by the Russian Foundation for Fundamental Research, Project 03-01-00726.

## References

[1] D.A. Ovsyannikov, Modeling and Optimization of Charged Particle Beam Dynamics, Leningrad State University Publishing House, Leningrad, 1990 312pp.
[2] E.D. Kotina, A.D. Ovsyannikov, On simultaneous optimization of programmed and perturbed motions in discrete systems, in: Proceedings of the 11th International IFAC Workshop, vol. 1, Pergamon Press, Oxford, UK, 2001, pp. 187–189.
[3] E.D. Kotina, Control discrete systems and their applications to beam dynamics optimization, in: Proceedings of the International Conference on Physics and Control–PhysCon 2003, St. Petersburg, Russia, 2003, pp. 997–1002.
[4] E.D. Kotina, Probl. At. Sci. Technol. 1 (2004) 147.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 295–298
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Application of field and dynamics code to LEBT optimization

S.A. Kozynchenko[a,*], Yu.A. Svistunov[b]

[a]Saint Petersburg State University, Saint Petersburg, Russia
[b]Efremov Scientific Research Institute of Electrophysical Apparatus, Saint Petersburg, Russia

Available online 28 November 2005

**Abstract**

A code for computer simulation and optimization of beam dynamics in 3D or rotationally symmetric electrostatic fields is considered. It is based on a physical model that takes into account the beam space charge. The theoretical framework used for both formulation of the model and interpretation of the results of numerical experiments is a formalism of the charged particle dynamics in phase space. The code can be used as an effective tool for computer-aided design and optimization of electrostatic accelerating and focusing systems. The operation of the code is illustrated with a typical example.


*PACS:* 41.85.Ja; 07.05.Tp; 02.60.Pn

*Keywords:* 3D electrostatic fields; 3D $H^-$ ion beams; Electrode systems; Optimization problems

## 1. Introduction

Up until now, the methods of the optimal control theory have been widely applied for optimization of charged particle dynamics in accelerating and focusing structures. The optimization problem of accelerated particle's dynamics is well investigated if it reduces to minimization of some functional on the ordinary differential equations trajectories [1,2]. Usually the field problem is not solved in this case and the control functions are estimated according to a certain algorithm, which is determined by an applied method in the presence of physical limitations.

However, the problem becomes much more complicated if control functions are expressed in terms of physical parameters of the system realizing these functions. In other words, the field problem is solved simultaneously with the trajectory optimization. An attempt at simultaneous optimization of the particle trajectories in the electrostatic fields and the electrode system creating these fields may be considered to be the first step in solving the problem in point. An example of such an electrostatic system is the injector of a linear accelerator for electrons or ions [3], which is usually called a low-energy beam transport (LEBT) system.

The present paper considers a code for computer simulation and optimization of beam dynamics in 3D or rotationally symmetric electrostatic fields, taking into account the beam space charge. It can be used for the analysis and optimization of various LEBT systems consisting of electrodes of the thick washer type. An example of the optimization of a five-electrode LEBT illustrates possible applications of the code.

## 2. Computational models

In the Cartesian coordinate system $(X,Y,Z)$ the electrostatic potential distribution $u = u(x,y,z)$ in a domain $D$ with boundary $G$, composed of $N_G$ piecewise-smooth sections $G_i$ is related through the Laplace equation

$$\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 + \partial^2 u/\partial z^2 = 0, (x,y,z) \in D. \quad (1)$$

The boundary conditions are given by the formulae

$$u|_{G_i} = U_i,\ i = \overline{1, N_1^G}$$
$$\partial u/\partial \mathbf{n}|_{G_j} = U_j,\ j = \overline{N_1^G + 1, N^G}. \quad (2)$$

*Corresponding author.
*E-mail address:* Sergey_Kozyntchenko@hotmail.com (S.A. Kozynchenko).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.026

The Laplace Eq. (1) with boundary conditions Eq. (2) is solved by the finite-difference method with successive over-relaxation (SOR) [4,5] using a seven-point finite difference approximation on a grid of cubic meshes of step $h$ [4].

This paper considers non-relativistic elliptical or rotationally symmetrical ion beams. This allows the self-magnetic field of the space charge flow to be ignored. The dynamics of the real beam is simulated by the method of 'large' particles, namely by the ensemble of $N_{\mathrm{mod}} \ll N_{\mathrm{real}}$ uniformly charged balls, each of radius $\alpha_{\mathrm{mod}}$, mass $m_{\mathrm{mod}}$ and charge $q_{\mathrm{mod}}$ [2].

Using the method of 'large' particles, we assume the equality of total masses and charges of the ensembles of real and 'large' particles. Here we also consider the additional criterion

$$W^{\mathrm{pot}}_{\mathrm{real}} = W^{\mathrm{pot}}_{mod} \tag{3}$$

where $W^{\mathrm{pot}}_{\mathrm{real}}$ and $W^{\mathrm{pot}}_{\mathrm{mod}}$ are the potential energy of the ensembles of real and 'large' particles, respectively. For the method of 'large' particles—balls [2], Eq. (3) may not be valid.

For modeling in a given volume $V_0$ of the ensemble of $N_{\mathrm{real}}$ point charged particles by means of $N_{\mathrm{mod}}$ 'large' particles with the same total charge and mass and $N_{\mathrm{mod}} < N_{\mathrm{real}}$, $W^{\mathrm{pot}}_{\mathrm{mod}} < W^{\mathrm{pot}}_{\mathrm{real}}$ is valid. The dynamics of the ensembles of real and 'large' particles must be the same in the sense of validity of some given criterions, so

$$W^{\mathrm{pot}}_{\mathrm{real}} = K_0\, W^{\mathrm{pot}}_{\mathrm{mod}} \tag{4}$$

must be valid, where $K_0 > 1$. For Eq. (4) to be valid, we propose in this paper to consider the 'large' particles as point ones with mass $m_{\mathrm{mod}}$ and charge $q_{\mathrm{mod}}$, where coupled interactions between them are expressed by the formula

$$F_{ij} = K_0(1/4\pi\varepsilon_0)(q^2_{mod}/r^3_{ij})r_{ij}, \quad K_0 > 1. \tag{5}$$

Some numerical simulations of the 3D beam dynamics in a five-electrode LEBT (see Fig. 2) have been made using Eq. (5), with $K_0 \approx 1.12$.

For numerical simulation of beam dynamics in an electrostatic field, taking into account the beam space charge, the following initial value problem is considered:

$$\begin{gathered} \frac{\mathrm{d}\tilde{\mathbf{R}}^i(t)}{\mathrm{d}t} = \tilde{\mathbf{V}}^i(t),\ \frac{\mathrm{d}\tilde{\mathbf{V}}^i(t)}{\mathrm{d}t} = \frac{q_{mod}}{m_{mod}}\mathbf{E}^i + \frac{\mathbf{F}^i_{pI}}{m_{mod}} \\ \tilde{\mathbf{R}}^i(t_0) = \tilde{\mathbf{R}}^i_0,\ \tilde{\mathbf{V}}^i(t_0) = \tilde{\mathbf{V}}^i_0,\ i = \overline{1, N_{mod}} \end{gathered} \tag{6}$$

where $\tilde{\mathbf{R}}^i = (\tilde{X}^i, \tilde{Y}^i, \tilde{Z}^i)$, $\tilde{\mathbf{V}}^i = \left(\tilde{V}^i_x, \tilde{V}^i_y, \tilde{V}^i_z\right)$; $\tilde{\mathbf{R}}^i_0$, $\tilde{\mathbf{V}}^i_0$ are given for the input of the electrostatic system; $\mathbf{E}^i = (E^i_x, E^i_y, E^i_z)$ is the external electrostatic field strength; and $\mathbf{F}^i_{pI} = \left(F^i_{pI\_x}, F^i_{pI\_y}, F^i_{pI\_z}\right)$ is the total force acting on 'large' particle number $i$ from the remaining 'large' particles; For numerical solution of Eq. (6) the Runge–Kutta method [5] is applied.

## 3. General description of the code

The code is intended for joint simulation and optimization of the electrostatic fields and the charged particle dynamics in LEBT, taking into account the beam space charge. It consists of several subroutine modules (Fig. 1) combined to solve a variety of problems encountered in the analysis, design and optimization of a 3D or rotationally symmetrical electrostatic LEBT consisting of electrodes of the thick washer type (Fig. 2).

The code was designed in object Pascal language in Delphi 5/7 using the object-oriented approach. It consists of several interacting user-defined structures (records and classes in object Pascal). Each program module (Fig. 1) includes one or more structures. The main structures of the code are as follows. The structure NODE (field simulation module) describes the electrostatic field potential at a single node in coordinates $(x,y,z)$ of the 3D grid. The structure ION (beam dynamics simulation module) describes a single computational particle in the point of phase space. The structure FIELD (field simulation module) is used for both the simulation and description of the 3D or rotationally symmetrical electrostatic fields in a given LEBT. The structure MBD (beam dynamics simulation module) is involved in both the simulation and description of the behavior of the entire 3D beam in phase space. FIELD and MBD also contain subroutines necessary for graphical and numerical output of the results. The structures: BWmethod1cv, BWmethod2cv and BWmethod3cv (optimization module) contain subroutines and data for solving one-, two- and three-parameter optimization problems, respectively, by the Box–Wilson method [6]. The structures Rmethod1cv, Rmethod2cv and Rmethod3cv (optimization module) contain subroutines and data for solving one-, two- and three-parameter optimization problems, respectively, by the 'ravine' method. The structure CDmethod

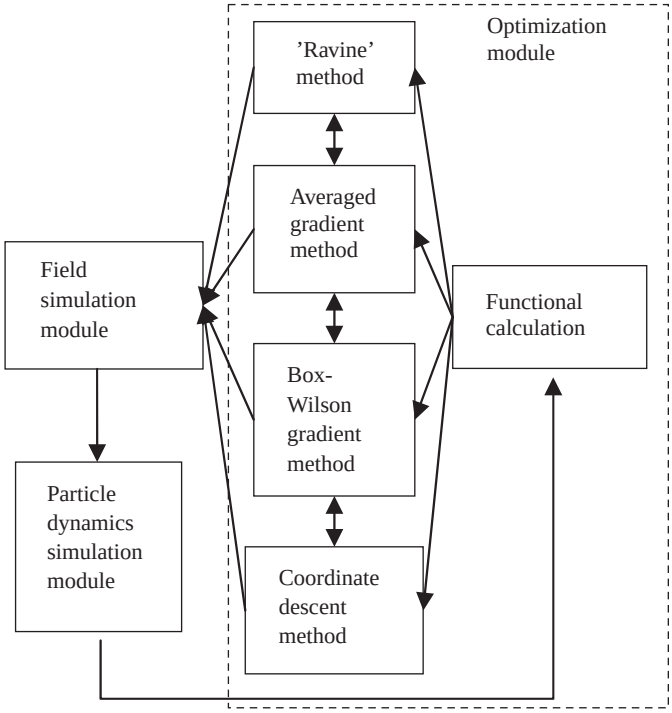



Fig. 1. Scheme for the computer code.

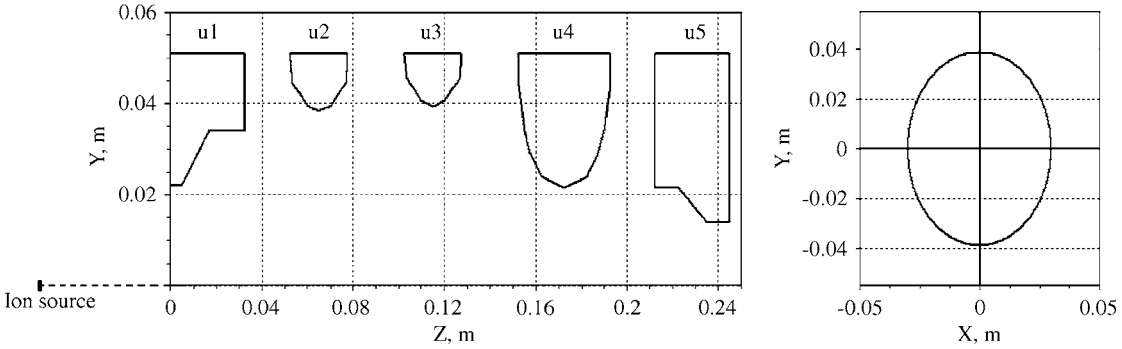



Fig. 2. Scheme for the optimized five-electrode LEBT.

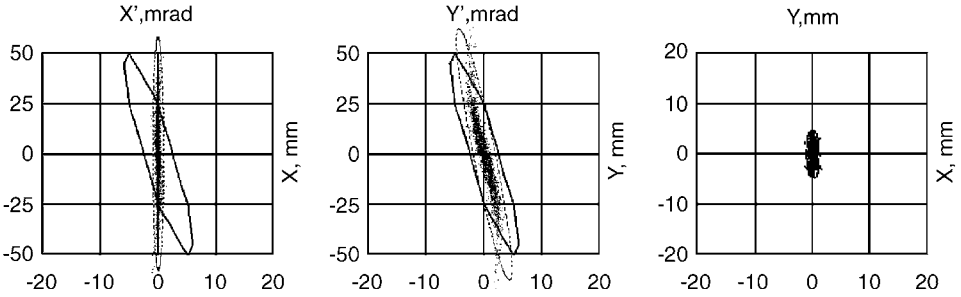



Fig. 3. Matching of the beam for five-electrode LEBT output and given RFQ acceptance. The solid sloped ellipse is the RFQ acceptance.

(optimization module) contains a subroutine and data for solving multi-parameter optimization problems of LEBT systems by the coordinate descent method. The structures: AGmethod2cv and AGmethod3cv (optimization module) contain subroutines and data for solving two- and three-parameter optimization problems, respectively, by the 'averaged' gradient method [1]. The code also contains a number of modules and structures containing auxiliary data and subroutines. The code was applied to numerical solution of a multi-parameter optimization problem on a Pentium 4/1.7 PC, as described in the next section.

## 4. Optimization problem solving

For optimization of the ion-optical performance of a given LEBT, the aim is to choose physical parameters so as to obtain desirable beam characteristics and to ensure high transmission efficiency. As an illustrative example, we present some results of the numerical experiments performed for analysis and optimization of an LEBT consisting of $n_e = 5$ elliptical electrodes of the thick washer type (Fig. 2). The aim of such optimization is the creation of an optimized LEBT system, producing a final beam with given characteristics and energy $W_{out} = 100\,\text{keV}$ for subsequent injection into the RFQ. The ion source [3] produces an elliptical beam of $H^-$ ions with energy of 17 keV and current of 15 mA. The quality of the LEBT is defined by the function $K_m = K_m(\mathbf{v}, \mathbf{p}, E_{out}^{xx'}, E_{out}^{yy'}, F_a, N_b^l, W_{out})$ where $E_{out}^{xx'}$ and $E_{out}^{yy'}$ are the normalized beam emittance for the LEBT output in the $XX'$ and $YY'$ planes, respectively; $\mathbf{v} = (v_1 \ldots v_{n_1})$ is the vector for the optimization parameters; $F_a$ is the function defining the RFQ acceptance (same for the $XX'$ and $YY'$ planes); $N_b^l$ is the beam losses in the LEBT; $W_{out}$ is the beam energy for the LEBT output; and $\mathbf{p} = (p_1 \ldots p_{n_2})$ are parameters different from $\mathbf{v}$, $E_{out}^{xx'}$, $E_{out}^{yy'}$, $N_b^l$, $F_a$ and $W_{out}$. The function $K_m$ characterizes matching of the beam emittance for the LEBT output and the given acceptance of the subsequent RFQ. In particular, it determines the number of particles $N_b^{RFQ}$ in the final beam, which may be involved in subsequent acceleration processes in the RFQ ($N_b^{RFQ} \leqslant N_b \leqslant N_b^0, N_b = N_b^0 - N_b^l$, where $N_b^0$ and $N_b$ are the number of beam particles in the LEBT input and output, respectively). The physical parameters of the LEBT, namely the electrode potentials $u_i$, the large $a_e^i$ and small $b_e^i$ semi-axes of the electrode minimal elliptical cross-sections, and the interelectrode distances $z_e^j$, $i = \overline{1, n_e}$, $j = \overline{1, n_e - 1}$ are considered as optimization parameters ($n_1 = 19$). Parameters of the LEBT different from $\mathbf{v}$ were fixed during the optimization.

Consider an optimization problem of the form

$$K_m^{opt} = \max_v K_m; \text{Eqs.(1)} \quad \text{and} \quad (2); \text{Eq.(6)}$$
$$|v_1| \leqslant A_1, \ldots |v_{19}| \leqslant A_{19}; \quad W_{out} = 100; \quad N_b^l \leqslant 0.1\ N_b^0$$
$$B_0 < N_b^{RFQ} \leqslant N_b^0; \quad \left|E_{out}^{xx'}\right| \leqslant B_1; \quad \left|E_{out}^{yy'}\right| \leqslant B_2; \quad |F_a| \leqslant B_3 \tag{7}$$

where $A_1, \ldots, A_{19}$, $B_0$, $B_1$, $B_2$ and $B_3$ are given values.

The time for solution of Eq. (7) on a Pentium 4/1.7 PC by the coordinate descent method is approximately 18 h. The 3D dynamics of the real beam in a 3D electrostatic field, together with the beam space charge, has been simulated with help of an ensemble of $N_b^0 = 2000$ 'large' particles—uniformly charged balls. The 3D electrostatic field was simulated on a space grid with approximately $N^{sg} = 2000000$ nodes.

The optimized LEBT configuration is shown in Fig. 3, where $u_1 = 12\,\text{kV}$, $u_2 = 7\,\text{kV}$, $u_3 = 100\,\text{kV}$, $u_4 = 45\,\text{kV}$ and $u_5 = 83\,\text{kV}$, which produces a final beam with $N_b^l = 0$ and 78% of particles involved in the subsequent acceleration process in the RFQ (Fig. 3).

## 5. Conclusion

The code presented allows determination of the optimal construction and physical parameters for injection and transport systems. The 3D field modeling subroutines of the code give the possibility of optimizing the injection systems for both rotationally symmetrical and 3D beams. However, there are limitations on $n_1$ and $N^{sg}$ when solving optimization problems on a PC.

## References

[1] V.I. Petrov, et al., Optimizing the parameters of initial part of the linear wave guide accelerator, in: Transactions of the 2nd All-Union Meeting on Charged Particle Accelerators, Moscow, 2 (1972) 159 (in Russian).
[2] D.A. Ovsyannikov, Modeling and Optimization of Charged Particle Beam Dynamics, Leningrad, 1990, 312pp. (in Russian).
[3] Yu.A. Svistunov, et al., Advances in NPK LUTS 433 MHz ion linac, in: Proceedings of the 18th International Linear Accelerator Conference, August 1996, Geneva, vol. 2, p. 869.
[4] V.P. Il'in, Numerical Methods of Solving Electro-optical Problems, Novosibirsk, Russia, 1974 (in Russian).
[5] W. Press, et al., Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, Cambridge, ISBN 0-521-43108-5, 1992.
[6] K. Hartmann, et al., Statistische Versuchsplanung und-auswertung in der Stoffwirtschaft, VEB Deutscher Verlag fur Grundstoffindustrie, Leipzig, 1974.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 299–302
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Measurement and correction of linear optics and coupling at tevatron complex

Valeri Lebedev[a,*,1], Vladimir Nagaslaev[a], Alexander Valishev[a], Vadim Sajaev[b,2]

[a]*Fermilab, MS 340, P.O. Box 500, Batavia, IL 60510, USA*
[b]*ANL, Argonne, IL 60439, USA*

Available online 28 November 2005

**Abstract**

The optics measurements have played important role in improving the performance of Tevatron collider. Until recently, most of them were based on the differential orbit measurements with data analysis, which neglects measurement inaccuracies such as differences in differential responses of beam position monitors, their rolls, etc. To address these complications we have used a method based on the analysis of many differential orbits. That creates the redundancy in the data allowing to get more detailed understanding of the machine. In this article, we discuss the progress with Tevatron optics correction, its present status and future improvements.


*PACS:* 29.20; 29.27; 41.85.−p

*Keywords:* Beam optics

## 1. Introduction

The commissioning of Tevatron Run II began in the spring of 2001 with the first luminosity seen in June. By the year's end the luminosity was in the range of $(5–10)\times10^{30}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Although the luminosity growth was significantly slower than expected, steady growth of luminosity has been demonstrated during last 3 years with the peak luminosity of $1.02\times10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ achieved in July 2004. This luminosity growth would not be possible without thorough understanding of the accelerator physics problems, which have restricted the machine operation. Understanding and tuning the Tevatron optics was one of the main contributors to the success. Two problems have been encountered: the emittance growth due to optics mismatch at injection, and optics mismatch for the collision optics with subsequent increase of beta-functions in the interaction points (IP).

The collider is filled from Main Injector (MI) by protons and antiprotons through two different transfer lines. To maximize the luminosity the optics of each line has to be matched to both rings. Although, in principle, emittance growth related to each transfer could be measured using the emittance monitors of each ring, in reality, this method does not work because of uncertainty in calibration of emittance monitors. To exclude this uncertainty the round trip emittance measurement has been used. In this case, the proton beam is sent from MI to Tevatron through the proton transfer line and then sent back to MI through the antiproton line. The MI emittance monitor is used to measure the total emittance growth for both transfers. The measurements performed in 2002 exhibited the round trip emittance growth of about 50%. After linear optics correction in both transfer lines this value was only slightly improved and still stayed well above the emittance growth related to the betatron oscillations due to injection errors. Initial estimates of possible coupling effects yielded that small coupling cannot be a reason of such emittance growth. Soon we learned that the coupling in Tevatron is

*Corresponding author.
*E-mail address:* val@fnal.gov (V. Lebedev).
[1]Work supported by the Universities Research Assosciation Inc., under Contract DE-AC02-76CH03000 with the US Department of Energy.
[2]Work supported by US Department of Energy, Office of Basic Energy Sciences under Contract No. W-31-109-ENG-38.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.027

not small and leads to the significant emittance growth. The source of the coupling has been the regular skew-quadrupole component in all Tevatron dipoles. It originated from settling down the superconducting coil relative to the iron core due to compression of thermo-insulating coil support by $\sim$150 μm. Although the value of skew quadrupole gradient, $G_s$, for a single dipole does not look large ($G_sA/B_0 \sim (1.5\text{–}2)\times 10^{-4}$, where $A = 2.54$ cm and $B_0$ is the dipole field) summing the contributions of all 772 dipoles yields the uncompensated tune split $\sim$0.3. This value is usually compensated to better than $5\times 10^{-3}$ by machine skew quadrupoles but it still leaves large local coupling through the entire machine. Hundred and twelve dipoles, which did not have nearby machine skew quadrupoles, were main source of this local coupling. During the 2003 shutdown their skew quadrupole fields were corrected by restoring thickness of the thermo-insulating supports (shimming of the dipoles). That resulted in a significant reduction of $x$–$y$ coupling on the emittance growth.

Optics measurements performed in 2003 showed that the Tevatron low-beta optics has been strongly mismatched. That resulted in a beta-wave through the entire machine and increased beta-functions in IPs with the corresponding luminosity loss of $\sim$15–20%. Optics correction in IPs was performed in the spring of 2004. It also yielded some reduction of optics mismatch in arcs but further improvements are still required.

## 2. Emittance growth due to $x$–$y$ coupling

To find the emittance growth due to $x$–$y$ coupling at injection we will parameterize the eigen vectors of coupled betatron motion in the following form:

$$\mathbf{v}_1 = \begin{bmatrix} \sqrt{\beta_{1x}} \\ -\frac{i(1-u)+\alpha_{1x}}{\sqrt{\beta_{1x}}} \\ \sqrt{\beta_{1y}}e^{i\nu_1} \\ -\frac{iu+\alpha_{1y}}{\sqrt{\beta_{1y}}}e^{i\nu_1} \end{bmatrix},\ \mathbf{v}_2 = \begin{bmatrix} \sqrt{\beta_{2x}}e^{i\nu_2} \\ -\frac{iu+\alpha_{2x}}{\sqrt{\beta_{2x}}}e^{i\nu_2} \\ \sqrt{\beta_{2y}} \\ -\frac{i(1-u)+\alpha_{2y}}{\sqrt{\beta_{2y}}} \end{bmatrix}. \tag{1}$$

where $\beta_{nx,ny}$ and $\alpha_{nx,ny}$ are the generalized beta- and alpha-functions, and parameters $u$, $\nu_1$ and $\nu_2$ are determined by the symplecticity conditions. Then, the particle motion can be written as,

$$\mathbf{x} = \frac{1}{2}\left(\sqrt{2\varepsilon_1}(\mathbf{v}_1 e^{i\mu_1} + \mathbf{v}_1^* e^{-i\mu_1}) + \sqrt{2\varepsilon_2}(\mathbf{v}_2 e^{i\mu_2} + \mathbf{v}_2^* e^{-i\mu_2})\right), \tag{2}$$

where $\varepsilon_1$ and $\varepsilon_2$ are the rms single particle emittances, and $\mu_1$ and $\mu_2$ are the betatron phase advances. Multiplying each side of Eq. (2) by $\mathbf{v}_1^+\mathbf{U}$ or $\mathbf{v}_2^+\mathbf{U}$ and using orthogonality conditions for eigen vectors we obtain

$$\varepsilon_1 = \frac{1}{2}(\mathbf{v}_1^+\mathbf{U}\mathbf{x})^2,\ \varepsilon_2 = \frac{1}{2}(\mathbf{v}_2^+\mathbf{U}\mathbf{x})^2. \tag{3}$$

Here $+$ denotes the Hermite conjugate vector, and $\mathbf{U}$ is the symplectic unit matrix:

$$\mathbf{U} = \begin{bmatrix} 0 & \mathbf{u} \\ \mathbf{u} & 0 \end{bmatrix},\ \mathbf{u} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \tag{4}$$

For Gaussian distribution the beam distribution function at the injection point can be written in the following form:

$$f(\mathbf{x}) = \frac{1}{4\pi^2\varepsilon_1\varepsilon_2}\exp\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\boldsymbol{\Xi}\mathbf{x}\right), \tag{5}$$

where matrix $\boldsymbol{\Xi}$ is determined by the eigen vectors, $\mathbf{v}_{t1}$ and $\mathbf{v}_{t2,}$ of the incoming beam

$$\boldsymbol{\Xi} = \mathbf{U}\mathbf{V}_t\boldsymbol{\Xi}'\mathbf{V}_t^{\mathrm{T}}\mathbf{U},\ \boldsymbol{\Xi}' = \begin{bmatrix} \varepsilon_1^{-1}\mathbf{I} & 0 \\ 0 & \varepsilon_2^{-1}\mathbf{I} \end{bmatrix},\ \mathbf{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$
$$\mathbf{V}_t = [\mathrm{Re}\,\mathbf{v}_{t1},\quad -\mathrm{Im}\,\mathbf{v}_{t1},\quad \mathrm{Re}\,\mathbf{v}_{t2},\quad -\mathrm{Im}\,\mathbf{v}_{t2}]. \tag{6}$$

If the injected beam is not matched to the ring lattice the decoherence of beam envelope oscillations leads to the emittance growth. After few thousand turns the system comes to equilibrium, and the emittance of the injected beam is determined by the following equation:

$$\varepsilon_i' = \frac{1}{8\pi^2\varepsilon_1\varepsilon_2}\int dx^4(\mathbf{v}_i^+\mathbf{U}\mathbf{x})^2\exp\left(-\frac{1}{2}\mathbf{x}^{\mathrm{T}}\boldsymbol{\Xi}\mathbf{x}\right),\quad i = 1,2. \tag{7}$$

For initially uncoupled beam characterized by $\beta_x$, $\alpha_x$, $\beta_y$ and $\alpha_y$ that yields

$$\begin{aligned}
\varepsilon_1' &= \varepsilon_1 A_{11} + \varepsilon_2 A_{12} \\
\varepsilon_2' &= \varepsilon_1 A_{21} + \varepsilon_2 A_{22} \\
A_{12} &= \frac{1}{2}\left(\frac{\beta_y}{\beta_{1y}}[u^2+\alpha_{1y}^2] + \frac{\beta_{1y}}{\beta_y}[1+\alpha_y^2] - 2\alpha_{1y}\alpha_y\right), \\
A_{11} &= \frac{1}{2}\left(\frac{\beta_x}{\beta_{1x}}[(1-u)^2+\alpha_{1x}^2] + \frac{\beta_{1x}}{\beta_x}[1+\alpha_x^2] - 2\alpha_{1x}\alpha_x\right), \\
A_{21} &= \frac{1}{2}\left(\frac{\beta_x}{\beta_{2x}}[u^2+\alpha_{2x}^2] + \frac{\beta_{2x}}{\beta_x}[1+\alpha_x^2] - 2\alpha_{2x}\alpha_x\right), \\
A_{22} &= \frac{1}{2}\left(\frac{\beta_y}{\beta_{2y}}[(1-u)^2+\alpha_{2y}^2] + \frac{\beta_{2y}}{\beta_y}[1+\alpha_y^2] - 2\alpha_{2y}\alpha_y\right).
\end{aligned} \tag{8}$$

The developed Tevatron optics model is based on the results of differential optics measurements and takes coupling into account. In addition to coherent skew-quadrupole component in dipoles it also includes a few dozen local focusing and coupling errors scattered through the entire ring. Table 1 presents coupled Twiss parameters predicted by the model (see Section 3). Substituting these parameters into Eq. (8) yields that the emittance growth due to coupling was about 15% for each of two transfers before shimming of dipoles. This value dropped to about 3% after the shimming of dipoles.

Table 1
Twiss parameters at proton and antiproton injection points before shimming of dipoles

| | Proton injection | | Antiproton injection | |
|---|---|---|---|---|
| | Mode 1 | Mode 2 | Mode 1 | Mode 2 |
| $\beta_x$ | 102.0 m | 4.71 m | 84.34 m | 4.04 |
| $\alpha_x$ | −0.8348 | −0.0321 | −0.6825 | −0.03233 |
| $\beta_y$ | 2.96 m | 65.42 m | 3.77 m | 75.64 m |
| $\alpha_y$ | 0.0126 | 0.2019 | 0.0294 | 0.5227 |
| $\nu$ | −110.9° | −112.47° | −128.5° | −126.9° |
| $u$ | −0.0609 | | 0.0539 | |

## 3. Optics measurements and correction

Optics correction have been complicated by limited time for optics measurements and the poor performance of the 30-year-old beam position monitors (BPM), in particular, by their low accuracy (~150 μm rms resolution) and malfunctioning ring-wide turn-by-turn mode. Presently, the only reliable way for optics measurements is the differential optics measurements, i.e. the orbit response to a single corrector bump. During the last 2 years we mainly used the fast measurements, where only four correctors and an energy change are exercised. In this case, the measurement takes about 5 min and getting time for the measurements is not a problem. The measurements are fully automated. Software records the reference orbit and its change in both planes due to a perturbation. Two correctors in each plane are chosen so that the betatron phase advance between them would be close to $\pi/2$. Such set of measurements provides minimum information sufficient to restore all details of linear details. Off-line analysis yielded quad and skew-quad optics corrections matching the measured and computed orbits. The accuracy of the model has been tested by measurements of tune shifts due to small focusing change of a few single power supply quads. An agreement within ~15% has been found for beta-function measurements. Comparatively large discrepancy is related to the fact that Tevatron tunes ($Q_x = 20.585$, $Q_y = 20.575$) are quite close to half integer resonance. That additionally amplifies effects of optics errors on the beta-functions.

Data analysis at injection yielded the following results. There is a systematic difference between main bus superconducting dipoles and quads. As compared to 30-year old magnetic measurements the quads are ~0.15% stronger. There is aforementioned systematic skew-quad field in dipoles $G_sA/B_0 \sim 1.4 \times 10^{-4}$ for $A = 2.54$ cm. It is in good agreement with the measured displacement of coils, ~150 μm. There are significant non-systematic (point-like) focusing and skew-focusing errors scattered through the entire machine. We used ~30 quad and/or skew-quad corrections with strengths [0.5–2%] of the main bus quad strength. The most striking finding was that focusing errors are about 0.5% for the final focus quadrupoles.

Similar to the injection we performed optics measurements for collision optics (low beta), where errors of focusing in the interaction region quads dominate errors in the sectors. The following conclusions were drawn out. There is a systematic difference between main bus superconducting dipoles and quads (~0.18%) consistent with measured at the injection energy. Systematic skew-quad field in dipoles $G_sA/B_0 \sim 2.1 \times 10^{-4}$ for $A = 2.54$ cm which is 1.5 times higher than at injection. The origin of this discrepancy is unknown. Similar to the injection energy there are significant non-systematic (point-like) focusing and skew-focusing errors scattered throughout the entire machine. Although there is good correlation for large optics errors it is clearly different in details. Interaction region quads need to be fudged up to 1%. It is well above what one would expect. Additionally, there is about 0.1% difference for quads of the same design. Model exhibited that due to optics mismatches the beta-functions in IPs were ~30% above design value and there was significant betatron function mismatch through entire machine.

Optics measurements proved that the resonance perturbation of beta-function dominated other optics discrepancies. Taking into account that analysis of differential orbit data is slow for on-line optics correction we used the optics measurements based on the measurements of tune shifts due to strength changes of one of four designated quads while the optics correction was performed using four other quads. Injection optics was corrected in 2003. The first step of the collision optics correction was carried out to correct optics mismatches in the IPs but there is still considerable mismatch in arcs.

## 4. Extended differential orbit measurements

Although the differential optics measurements are fast, data analysis is tedious and the results are not sufficiently accurate. Recently, we introduced the extended measurements using Linear Optics from Closed Orbit (LOCO) technique [1–3], where the single corrector orbit bumps are produced using about half of available correctors. That large redundancy in the data results in significant improvement of optics model accuracy. In this case, a single measurement takes much longer (1–2 h) and getting time for measurements is a problem but it is paid by improvement of the machine model. The effort has been built as collaboration between FNAL and ANL and aimed to upgrade the ANL software developed for APS [3]. The major objectives are software integration with FNAL data structures and optics software, taking into account strong $x$–$y$ coupling, and fitting data of dispersion measurements.

The idea of the method is to acquire large amount of data so that the systematic errors in the measurements could be determined. The unknowns are strengths and rolls of quads, strengths and rolls of correctors, and BPM responses and rolls. That amounts to about 800 unknowns. To have sufficient redundancy we perform measurements with ~100 correctors which yields ~20,000 equations. The

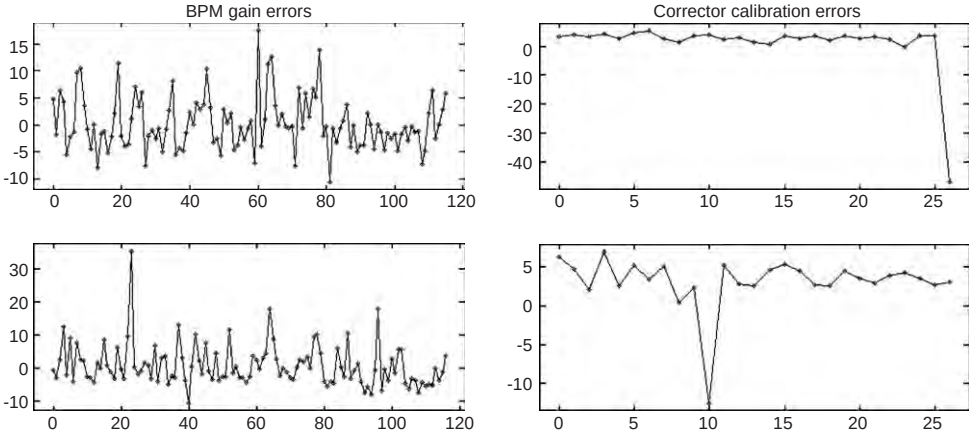



Fig. 1. Computed errors for differential BPM response (left) and corrector calibration errors (right); top pictures correspond to the horizontal plane, and bottom pictures to the vertical plane.

singular value decomposition [4] (SVD) is used to find unknowns from the measured data.

The first results of data analyses have been recently obtained. Further work is required to finish the project. Data analysis performed for collision optics yielded significant improvement of the fit accuracy. Compared to the model obtained before with differential orbit measurements the rms difference between measured and predicted differential orbits has been decreased from 10% to 1.5% of orbit oscillation amplitude. Fig. 1 presents the computed errors for differential BPM response and corrector calibration errors.

## 5. Conclusions

We built a reliable model of Tevatron optics based on the differential orbit measurements. The model includes actual power supply currents and has global and local focusing corrections to match the model to the measurements. It has been used to correct Tevatron optics and $x$–$y$ coupling. That resulted in a smaller emittance growth at injection (~10–15%), and reducing beta-functions in IPs from ~45 cm to the design values of 35 cm. These optics corrections contributed to collider luminosity growth of ~20–30%. Presently, further improvements of the model are limited by poor accuracy of BPMs and sufficiently long measurement time. To address it we are upgrading BPM electronics. That will boost BPM accuracy from ~150 μm to 20 μm rms and will allow to acquire turn-by-turn data for all BPMs. Turn-by-turn measurements will significantly reduce the measurement time, but their effective use implies development of software to fully utilize redundancy in their data. In addition the turn-by-turn measurements will allow us to measure lattice non-linearities.

## Acknowledgements

The authors are grateful to Yu. Alexahin, V. Shiltsev and M. Syphers for multiple discussions and support. We also would like to thank the entire Tevatron department for their help in carrying out of the measurements and to thank K.-J.-Kim and K. Harkay for stimulating this work and support.

## References

[1] W.J. Corbett, M. Lee, V. Ziemann, PAC"93, Washington, DC, 1993, p. 108.
[2] J. Safranec, Nucl. Instr. and Meth. A 388 (1997) 27.
[3] V. Sajaev, L. Emery, EPAC"02, Paris, France, 2002, p. 742.
[4] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes, Cambridge University Press, Cambridge, 1986.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 303–307
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Simulation of crystalline beams in storage rings using molecular dynamics technique

I. Meshkov[a], T. Katayama[b], A. Sidorin[a], A. Smirnov[a,*], E. Syresin[a], G. Trubnikov[a], H. Tsutsui[b]

[a]*Joint Institute for Nuclear Research (JINR), Joliot Curie 6, Dubna 141980, Russian Federation*
[b]*Institute of Physical and Chemical Research (RIKEN), Hirosawa 2-1, Wako 351-0198, Japan*

Available online 28 November 2005

**Abstract**

Achieving very low temperatures in the beam rest frame can present new possibilities in accelerator physics. Increasing luminosity in the collider and in experiments with targets is a very important asset for investigating rare radioactive isotopes. The ordered state of circulating ion beams was observed at several storage rings: NAP-M [Budker, et al., in: Proceedings of the 4th All-Union Conference on Charged-Particle Accelerators [in Russian], vol. 2, Nauka, Moscow, 1975, p. 309; Budker et al., Part. Accel. 7 (1976) 197; Budker et al., At. Energ. 40 (1976) 49. E. Dementev, N. Dykansky, A. Medvedko et al., Prep. CERN/PS/AA 79-41, Geneva, 1979] (Novosibirsk), ESR [M. Steck et al., Phys. Rev. Lett. 77 (1996) 3803] and SIS [Hasse and Steck, Ordered ion beams, in: Proceeding of EPAC '2000] (Darmstadt), CRYRING [Danared et al., Observation of ordered ion beams in CRYRING, in: Proceeding of PAC '2001] (Stockholm) and PALLAS [Schramm et al., in: J.L. Duggan (Eds.), Proceedings of the Conference on Appl. of Acc. in Research and Industry AIP Conference Proceedings, p. 576 (to be published)] (Munich). In this report, the simulation of 1D crystalline beams with BETACOOL code is presented. The sudden reduction of momentum spread in the ESR experiment is described with this code. Simulation shows good agreement with experimental results and also with the intrabeam scattering (IBS) theory [Martini, Intrabeam scattering in the ACOOL-AA machines, CERN PS/84-9 AA, Geneva, 1984]. The code was used to calculate characteristics of the ordered state of ion beams for the TARN-II [Katayama, TARN II project, in: Proceedings of the IUCF workshop on nuclear physics with stored cooled beams, Spencer, IN, USA, 1984].


*PACS:* 29.27. Bd

*Keywords:* Electron cooling; Crystalline ion beams

## 1. Introduction

### 1.1. Motion equations

To calculate crystalline beams, pairs of canonical variables are chosen. Transverse motion is described by the traditional variables: coordinate and transverse momentum normalized on the longitudinal momentum. Longitudinal motion is defined as arrival time and longitudinal momentum as spread. The vector of canonical variables is as follows:

$$X = \left(x, p_x = \frac{P_X}{P_S},\ y,\ p_y = \frac{P_Y}{P_S},\right.$$
$$\left. z = -(t - t_0)\beta_0 c,\ p_z = \frac{E - E_0}{P_S \beta_0 c}\right), \quad (1)$$

where $x$, $y$ are the horizontal and vertical positions, $p_x$, $p_y$ are the corresponding normalized momenta, $z$ is the arrival time of the particle times $-c\beta_0$, $p_z$ is the normalized momentum spread $p_z = \Delta P/P_S$.

*Corresponding author. Tel.: +7 09621 564479; fax: +7 09621 65322.
*E-mail address:* smirnov@jinr.ru (A. Smirnov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.028

A Hamiltonian for ordered state simulation is used with independent variable $s$:

$$H = -\frac{xp_z}{\rho}+\frac{p_z^2}{2\gamma_0^2}+\frac{p_x^2+p_y^2}{2}+\frac{1}{2}\left(\frac{1}{\rho^2}+K\right)x^2-\frac{K}{2}y^2 + \frac{\lambda}{6}(x^3-3xy^2) + \frac{r_{\text{ion}}}{\gamma_0^2\beta_0^2}\sum_i \frac{1}{\sqrt{(x-x_i)^2+(y-y_i)^2+\gamma_0^2(z-z_i)^2}}, \quad (2)$$

where $\rho$ is the curvature radius of the reference orbit. The effect of multipole components of the magnets is included with parameters $K$ and $\lambda$, which are defined by

$$K \equiv \frac{1}{B\rho}\frac{\partial B_y}{\partial x}, \quad \lambda \equiv \frac{1}{B\rho}\frac{\partial^2 B_y}{\partial x^2}, \quad r_{\text{ion}} = \frac{Z^2}{A}\frac{e^2}{4\pi\varepsilon_0 m_u c^2}, \quad (3)$$

where $B\rho = p_0/q$ is the magnetic rigidity.

### 1.2. Molecular dynamics

The last element in Eq. (2) describes the Coulomb interaction between particles. Since the calculation of the space–charge effect is time consuming, a molecular dynamics (MD) technique, with periodic boundary condition, is used. When ions of charge $q = Ze$ are put at position $(s, r) = (0,0),(\pm L,0),(\pm 2L,0),\ldots$ (Fig. 1), the last element in the region $|s| < L/2$ is [1–8]

$$U_{\text{sc}}(s,r) = \frac{1}{4\pi\varepsilon_0}\left(\frac{q}{a}+\frac{2q}{L}\int_0^\infty \frac{J_0(kr/L)\cosh(ks/L)-1}{\exp(k)-1}\,\mathrm{d}k\right), \quad (4)$$

where $q$ is particle charge, $a = \sqrt{s^2+r^2}$ is distance between particles, $s = z - z_i$, $r = \sqrt{(x-x_i)^2+(y-y_i)^2}$, $L$ is MD cell size, $J_0$ is the Bessel function of 0th order. The space–charge force vector from the $i$th particle, with the MD technique, can be derived from Eq. (4)

$$F_{px} = \frac{1}{4\pi\varepsilon_0}\frac{q^2(x-x_i)}{m_0c^2\gamma_0^2\beta_0^2}\left(\frac{1}{a^3}-\frac{2I_1}{rL^2}\right),$$
$$F_{py} = \frac{1}{4\pi\varepsilon_0}\frac{q^2(y-y_i)}{m_0c^2\gamma_0^2\beta_0^2}\left(\frac{1}{a^3}-\frac{2I_1}{rL^2}\right),$$
$$F_{pz} = \frac{1}{4\pi\varepsilon_0}\frac{q^2(z-z_i)}{m_0c^2\gamma_0^2\beta_0^2}\left(\frac{1}{a^3}-\frac{2I_0}{sL^2}\right), \quad (5)$$

where

$$I_1 = -\int_0^\infty \frac{kJ_1(kr/L)\sinh(ks/L)-1}{\exp(k)-1}\,\mathrm{d}k,$$
$$I_0 = -\int_0^\infty \frac{kJ_0(kr/L)\cosh(ks/L)-1}{\exp(k)-1}\,\mathrm{d}k,$$

$J_1$ is the Bessel function of first order, $m_0c^2$ is ion rest mass. Integrals $I_1$ and $I_0$ were numerically calculated and are used in the program as table values. With this MD technique, the dynamics of the ordered state of ions in a storage ring is usually simulated with $N_{\text{p}} = 10/100$ particles.

### 1.3. Crystallization conditions

The main criterion for beam crystallization (orderliness) is a decrease in particle temperature, lower than inter-particle potential energy, which can be described by the plasma parameter:

$$\Gamma = \frac{U}{T} = \frac{1}{4\pi\varepsilon_0}\frac{Z^2e^2}{aT} > 1, \quad (6)$$

where $U$ and $T$ are potential energy and temperature of the ion beam, respectively. $Ze$ is the charge of particles, $a$ is average inter-particle distance. 3D crystals are obtained for plasma parameter $\Gamma \gg 150$.

The next condition is related to the optical structure of the storage ring. The storage ring must be alternating-gradient (AG) focusing and the beam energy must be less than the transition energy of the ring [9]:

$$\gamma < \gamma_{\text{T}}. \quad (7)$$

Another condition defines the periodicity of the ion storage ring. The ring lattice periodicity should be at least four times as high as the maximum betatron value [9]:

$$4\max\{Q_x, Q_y\} < \text{Periodicity}. \quad (8)$$

These conditions (Eqs. (6)–(8)) need to be satisfied for 3D crystalline beams. In the case of 1D ordered beam, other criteria were formulated [10]. The main criterion defines the situation where particles cannot pass each other in a longitudinal direction:

$$\Gamma_2 = \frac{1}{4\pi\varepsilon_0}\frac{Z^2e^2}{T_{\|}\sigma_\perp} > \pi, \quad (9)$$

where $T_{\|}$ is longitudinal temperature and $\sigma_\perp$ is transverse size of the ion beam. As will be shown below, the criterion $\Gamma_2$ plays a very important role in the process of beam ordering.

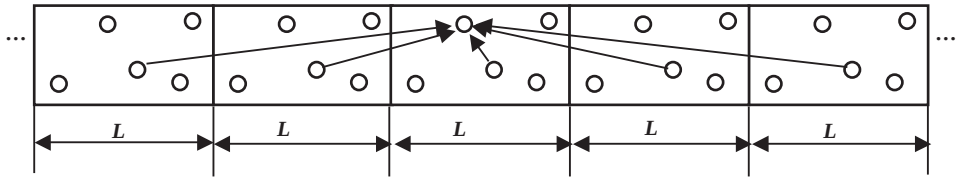



Fig. 1. Periodic distribution of five particles in MD cells.

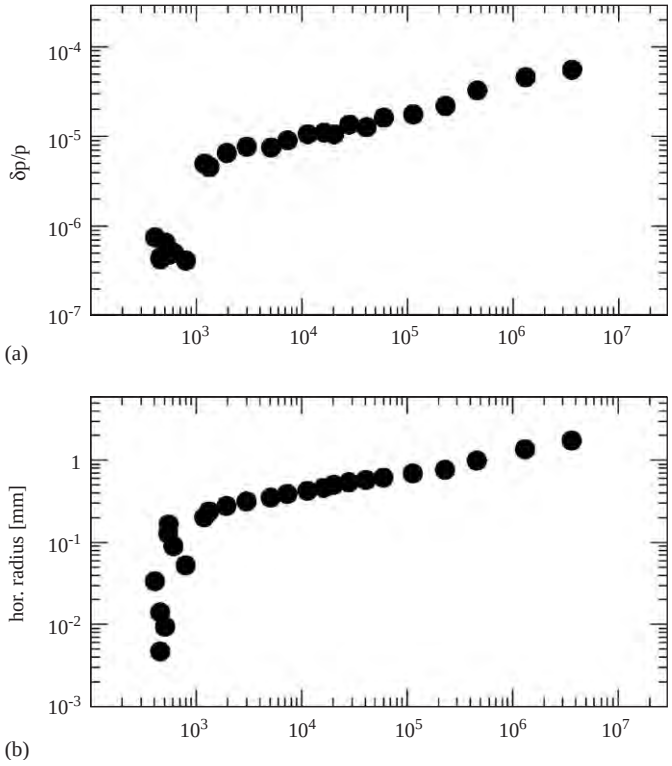



Fig. 2. Beam behavior in the ESR experiment showing the anomalous temperature reduction at low intensity. Momentum spread and beam size are given as functions of number of stored ions. $^{238}U^{92+}$ 400 MeV/u.

## 2. Ordered beams in storage rings

### 2.1. ESR experiments

Since the sudden reduction in momentum spread of a circulating proton beam was observed in NAP-M [1], the ordered state of ion beams has also been achieved on several storage rings. The most extensive experimental program was perfomed on the ESR [2]. The momentum spread reduction was observed for a wide range of ion species, except for very light ions.

On the ESR, the momentum spread of a uranium beam at 400 MeV/u, cooled by a 0.25 A electron beam, drops around 1000 stored ions, from a value of $\Delta p/p = 5 \times 10^{-6}$ to $\Delta p/p = 5 \times 10^{-7}$, corresponding to a change in longitudinal temperature of two orders of magnitude (Fig. 2) [11]. The scraper measurement allows derivation of the momentum spread and the beam radius as a function of the number of stored ions (Fig. 2). For the horizontal degree of freedom, a radius reduction from 0.2 mm to less than 0.01 mm is obvious. Despite the limited resolution, even for the vertical beam radius, a reduction is suggested by the data points at the transition point of the longitudinal momentum spread and the horizontal radius.

The behavior of ion beam values can be explained using 3D diagrams of beam parameter growth rates due to intrabeam scattering, which were calculated using a generalized Piwinski model [6]. The model presumes a sufficient number of particles and Gaussian distribution of the particles in all degrees of freedom. Horizontal and vertical emittances were chosen to be equal.

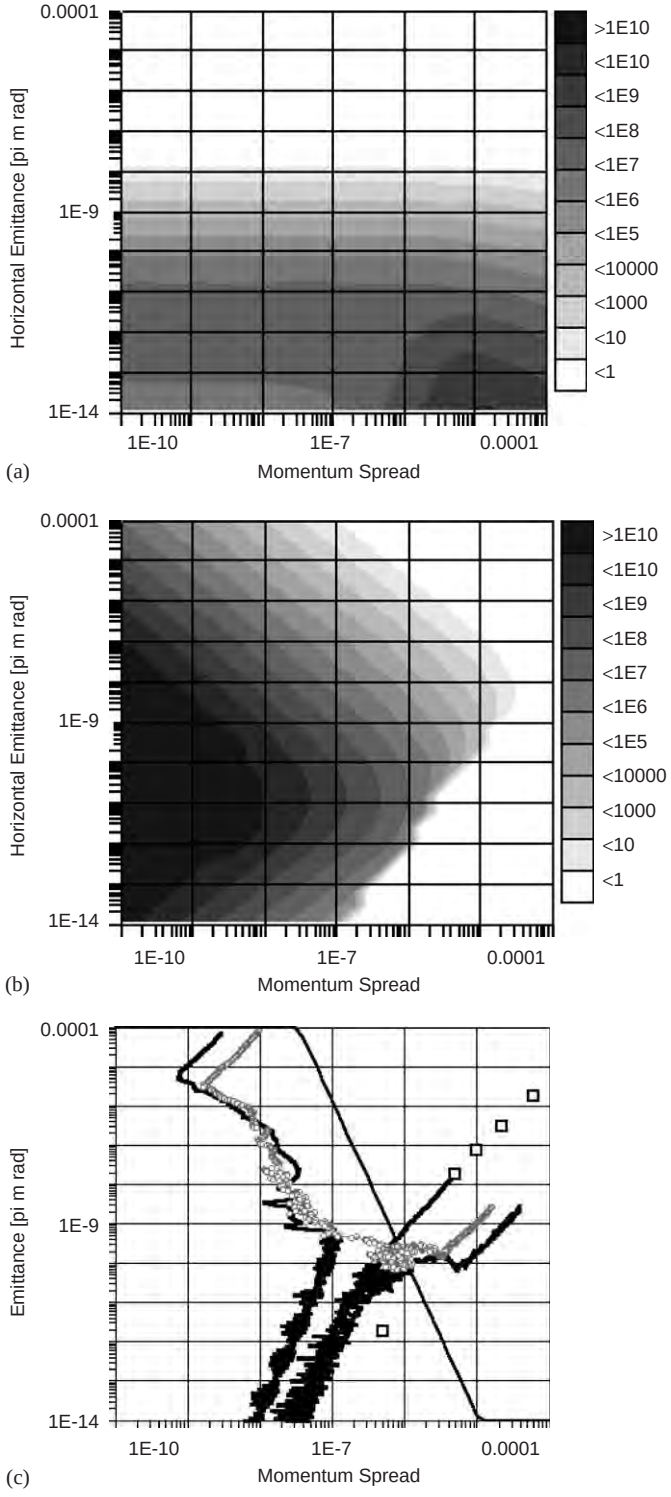



Fig. 3. Theoretical dependence of IBS growth rates on beam emittance and momentum spread for ESR: (a) horizontal and (b) longitudinal components of IBS growth rates. Results of MD calculations (c): evolution of ion beam parameters during cooling. Solid black line corresponds to cooling rate of $4 \times 10^4$ Hz. Grey circles correspond to cooling rate of $10^4$ Hz. Straight line is criterion $\Gamma_2 = \pi$ ($N = 5 \times 10^5$). Open squares are ESR experiment (see Fig. 2).

The condition $\Gamma_2 = \pi$ (Eq. (9)) describes the following relationship between beam emittance and momentum spread: $\varepsilon \sim (\Delta p/p)^{-4}$. In the twice-logarithmic scale, this dependence is presented in Fig. 3(c) as a black solid line, where the theoretical dependence of intrabeam scattering (IBS) on beam values at high particle numbers is compared with the results of numerical simulation using the MD method [12]. In ESR experiments, the magnitude of beam emittances and momentum spread before momentum reduction are defined by equilibrium between IBS growth and cooling rates (series of square points on Fig. 3(c)). The experimental points lie on the line, which approximately corresponds to equal values of longitudinal and transverse growth rates. Particle numbers decrease with time and, when the $\Gamma_2$ value exceeds $\pi$, cooling forces suppress the IBS forces and the beam reaches an ordered state (last experimental point in Fig. 3(c)).

The calculated results, using the MD technique (Fig. 3(c)), are in good qualitative agreement with the 3D diagram (Figs. 3(a) and (b)). Calculations were performed using the same particle numbers, i.e. $5 \times 10^{-5}$, which is three orders of magnitude larger than in the ESR experiments. Cooling rates were also chosen as three orders larger, in comparison with real electron cooling systems. In the first stage of beam cooling, all the lines have the same angular inclination, determined by the ratio between cooling rates in transverse and longitudinal degrees of freedom. In the case of uniform cooling ($\varepsilon \sim (\Delta p/p)^2$), one can see this dependence in the initial part of the beam phase trajectory, independent of initial point. This means that at a large initial phase, space of the beam IBS process does not effect the cooling process.

Prior to the ordered state, MD calculations are in good agreement with the position of maximum IBS growth rates, as predicted by the Piwinski model. No other additional heating was used in these simulations. In the ordered state, IBS growth rates, calculated using the MD method, are substantially less than predicted by the Piwinski model. At a cooling rate of $4 \times 10^4$ Hz and higher, beam emittance and momentum spread decrease to very small values.

### 2.2. Simulation for TARN-II

The 3D phase diagrams of IBS heating rates, in accordance with the Martitni model for the TARN-II ring [7], display the same behavior as the ESR ring (Figs. 3(a) and (b)). To study intrabeam scattering in the ordered state, the 3D phase diagram of heating rates was numerically simulated using the MD technique. Initial distribution was generated with the same distance between particles in a longitudinal direction. Transverse emittance and momentum are generated with a Gaussian distribution. This means that particles initially have only kinetic energy in the longitudinal plane. Growth rates are calculated after a few hundred cycles, when relaxation between kinetic and potential energies in the longitudinal plane occurs. The solid black line in Fig. 4 corresponds to

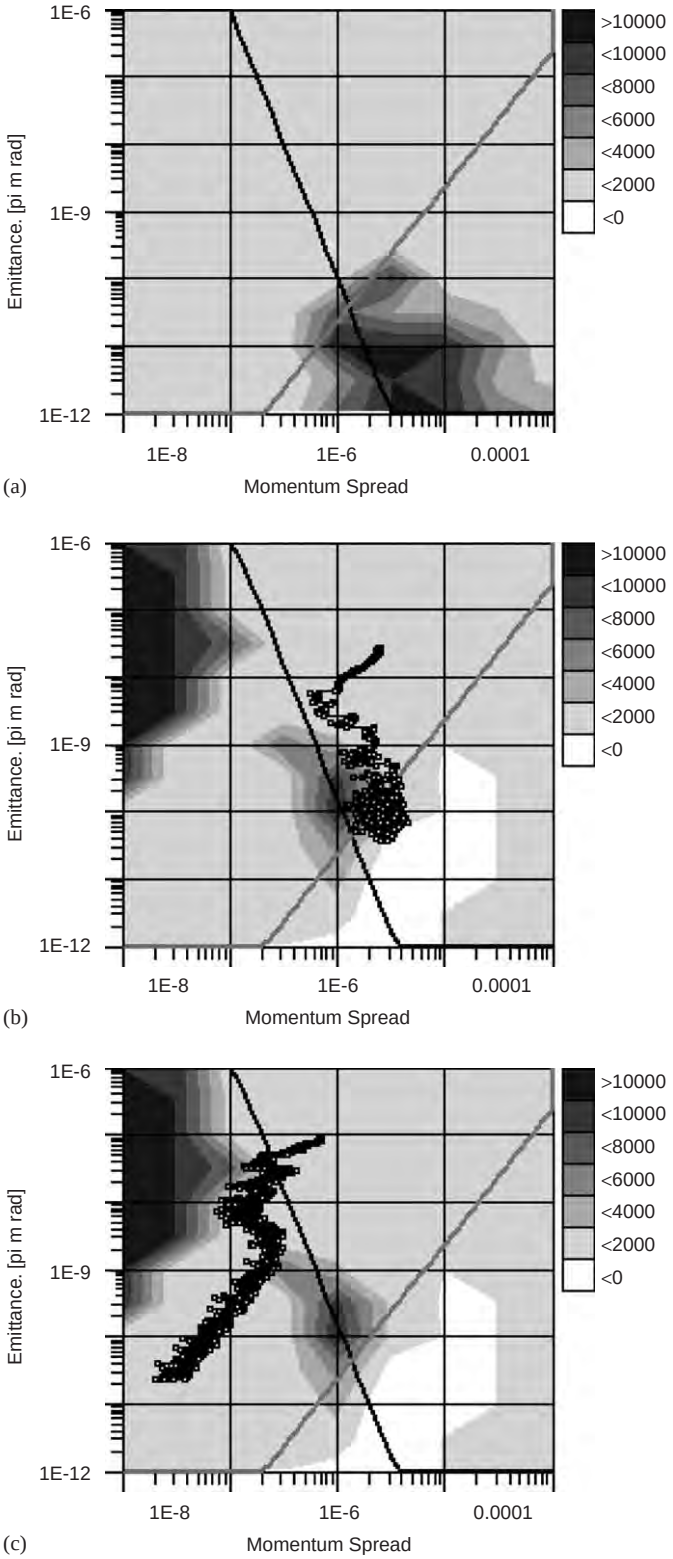



Fig. 4. MD simulation: (a) Horizontal component of IBS growth rates, (b) longitudinal component of IBS and (c) evolution of beam parameters during cooling process. $I_{ecool} = 5$ A, $^{132}Sn^{50+}$ 220 MeV/u, $N = 10^5$.

criterion $\Gamma_2 = \pi$; the gray line is the equilibrium between longitudinal and transverse temperatures. The intersection between criterion $\Gamma_2$ and temperature equilibrium is the final point before the sudden reduction of momentum

spread, which has been experimentally observed in a few storage rings.

The simulation result (Fig. 4) displays a number of differences from the theoretical model of IBS heating (Figs. 3(a) and (b)). Transverse components have very low heating rates for the range of momentum spread, which are located below the temperature equilibrium. Another big difference is the shape of the longitudinal component of IBS heating rate. It is divided into two parts. The first is defined by heating from an optic structure. The second part resembles an island whose height is linearly dependent on particle number and is well described by the IBS theory. The maximum of the IBS island is located near the intersection between criterion $\Gamma_2 = \pi$ and temperature equilibrium.

The longitudinal component of IBS heating breaks up when heating rates have very small values compared to the theoretical prediction. If the initial ion beam values are chosen near the break-up, then the ordered state for a large number of particle $N = 10^5$ can be achieved for a real cooling system with electron beam current $I_{\mathrm{ecool}} = 5\,\mathrm{A}$ (Fig. 4(c)).

To achieve ordered ion beams with a large number of particles and a realistic cooling force, a special strategy for the cooling process should be elaborated. When the ion beam remains in equilibrium between IBS and cooling, additional heating may be applied in the transverse direction. For example, heating by an RF-kicker can be used. Initially, the momentum spread will continue to decrease and emittances will increase. When the beam parameters have to satisfy the condition $T_\perp \gg T_\parallel$, the additional heating can be switched off and the ion beam will continue to cool down to the ordered state. The same idea is projected for other storage rings. Experimental verification of the new strategy for achieving an ordered ion beam with a large density can open new possibilities in accelerator physics.

## 3. Conclusion

Ion beam parameters, before momentum reduction, are defined by the equilibrium between cooling force and IBS. When cooling rate exceeds maximum IBS growth rate, the ion beam reaches an ordered state. The ESR case is studied in detail and simulation shows very good agreement with experimental results. Numerical simulation shows a large difference in the behavior of IBS heating in the ordered-state compared with the theoretical model. To increase particle numbers in the ordered state, a new strategy of cooling process is proposed. Additional heating in the transverse direction should be applied during the cooling process. For real cooling systems, the ordered state can be achieved with number of particles up to $10^5$.

## Acknowledgements

This work was supported by RFBR grant #02-02-16911, INTAS grant #03-54-5584.

## References

[1] G.I. Budker, N.S. Dikansky, V.I. Kudelainen, et al., in: Proceedings of the Fourth All-Union Conference on Charged-Particle Accelerators [in Russian], vol. 2, Nauka, Moscow, 1975, p. 309;
G.I. Budker, N.S. Dikansky, V.I. Kudelainen, et al., Part. Accel. 7 (1976) 197;
G.I. Budker, N.S. Dikansky, V.I. Kudelainen, et al., At. Energy 40 (1976) 49;
E. Dementev, N. Dykansky, A. Medvedko, et al., Prep. CERN/PS/AA 79-41, Geneva, 1979.
[2] M. Steck, K. Beckert, H. Eickhoff, et al., Phys. Rev. Lett. 77 (1996) 3803.
[3] R.W. Hasse, M. Steck, Ordered ion beams, in: Proceeding of EPAC 2000.
[4] H. Danared, A. Kallberg, K.G. Rensfelt, A. Simonsson, Observation of ordered ion beams in CRYRING, in: Proceeding of PAC '2001.
[5] U. Schramm, T. Schätz, D. Habs, Laser-cooling of ions and ion acceleration in the RF-quadrupole ring trap PALLAS, in: J.L. Duggan (Ed.), Proccedings of the 16th International Conference on the Application of Accelerators in Research and Industry (CAARI), Denton, TX, USA, 11/2000, AIP Conference Procedings, 〈http:\\proceedings.aip.org/proceedings/confproceed/576.jsp〉 576, 667 (2001).
[6] M. Martini, Intrabeam scattering in the ACOOL-AA machines, CERN PS/84-9 AA, Geneva, 1984.
[7] T. Katayama, TARN II project, in: Proceedings of the IUCF Workshop on Nuclear Physics with Stored Cooled Beams, Spencer, IN, USA, 1984.
[8] V. Avilov, Solid State Commun. 44 (4) (1982) 555.
[9] J. Wei, A. Draeseke, M. Sessler, X. Li, Diverse topics in crystalline beams, in: Proceedings of the 31st Workshop of the INFN Eloisatron Project, Erice, Italy, 12–21 November 1995, p. 229.
[10] T. Katayama, I. Meshkov, A. Sidorin, A. Smirnov, E. Syresin, Ordered state of ion beams, Preprint RIKEN-AF-AC-40, 2002.
[11] M. Steck, K. Beckert, P. Beller, B. Franzke, F. Nolden, J. Phys. B 36 (2003) 991.
[12] T. Katayama, I. Meshkov, A. Sidorin, A. Smirnov, E. Syresin, H. Tsitsui, D. Mohl, Simulation study of ordered ion beams, Preprint RIKEN-AF-AC-42, 2003.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 308–310
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Derivation of azimuthal harmonics series for stability analysis of the coherent oscillations of bunched beams

N.V. Mityanina

*Budker INP, Academician Lavrentyev 11, Novosibirsk 630090, Russia*

Available online 22 November 2005

**Abstract**

Stability analysis of the coherent motion of multibunch beams (including counter-rotating beams) should involve expressions analogous to the effective impedance—the series over harmonics of the revolution frequency of the impedance of the RF structure at the side frequencies to these harmonics—with certain factors depending on the harmonic number, such as the bunch line-density spectrum, the phase factor and the factor describing the order of multipole synchrotron oscillations. In this paper, we present a method for analytical summation of these series for resonant impedance, which does not seem to have been carried out previously for the general case including all the factors mentioned. Comparison of the expressions obtained with formulae used in previous papers shows the limits of validity of simpler approaches. The expressions obtained are used in the computer codes MBIM1 and MBIM2 presented at this conference, which calculate coherent oscillation stability for arbitrary multibunch beams.


*PACS:* 29.27.Bd

*Keywords:* Beam dynamics; Coherent instability; Multibunch instability

## 1. The series to be considered

The following series is summarized:

$$S(N,\theta)=\sum_{m=-\infty}^{\infty} m^N \mathrm{e}^{\mathrm{i}m\theta}\mathrm{e}^{-m^2\phi_0^2}Z_m^+ \tag{1}$$

where $Z_m^+=Z(-\mathrm{i}(m\omega_0+\Omega))$, $Z(-\mathrm{i}\omega)$ is the RF cavity impedance, $\Omega$ is the frequency of coherent oscillations, $\omega_0$ is the revolution frequency, $\theta$ is the angular distance between two bunches of the beam, and $\phi_0$ is the angular root mean square (rms) length of a Gaussian bunch. The power order $N$ is a positive integer, and for different problems can be both odd and even.

Such a series differs from those considered, for example, in Ref. [1], first, by the factor $\mathrm{e}^{-m^2\phi_0^2}$, which describes the square of the spectral density of a Gaussian bunch, and second, by the fact that the order of the numerator of the algebraic factor of a term of the series is higher than that of the denominator.

The formulae for summation of the series (1) are given below for the impedance with characteristic resistance $\rho$, resonant frequency $\omega_\mathrm{r}=\omega_0 m_\mathrm{r}$ and quality factor $Q$, as follows:

$$Z(s)=\frac{\rho s\omega_\mathrm{r}}{(s-s_1)(s-s_2)}=Z^1(s)+Z^2(s) \tag{2}$$

$$Z^{1,2}(s)=\frac{\rho\omega_\mathrm{r}s_{1,2}}{(s_{1,2}-s_{2,1})(s-s_{1,2})}$$

$$s=-\mathrm{i}\omega,\quad s_{1,2}=\omega_\mathrm{r}\left(\pm\mathrm{i}\sqrt{1-(1/2Q)^2}-(1/2Q)\right).$$

The impedance equation (2) contains two resonant summands. It is sufficient to carry out summation for the first of them, $Z^1(s)$.

The details of these derivations are given in Ref. [4].

*E-mail address:* mityanina@inp.nsk.su.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.029

## 2. Separation of the series into convenient parts

For summation, the series in Eq. (1) should be separated into two parts:

$$S^1(N,\theta) = \sum_{m=-\infty}^{\infty} m^N e^{im\theta} e^{-m^2\phi_0^2} (Z^1)_m^+ = S_P^1(N,\theta) + S_W^1(N,\theta)$$

$$S_P^1(N,\theta) = iR_1 \sum_{m=-\infty}^{\infty} (m^N - m_1^N) \frac{e^{im\theta - m^2\phi_0^2}}{m - m_1} \quad (3)$$

$$S_W^1(N,\theta) = iR_1 \sum_{m=-\infty}^{\infty} m_1^N \frac{e^{im\theta - m^2\phi_0^2}}{m - m_1} \quad (4)$$

where $m_{1,2} = i(s_{1,2} + i\Omega)/\omega_0$, $R_1 = \rho s_1 m_r/(s_1 - s_2)$.

The series $S_W^1(N,\theta)$, containing all the poles and decreasing sufficiently quickly at harmonic numbers $|m| \to \infty$, can be summed analytically, similar to the Watson–Sommerfeld transformation [2]. The second series, $S_P^1(N,\theta)$, containing no poles, with the help of the Poisson formula [3], can be transformed into an exponentially converging series, from which in most cases it is possible to retain only one or two terms.

## 3. Summation of the term containing poles

The series in Eq. (4) can be summed with the help of the Watson–Sommerfeld transformation [2] with some modifications. For this, auxiliary functions $\phi^{\pm}(z) = f(z)e^{\pm i\pi z}/\sin(\pi z)$ should be considered, where $f(z) = e^{iz\theta - z^2\phi_0^2}/(z - m_1)$, and their integrals over the contours of integration $C^{\pm}$, consisting of the real axis and semicircles of infinite radius in the upper and lower semi-planes $S^{\pm}$. In Ref. [2], the integrals over $S^{\pm}$ had zero limits at infinity, but here they have finite limits, which were calculated in Ref. [4]. For $0 \leqslant \theta < 2\pi$, we obtain

$$\sum_{m=-\infty}^{\infty} \frac{e^{i\theta m - m^2\phi_0^2}}{m - m_1} = \pi e^{-m_1^2\phi_0^2} e^{i\theta m_1} \times \left\{ -(\mathrm{ctg}(m_1\pi) - i) - i \cdot \mathrm{erfc}\left( im_1\phi_0 + \frac{\theta}{2\phi_0} \right) + \frac{1}{2} e^{2i\pi m_1} \mathrm{erfc}\left( im_1\phi_0 + \frac{(2\pi + \theta)}{2\phi_0} \right) - \frac{1}{2} e^{-2i\pi m_1} \mathrm{erfc}\left( -im_1\phi_0 + \frac{(2\pi - \theta)}{2\phi_0} \right) \right\}. \quad (5)$$

Usually, $\phi_0$ is much less than $\theta_{\min}$, the minimal angular distance between bunches, which is why, with erf expansion ([5], Section 7.1.6) at $\phi_0 \ll \theta_{\min}$, we can simplify Eq. (5) and finally obtain

$$S_W^1(N,\theta) = -\pi R_1 m_1^N \left\{ e^{i\theta m_1 - m_1^2\phi_0^2} (\mathrm{ctg}(m_1\pi) - i(1 - \delta_{\theta,0})) + \delta_{\theta,0} \frac{2m_1\phi_0}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-2m_1^2\phi_0^2)^k}{(2k+1)!!} \right\}. \quad (6)$$

In the limit $\phi_0 \to 0$, at $N = 0$, the expression obtained (Eq. (6)) coincides with that obtained in Ref. [1]. Eq. (6) shows that at $\theta = 0$, the correcting term (in comparison with [1]) obtained here depends on the ratio of the bunch length to the wavelength of the resonant impedance considered. For wavelengths comparable to the bunch length, it is necessary to take into account this correcting term. The sum over $k$ in Eq. (6) converges very quickly and usually only a few of its terms can be retained.

## 4. Summation of the term containing no poles

The series in Eq. (3) can be written in the form

$$S_P^1(N,\theta) = iR_1 \sum_{k=0}^{N-1} m_1^{N-k-1} C_k \quad (7)$$

$$C_k = C_k(\theta,\phi_0) = \sum_{m=-\infty}^{\infty} m^k e^{-m^2\phi_0^2} e^{i\theta m}. \quad (8)$$

For calculation of Eq. (8) we can apply the Poisson formula (see, for example, Ref. [3, Eq. (11.1)], for which the conditions of applicability are fulfilled. According to Section 2.3.15.9 in Ref. [6],

$$C_k = \sum_{n=-\infty}^{\infty} P(\theta + 2\pi n, k) \quad (9)$$

$$P(\theta + 2\pi n, k) = \int_{-\infty}^{\infty} x^k e^{ix(\theta + 2\pi n) - x^2\phi_0^2} dx = k! \frac{\sqrt{2\pi}}{(\phi_0\sqrt{2})^{k+1}} e^{-(\theta + 2\pi n)^2/4\phi_0^2} \times \sum_{l=0}^{[k/2]} \frac{(-i(\theta + 2\pi n)/(\phi_0\sqrt{2}))^{k-2l}}{(2l)!!(k-2l)!!}. \quad (10)$$

As $\phi_0 \ll 2\pi$, at $\theta = 0$ it is possible to drop from Eq. (9) the terms with $n \neq 0$, and at $\theta > 0$, all terms except those for $n = 0, -1$.

The total expression for the sum of the series in Eq. (1) is calculated by substituting Eqs. (9) and (10) into Eq. (7), adding Eq. (6) (or Eq. (5)) and summing the terms for $Z^1(s)$ and $Z^2(s)$.

## 5. Additional notes

1. The same method can be applied to series in terms of normal symmetric modes, analogous to Eq. (1), but with replacement of $\sum_m f(m) \to \sum_p f(pN_b + k)$ ($N_b$ is the number of bunches in a symmetric beam and $k$ is the

number of the normal mode). The details are given in Ref. [4].

2. The same method can be applied in the case of transverse oscillations for $N = 0$ and for the transverse impedance $Z_t(s) = i\rho_t\omega_r^2/(s - s_1)(s - s_2)$. In this case, first, the size of a bunch can be neglected because the impedance quickly decreases with increasing $|m|$, and second, the summable function decreases sufficiently quickly at $|m| \to \infty$, so summation formulae similar to those given in Ref. [1] can be applied, as well as direct application of the Watson–Sommerfeld formula.

## 6. Contribution of the non-resonant summand

The expression for the sum of the series in Eq. (1) contains two different parts: the resonant summand, containing $(\mathrm{ctg}(m_j\pi) - i(1 - \delta_{\theta,0}))$ (see Eq. (6)) and the non-resonant summand (all other terms).

Usually, papers devoted to the summation of similar series take into account only the first of these summands, as, for example, in Ref. [7].

Using the formulae given in this paper, we can show that the contribution of the second (non-resonant) summand at $N > 0$ should not be neglected if $m_r$ essentially differs from integer numbers and at low quality factors.

## References

[1] B. Zotter, On the summation of infinite algebraic and Fourier series, CERN/ISR-TH/78-9, CERN, Geneva, April 1978.
[2] M. Month, R.F. Peierls, Nucl. Instr. and Meth. 137 (1976) 299 (Appendix A).
[3] M.V. Fedoryuk, Asymptotic Integrals and Series, 1987 (in Russian).
[4] N. Mityanina, Summation of series over azimuthal harmonics for the analysis of coherent oscillation stability, Budker INP 2004-34, BINP, Novosibirsk, 2004, preprint (in Russian).
[5] M. Aramovitz, I. Stegun (Eds.), Handbook of Mathematical Functions, 1979.
[6] A.P. Prudnikov, Yu.A. Brychkov, O.I. Marichev, Integrals and Series: Elementary Functions, 1981 (in Russian).
[7] K. Balewski, R.D. Kohaupt, Analytic evaluation of the effective impedance for coupled bunch instabilities, DESY Document 90-152, DESY, Hamburg, 1990.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 311–313
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Approximation technique for solving the Vlasov–Poisson problem

Z. Parsa[a,*], V. Zadorozhny[b,1]

[a]Brookhaven National Laboratory, Physics Department 510A, Upton, NY, 11973–5000, USA
[b]Institute of Cybernetics, Kiev, Ukraine

Available online 28 November 2005

**Abstract**

In this paper, we describe the beam distribution in particle accelerators in the framework of a Vlasov–Poisson scheme. A new approach to the investigation and numerical simulation is based on the property of an universality of Maxwell equations, and Ł-*moment problem*. In this scheme, it is possible to reduce a problem of an optimal stabilization of the given motion to the Ł-moment problem using its regular approximation technique and well developed computational procedures.


*PACS:* 23.23.+x; 56.65.Dy; 29.17.w

*Keywords:* *L*-moment problem; Vlasov–Poisson equation

## 1. The problem

A new approach to studying a nonlinear bunched beam dynamics based on the self-consistent Vlasov–Poisson equations

$$\partial_t f + v \cdot \partial_x f + F \cdot \partial_v f = 0 \tag{1}$$

where

$$F = \frac{q}{m} \operatorname{grad} U + \frac{e}{mc}(v \times H)$$

is the self-consistent force.

As known, the Vlasov equation is used for a description of collisionless plasma in which charged particles motion is given by a set of Maxwell equation for the electromagnetic field. Many special and approximate (wave-like) solution to the Vlasov equation are known and they describe important physical effects. One of the most well-known effects is the Landau damping which is found from the dispersion equation. The Landau damping is thermodynamically reversible effect, and it is not accompanied with an entropy increase.

Here, it is assumed that the magnetic field $H$ is small in comparison with the electric field $\operatorname{grad} U$ so that we can ignore it. Because we consider Poisson's equation as follows:

$$\operatorname{div} \operatorname{grad} U = -4\pi q \int f(x, v)\, \mathrm{d}v \tag{2}$$

where $q$ is a charge of a particle.

Thus the system (1)–(2) is Vlasov–Poisson system and looking for nontrivial solution of it in the form

$$f = \sum_0^N C_k \psi_k(x, v) \mathrm{e}^{\mathrm{i}\omega t}, \quad N \leqslant \infty \tag{3}$$

where $C_k$ are some constant that can be found from boundary conditions.

Next, for convenience, write down Eq. (1) in the form

$$\partial_t f + Lf = 0. \tag{4}$$

Let a spectrum $\sigma(L)$ of the operator $L$ be such that $\omega_0 = 0 \in \sigma(L)$. Assume that there exists a function $f_0 = \sum_0^N C_k \psi_k$ satisfying Eq. (4). Obviously, it is a stationary solution to Eq. (4). It is therefore, reasonable to ask for which classes of perturbations the one will be stable.

The solution $f_0$ is called stable in respect to the norms $\|\cdot\|_i$, $i = 1, 2$ if for any $\varepsilon 0$ and each $t_0 0$ it is possible to find

*Corresponding author.
*E-mail addresses:* parsa@bnl.com (Z. Parsa), zvf@compuserv.com.ua (V. Zadorozhny).
[1]Supported by STCU grant #1746.

0168-9002/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.030

a number $\delta 0$ such that for $\|f_0\|_1 \leqslant \delta$ we have $\|f - f_0\|_2 \leqslant \varepsilon$ for $t \geqslant t_0$, where $f = \sum_1^n C_k \psi_k e^{i\omega t}$ is solution of Eq. (1) satisfying the initial condition $f(0) = f(t_0, x_0, v)$. Here $\|\cdot\|_i$, $i = 1, 2$ are same given norms and $x_0 = x(t_0)$ such that $\sum_1^3 x_i^2(t_0) \leqslant \alpha$ and $\alpha$ is a radius of the domain for initial value of $x_0 = \{x_1(t_0), x_2(t_0), x_3(t_0)\}$.

We will use some arguments from optimal control theory for construction of an optimal electric field. An interesting property of Maxwell equations is employed in the following. Given a specified beam motion in $R^3$ there exist electric fields which realize this motion. This property along with some mathematical aspect of optimal control theory with a given quality criterion allows to construct in certain cases a solution to the problem of focusing and acceleration for charged particle beams.

This reasoning leads to the following main steps:

- We have to find the Vlasov distribution function $f(t, x, v)$ so that the given transport conditions are satisfied. Those may include requirements on an acceleration, current density, value of focus, minimum of some criterion and so on. The idea is to look at this problem as to some optimal control problem [1]. The Vlasov equation is transformed into a Fredholm equation [2,3] and here we find $f(t, x, v)$ and dispersion waves.
- Now we can construct a beam density $\rho(t, x)$ and a beam current density $j(t, x)$ according to equations. Thus, there exist such fields $E$ Poisson which provide a motion of a beam according to the fixed law of motion.

Based on results of e.g. Zubov [1] and Halmos this approach makes it feasible to apply the direct Lyapunov method to nonlinear problems for which an empirical method of constructing a Lyapunov function generates a certain kernel operator in a domain of its stability, and then the Lyapunov equation yields a well known Fredholm equation. The self-focused and accelerating particle beams are studied using an analytic solution to the self-consistent Vlasov equation. A Lorentz force is treated as a control parameter (a control vector describing control fields), where a problem of optimal control is resolved. According to the Poisson equations charge and current densities are studied in the framework of a programming problem in an usual form.

## 2. Main result

The most convenient for the kinetic description of a beam particle behavior in an electrostatic field is the Vlasov–Poisson system (VPS) (1). The VPS problem consists in proving the existence of a $C^1$ or $L^p$ solution $f(t, x, v)$ for all $t \geqslant 0$ where $f(0, x, v) = \xi(x, v)$ is a given function. Here $f = f(t, x, v)$ is a distribution of particles in a phase space $\{x, v\}$ depending on the time $t$; $x \in R^3$, $v \in R^3$, $E = E(t, x)$ is an electric field, and

$$\rho = q \int f(t, x, v)\, dv, \quad j = q \int v f(t, x, v)\, dv$$

are the charge density and the current density respectively.

The standard argument (e.g. see Ref. [7] and references therein) implies that we have characteristic of the nonlinear dynamical system. Thus the distribution have a random behavior which is not well defined [8]. Here we only briefly describe how the $L$-moment problem (e.g. see Refs. [4,9,10]) reduces questions of choice of needed field $E(x)$ to an approximation problem.

Let us represent the solution $f$ as a sum

$$f = \sum_1^N C_k \psi_k(x, v) e^{i\omega t}, \quad k = 1, 2, \ldots, \quad N \leqslant \infty$$

where $\psi_k$, $k = 1, 2, \ldots$ are functions that be given as follows:

$$i\omega\psi_k + E\partial_v\psi_k = \int \partial_x \Phi(x - y)\psi_k(y)\, dy$$

where $\Phi(x - y)$ is a Chezaro kernel. It is an integral equation with the kernel $\partial_x \Phi(x - y) = \sum \mu_n i n e^{in(x-y)}$. The calculation can be carried out and a solution will be a unique one iff

$$\int_{\Omega_x} E\partial_v\psi_k \cdot \psi_k^*\, dy = 0. \tag{5}$$

Here $\psi_k^*$ is a conjugate function to $\psi_k$, the latter is a solution to a uniform equation

$$i\omega\psi_k = \int v \sum_1^N i n e^{in(x-y)}\psi_k(y)\, dy$$

where $\psi_k = \sum h_n^k e^{inx}$, $h_n^k = \int e^{-inx}\psi_k(x, v)\, dx$. This reasoning yields a simple equation $(i\omega - inv)h_n^k(v) = 0$.

Putting $n = e \cdot n_1$, $e \cdot e = 1$, for the sake of simplicity. Consequently $h_n^k(v) = \delta(\omega - k_1 v)$ is the delta function and $\omega$ is a point of the continuous spectrum.

In order to construct the $E(x)$ we can use relation (5). $L$-moment problem for a continuous medium the physical system is considered at a defined mesh point $\{\Theta_\alpha\}$, $\alpha = \overline{1, M}$ may be written down as follows:

$$\sum_{\alpha=1}^N E(\Theta_\alpha)\eta_k(\Theta_\alpha) = 0, \quad k = \overline{1, N}$$

$$\sum_{\alpha=1}^N E(\Theta_\alpha) f_0(\Theta_\alpha) = 1.$$

Here $\{\Theta_\alpha\}$ is a set of special given random points, $N \leqslant \infty$, $\eta_k(\Theta_\alpha) = \psi_k^* \partial\psi_k/\partial x$, and $f_0$ is some optimal process, such that

$$\partial_t = 0 \Rightarrow \partial_x f_0 = -\partial_v f_0.$$

This technique allows to provide a precise numerical calculation of the dynamics of charged particle beams. The $L$-moment method allows studying the detailed characteristics of bunched beams, taking into account a

distribution of particles, real self and external fields, construct optimal fields, and others. It is obvious that function the $E(x)$ is such that the known equations are valid. More precisely, the function $E(x)$ given above must be such that the following conditions hold true

$$\nabla E(x) = 4\pi q \int f(t,x,v)\,\mathrm{d}v \tag{6}$$

$$\mathrm{rot}\,E(x) = 0 \tag{7}$$

are satisfied. This system is overdetermined and for this reason we consider the following approach [5].

The second equation, the system (6), (7) is unresolved for any vector field $E$. Indeed, let $\mathrm{rot}\,E = \varpi$, $\varpi \in L^2$, $\int f\,\mathrm{d}v \in L^2$ and from here we have got the following condition $\mathrm{div}\,\mathrm{rot}\,E = 0$ thus we go to equation $\mathrm{div}\,\varpi = 0$. But all fields in $L^2$ form a subspace $S \subset L^2$. Thus the system (5), (6) may be resolved in the subspace $S \times L^2$ of the space $L^2 \times L^2$ only. But an orthogonal supplement to subspace $S$ in $L^2$ is the gradient functions which is equal zero on the boundary $\partial\Omega$. In this connection we shall associate some scalar function $P$. This reasoning yields the following system

$$\nabla E(x) = 4\pi q \int f(t,x,v)\,\mathrm{d}v$$

$$\mathrm{rot}\,E(x) + \mathrm{grad}\,P = 0 \tag{8}$$

with condition $P_{\partial\Omega} = 0$ and other condition on the boundary $\partial\Omega_x : \beta E_{\partial\Omega_x} = \alpha$. The system (8) is the elliptic system and it can be resolved (e.g. see Ref. [6]).

## Acknowledgement

The work of Z.P. was authored under contract no. DE-AC02-98CH1086 with the US Department of Energy. Z.P. has been the US Group-Leader/Collaborator on the STCU project #1746, the project is completed.

## References

[1] V.I. Zubov, Oscillations and Waves, Leningrad State University, 1989, 405p.
[2] Z. Parsa, V. Zadorozhny, Focusing and Acceleration of Bunched Beams, in: B.J. King (Ed.), Colliders and Collider Physics at the Highest Energies, AIP CP 530, 1999, pp. 249–259.
[3] Z. Parsa, V. Zadorozhny, Nonlinear Anal. 47 (2001) 4897.
[4] M.G. Krein, A.A. Nudel'man, The Problem of Moments and the Extremal Problems, Nauka, Moscow, 1973.
[5] S.G. Krein, Linear Equation in Banach Space, Nauka, Moscow, 1971.
[6] Yu.S. Sigov, Physica Scripta T2/2 (1982) 367.
[7] R.C. Davidson, H. Qin, Physics of Intense Charged Particle Beams on High Energy Accelerators, World Scientific, Singapore, 2001.
[8] A.A. Vlasov, The Statistically Distribution Functions, Nauka, Moscow, 1966, 355p.
[9] Z. Parsa, P. Zenkevich, Application of moments method, in: AIP CP No. 405, American Institute of Physics Publishing, New York, NY, 1997, pp. 183–188.
[10] W. Lysenko, Z. Parsa, Beam maching and halo control, in: Proceedings of 1997 Particle Accelerator Conference: Accelerator, Technology and Applications, Vancouver, BC, May 1997, IEEE, Piscataway, NJ, 1998, pp. 1917–1919 (This paper uses moments and provides a comparison with weighted moments and other variables in a Lie–Poisson's formulation. Also see in 'Beam Stability and Nonlinear Dynamics' Book edited by Z. Parsa, CP No. 405: 211–222, (AIP, NY, 1997), and references therein).

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 314–316
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Simulations of the experiment on efficient plasma wakefield acceleration

Pavel Logatchov[a,b], Konstantin Lotov[a,b,*], Alexey Petrenko[a,b]

[a]*Budker Institute of Nuclear Physics, Novosibirsk, Russia*
[b]*Novosibirsk State University, Novosibirsk, Russia*

Available online 28 November 2005

**Abstract**

We report the results of simulations made as a part of the plasma wakefield acceleration project based on VEPP-5 injection complex at Budker INP. Beam dynamics in the transport line and in 1-m-long plasma section is studied. At the designed facility, the efficient regime of plasma wakefield acceleration can be demonstrated. This regime is realized in the case of two properly shaped electron microbunches with a high peak current. To find the required beam shape and plasma parameters, a multiparametric optimization is carried out, which includes end-to-end simulations of beam dynamics in the plasma section as elementary steps. With the drive beam from the VEPP-5 injection complex, it will be possible to accelerate, in 1-m-long plasma, up to $3 \times 10^9$ particles with 30% driver-to-witness efficiency, energy spread lower than 10%, and acceleration rate 600 MeV/m.


*PACS:* 52.40.Mj; 41.75.Lx; 52.35.Mw

*Keywords:* Plasma wakefield acceleration; Simulation; Relativistic beam modulation

## 1. Introduction

Plasmas can sustain very large electric fields that are many orders of magnitude higher than those in conventional accelerating structures. This property is used in plasma wakefield accelerators (PWFA), in which one electron beam drives the high amplitude field in the plasma, and another beam (witness) is accelerated by this field (see reviews [1,2] and references therein).

In search of a good PWFA regime, several beam configurations were proposed and studied. One of them is the blowout regime [2,3] in which all the plasma electrons are ejected off the beam propagation channel, and an electron-free region (the cavern) is formed around the drive beam. For high beam currents and moderate beam lengths, the efficient blowout regime [4] is realized, at which high efficiency of the beam-to-beam energy exchange, high acceleration rate, and low energy spread can be achieved simultaneously.

To demonstrate the efficient blowout regime, a high-quality high-energy electron beam from a conventional accelerator must be longitudinally compressed, properly shaped, and injected into a plasma section that is long enough to completely decelerate some parts of the driver. All this possibilities will be available at the experimental facility based on the VEPP-5 injection complex [5]. The results of simulations made as a part of the experimental project are reported.

VEPP-5 injection complex is currently under construction at Budker Institute of Nuclear Physics to provide VEPP-4 and VEPP-2000 colliders with high-quality electron and positron beams. The complex consists of 510 MeV linear accelerator followed by the damping ring. Beam parameters after the damping ring are shown in Table 1. These beams can be also used for experiments on plasma wakefield acceleration.

For excitation of a high-amplitude plasma wave, the beam is to be compressed and shaped before the plasma chamber (Fig. 1). To this end, a linear correlation between energy and longitudinal position is induced in the beam by passing an RF structure at the zero-crossing phase. Then follow two 45° bending magnets where particles with

*Corresponding author. Budker Institute of Nuclear Physics, Novosibirsk, Russia.
*E-mail address:* K.V.Lotov@inp.nsk.su (K. Lotov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.040

Table 1
Beam from VEPP-5 injection complex

| | |
|---|---|
| Energy, $W_0$ | 510 MeV |
| Number of particles in bunch, $N_b$ | $(2 \div 5) \times 10^{10}$ |
| rms bunch length, $\sigma_z$ | 4 mm |
| Transverse rms size, $\sigma_x \times \sigma_y$ | $1.5 \times 0.03$ mm |
| $x$-Emittance, $\varepsilon_x$ | $2.3 \times 10^{-3}$ mrad cm |
| $y$-Emittance, $\varepsilon_y$ | $0.5 \times 10^{-3}$ mrad cm |
| Energy spread, $\delta W/W_0$ | 0.05% |

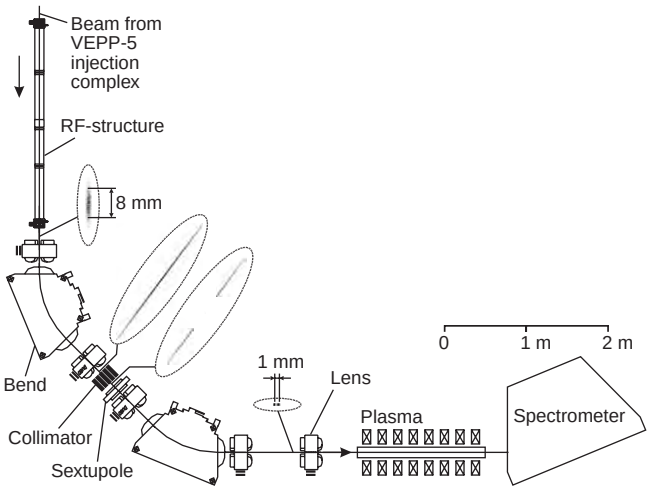



Fig. 1. General layout of designed facility. Insets show the beam portraits at different positions.

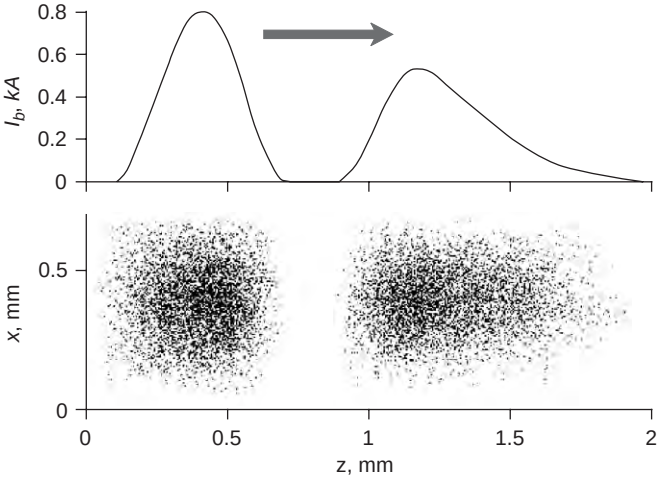



Fig. 2. Beam line tracing example. Arrow shows the direction of beam propagation.

different energies have different path lengths, so the bunch is compressed longitudinally. In the region of maximum dispersion, a collimator is placed to cut out some parts of the beam, so the bunch is modulated. Beam line tracing (Fig. 2) shows that this system can produce electron or positron bunches of various shapes, for example, a single short bunch (0.2 mm-long) with peak current up to 3 kA, two 0.2 mm-long microbunches with peak currents of 1–2 kA, or a long train of microbunches with the peak current about 100 A. The tracing code does not take into account coherent synchrotron radiation (CSR) effects and edge magnetic fields. These effects are important only for maximum compression rates that are not necessary for the regimes discussed below.

## 2. Simulation results

The designed facility will be flexible enough to study various regimes of plasma wakefield acceleration. Here we describe only the efficient two-bunch regime which is of a prime interest for possible collider applications of PWFA.

In the efficient blowout regime, the plasma response to the beam is essentially nonlinear and allows no detailed analytical study. The account of beam dynamics complicates the problem further. Therefore, for optimization of the system, we use two-dimensional hybrid code LCODE [6,7] to make end-to-end simulations of beam propagation through the whole plasma section.

In simulations, we take the initial beam density at the entrance to the plasma in the form

$$n_b = A\, e^{-r^2/2\sigma_r^2}\left[1 - \cos\left(\frac{2\pi z}{L}\right)\right], \quad 0 < z < L, \quad L = 2\sqrt{2\pi}\sigma_z$$
$$A = \frac{N_b(g(z,z_2) + 2 - g(z,z_1))}{4(2\pi)^{3/2}\sigma_r^2\sigma_z} \tag{1}$$

$$g(z,z_i) = \begin{cases} 2, & z < z_i - \delta z \\ 1 - \sin\dfrac{\pi(z - z_i)}{2\delta z}, & |z - z_i| < \delta z \\ 0, & z > z_i + \delta z \end{cases}$$

where $z$ is the longitudinal coordinate, $z_1$ and $z_2$ are the locations of driver and witness inward fronts, correspondingly, $\delta z = 0.125$ mm is the blur of these fronts determined by initial energy spread $\delta W$ of the beam; other notation is common.

The initial angular spread is $\Delta\alpha = \varepsilon_x/\sigma_r$ for both transverse coordinates to stay within the axisymmetric approximation, so the six-dimensional distribution function of beam particles is

$$f(\vec{r},\vec{p}) \propto n_b(r,z)\exp\left(-\frac{p_r^2 + p_\varphi^2}{2\Delta\alpha^2 p_{z0}^2}\right)\delta(p_z - p_{z0})$$

where $p_{z0}$ is the initial beam momentum ($p_{z0} = 1000mc$), and $\delta$ is the delta-function.

The process of optimization involves the adjustment of $\sigma_r$, $\sigma_z$, $z_1$, $z_2$, as well as plasma density $n_i$ and length $L$ to maximize the witness energy gain and witness charge at moderate ($\lesssim 10\%$) energy spread.

The result of optimization is shown in Fig. 3. The beam that initially comprises $2 \times 10^{10}$ electrons is to be compressed to $\sigma_z = 0.3$ mm, focused to $\sigma_r = 0.026$ mm, and

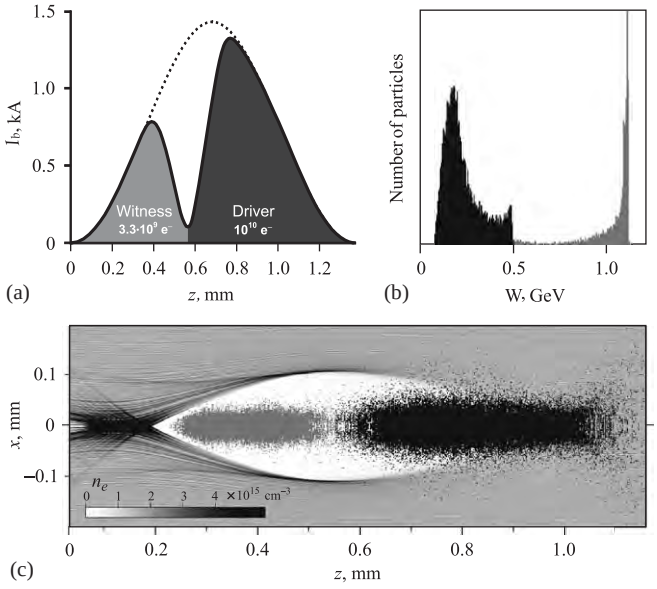



Fig. 3. Results of simulations: (a) the profile of beam current before the entrance to the plasma, dashed line shows the current profile with no collimation, the beam propagates toward positive $z$-direction; (b) energy spectrum of the beam after the plasma; (c) plasma wave structure and the beam portrait at the midway of the plasma section. Black dots show driver particles, gray dots show witness particles.

collimated to create the double-hump density distribution (Fig. 3a). After passage of 95 cm in the plasma of density $1.7 \times 10^{15}\,\mathrm{cm}^{-3}$, the first bunch ($10^{10}$ electrons) loses 54% of its energy and decelerates in average from 510 to 240 MeV (Fig. 3b). The second bunch takes 63% of the plasma energy and accelerates to 1.1 GeV. This corresponds to 34% bunch-to-bunch efficiency and the acceleration rate of 600 MeV/$m$. The final energy spread of accelerated bunch is 9% (Fig. 3b), the angular spread is lower than 0.3°. Fig. 3c shows the plasma wave structure and beam portrait at the midway of the plasma section. Here we can see that the tail of the witness is already defocused by plasma wake and the beam becomes shorter.

## Acknowledgements

This work was supported by Science Support Foundation, SB RAS Lavrent'ev Grant for young researchers, Russian Foundation for Basic Research (Grant 03-02-16160a), and Russian Ministry of Science (Grant NSh-229.2003.2).

## References

[1] E. Esarey, P. Sprangle, J. Krall, A. Ting, IEEE Trans. Plasma Sci. 24 (2) (1996) 252.
[2] C. Joshi, B. Blue, C.E. Clayton, E. Dodd, C. Huang, K.A. Marsh, W.B. Mori, S. Wang, M.J. Hogan, C. O'Connel, R.H. Siemann, D. Watz, P. Muggli, T. Katsouleas, S. Lee, Phys. Plasmas 9 (5) (2002) 1845.
[3] J.B. Rosenzweig, B. Breizman, T. Katsouleas, J.J. Su, et al., Phys. Rev. A 44 (10) (1991) 6189.
[4] K.V. Lotov, Phys. Plasmas 12 (5) (2005) 053105.
[5] A.V. Burdakov, A.M. Kudryavtsev, P.V. Logatchov, K.V. Lotov, A.V. Petrenko, A.N. Skrinsky, Plasma Phys. Rep. 31 (4) (2005) 292.
[6] K.V. Lotov, Phys. Rev. ST Accel. Beams 6 (6) (2003) 061301.
[7] K.V. Lotov, Phys. Plasmas 5 (3) (1998) 785.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 317–324
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Particle-in-cell method for numerical simulation of beam and plasma dynamics

Grigori Shirkov*

*Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russian Federation*

Available online 29 November 2005

**Abstract**

The finite particle methods are widely applied for the numerical simulation of a continuous medium including continuum electrodynamics and simulation of charged particle beams. This method allows studying the detailed beam and plasma characteristics, the distribution function of particles, the nonlinear self and external fields. The so-called ''particle-in-cell'' method is the most effective for simulations in the physics of plasma and charged particle beams.

Two examples illustrate the application of the particle-in-cell method in beam dynamics and plasma physics. The first is a program library for numerical simulation and optimization of multi-component ion beam transportation from ion sources where RFQ accelerators are presented. The latest version of the library was applied for simulations of transverse dynamics of an argon ion beam in the beam line from the 18 GHz ECR Ion Source to the linear RFQ accelerator of the RIKEN Beam Factory. A partial neutralization of ion beam charge due to secondary electrons was assumed to realize the real experimental conditions in the beam line and was used in simulation. The second is a summary of recent development of physical and mathematical basements and the first version of a computer code library aimed for the three-dimensional (3D) simulation of the ECR plasma and ion production in the ECR ion source.


*PACS:* 02.60.Cb; 29.25.Ni; 52.50.−b; 52.20.−j

*Keywords:* Numerical simulation; Ion sources; Beam transport; Plasma production and heating; Elementary processes in plasma

## 0. Introduction

Mathematical modeling and numerical simulation of beam dynamics and multicomponent plasma with consideration of space charge effects are of great important in connection with modern trends of investigations in the physics of particle accelerators and ion sources. These problems belong to the electrodynamics of a continuous medium.

The distribution of particles within the phase–space of co-ordinate $\vec{r}$; and velocities $\vec{v}$; is described by the distribution function $f(\vec{r};,\vec{v};,t)$. In consideration of the particles' electromagnetic space fields and of impact processes that take place in ion sources and beams, the movement of charged particles is described in a generalized form by a system of self-consistent kinetic equations completed by Maxwell equations [1].

The particle-in-cell is one of the finite particle methods for the simulation of multicomponent plasma and particle beams [2]. These methods are the most powerful methods for the numerical simulation of motion of continuous medium, gas and plasma dynamics, dynamics of charged particle beams. The finite particle method allows studying the detailed characteristics of continuous medium, taking into account the distribution functions of particles (spatial, velocity and energy distributions), real self and external fields, particle–particle interactions and many other effects. This technique allows providing very precise numerical simulations of the plasma and beam dynamics.

The electrostatic potential $U$ for the medium simulation performed in quasi-static approximation is described by the

*Tel.: +7 0962165136; fax: +7 0962165379.
*E-mail address:* shirkov@sunse.jinr.ru.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.084

Poisson's equation:

$$\Delta U = -\frac{\rho}{\varepsilon_0}, \quad (1)$$

where $\rho$ is the particle density.

Particles of all masses and charge states are available in the consideration—ions of all charge states of all elements, different kinds of neutrals (atoms ore molecules) or electrons. Each finite particle represents the group of physical particles. Thus, every finite particle species has a mass and charge accordingly to the type of ions, neutrals or electron.

## 1. Beam dynamics

A program library based on the macro particle method (MCC code) [3] was applied for simulations of transverse dynamics of an argon ion beam in the beam line from the 18 GHz ECR Ion Source to the linear RFQ accelerator of the RIKEN Beam Factory [4]. The line consists of an Einzel lens (EL), a horizontal bending magnet (BM) with poles, removable Faraday cup (FC) and a solenoid (SRFQ), and it is presented in Fig. 1.

The results of numerical simulation of the multi-component ion beam dynamics in this channel were published in the Ref. [5]. A new beam line with a solenoid S2 instead of EL is proposed, simulated and optimized. A partial neutralization of ion beam charge due to secondary electrons was assumed to realize the real experimental conditions in the beam line, and was used in the simulation.

The particle motion in a solenoid field was integrated in the rotating co-ordinate system $X_r$ to improve the accuracy of the calculations:

$$X_r = \begin{pmatrix} x_r \\ y_r \end{pmatrix}; \quad X = \begin{pmatrix} x \\ y \end{pmatrix}; \quad Y_r = QX,$$

$$Q = \begin{pmatrix} \cos\Phi(z) & \sin\Phi(z) \\ -\sin\Phi(z) & \cos\Phi(z) \end{pmatrix}, \quad (2)$$

$$X_r' = QX' + Q'X; \quad \Phi' = k(z) = -\frac{1}{2}\frac{ZeB_z}{cp_z},$$

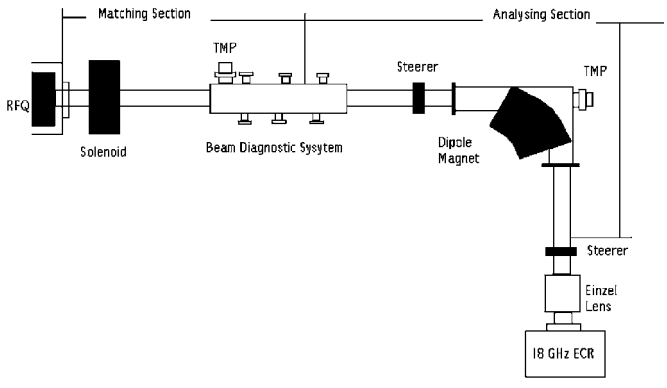



Fig. 1. Structure of RIKEN ion beam line.

where symbol "′"means the derivative on longitudinal co-ordinate $z$, $p_z$ is the longitudinal particle momentum and $B_z$—the longitudinal component of the induction of magnetic field. The equations of particle motion in this co-ordinate system are as follows:

$$X_r'' + k^2 X_r = \frac{Ze}{cp_z\beta_z} QE_s; \quad E_s = \begin{pmatrix} E_{xs} \\ E_{ys} \end{pmatrix}.$$

Here $\beta_z = v_z/c$ is the relative velocity of ion and $E_s$—the space electric field of the ion beam in laboratory co-ordinate system. Since ions with different charges have different angular velocities, so the space electric field of the beam is calculated in the laboratory co-ordinate system.

Ions with different charge states move along different orbits. The ion of charge state $Z$ is shifted from the equilibrium orbit for the charge state $Z_0$ in the $x$ direction according to the equation

$$x'' + \frac{1}{\rho_0^2}x = \frac{1}{\rho_0} - \frac{1}{\rho} + \frac{Ze}{cp_z\beta_z}E_{xs},$$

$$\frac{1}{\rho_0} = -\frac{Z_0 eB}{cp_{z0}},$$

$$\frac{1}{\rho} = -\frac{ZeB}{cp_z}.$$

Here $p_{z0}$ is the longitudinal momentum for equilibrium particle and $B$—the induction of magnetic field in BM.

It should be noted here that the orbit of an equilibrium particle is shifted after BM. The unsymmetrical influence of the beam space charge on an equilibrium trajectory is a cause of this effect during separation if the beam current is not negligible. These shifts of the center of mass and the angle of trajectory bending for the equilibrium particle are corrected automatically on the exit of BM in our calculations. This correction can be done by a small tuning of magnetic field induction $B$ or by using an additional dipole magnet in the real experimental conditions.

The simulation of ion beam dynamic was carried out for the beam of argon ions of five species—$Ar^{6+}$–$Ar^{10+}$. The experimental charge state distribution of the ion beam measured after BM was used as the initial one. The ion beam emittance at the channel entrance was assumed to be equal 100 mm mrad and the initial momentum spread $\Delta p/p = 0$. Particles in the phase planes were generated according to the micro-canonical distribution. The kinetic energy of ions was 11 $Z$ keV (here $Z$ is the ion charge). The magnetic field in BM corresponds to the equilibrium orbit for $Ar^{8+}$.

Secondary electrons are accumulated in the beam if there is no external transverse magnetic or electric field in the beam line [5]. The beam neutralization is not full due to the initial energy of secondary electrons. The process of neutralization stops when the beam potential becomes less that the electron energy. Nevertheless, according to simulation and all experimental data, an ion beam produces a channel of secondary electrons in the linear beam line sections without external transverse fields and,

in this way, the beam fields are reduced and the beam transition is improved. A so-called factor of neutralization FN was introduced in simulations for taking into account the influence of secondary electrons on the beam transmission.

The optimization of S2 parameters (location and magnetic field) was carried out to achieve the maximum capture of beam current into the RFQ acceptance. The dependence of accepted current on the neutralization factor FN was studied. The magnetic field induction was varied in the limits—0.15,…,0.3 T in S2 and 0.25,…,0.35 T in SRFQ. The results of simulation for $I_{ext} = 6$ mA and various NF are presented in Figs. 2–4. Fig. 2 corresponds to the results for the existent beam line (without S2). The following data are presented at the figures:

- *Main windows*: particle trajectories in vertical (above) and in horizontal (below) planes, left curve is the residual magnetic field of ECRIS;
- *Upper window*: schematic drawing of the channel, boxes 1 are the regions with magnetic field of the ECRIS solenoids and S2 (left) and SRFQ (right), box "0" is the BM;
- *Left and right panels*: initial and final ion beam parameters correspondently.

The simulations have shown that the experimental ratio (~80%) of beam current captured into the RFQ acceptance and beam current on FC is also achieved for FN = 0.9. The focus spots can differ in the horizontal and vertical planes, and it is difficult to obtain and adjust poperly the RFQ acceptance with the beam emittance in this case. The maximum value of $I_{RFQ}$ is possible only if the beam envelope has the double waist near FC. It is possible to avoid such a situation by a small (<1%) variation of the magnetic field of BM in practice.

### 1.1. Results of beam line simulation and optimization

- The results of numerical simulation with *NF*~0.9 of the beam are in the good coincidence with experimental data.

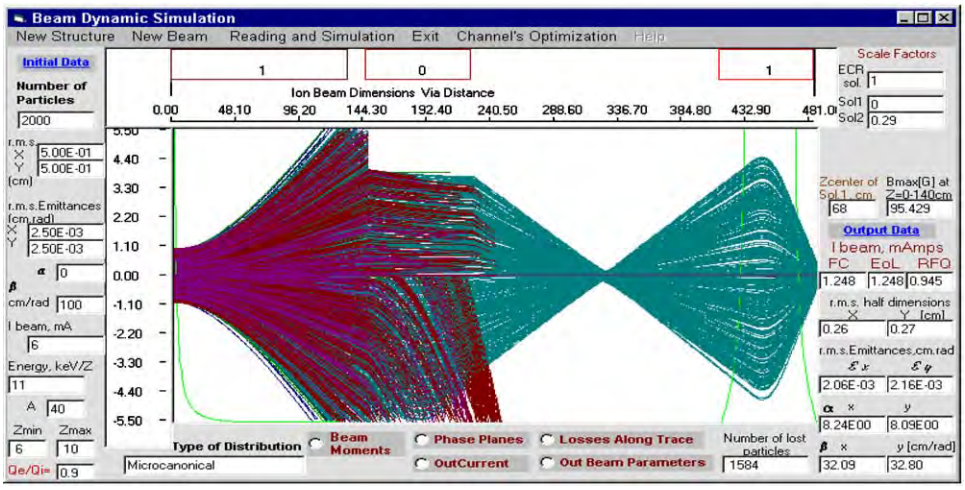



Fig. 2. Beam trajectory for $B_{sol.1} = 0$; $Ar^{8+}$ ion currents are 1.25 mA at FC and 0.95 mA at RFQ.

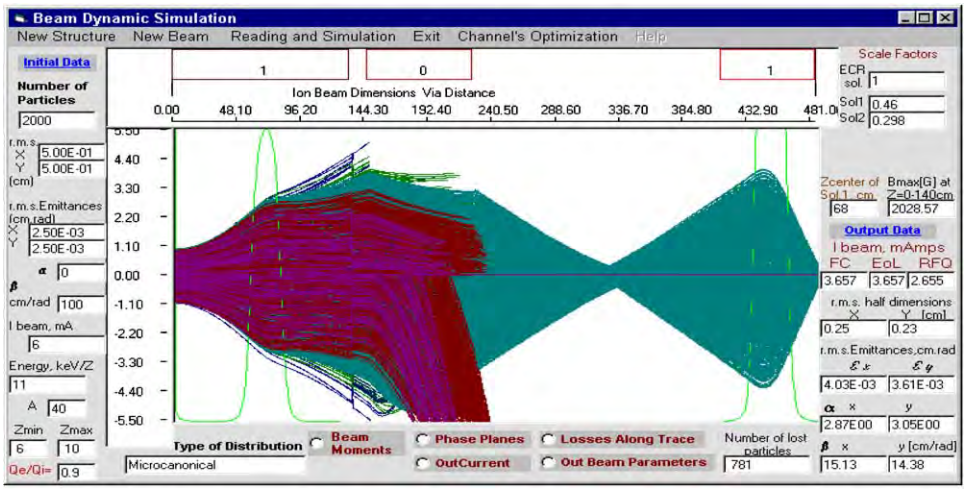



Fig. 3. Beam trajectory for $B_{sol.1} = 0.203$ T at 68 cm; $Ar^{8+}$ ion currents are 3.7 mA at FC and 2.7 mA at RFQ.

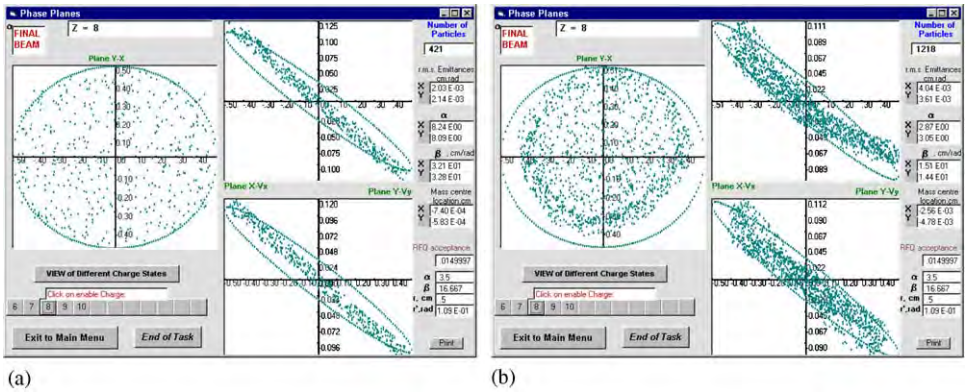



Fig. 4. (a) and (b) Beam cross-sections and phase planes.

- The maximum acceptance of ion beam current into the RFQ is about 70–80% of the beam current on FC.
- The maximum beam current into the RFQ acceptance can be provided by the beam with double waist envelope in the FC region of the line.
- The solenoid S2 instead of Einzel lens EL is able to increase the beam current in the RFQ acceptance by a factor of 2.5–3. The first test of S2 without any optimization has improved the beam current by a factor of 1.6.

## 2. Multicomponent plasma of ECR ion source

The project "Numerical simulation and optimization of ion accumulation and production in multicharged ion sources" was fulfilled in 1999–2001 at the International Science and Technology Center (ISTC). This project was carried out in the frame of the Radioactive Isotope (RI) Beam Factory at RIKEN [6]. The factory consists of a cascade of ion accelerators and is aimed at providing RI beams over the whole atomic mass range with a high level of intensity in a wide range of energies up to several hundreds MeV/nucl. The first section of the accelerator chain is a new superconducting ECR ion source. The physical substantiation of the ISTC project was prepared and published in RIKEN in 1997 [2].

The goal of the project is the creation of a new generation 3D model and codes for the numerical simulation based on the "particle-in-cell" method. This technique allows providing very precise numerical simulations of the ECR plasma and optimizations of ECR ion sources in this way.

### 2.1. Physical and mathematical models [7]

The velocities of electrons and ions are different by 2–3 orders of magnitudes due to different masses and energies in the plasma. The integration of the equations of motion for the electron component requires a time step of about $10^{-12}$–$10^{-11}$ s. This is the main limitation for the computer simulation of ion accumulation in plasmas with a time scale of 10 ms.

The electron motion consists of fast Larmor rotation with frequencies of about $10^{10}$ Hz and relatively slow drift of the electron trajectory in the magnetic trap of the ECR source. The Larmor rotation does not determine peculiarities of slow rate processes of ion production and accumulation in the ECR source. Therefore, it is possible to exclude fast rotation of a particle around a small Larmor circle from the equations of motion, with the aim to sufficiently increase the integration step of the problem to be solved. This can be done by substitution of particle motion by the motion of the center of Larmor rotation.

The theory providing the description of charged particle motion in an electromagnetic field in that way, called "*guiding center theory*" or "*guiding center approximation*" [8], was developed and applied in computer codes for simulation of the ECR plasma.

Electrical and magnetic fields are created by a system of coils, permanent magnets and special electrodes in the ECRIS. These fields can be calculated analytically or numerically, or can be fixed as an experimental data table. A quadratic interpolation is used to calculate fields in every point inside of cells. The following models and procedures have been chosen for the simulations:

- *Analytical equations to calculate the fields of typical elements*: solenoids, quadrupole or multipole lenses, horizontal and vertical bending magnets, accelerating and drift cavities;
- *POISSON/SUPERFISH Group of Codes* [9] *to calculate the fields from user defined sources*: permanent magnets, solenoids with iron, "Einzel" lenses, electrodes of a complex configuration, etc.
- Data tables of experimental fields.

The physical model is based on differential equations of motion for the charged particles and the Poisson equations for the self and external fields [2]. The self-field of particles is obtained from a direct solution of the Poisson's equation for an electrostatic potential at every time step, in a

cylindrical system of co-ordinates with Dirichlet boundary conditions at the conductive surface, and with Neumann conditions at the axis. The double Fourier transformation is used to solve the Poisson's equation. The procedure of quadratic interpolation is used to calculate the values of self-fields in every point of every particle position. The density distributions of every plasma component are calculated at every integration step on these mesh cells, in order to provide self-consistent solutions of equations of particle motion and fields in the plasma.

Particles of all masses and charge states are available in the consideration—ions of all charge states of all elements, different kinds of neutrals, electrons, etc. Each finite particle represents the group of physical particles. Thus, every finite particle species has a mass and charge according to the type of ions, neutrals or electrons.

Simulation of processes in a multicomponent plasma must take into account the geometry of the real set-up and the parameters of its elements. The data to describe an initial condition of a system of interacting particles should be given as well.

The following model conceptions of ECR plasma components are used in Refs. [10,11]:

- *The electrons have a complicated distribution of a few components*: cold primary electrons of tens or hundreds of eV; a main electron component of keV energy which produces highly charged ions and, according to the experimental data, a component of superhot electrons of tens or hundreds of keV;
- Electrons undergo the RF heating on a surface with ECR conditions in the magnetic field;
- ECR plasma is neutral or quasi-neutral in every point of the volume.
- All ion components have Maxwellian energy distribution, with a common temperature in the plasma due to intensive elastic Coulomb collisions.

The basic inelastic atomic processes providing the change of a charged particle distribution are ionization of neutral atoms and ions by electron impact, charge exchange of ions with neutral atoms and recombination among ions and electrons.

The balance equations for ion densities are a basis of the algorithm to simulate inelastic processes in the ECR plasma. The balance equations are a consequence of the kinetic equations for the ion distribution function. These equations take into account all processes determining production and loss of ions and have a probabilistic character. It is quite straight forward to use a Monte-Carlo method of statistical modeling to find a solution of the balance equations in each cell of the mesh in simulations.

The complete amount of finite size particles in the mesh should be much greater than the amount of cells. Supposing that the amount of macro particles—ions of each charge state in an elementary cell is equal to 10–15, and that the amount of various ions species does not exceed 10–20, we fixed that the macro particle amount must be in the order of $(2–3) \times 10^5$ for the adequate plasma description in the 2D model. The relative error in the definition of a plasma field should not exceed the value of $1/N^{1/2} = 1/[(2–3) \times 10^5]^{1/2} \approx 2 \times 10^{-3}$, i.e. less than 1%. In a 3D case, the number of the finite size particles will not be less then $(3–5) \times 10^6$.

Each finite particle has six co-ordinates in total—three space co-ordinates (*x, y, z* in Cartesian co-ordinates) and three velocity or momentum co-ordinates. Equations of motion describe the changes in time of the entire set of phase co-ordinates. The equations of motion can be integrated by numerical methods, the most economic with regard to computing resources, being a "leap-frog" method of the second order.

It takes at least several thousand integration steps to perform the integration of one ion oscillation due to the large difference in Larmor frequencies of electrons and ions in the plasma. This is one of the main difficulties of plasma simulation that can be solved via the macroparticle method, at least in two ways—a computation model with increased electron mass, and the averaging of fast oscillating motion of electrons in a magnetic field (see above).

### 2.2. Testing of the library of computer codes

The code library with the name ECR source particle code (EPC) to simulate ECR sources has been prepared. The complex includes earlier developed blocks of programs [7] for realizing procedures of different processes mentioned above using Delphi 5.

The common WINDOWS interface of the library is aimed to set parameters and initial conditions, to start the program of the ECR ion source simulation and to control the calculation process. The main menu of the program (Fig. 5) includes the following items—*parameters, initial conditions, viewand calculations.*

The programs have been united by the user interface that allows investigating various phases of processes proceeding

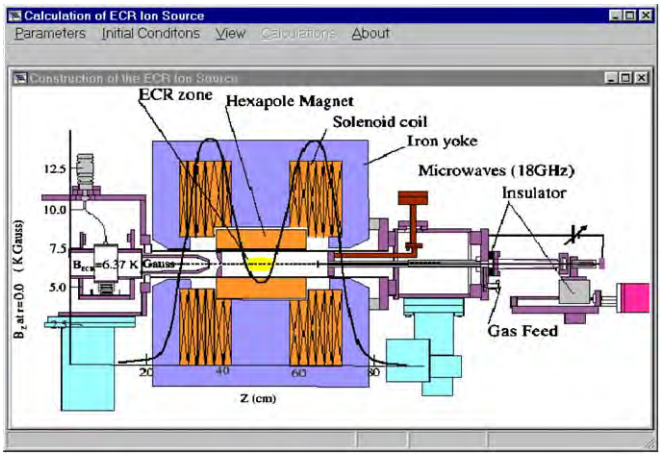



Fig. 5. Main menu of the library with schematic view of ECR source.

in ECR ion source. First tests of the computer code library are presented in Figs. 5–8. These figures show results of a model simulation of nitrogen ion production and accumulation during about 50 μs. These figures have been taken directly from the computer screen and can give some impressions of the user interface—numbers of macroparticles, items of the user menu, charge states of ions, etc. The initial data for this numerical experiment for the 18 GHz ECR source were as follows:

- density of nitrogen neutrals was equal to $2 \times 10^{10}\,\mathrm{cm}^{-3}$, and kept constant;
- average initial energies of electrons and neutrals were 10 and 0.03 eV, correspondingly;
- longitudinal magnetic field with trap longitudinal configuration ($B_z = 0.47\,\mathrm{T}$ at the center of chamber—$Z = 0$ and $B_z = 1.5\,\mathrm{T}$ at $Z = \pm 12\,\mathrm{cm}$), and sextupole magnetic fields were used in computations;
- initial amount of macroparticles was 50,000;
- mesh sizes on $X$-, $Y$- and $Z$-axes were $4\,\mathrm{mm} \times 4\,\mathrm{mm} \times 8\,\mathrm{mm}$.

The typical rate of calculations was 1 μs of plasma time for 1 h of simulation time using a Pentium III PC.

The spatial distributions of two electron components are shown in Fig. 6: (a) for $X$-$Y$ plane and (b) for $X$-$Z$ plane. The blue color of macroparticles corresponds to a low energy electron component and the green color corresponds to high-energy species. Fig. 7 presents the ionization dynamics of nitrogen ions. Fig. 8 shows spatial distributions in the $X$-$Y$ plane for nitrogen ions of different charged state.

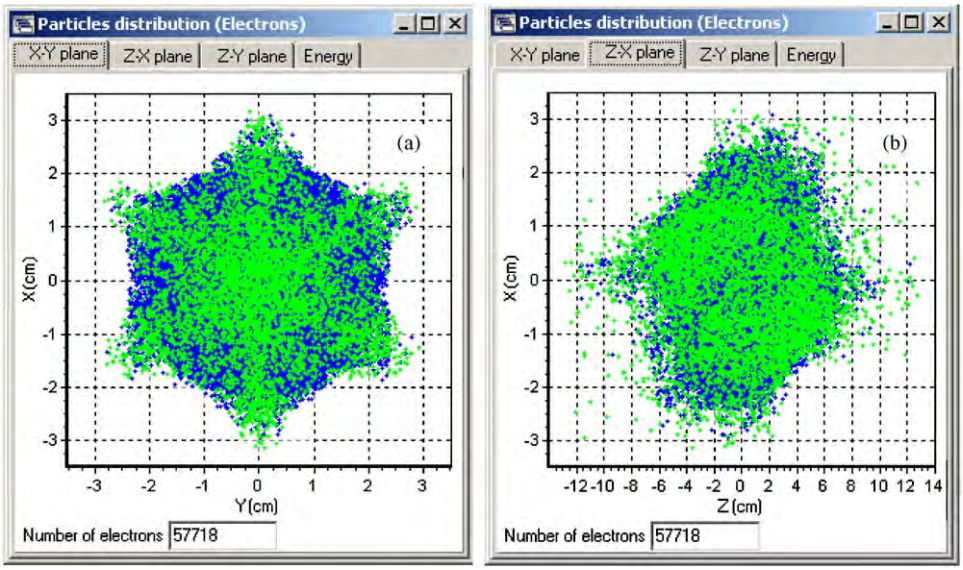



Fig. 6. Distributions of electrons in $XY$ (a) and $XZ$ (b) planes.

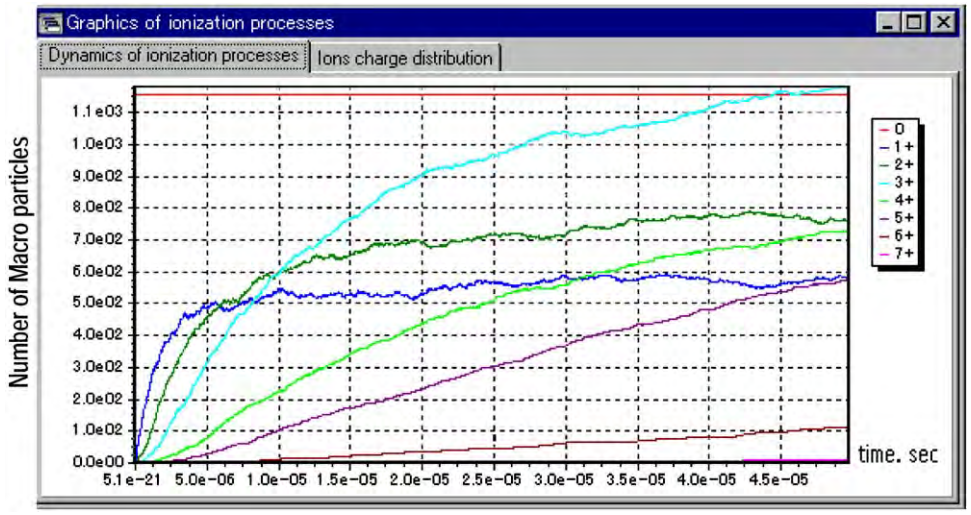



Fig. 7. Ionization dynamics of nitrogen ions.

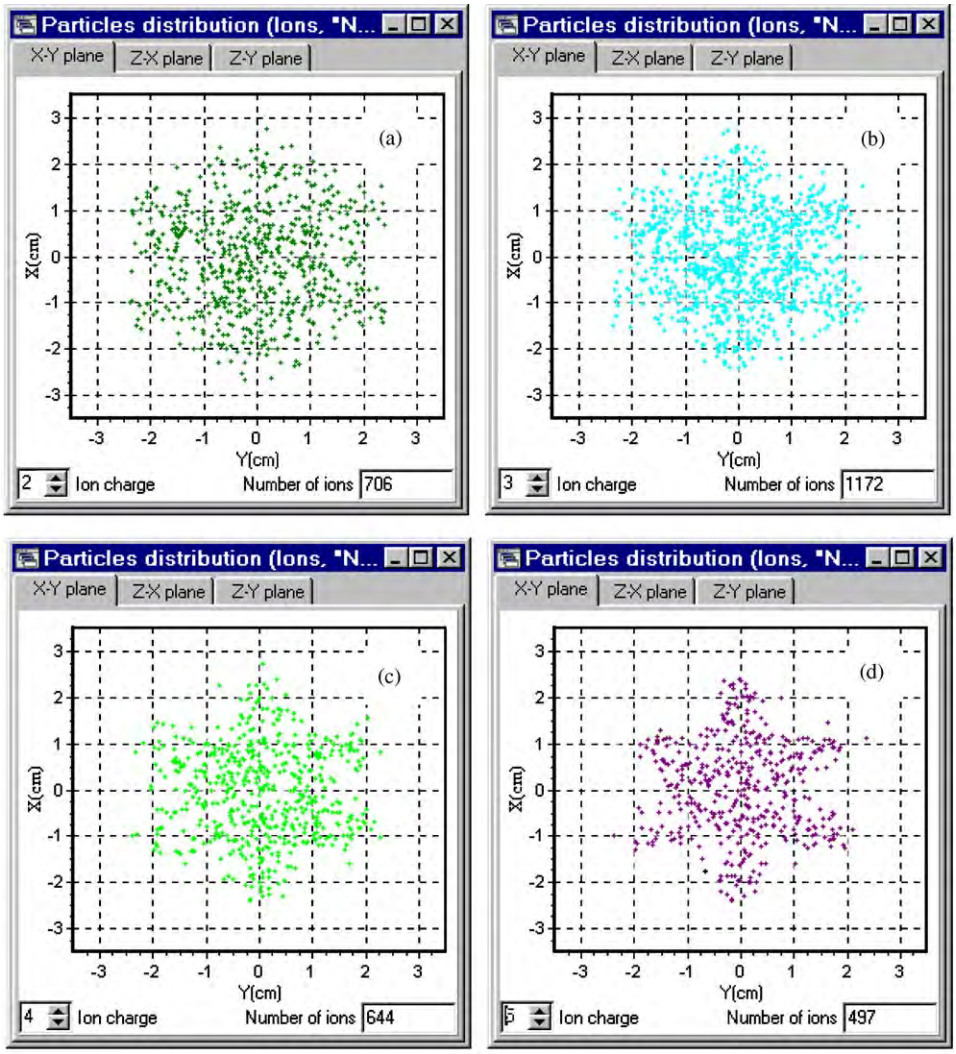



Fig. 8. Final distributions of ions with second (a), third (b), fourth (c), fifth (d) charged states in $XY$ plane.

## 3. Results of ECRIS plasma simulations

The physical basement and mathematical methods for simulation of ECR plasmas with the particle-in-cell model have been developed. A first version of a computer code library was created and tested in simulations of processes in a real ECR ion source. It has been shown:

- detailed and exhaustive spatial description of fields and particle motion and interaction in the ECR source plasma requires a full-scale 3D task to be solved;
- computer executing time of simulation is mostly determined by the time step of integration ($\sim 10^{-10}$ s) of electron equations of motion and by the number of macroparticles for the electron component representation;
- time range of production and accumulation of heavy highly charged ions is about 1–10 ms and it requires about $10^8$ time steps of simulation;
- 2D task requires at least $2–3 \times 10^5$, and full-scale 3D task requires at least of $3–5 \times 10^6$ macroparticles in the simulation.

The most significant difficulty of 3D problems is a large computation time. High-power multiprocessor computers with parallel computation only are able to satisfy the requirements of a 3D simulation. The development and adaptation of numerical methods and computer codes for parallel computations at a multiprocessor computer are necessary for the full-scale 3D task.

## References

[1] E.M. Lifshitz, L.P. Pitaevski, in: Theoretical Physics. vol.10, Physical Kinetic, Nauka, Moscow, 1979.
[2] G. Shirkov, Y. Batygin, Y. Yano, Preprint RIKEN-AF-AC-5, 1997, 36pp.
[3] V. Alexandrov, Yu. Batygin, V. Shevtsov, G. Shirkov, in: Proceedings of the Sixth European Particle Accelerator Conference, June

1998, Stockholm, Preprint JINR, E9-98-148, JINR, Dubna, 1998, p. 1239.
[4] T. Nakagawa, T. Kageyama, A. Goto, N. Inabe, O. Kamigato, M. Nagase, T. Chiba, M. Hemmi, M. Kase, Y. Myazawa, et al., RIKEN Acceleration Progress Report 1994, RIKEN 1995, vol.28, p. 165.
[5] V. Alexandrov, T. Nakagawa, V. Shevtsov, G. Shirkov, Rev. Sci. Instrum. 71 (2) (2000) 1094;
V. Alexandrov, N. Kazarinov, V. Mironov, V. Shevtsov, G. Shirkov, Y.Yano, T. Nakagawa, Preprint RIKEN, 2001, RIKEN-AF-AC-27, 14pp.
[6] Y. Yano, A. Goto, T. Katayama, in: Collected papers on RIKEN RI Beam Factory Project, December 1998, RIKEN.
[7] G. Shirkov, V. Alexandrov, V. Preisendorf, V. Shevtsov, R. Komissarov, M. Koryovkina, V. Mironov, E. Shirkova, O. Strekalovsky, N. Tokareva, et al., Preprint RIKEN, 2000, RIKEN-AF-AC-24, 58pp.
[8] H. Alfven, Cosmical Electrodynamics, Oxford, New York, 1950.
[9] M.T. Menzel, H.K. Stokes, POISSON/SUPERFISH Group of Codes, LA-UR-87-115, Los Alamos, NM, 1987.
[10] G. Shirkov, G. Zschornack, Electron Impact Ion Sources for Highly Charged Ions, GWV Verlagsgesellschaft, Vieweg & Sohn, Wiesbaden, 1996 310pp.
[11] G. Shirkov, Plasma Source. Sci. Technol. 2 (1993) 250.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 325–328
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# BETACOOL program for simulation of beam dynamics in storage rings

A.O. Sidorin*, I.N. Meshkov, I.A. Seleznev, A.V. Smirnov, E.M. Syresin, G.V. Trubnikov

*Joint Institute for Nuclear Research, Dubna 141980, Russian Federation*

Available online 28 November 2005

**Abstract**

The BETACOOL program developed by the JINR electron cooling group is a package of algorithms based on a common format for input and output files. The program is oriented to simulation of the ion beam dynamics in a storage ring in the presence of cooling and heating effects. The version presented in this report includes three basic algorithms: simulation of RMS parameters of the evolution of the ion distribution function over time; simulation of the distribution function evolution using a Monte-Carlo method; and a tracking algorithm based on a molecular dynamics technique. General processes to be investigated with the program include intrabeam scattering in the ion beam, electron cooling, and interaction with the residual gas and the internal target.


*PACS:* 34.70.+e; 36.10−k; 29.27.Bd

*Keywords:* Storage ring; Beam cooling; Electron cooling

## 1. Introduction

Electron cooling is widely used for ion-beam parameter control in storage rings. At present there are approximately 20 storage rings in operation and under construction that are equipped with electron cooling devices. The BETACOOL program developed for simulation of electron cooling processes [1] is actively used for the design and simulation of electron cooling systems in several research centers: JINR, ITEP, Jülich-FZ, GSI, RIKEN and BNL [2–6]. BETACOOL is programmed with an object-oriented method using the C++ language. The interface with the Windows operating system was developed on the basis of a BOLIDE system (Builder Object Library & Interface Development Environment), which is dedicated to fast elaboration of the physics and mathematics applications.

The general goal of the BETACOOL program is to simulate long-term processes (in comparison with the ion revolution period) leading to variation of the ion distribution function in six-dimensional phase space. The ion beam motion inside the ring is supposed to be stable and is treated by a linear approximation.

The structure of the program allows simulation of the evolution of the ion distribution function using a few independent numerical algorithms. Each algorithm simulates the ion beam dynamics for the same input beam and ring parameters and uses the same set of effects acting on the beam distribution function in simulations.

This report discusses the latest version of the BETACOOL program, which includes three algorithms for beam dynamics simulation and takes into account the following processes: electron cooling, intrabeam scattering, ion scattering on residual gas atoms, and interaction of the ion beam with the internal target, among others.

## 2. BETACOOL algorithms

Three basic algorithms for simulation of the evolution of the ion distribution function are now included in the program:

- RMS dynamics simulation;
- simulation of distribution function evolution using a Monte-Carlo method (model beam algorithm); and
- multi-particle tracking based on a molecular dynamics technique.

*Corresponding author.
*E-mail address:* sidorin@jinr.ru (A.O. Sidorin).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.041

The physical model used in the RMS dynamics simulation is based on the following general assumptions:

(1) The ion beam has Gaussian distribution over all degrees of freedom, and is not changed during the process.
(2) The algorithm for analysis of the problem is considered as a solution of the equations for RMS values of volumes of the beam phase space for three degrees of freedom.
(3) The maxima of all the distribution functions coincide with the equilibrium orbit.

The evolution of the ion beam parameters during its motion inside the storage ring is described by the following system of four differential equations:

$$\frac{dN}{dt} = N\sum_j \frac{1}{\tau_{\text{life},j}}, \quad \frac{d\varepsilon_h}{dt} = \varepsilon_h \sum_j \frac{1}{\tau_{h,j}},$$
$$\frac{d\varepsilon_v}{dt} = \varepsilon_v \sum_j \frac{1}{\tau_{v,j}}, \quad \frac{d\varepsilon_{lon}}{dt} = \varepsilon_{lon} \sum_j \frac{1}{\tau_{lon,j}} \tag{1}$$

where $N$ is the particle number, and $\varepsilon_h$, $\varepsilon_v$ and $\varepsilon_{lon}$ are the RMS values of the horizontal, vertical and longitudinal beam emittance, respectively. Characteristic times are functions of all three emittance values and the particle number and have a positive sign for a heating process and a negative sign for a cooling process. A negative sign for the lifetime corresponds to particle loss and the sign of the lifetime can be positive in the presence of particle injection, when the particle number increases. Index $j$ in Eq. (1) indicates the number of processes involved in the calculations. The algorithm structure is designed in such a way that allows the inclusion of any process into calculation that can be described by cooling or heating rates. The numerical solution of Eq. (1) is performed using a Euler method with automatic step variation. The result of the simulation gives the time dependence of the emittance and the particle number. An ion ring optic structure is necessary only for intrabeam scattering (IBS) simulation. The IBS growth rates are calculated in accordance with an analytical model using ring lattice functions imported from the output file of the MAD program [7].

Step of the integration of Eq. (1) over time is determined by the characteristic times of the effects investigated and the calculation speed can be very fast. However, in some cases the basic physical model cannot provide realistic simulation, mainly due to basic assumptions about the Gaussian shape of the ion distribution function. This assumption is more or less realistic in an equilibrium state of the ion beam, when the equilibrium is determined by many processes of a stochastic nature. If equilibrium does not exist due to fast particle loss or in the initial stage of beam cooling, the ion distribution function can be far from Gaussian. The same situation occurs in an experiment with internal targets for which the dimensions do not coincide with the ion beam dimensions. In addition, the ionization energy losses of the ion beam in the target cannot be correctly calculated within the framework of this model.

Investigation of the ion beam dynamics for an arbitrary shape of the distribution function is performed using multi-particle simulation within the framework of the model beam algorithm. In this algorithm the ion beam is represented by an array of model particles. The heating and cooling processes involved in the simulation lead to changes in the particle momentum components and particle number, which are calculated in accordance with the dynamics simulation step over time. Each effect is located at some position of the ring characterized by the ring lattice functions. Transformation of the beam inside the ring is carried out using a linear matrix for random phase advance between the effect locations. The simulation results can be presented both as a beam profile evolution over time or as the time dependence of the beam emittance and particle number.

The real ion-ring optic structure is only necessary for IBS diffusion power calculation. The change in particle momentum due to IBS is calculated on the basis of one of the analytical models, as in the case of RMS dynamics simulation.

For simulation of the IBS process through Coulomb interaction between ions, the tracking algorithm is used. One of the goals of this algorithm is to simulate the formation of a crystalline state of the ion beam. In this crystalline state, the IBS process cannot be treated within the framework of analytical models, which are based on the assumption of an ion distribution function of Gaussian shape. To speed up calculations in the tracking algorithm, IBS simulations are performed using a molecular dynamics technique. In this case the motion equations are solved for a small number of particles located inside a short cell. The influence of all other particles is taken into account through periodic boundary conditions in the longitudinal direction for the particle distribution function, and the use of Evald's sum for calculation of the Coulomb forces. Therefore, this algorithm can only be used for a coasting beam.

Within the framework of the tracking algorithm, the particle motion equations are integrated in the real optical structure of the ring. The ring structure is imported from the MAD input file. Each cooling or heating effect involved in the calculations, together with the IBS, is located in some optical element. Calculation of the variation of particle co-ordinates due to any effect is achieved using the MAP effect. The effect position in the ring is described in the MAD input file using special marks.

The structure of basic objects in the BETACOOL program, namely, the models of the ion ring and the ion beam, are developed in such a way that allows realization of all three algorithms with the same input parameters. The heating and cooling effects are realized on the basis of a common standard and the same parameters can be used in each algorithm.

## 3. Structure of the effects

In the present version of the program, the ion beam dynamics can be investigated taking into account one or a few effects from the following list:

(1) Electron cooling;
(2) Intrabeam scattering;
(3) Scattering on residual gas;
(4) Interaction with the internal target;
(5) Collisions with another beam in the collider mode of the ring operation; and
(6) External heating of the ion beam.

Algorithms for simulation of beam–beam effect, stochastic and laser cooling are under development.

The effect structure permits the uniform use of each effect in all basic algorithms. For this purpose, each effect is represented by three models: a transformation map, a kick of the ion momentum, and characteristic time calculation.

The effect used as a transformation map is associated with some optical element of the ring and its position is marked in the input file. The map transforms the particle co-ordinates from the entrance to the exit of the element and calculates the particle loss probability.

On the basis of the transformation map for each effect the procedures for calculation of the particle momentum kick and of the characteristic time are developed. Calculation of the momentum kick is used in the model beam algorithm and characteristic times are necessary for RMS dynamics simulation.

## 4. Electron cooling simulation

The structure of the effects can be illustrated in an example of an electron cooler model.

Usually the action of electron cooling on the ion dynamics inside a storage ring is described using a few standard simplifications:

(1) Angular deviation of the longitudinal magnetic field line is substantially less than the ion beam angular spread.
(2) Ion transverse displacement inside the cooling section is substantially less than the electron beam radius.
(3) The ion beam temperature is substantially greater than the electron temperature and ion diffusion in the electron beam can be neglected.
(4) The electron beam has a round cross-section and a uniform density distribution in the radial direction.

Under these assumptions and using an asymptotic of the analytical friction force, formulae were obtained for the characteristic times for the emittance and momentum spread decrease on electron cooling [8]. In the first version of the BETACOOL program, electron cooling was simulated in accordance with this model [1]. This model is also used in a few programs dedicated to electron cooling simulation. However, this model cannot cover all possible versions of electron cooling designs.

Recently, modifications of the usual configuration of the electron cooling system were proposed. To avoid instability of the ion beam related to the extremely high density of the cooled beam, use of a so-called "hollow" electron beam—a beam with low density in the central part—was proposed. Extension of the electron cooling method in the region of electron energy of a few MeV is related to RF acceleration of the electrons. In this case a Gaussian distribution of the electrons can be expected in the radial plane and, if the electron bunch is shorter than the ion bunch, in the longitudinal direction too. Calculation of the cooling times in this case requires modification of both the electron beam model and the base physical model.

Another expected peculiarity of a medium-energy cooling system is the great length of the cooling section of up to approximately 20–50 m. To obtain very high accuracy of the magnetic field is a difficult technical task and the cost of the cooling system will strongly depend on the required level of accuracy. Therefore, before design of the cooling section solenoid, the influence of the magnetic field line curvature on the cooling process needs to be investigated. All the effects can be taken into account by numerical solution of the ion equations of motion in the cooling section.

To solve all the problems related to simulation of the cooling process, a hierarchy of objects was developed within the framework of the BETACOOL program. The structure of the electron cooler presentation allows the extraction of procedures of different levels and their inclusion in calculations for cooling processes in other programs. The cooling simulation is based on a friction force calculation in the particle rest frame. The friction force can be calculated in accordance with an analytical model from a library or using results of numerical calculations imported from an external file. The next layer of the simulation is related to a cooler representation as a map, transforming particle coordinates from the entrance to the exit of the cooling section and calculating the ion loss probability due to recombination with electrons. Calculation of the cooler map is based on a model of the electron beam that provides transformation of the ion velocity to the frame related to the electron beam and takes into account the real geometry of the cooler. Currently, BETACOOL has three electron beam models available for simulations: a uniform cylinder, a Gaussian cylinder and a Gaussian bunch. A model of a hollow beam will be realized in the near future. The cooler model takes into account variation of the magnetic field in the cooling section. For this aim the co-ordinates of the electron beam trajectory inside the cooling section are input from an additional file and the ion equations of motion inside the cooler are solved numerically.

The map of the cooler can be used directly within the framework of the molecular dynamics algorithm or another tracking procedure. On the basis of the map, the kick in ion momentum after crossing the cooling section can be calculated, which is necessary for simulation of the ion distribution evolution in the model beam algorithm. The map of the cooler is also used for the cooling rate calculation necessary for RMS dynamics simulation. The cooling rate calculation can be carried out using two models of the ion beam: the cooling rates for "rms particles", or cooling rates for an ion beam with Gaussian distribution in all degrees of freedom.

## 5. Software structure

The BETACOOL program is a part of the software developed for electron cooling simulation. The software is divided into two independent parts: a physical code, which is written using only standard C++ syntax; and an interface, which is an executable program working under the Windows environment. Connection between the two parts of the program is provided using three types of files: input, output and files used for control of the calculation process. On one hand, such a structure allows the program to be used on a PC to control the calculation process and analyze results during simulations. On the other hand, the physical part of the program can be compiled for a UNIX operation system and used for calculations independently of the interface. The interface in this case can be used for preparation of the input file and result visualization after completion of the calculations. All input and output files are in text format. The parameters in the input file are divided into groups according to the structure of the BETACOOL objects.

The interface part of the software consists of an executable file (Bolide.exe), *.dfm files containing information about the BETACOOL exterior, and input files for post-processing of the calculated data. Development of the BETACOOL exterior is possible without recompilation of the Bolide.exe file. The interface is also used to work with the file structure on disc.

The physical part of the software consists of the executable file Betacool.exe compiled for Windows or a UNIX operation system and files of input parameters. For intrabeam scattering calculation, use of a file of lattice parameters is required, for instance, a MAD file. Additional input files are used for electron cooling simulation under friction force calculated by other programs and for input of magnetic field errors in the cooling section.

The source code of the physical part of the software consists of three relatively separate parts:

- The interface, which supports the format of input and output files common to the Bolide system;
- A library of base numerical algorithms, including the description of dimensional variables, templates of the program self-counters, procedures for matrix algebra, and algorithms for the numerical solution of differential equations; and
- Physical codes describing the objects of the program and procedures for them.

The structure of the BETACOOL program exterior corresponds to the structure of general objects in the source code and correspondingly to the structure of the input file.

## Acknowledgements

This work was supported by RFBR Grant #02-02-16911 and INTAS Grant 03-54-5584.

## References

[1] A. Lavrentev, I. Meshkov, The computation of electron cooling process in a storage ring, Preprint JINR E9-96-347, 1996.
[2] A. Smirnov, A. Sidorin, E. Syresin, I. Meshkov, T. Tanabe, T. Katayama, Computer simulation of ECOOL and IBS process in ACR and DSR using BETACOOL program, RIKEN-AF-AC-21, April 2000.
[3] E. Syresin, N. Alekseev, V. Bykovsky, D. Koshkarev, Y. Korotaev, I. Meshkov, et al., Phys. Sci. T104 (2003) 160.
[4] V.A. Mikhailov, A.V. Butenko, I.B. Issinsky, I.N. Meshkov, et al., Progress in the nucleotron booster design, in: Proceedings of EPAC 2002, Paris.
[5] I.N. Meshkov, R. Maier, et al., Electron cooling application for luminosity preservation in an experiment with internal targets at COSY, JINR, Dubna, 2001, Jul-4031.
[6] I.N. Meshkov, A.O. Sidorin, A.V. Smirnov, E.M. Syresin, G.V. Trubnikov, P.R. Zenkevich, Simulation of electron cooling process in storage rings using BETACOOL program, in: Proceedings of Beam Cooling and Related Topics, Bad Honnef, Germany, 2001.
[7] F. Christoph Iselin, The MAD Program. Physical Method Manual, Geneva, Switzerland, September 1994.
[8] I. Meshkov, Phys. Part. Nucl. 25 (6) (1994) 631.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 329–332
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Computing at the Dubna gas-filled recoil separator

Yuri S. Tsyganov*, Alexandr N. Polyakov

*Flerov Laboratory of Nuclear Reactions, JINR, 141980 Dubna, Russian Federation*

Available online 1 December 2005

**Abstract**

Simulation codes for the spectra of heavy implanted nuclei, applications for online data visualization and real time PC-based algorithms are considered. Special attention is paid to the application of real time techniques for radical suppression of background products in heavy-ion-induced nuclear reactions at the U-400 cyclotron of the Flerov Laboratory of Nuclear Reactions. The detection system of the Dubna gas-filled recoil separator (DGFRS) is also briefly described. Calculated heavy recoil spectra are compared with those measured in heavy-ion-induced nuclear reactions.


*PACS:* 02.70.Lq; 07.05.Tp; 07.81.+a; 29.30.Ag

*Keywords:* Recoil separator; Heavy ion reaction; Silicon radiation detector; Computer simulation

## 1. Introduction

The problem of experimental verification of the existence of the hypothetical domain of super-heavy nuclides is one of the fundamental outcomes of the nuclear shell model that has been extensively discussed in recent years. Another problem is the existence of even stronger spherical shells beyond $^{208}$Pb, in the domain of heavier neutron-rich nuclei with $Z = 114$ (possibly 120, 122 or 126) and $N = 184$. Since 1998, the Dubna gas-filled recoil separator (DGFRS) heavy-element research group has attempted to verify this non-trivial theoretical hypothesis [1–6]. This region is not easily reached with stable partner nuclei. In order to approach the $N = 184$ shell, maximum neutron excess is needed in both the target and projectile nuclei. With this aim, we used targets of enriched isotopes such as Pu, Am, Cm and Cf, and the rare and expensive isotope $^{48}$Ca as a projectile. We used the main U-400 cyclotron of the Flerov Laboratory of Nuclear Reactions (FLNR) to accelerate $^{48}$Ca ions and the DGFRS to separate in flight the reaction products under investigation from different backgrounds. The separator was filled with hydrogen at a pressure of approximately 1 Torr [7]. Evaporation residue (ER) recoil passed through a time-of-flight (TOF) system, and was implanted into a $4 \times 12\,\mathrm{cm}^2$ semiconductor detector array with 12 vertical position-sensitive strips [8–11]. This detector was surrounded by eight $4 \times 4\,\mathrm{cm}^2$ detectors to provide detection efficiency for alpha decay registration of up to 87% of $4\pi$. The full-width at half-maximum (FWHM) for position resolution of the signals from correlated decays of nuclei implanted in the detectors was 0.8–1.3 mm for ER-alpha signals and 0.5–0.8 mm for ER-spontaneous fission (SF) signals. The PC-based DGFRS data acquisition system provides not only data storage event by event, allowing the accumulation of more than hundred working histograms, but also the visualization and control of definite parameters related to the detection module and the separator set-up.

## 2. Background products

Although the DGFRS (and its analogues, see Ref. [7]) is a highly effective set-up for ER separation, most of the signals detected at the focal plane of the separator during long-term experiments are of a background nature. This is due to different reasons, such as an extremely low cross-section of the product under investigation (sometimes units of picobarns or lower), very intense beams of heavy ions coming from the cyclotron, transfer reaction products,

*Corresponding author. Tel.: +7 09621 64246; fax: +7 09621 65083.
*E-mail address:* tyura@sungns.jinr.ru (Y.S. Tsyganov).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.044

neutron-induced signals in a passivated implanted silicon (PIPS) detector, etc. [9].

The typical counting rate is approximately tens of events per second in the focal plane detector, depending on the cyclotron tuning and reaction conditions. This means that more than $1.0 \times 10^8$ background signals are measured by the detection system during a 1 month experiment, whereas only single detected events can be attributed to the complete fusion products. A significant proportion of the background particles are charged (alpha decay events being the most significant part of the experimental data flow) and, therefore, can be detected by a TOF and/or a veto detector [9].

Of course, neutron-like backgrounds cannot be suppressed with a TOF detector and in this sense, can simulate signals such as alpha particles in a silicon detector. A typical value for the latter parameter is approximately $0.2\,\mathrm{min}^{-1}$ per strip for signals in the energy interval 9.6–11.0 MeV.

## 3. Real time detection mode for recoil–alpha-correlated sequences

In many experiments we used a special detection mode to detect two or more sequential decays. The beam was switched-off after a recoil signal was detected with parameters of implanted energy and TOF expected for the evaporation residues, followed by an alpha-like signal within preset energy and time intervals in the same strip, within a position window corresponding to the position resolution. Thus, all the expected sequential decays of the daughter nuclides should be observed in the absence of a beam-associated background. We used a more complicated form of beam switching based on recoil–alpha correlation chain real time detection, except for single recoil detection [18] to minimize total experimental efficiency losses for correlation times up to tens of seconds. In a more detailed form, it can easily be shown that the equation relating to these two approaches from the viewpoint of equivalence of total experimental efficiency losses is

$$\tau_{R\alpha} = \sqrt{\frac{\tau_R}{v_\alpha}}, \tag{1}$$

where $\tau_{R\alpha}$ is the recoil–alpha correlation time, $\tau_R$ is the recoil time (duration of pause, generated by recoil detection [18]) and $v_\alpha$ is the signal rate, such as alpha particles per strip and per actual vertical position element.

The real time algorithm for searching recoil–alpha sequences operates in parallel with data capture and file writing. It should be noted that such parallelism is achieved by applying a specially designed autonomous CAMAC crate controller [12] operating together with intermediate buffer memory for the main data flow, whereas an intellectual KK-012 crate controller [13] operates with the imminent event to find a correlated recoil–alpha pair. The basic idea of the algorithm, in brief, is that it uses the discrete representation of the PIPS detector separately for recoils and alpha particle signals. Thus, the real detector is presented in the form of two matrixes, one for "ER" and the other for "alpha particles" [9,14,15], with elapsed time as a matrix element in both cases, whereas the matrix indexes correspond to strip number and vertical position (in discrete representation). In addition [9], to prolong the "beam OFF" interval up to tens of minutes or hours, information about the strip number is used, as well as extraction of the next (after switch-off) alpha particle with its vertical position, although with a wider position interval in comparison with one generating the beam switch-off. Thus, modified in comparison with Ref. [9], the working system of equations is as follows:

$$j = \mathrm{int}\left\{ N_{\max}\left[\frac{a_i N_\alpha + b_i}{a_{yi} + b_{yi}}\left(\frac{R0_i}{R_i} + 1\right) - \frac{R0_i}{R_i} + \delta_i^{\alpha,\mathrm{esc,ER}}\right]\right\}, \tag{2}$$

$$t_{i,j}^{a,\mathrm{ER}} = t(\mathrm{elapsed}), \tag{3}$$

$$t_{i,j+k}^{a1,a2} - t_{i,j}^{\mathrm{ER}} \leqslant T_{\mathrm{pr}}^{1,2}(E_{\alpha_{1,2}})\ (\mathrm{recoil-alpha}), \tag{4}$$

(Usually $T_{\mathrm{pr}}$ is a constant, but in the same experiments a relation of the form $T \approx 10[(aZ+b)Q^{-1/2} + cZ + d]$ is used, where $a = 1.78$, $b = -21.398$, $c = -0.25488$ and $d = -28.423$. $Q$ is the estimated decay energy.)

or:

$$t_{i,j+k}^{\mathrm{a2}} - t_{i,j}^{\mathrm{a1}} \leqslant T_{\mathrm{pr}}^{\mathrm{a}}(E_{\alpha_{1,2}})\ (\mathrm{alpha-alpha}), \tag{5}$$

$$t_{i,j+2k}^{\mathrm{a\_OFF}} - t_{i,j}^{\alpha 1,2} \leqslant \delta T_{\mathrm{OFF}}$$
$$\Rightarrow \delta T_{\mathrm{OFF}} = K\delta T_{\mathrm{OFF}}\ (\mathrm{beam\text{-}OFF prolongation}). \tag{6}$$

In these equations, $T_{\mathrm{pr}}(E_{\alpha_{1,2}})$ are the preset time intervals for the first and second (if the first escapes with a position signal below the 500–1000 keV threshold) alpha particles, respectively; $N_{\max} = 170$ is the maximum value of cells per strip in a discrete representation; $i = 1, \ldots, 12$ is the strip number; $k = -3, \ldots, +3$ is the cell index [9,14,15]; $K = 10$–$20$ is the prolongation factor; and $\delta T_{\mathrm{OFF}}$ is the duration of the beam-OFF interval after detection of the first correlation chain. Subscript a_OFF indicates that the appropriate alpha particle is detected in the beam-OFF interval. Other parameters in Eqs. (2)–(6) are calibration constants that were extracted from the calibration reactions and parameters of the PIPS detector. The calibration parameter $\delta$ is equal to zero in the ER case, and is a small, non-zero value in the case of alpha decay in the focal plane detector and for an alpha particle escaping this detector, detected in the backward detector and having a non-zero component in the focal plane detector (superscripts $\alpha$ and esc, respectively). It is extracted from the calibration reactions $^{\mathrm{nat}}\mathrm{Yb} + {}^{48}\mathrm{Ca}$ and $^{207}\mathrm{Pb} + {}^{48}\mathrm{Ca}$. Subscripts 1 and 2 for $\alpha_{1,2}$ indicate the first and second (forthcoming) alpha particles that cause the beam-OFF pause.

Figs. 1(a) and (b) shows two time sequences in the decay chains observed at $^{48}$Ca energy values of (a) $E_L = 248$ MeV and (b) $E_L = 253$ MeV, as well as the energy measured,

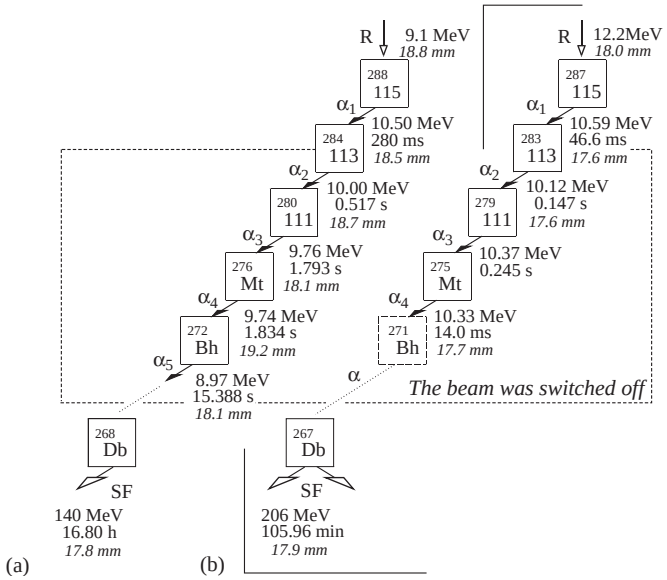



Fig. 1. Two sequences of the decay chains observed at different $^{48}$Ca energy values: (a) $E_L = 248$ MeV and (b) $E_L = 253$ MeV.

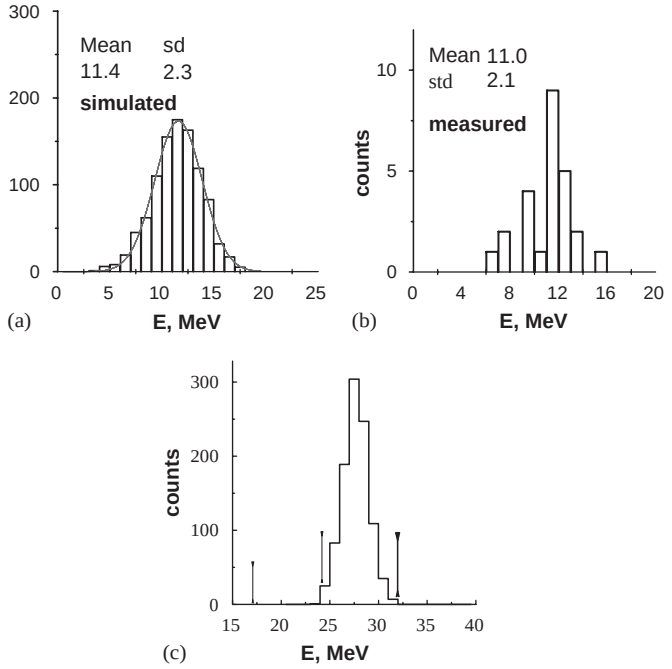



Fig. 2. (a) Simulated and (b) measured ($Z = 114$–$116$) spectra of evaporation residue recoils. (c) Simulated spectrum of super heavy recoils for $^{208}Pb + {}^{70}Zn \rightarrow 112 + 1n$ nuclear reaction. Detected events [17,18] are shown by arrows.

time intervals and vertical positions, with respect to the top of the strips for the decay events observed [1]. During the last 3 yr, more than 50 decay chains for $Z = 110$–$116$ were detected under low-background conditions in long-term experiments.

## 4. Super-heavy recoil spectra calculation in a silicon radiation detector

When applying the above technique, a reasonable question is how to estimate the shape and position of an ER spectrum due to the relatively high pulse-height defect (PHD) in a silicon radiation detector. There are several ways that are actually complementary to each other. Estimation based on the measured mass of the ER is one of these. Of course, of great interest is the possibility of computer simulation based on knowledge of the energy losses and broadening in different media [16], the neutron evaporation contribution to broadening, and PHD and its fluctuations in silicon. The code described in Ref. [16] allows a simulation that takes into account the reasons mentioned for transformation of the spectra originating in the target to that registered by a PIPS detector. Figs. 2(a)–(c) shows calculated and measured spectra of registered energy for the two reactions. Note that Fig. 2c corresponds to the GSI experiment aimed at the synthesis of the $Z = 112$ element some years ago [17] and that Fig. 2a included no free parameters in the simulation procedure. Special attention should be paid to the fact that the event marked by the left arrow (2c) on the histogram was eliminated, as described in [18], after careful analysis of the raw data. In May 2000 the experiment $^{70}Zn + {}^{208}Pb \rightarrow 112 + 1n$ was repeated by the same group [18]. One additional event of element 112 with ER energy of 24.1 MeV was reported. This event is shown in Fig. 2c with an appropriate comment. The remaining questions are the formation mechanism for PHD, especially its recombination component, and a more exhaustive discussion of the applicability and limits of the overall "one event–one element" philosophy, but such a discussion is outside the scope of this paper.[1]

## 5. Conclusions

Together with the development and improvement of both accelerating and separation techniques, computer codes applied in active form not only allow data capture and visualization, but also play an even more significant part in improving the overall experimental conditions from the viewpoint of effect/background ratio. This resulted in the possibility of establishing genetic links between alpha particles for decay times of up to tens of hours or even days. We plan to use this approach in forthcoming experiments with an estimated cross-section for complete fusion reaction product below 1 pb.

When preparing this manuscript, an experiment [19] on chemical identification of the $^{268}$Db isotope as a product terminating decay chains of the $Z = 115$ element from the reaction $^{243}Am + {}^{48}Ca \rightarrow {}^{288}115 + 3n$ was successfully performed. This confirmed the decay properties of $^{268}$Db measured previously [1].

[1] In the case of large (tens of percent) recombination components, it should be measured, although for model recoil of approximately $Z = 100$ with comments on a reasonable type of recombination in silicon.

## Acknowledgements

The authors are indebted to Drs. V.K. Utyonkov, A.M. Sukhov, I.V. Shirokovsky and S. Iliev for their assistance and fruitful discussions on the item reported in this paper.

## References

[1] Yu.Ts. Oganessian, et al., Phys. Rev. C 69 (2004) 021601(R).
[2] Yu.Ts. Oganessian, et al., Phys. Rev. C 69 (2004) 054607.
[3] Yu.Ts. Oganessian, et al., Eur. Phys. J. A 15 (2002) 201.
[4] Yu.Ts. Oganessian, et al., Nucl. Phys. A 734 (2004) 109.
[5] Yu.Ts. Oganessian, V.K. Utyonkov, Yu.V. Lobanov, et al., Rev. Mex. Fis. 46 (Suppl. 1) (2000) 35.
[6] Yu.Ts. Oganessian, et al., Phys. Rev. C 64 (2001) 054606.
[7] K. Subotic, et al., Nucl. Instr. and Meth. Phys. Res. A 481 (2002) 71.
[8] Yu.S. Tsyganov, V.G. Subbotin, A.N. Polyakov, et al., Nucl. Instr. and Meth. Phys. Res. A 392 (1997) 197.
[9] Yu.S. Tsyganov, V.G. Subbotin, A.N. Polyakov, et al., Nucl. Instr. and Meth. Phys. Res. A 525 (2004) 213.
[10] Yu. Tsyganov, JINR Commun. P-10-98-20, Dubna, 1998 (in Russian).
[11] Yu. Tsyganov, JINR Commun. E-10-99-36, Dubna, 1999.
[12] Yu.S. Tsyganov, et al., in: Proceedings of High Performance Computing on the Information Superhighway, HPC Asia-97, 28 April–2 May 1997, Seoul, Korea, IEEE Computer Society Press, Los Alamos, CA, 1997, p. 651.
[13] I.N. Churin, et al., JINR Commun. P10-90-589, Dubna, 1990 (in Russian).
[14] Yu.S. Tsyganov, A.N. Polyakov, A.M. Sukhov, Nucl. Instr. and Meth. Phys. Res. A 513 (2003) 413.
[15] Yu. Tsyganov, J. Phys. G Nucl. Part. Phys. 25 (1999) 937.
[16] Yu. Tsyganov, Nucl. Instr. and Meth. Phys. Res. A 378 (1996) 356.
[17] S. Hofmann, et al., Zh. Phys. A 354 (1996) 229.
[18] S. Hofmann, et al., Eur. Phys. J. A 14 (2002) 147;
V. Ninov, K.E. Gregorich, W. Loveland, et al., Phys. Rev. Lett. 89 (3) (2002) 039901(E).
[19] Yu.Ts. Oganessian, et al., Phys. Rev. C 72 (2005) 034611.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 333–335
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# The features of high intensity beam dynamics in low energy super-conducting linear accelerator

N. Vasyukhin*, R. Maier, Y. Senichev

*FZJ/IKP, Jülich, Germany*

Available online 28 November 2005

**Abstract**

The super-conducting linear accelerator is considered as a candidate for the high intensity beam acceleration in low energy region from 2–5 MeV and up to 50–100 MeV. The conventional room temperature linear accelerator was struck off the list of effective accelerators for this purpose, because of the high power RF and the appropriate cooling systems are needed to produce high gradient with duty factor near to 100%. The super-conducting linear accelerator has a number of advantages in comparison with room temperature accelerators. However, due to RF requirements the geometry of super-conductive cavity has to be simplified as much as possible. Also cavities belonging to one family should have identical geometry. Therefore in such accelerators the longitudinal and transverse dynamics of high intense beam are more complicated. In this paper, we consider 3D beam dynamics together with space charge effect. The latter is investigated in the system without synchronism, when focusing and defocusing factors in longitudinal and transverse planes are the time dependent functions.


*PACS:* 29.27.a; 29.17.+w; 52.59.Sa; 52.65.Rr

*Keywords:* Linear accelerators; Space-charge-dominated beams; Particle-in-cell method (simulations)

## 1. Halo creation

Super-conducting linear accelerators have a number of advantages in comparison with room temperature accelerators: high accelerating gradient, low energy losses and high duty cycle. But despite of these benefits some specific features of super-conducting accelerators like high RF defocusing factor or long focusing period makes the working point smeared in the Smith–Gluckstern stability diagram for transverse plane. Additionally compensation of linear space charge detuning by quadrupoles moves halo particles out of the stable region due to non-linear detuning of these particles. Attempts to describe qualitatively halo creation mechanism have been undertaken, but in this case an analytical solution is possible only with a significant simplification of model. It is generally accepted that it is a more or less resonant process, in which some particles increase amplitude of transverse oscillations. It was shown [1], that the mechanical concept is more appropriate than a thermodynamical one for linear accelerator description. The point was to describe qualitatively non-linear parametric resonances due to mismatched RMS size of bunch and forced envelope oscillations in alternative gradient focusing structure. These resonance conditions should increase the amplitude of transverse oscillations. To solve the equation analytically a number of approximations were undertaken. Following this idea one can obtain resonant Hamiltonian for particles motion in case of mismatched beam

$$H_{\mathrm{r}} = -\frac{(\delta I)^2}{2} R_c C_{\mathrm{sc}} \varepsilon^{-2} \cdot \left\{ R_c \sum_{k=2}^{m} (k-1) c_k I_0^{k-2} G_{2k} + (n-1) r_m c_n I_0^{n-2} \left( b_{n0} + \frac{1}{2} d_{n0} \right) \cos(2n\psi) \right\}$$

where $\beta$ is the beta function; $\tilde{\mu}$ the frequency of beta function oscillation due to mismatching; $n$ the order of resonant term in binomial distribution; $m$ the number

*Corresponding author.
*E-mail address:* n.vasyukhin@fz-juelich.de (N. Vasyukhin).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.042

of terms in binomial distribution; $c_k$ the coefficient in distribution; $C_{sc}$ the space charge coefficient; $\varepsilon$ the emittance; $G_{2k} = (2k)!/2^{2k}(k!)^2$; $d_{n0} = 1/2^{2n-2}$; $b_{n0} = 1/2^{2n-1}$; $r_m$ the amplitude of mismatching; $\psi$ the slow phase; $I_0$ the meaning of invariant in resonance vicinity; $\delta I$ the deviation of invariant because of nonlinear terms of space charge.

The first term is linear detuning, while the second term describes the deviation of non-linear detuning and the third term is the resonant term. But this is valid only for a qualitative understanding of the phenomenon.

## 2. Superconducting accelerator features

From the simulation point of view the most important features of super-conducting linacs are:

- *High accelerating gradient*. In a super-conducting cavity one can achieve very high accelerating fields. Therefore at low energy the velocity gain is high and any field approximation leads to incorrect results. The defocusing factor of accelerating field is phase dependant. This leads to a smearing of the working point in the stability diagram.
- *Absence of synchronism*. It is desirable to have the super-conducting cavity with constant phase velocity geometry. In contrast to the normal-conducting linac, in the super-conducting linac the equivalent phase velocity is adjusted via phasing of the cavities [2]. Therefore effective separatrix decreases.
- *Relatively long focusing period*. In some cases of super-conducting linacs the focusing elements are located between cryomodules only. The transition from the cooled cavities to warm quadrupoles takes some distance. However the space charge effect is proportional to the square of focusing period. For instance in the COSY injector design [3] the focusing period is 1.7 m.

Besides, the compensation of linear space charge leads to an additional decreasing of the stable region in the case of a high intensity beam. The particles in the halo, which are affected by non-linear tune shift, are subject to resonances and can be lost. In theory at low energy it is difficult to take into account these features. Therefore for investigation of beam dynamics the particle-in-cell code with space charge was developed at FZJ. The electrical and magnetic 3D RF field were calculated by CST MicroWave Studio. The space charge force is computed in 3D by the FFT method with rectangular boundary conditions.

## 3. Simulation results

As an example COSY injector linear accelerator design is considered. The focusing period consists of four super-conducting half wave cavities [4] in one cryo-module and quadrupole doublet. In this design it is possible to accelerate 30 mA average current with 100% duty factor with losses 5 W/m. In 11 periods the proton beam is accelerated from 2.5 MeV up to 50 MeV. The first 5 periods operate at 160 MHz and the remaining 6 modules operate at 320 MHz. The transverse RMS sizes are shown in Fig. 1. In this case it is difficult to make matching because the space charge linear tune shift is about 40%. However, the halo particles have a different tune shift. Therefore the fixed level of losses is the criterion for optimisation. The transverse RMS emittance grows by 30%.

To investigate how the focusing period length affects on peak current value the simulation has been done for the modified structure, where there are only 2 resonators in one cryo-module and focusing period is double shorter. During simulation we have concluded that maximum current limit raised up to 160 mA with the same level of losses. However, from the technical point of view this scenario is more difficult, since so short focusing period is possible only with super-conducting quadrupoles or solenoids located inside the cryo-module.

In order to compare the numerical model with theory at resonant conditions the real accelerator was simulated. The resonance due to mismatching cannot be clearly expressed in phase space, because the space charge forces change while the particle distribution changes. In the analytical formula the terms responsible for that should be expressed through the high order term in the particle distribution. Obviously this leads to an insignificant redistribution. Another resonance is the forced resonance caused by alternate gradient focusing element. As an example the simulation has been done for the accelerator with resonant conditions (see Fig. 2), when the phase advance for period is near to $\pi/2$. It means the lower harmonic of $q$-function has the frequency by factor four higher of eigen frequency, and we see the 4th order resonance. In reality it is desirable to avoid this case.

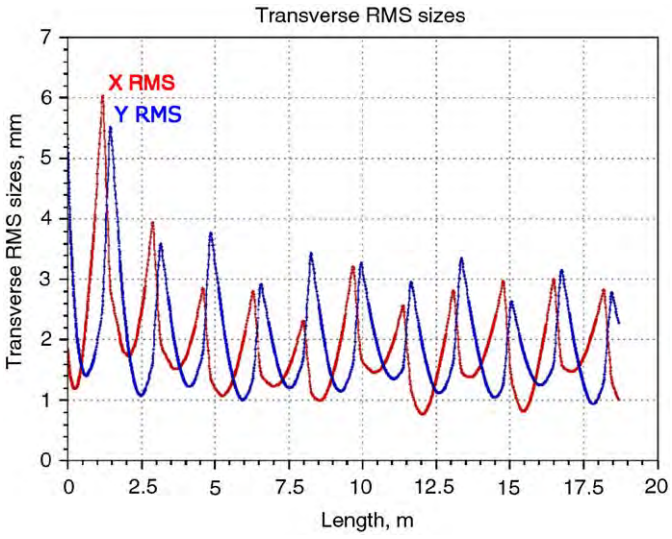



Fig. 1. Transverse RMS vs. linac length in case of high intense beam in COSY-injector design.

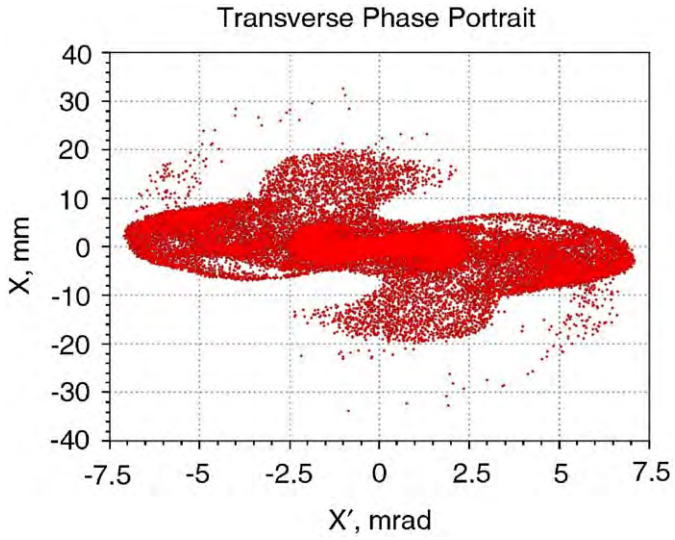



Fig. 2. Transverse phase portrait in case of resonance for $q$ function with particle motion.

## 4. Discussions and conclusions

The high intensity accelerators based on super-conducting structures can be used for low energy part as well. However, the specific features have to be taken into account at the design stage. First of all in the case of high current and at low energy it is very important to make 3D matching. The longer focusing period leads to higher sensitivity of space charge that is bigger coefficient $C_{sc}$. This fact causes various resonances. They should be avoided to minimize the problems with the transverse emittance growth.

## Acknowledgements

We thank R. Tölle for several informative discussions. We acknowledge the support of the European Community-Research Infrastructure Activity under the FP6 "Structuring the European Research Area" program (CARE, contract number RII3-CT-2003-506395).

## References

[1] Yu. Senichev, Mechanical and thermodynamical approach to halo creation problem in high intense beam, HEACC-98, Dubna, 1998.
[2] Yu. Senichev, A. Bogdanov, R. Maier, Phys. Rev. ST AB 6 (2003) 124001.
[3] R. Tölle, et al., A Super-conducting injector linac for COSY, EPAC 02, Paris, 2002, p. 966.
[4] ⟨http://www.fz-juelich.de/ikp/publications/AR2003/CHAP6/607.pdf⟩

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 336–339
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Precision PEP-II optics measurement with an SVD-enhanced Least-Square fitting

Y.T. Yan*, Y. Cai

*Stanford Linear Accelerator Center, Stanford, CA 94309, USA*

Available online 1 December 2005

**Abstract**

A singular value decomposition (SVD)-enhanced Least-Square fitting technique is discussed. By automatic identifying, ordering, and selecting dominant SVD modes of the derivative matrix that responds to the variations of the variables, the converging process of the Least-Square fitting is significantly enhanced. Thus the fitting speed can be fast enough for a fairly large system. This technique has been successfully applied to precision PEP-II optics measurement in which we determine all quadrupole strengths (both normal and skew components) and sextupole feed-downs as well as all BPM gains and BPM cross-plane couplings through Least-Square fitting of the phase advances and the Local Green's functions as well as the coupling ellipses among BPMs. The local Green's functions are specified by 4 local transfer matrix components R12, R34, R32, R14. These measurable quantities (the Green's functions, the phase advances and the coupling ellipse tilt angles and axis ratios) are obtained by analyzing turn-by-turn Beam Position Monitor (BPM) data with a high-resolution model-independent analysis (MIA). Once all of the quadrupoles and sextupole feed-downs are determined, we obtain a computer virtual accelerator which matches the real accelerator in linear optics. Thus, beta functions, linear coupling parameters, and interaction point (IP) optics characteristics can be measured and displayed.


*PACS:* 41.85.−p

*Keywords:* Optics measurement; Least-Square fitting; Electron–positron storage rings

## 1. Introduction

For a system with sufficient known constraints and specific quantities that can be accurately measured, one may be able to build a computer model that can simulate the system. One needs to identify efficient variables as inputs to the model that can generate response outputs that are to be identified to be equal to their corresponding system-measured specific quantities. This requires training process that re-identifies new efficient variables and eliminates degenerate or unnecessary (low efficient) variables through Least-Square fitting of those well-chosen responses from the model to those corresponding measured quantities. In many occasions, this training process may also involve identifying new responses with measurable corresponding quantities from the system. To interpret the above in short, one can have a simple mathematical formulae,

$$\vec{Y}(\vec{X}) = \vec{Y}_m \tag{1}$$

where all variables are represented by an array (a vector) $\vec{X}$; responses and their corresponding measurable quantities are represented by an array $\vec{Y}$ and an array $\vec{Y}_m$, respectively. Note that $\vec{Y}$ is the response to the $\vec{X}$ and therefore is a vector function of $\vec{X}$ as is explicitly shown in the equation. Also note that the array length of $\vec{Y}$ and $\vec{Y}_m$ must be the same and must be larger than the array length of $\vec{X}$ to avoid any degeneracy. Indeed, the array length of $\vec{Y}$ is preferred to be significantly larger than the array length of $\vec{X}$ to make sure an over determined Least-Square fitting for adding on accuracy and most of the time adding on convergence, too. The Least-Square fitting is to update $\vec{X}$ through iteration such that the residual of $\vec{Y}(\vec{X}) - \vec{Y}_m$ converge to a minimum that is sufficiently small.

*Corresponding author.
*E-mail address:* yan@slac.stanford.edu (Y.T. Yan).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.071

## 2. The SVD-enhanced Least-Square fitting

To perform Least-Square fitting, let us first denote the iteratively updated or the initially reasonably guessed variable values to be $\vec{X}_0$ and let $\vec{X} = \vec{X}_0 + \vec{x}$. Then, after performing Taylor expansion of $\vec{Y}(\vec{x}_0 + \vec{x})$, Eq. (1) can be written as:

$$Y(\vec{X}) = Y(\vec{X}_0 + \vec{x}) = \vec{Y}(\vec{X}_0) + M\vec{x} + \vec{\eta}(\vec{x}) = \vec{Y}_m \tag{2}$$

where the Taylor expanded nonlinear term $\vec{\eta}(\vec{x})$ is much smaller than the linear term $M\vec{x}$ and finally becomes negligible once $\vec{X}_0$ is getting close to $\vec{X}$ (and so $\vec{x}$ is very small) through convergent Least-Square fitting iterations. Thus, for fitting iteration purpose, Eq. (2) can be written as

$$M\vec{x} = Y_m - Y(\vec{X}_0) = \vec{b} \tag{3}$$

where $\vec{b}$ is the residual after a given fitting iteration. The task is to take a limited fitting iterations to achieve a small enough residual $\vec{b}$.

If one were to consider the regular Least-Square fitting, then each iterative equation would be simply

$$\vec{x} = (M^{\mathrm{T}}(\vec{X}_0)M(\vec{X}_0))^{-1}M^{\mathrm{T}}(\vec{X}_0)\vec{b}$$

where $\vec{X}_0$ and $\vec{b}$ are updated from the last iteration by taking $\vec{X}_0 = \vec{X}_0 + \vec{x}$ and $\vec{b} = Y_m - Y(\vec{X}_0)$. However, such regular Least-Square fitting cannot take care of degeneracies that ultimately cause the iteration to diverge. To overcome the degeneracy effect, we consider an SVD-enhanced Least-Square fitting by identifying and eliminating those degeneracy modes in the iteration so as to always get a convergent iteration process. In practice, by identifying the dominant SVD modes, we actually select those efficiently convergent modes for an optimized converging iterations.

Let us perform a singular value decomposition for the derivative matrix $M$ as follows:

$$M = U\Lambda V^{\mathrm{T}}, \tag{4}$$

where $\Lambda$ is the singular value diagonal matrix with singular values, $\vec{\lambda}$, given in an order from a large to a small magnitude. Then the Least-Square fitting solution becomes

$$\vec{x} = V\Lambda^{-1}U^{\mathrm{T}}\vec{b} = \sum_{i=1}^{n}\frac{1}{\lambda_i}(\vec{V}_i\vec{U}_i^{\mathrm{T}})\vec{b}. \tag{5}$$

Since the larger singular modes of $M^{\mathrm{T}}M$, which is proportional to $\lambda_i^2$, are more efficient, the SVD modes are re-arranged in the order of the magnitudes of $\lambda_i^3\vec{U}_i^{\mathrm{T}}\vec{b}$. The first $k<n$ modes are automatically tested for efficiency and then chosen for each iteration or sub-iteration, i.e.

$$\vec{x} = \sum_{j=1}^{k}\frac{1}{\lambda_j}(\vec{V}_j\vec{U}_j^{\mathrm{T}})\vec{b}.$$

We have successfully applied such an SVD-enhanced Least-Square fitting to the study of PEP-II optics measurement.

## 3. Application to PEP-II optics measurement

With the above SVD-enhanced Least-Square fitting technique applied for PEP-II optics studies, we consider all quadrupole strengths and sextupole feed-downs in the

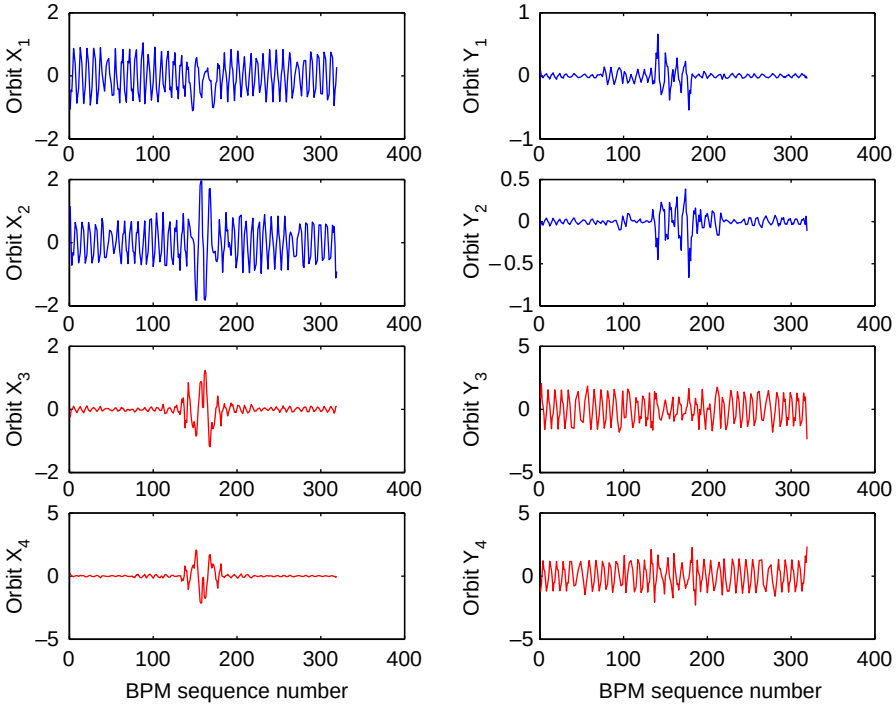



Fig. 1. Four independent orbits extracted from PEP-II LER BPM buffer data taken on January 13, 2004. The first two orbits ($x1, y1$) and ($x2, y2$) are extracted from beam orbit excitation at the horizontal tune while the other two orbits ($x3, y3$) and ($x4, y4$) are from excitation at the vertical tune.

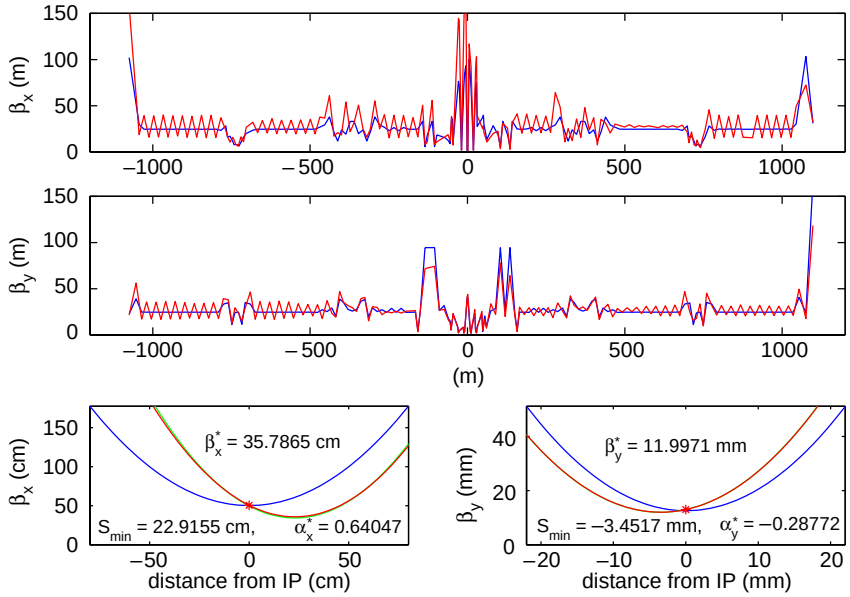



Fig. 2. Typical plots for comparing measured beta functions from the virtual accelerator (red color) to those of the designed lattice (blue color) for PEP-II LER on January 13, 2004. The top two plots show the beta functions for the whole ring and then the bottom plots for beta functions at IP, which show the $\beta^*$'s and the waists.

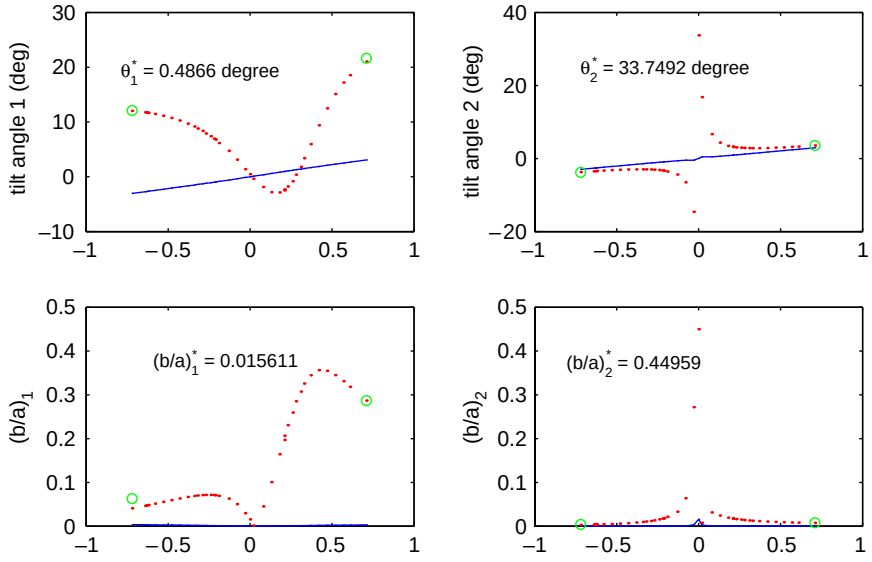



Fig. 3. Typical plots to show IP linear coupling characteristics. The eigen-mode ellipses' tilt angles (top plots) and axis ratios (bottom plots) are compared between measurement (red) and the design lattice (blue).

lattice model as well as all BPM gains and BPM cross-plane couplings as variables, i.e., the $\vec{X}$. $\vec{Y}(\vec{X})$, the to-be-calculated responses from the lattice model, are the Local Green's functions [1], the phase advances [2] as well as the eigen-mode coupling ellipses tilt angles and axis ratios among BPMs [3], $\vec{Y}_m$, the corresponding measurable quantities, are derived from orbit measurement using a model-independent analysis (MIA) [4]. Once the lattice model is fitted to the orbit measurement, we call this lattice model the computer virtual accelerator which matches the real accelerator in linear optics. To obtain the $\vec{Y}_m$, i.e., the Local Green's functions represented by the matrix components, $R_{12}, R_{14}, R_{32}, R_{34}$, the phase advances, and the eigen-mode coupling ellipses tilt angles and axis ratios among BPM locations, we make two resonance excitations at the horizontal and vertical tunes, respectively, to obtain two pairs of Fourier conjugate (Cosine-like and Sine-like) orbits, one pair (two orbits) for the horizontal-tune resonance excitation and the other pair for the vertical-tune resonance excitation as shown in Fig. 1 for a typical sample from PEP-II Low-Energy Ring (LER). The $Y_m$ are derived from these 4 independent linear orbits. Since the

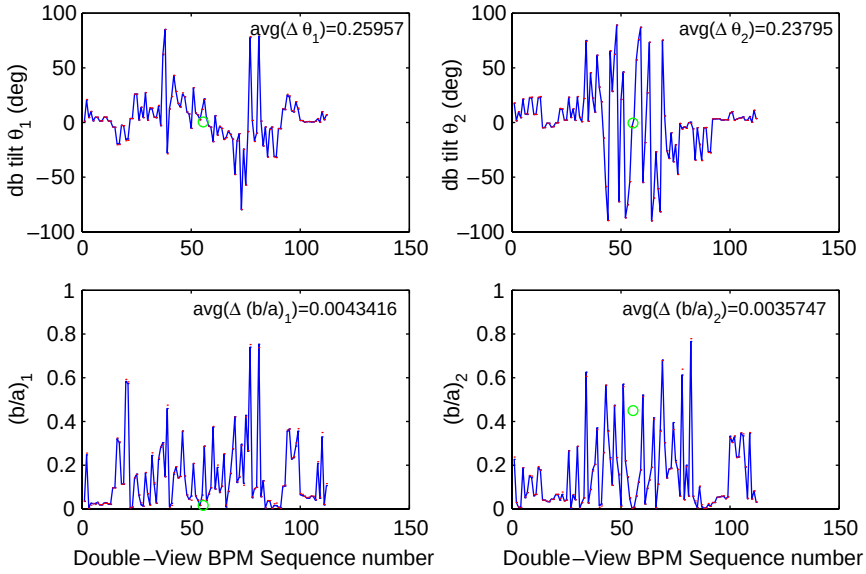



Fig. 4. Typical plots to show linear coupling characteristics for the whole ring. The eigen-mode ellipses' tilt angles (top plots) and axis ratios (bottom plots) are compared between measurement (red) and the design lattice (blue) at all double-view BPM locations.

linear optics is determined by 4 independent linear orbits, we have a complete set of constraints for the SVD-enhanced Least-Square fitting to determine the computer virtual accelerator. Fig. 2 shows a typical set of PEP-II linear optics measured with this SVD-enhanced Least-Square fitting, where the beta functions of the PEP-II LER for the whole ring are calculated from the fitted virtual accelerator comparing with the original ideal lattice model. The beta beating shown are subsequently corrected by first correcting the virtual accelerator and then applied to the real accelerator [5].Fig. 2 also shows the beta function and the waist shifts at IP. The other IP optics characteristics, i.e., the linear coupling parameters (eigen-mode ellipse tilt angles and axis ratios at IP) are shown in Fig. 3. These IP optics measurements provide valuable information about geometrical process of the e+–e− collisions that helps subsequent adjustment of the IP optics. The linear coupling parameters, the eigen-mode ellipses' tilt angles and axis ratios at double-view BPM locations around the whole ring are shown in Fig. 4.

## 4. Summary

We have developed a mature SVD-enhanced Least-Square fitting that has been successfully applied in the PEP-II linear optics studies as proved by improvement of the PEP-II optics [5]. The success basically comes from three key points: (a) the SVD-enhances Least-Square fitting can avoid degeneracies and has a fairly fast convergent rate allowing for application to a fair large system such as the PEP-II ring as a whole; (b) the PEP-II ring has a reasonable amount of BPMs allowing for extracting sufficient physical quantities for fitting; and (c) the linear Green's functions between any two BPMs can provide even more fitting constraints that add significantly on the convergence.

## Acknowledgements

We thank F.-J. Decker, S. Ecklund, J. Irwin, J. Seeman, M. Sullivan, J. Turner, U. Wienands for helpful discussions. The work is supported by Department of Energy contract DE-AC02-76SF00515.

## References

[1] J. Irwin, Y.T. Yan, Beamline model verification using model-independent analysis, SLAC-PUB-8515, 2000, in: EPAC2000 Conference Proceedings; Y. Cai, J. Irwin, M. Sullivan, Y.T. Yan, Application of model-independent analysis to PEP-II rings, in: PAC2001 Conference Proceedings.
[2] Y.T. Yan, Y. Cai, J. Irwin, M. Sullivan, Linear optics verification and correction for the PEP-II with model-independent analysis, SLAC-PUB-9368, 2002, in: Proceedings of the 23rd Advanced Beam Dynamics Workshop on High Luminosity e + e− Colliders, 2002.
[3] Y.T. Yan, Y. Cai, Precision measurement of coupling ellipses parameters in a storage ring, SLAC-PUB-10371, 2004.
[4] J. Irwin, C.X. Wang, Y.T. Yan, et al., Phys. Rev. Lett. 82 (1999) 1684.
[5] Y.T. Yan, et al., PEP-II beta beat fixes with MIA, SLAC-PUB-10369, in: Proceedings of the 30th Advanced Beam Dynamics Workshop on High Luminosity e + e− Colliders, 2003.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 340–345
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Computational modeling of magnets for electrophysical setups

I.P. Yudin*, E.E. Perepelkin, R.V. Polyakova, T.V. Shavrina

*Joint Institute for Nuclear Research, Dubna, Moscow region, Russia*

Available online 6 December 2005

**Abstract**

The design, construction and operation of magnetic systems of some electrophysical setups require preliminary mathematical modeling. While calculating the fields of the magnetic systems mentioned (based on solving a set of direct problems of magnetostatics), we attempted to solve the inverse problem of magnetostatics, namely, to find an optimal construction of the current elements and ferromagnetic yoke for a required distribution of the magnetic field. This work discusses the results of numerical modeling of the magnet field distribution for some modifications of the SP-94 spectrometric magnet used in some experimental setups.


*PACS:* 29.30.−h; 02.60.Cb

*Keywords:* Spectrometers; Magnetic systems; Numerical simulation; Direct and inverse problem of magnetostatics

## 1. Introduction

The significance of numerical modeling for investigating magnetic systems is emphasized not only by the importance of computational experiments, but also by the fact that measurement of magnetic fields is labor-intensive and expensive.

The self-process of the mathematical modeling of magnetic systems needs to be divided into two major stages. For the first stage involving the creation of the new magnetic system, it is advisable to have software for quick and operational modeling that has the properties of a ''logarithmic ruler'', i.e. with accessibility, simplicity of application and sufficient accuracy for numerical calculations. The second stage involves study of the configuration chosen for the magnetic system in more detail, i.e. making more accurate numerical calculations for two- and three-dimensional configurations.

In the present work we propose a method for refinement of the numerical solution of a magnetostatic problem for an area with boundaries containing the ''corner points''.

Using this method, we carried out numerical simulation of a set of modifications for the SP-94 magnetic system.

## 2. Geometry

The coordinate system (the right-hand triplet) is chosen in such a way that $Z$-axis is directed along the beam. The horizontal plain is $X0Z$. Fig. 1 shows half of the magnet yoke cross-section on the $X0Y$ plane (orthogonal to the beam), while Fig. 2 shows the $\frac{1}{8}$th cross-section on the vertical plane $Y0Z$. The domains $\Omega_f$ and $\Omega_c$ are filled with a ferromagnetic and a current source, respectively. Table 1 gives the linear size (in m) of the magnetic system (a basic version, without modifications).

The curve $\mu = \mu(B)$ corresponding to the steel used is given in Fig. 3.

## 3. Description of the program software

The first stage in the solution of the magnetostatics task involved the creation of the CPMMS-1.0 complex of programs [1] for the numerical modeling of magnetic systems. This includes software programs such as POISSON [2] for the numerical modeling of two-dimensional magnetic fields, and graphics packages such as GRAPHER, SURFER and PAW [3]. The CPMMS complex of

*Corresponding author. Tel.: +7 49621 64 922; fax: +7 49621 65 767, +7 09621 65 767.
*E-mail address:* yudin@jinr.ru (I.P. Yudin).

0168-9002/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.nima.2005.11.043

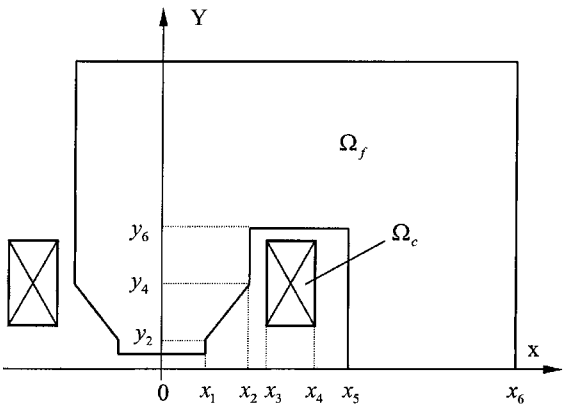



Fig. 1. Half of the magnet yoke cross-section on the X0Y plane.

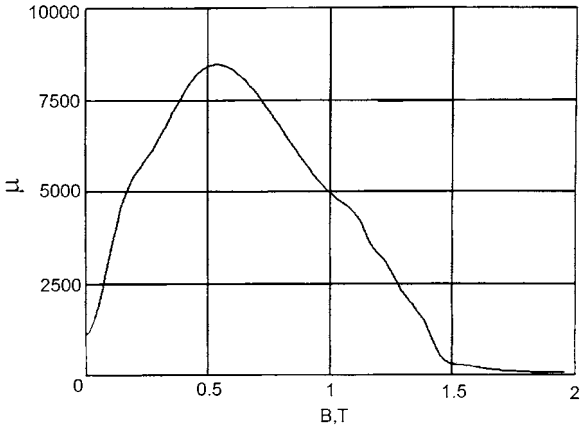



Fig. 3. The curve $\mu = \mu(B)$.

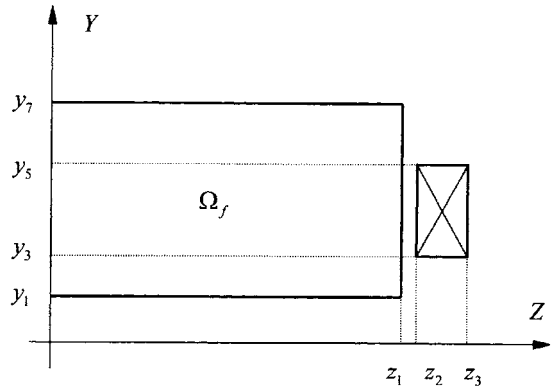



Fig. 2. $\frac{1}{8}$th cross-section on the vertical plane Y0Z.

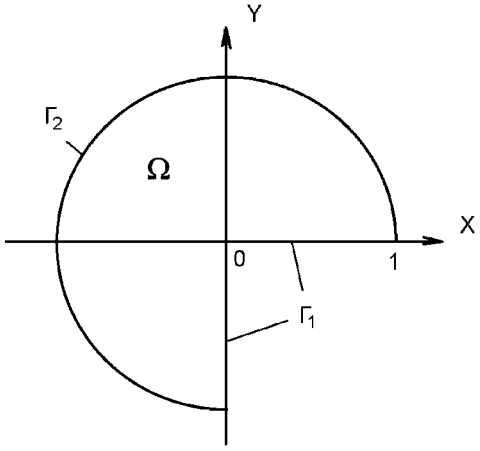



Fig. 4. A domain with a corner point.

Table 1
Linear size of the magnetic system (in m)

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $z_1$ | $z_2$ | $z_2$ |
|---|---|---|---|---|---|---|---|---|
| 0.15 | 0.3 | 0.36 | 0.49 | 0.605 | 1.175 | 0.65 | 0.62 | 0.81 |
| $y_1$ | $y_2$ | $y_3$ | $y_4$ | $y_5$ | $y_6$ | $y_7$ | | |
| 0.0645 | 0.114 | 0.144 | 0.302 | 0.46 | 0.49 | 1.06 | | |

programs includes a program with expert systems for optimal choice of the configuration of magnetic systems, and a program for polynomial representation of the components of a magnetic field, which is used in the tasks of the dynamics of charge particles. It also includes the COBRAM program for numerical simulation of some classes of non-linear inverse magnetostatic tasks for the modeling of ironless magnet systems with a rectangular aperture. The suggested CPMMS complex is formed as a library of programs, which can be used completely as is [4] or can involve replacement of the programs used.

The main criterion for good modeling results is comparison with experimental data with high accuracy. It should be noted that our results for numerical calculation used available experimental data on the $10^{-3}$–$10^{-4}$ level.

## 4. Numerical calculation at corner points

Very often in calculations for a particular magnetic system, the area for solution has a smooth boundary everywhere, with the exception of a finite number of "corner points", in the vicinity of which the boundary is formed by the intersection of two smooth curves.

In these cases, the solution of the problem or the derivatives of the solution have singularity. For example, the Dirichlet boundary problem for the Laplace equation may be used in the area represented in Fig. 4

$$\begin{cases} \Delta u(p) = 0, \quad p \in \Omega \\ u|_{\Gamma_1} = 0 \\ u|_{\Gamma_2} = \sin(2\varphi/3) \end{cases} \tag{1}$$

$(r, \phi)$ is the polar coordinate system; region $\Omega = \{(r, \phi) : 0 < r < 1, 0 < \phi < 3\pi/2\}$ boundary $\Gamma_1 = \Gamma' \cup \Gamma''$ where $\Gamma' = \{(r, \phi) : 0 \leqslant r \leqslant 1, \phi = 0\}, \Gamma'' = \{(r, \phi) : 0 \leqslant r \leqslant 1, \phi = 3\pi/2\}$ while the boundary $\Gamma_2 = \{(r, \phi) : r = 1, 0 < \phi < 3\pi/2\}$.

Problem (1) has the solution $u = r^{2/3} \sin(3\pi/2)$. We take the derivatives $u_x, u_y$ and obtain $u_x = -(\frac{2}{3})r^{-1/3} \sin(2\pi/3)$,

$u_y = (\frac{2}{3})r^{-1/3}\cos(2\pi/3)$. For $r \to 0$ the derivatives $u_x$ and $u_y$ increase without limit. Therefore, for solution of these problems by numerical methods, it is necessary to take into account the nature of the behavior in the vicinity of the corner point.

The present work involved study of the behavior of a magnetic field in the vicinity of the corner points using effective algorithms to improve the accuracy of the numerical solution of a magnetostatics problem.

Numerical calculation using such a method gives accurate results only for a mesh with a 4–5-fold greater number of nodes along every axis than the initial mesh.

The results obtained may be used to solve various nonlinear physics problems involving an equation of the form $\mathrm{div}[\mu(|\nabla u|)] = 0$, when the boundary contains corner points.

Using this method, we carried out simulation of the configuration of the SP-94 spectrometric magnet being used in experiments at the Laboratory of High Energies (LHE), the Joint Institute for Nuclear Research (JINR) and the Institute for High Energy Physics (IHEP), Protvino.

## 5. SP-94 magnet, IHEP, Protvino

The numerical results show (Fig. 5) that inhomogeneity of the magnetic field (on the 1-T level) in the 13-cm air gap is highest at the pole edge (the pole size is $x \times z = 30\,\mathrm{cm} \times 130\,\mathrm{cm}$). In the vicinity of the edge it can be more than 5% for $y = 5.2\,\mathrm{cm}$.

On the other hand, the same extent of inhomogeneity is also observed for the median plane ($y = 0$) when the coordinate $x > 10\,\mathrm{cm}$, i.e. at the pole edge for $x = 15\,\mathrm{cm}$. To increase the area of the homogenous field, it is natural to set up (in a reasonable way) the ferromagnet bars on the pole edge.

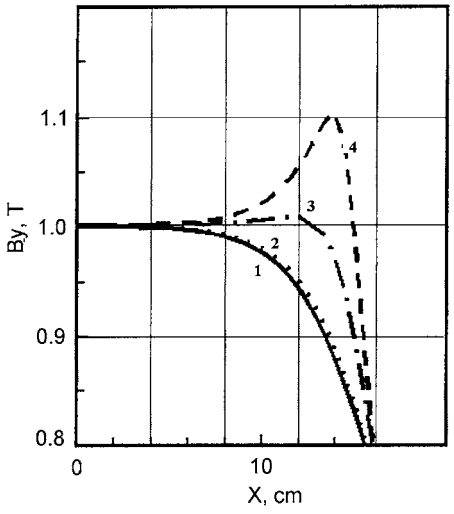



Fig. 5. Distribution of the magnetic field at the pole edge: (1) $y = 0\,\mathrm{cm}$, (2) $y = 1.3\,\mathrm{cm}$, (3) $y = 3.9\,\mathrm{cm}$ and (4) $y = 5.2\,\mathrm{cm}$.

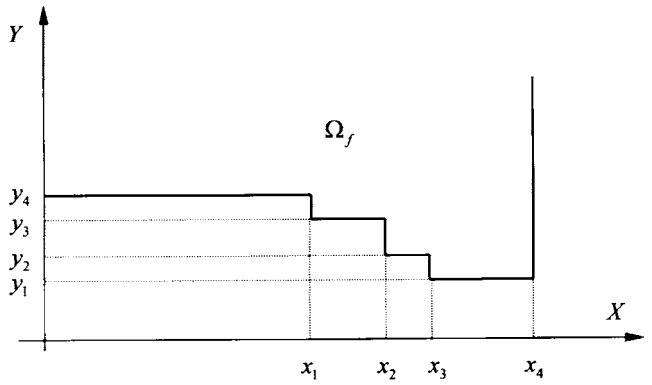



Fig. 6. $\frac{1}{8}$th cross-section on the vertical plane Y0Z.

Table 2
Linear size of the ferromagnet bars (in m)

| $x_1$ | $x_2$ | $x_3$ | $x_4$ | $y_1$ | $y_2$ | $y_3$ | $y_4$ |
|---|---|---|---|---|---|---|---|
| 0.108 | 0.1235 | 0.1345 | 0.15 | 0.053 | 0.05587 | 0.06162 | 0.0645 |

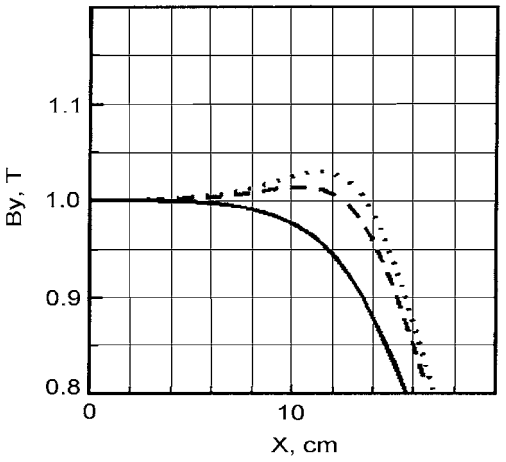



Fig. 7. $B_y(x)$: (1) without bars, (2) with ferromagnet bars (variant A) and (3) with ferromagnet bars (variant B).

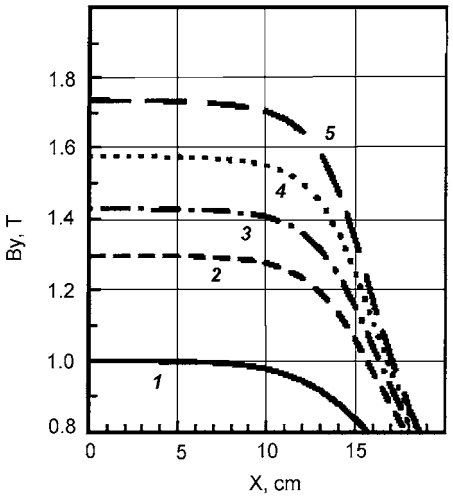



Fig. 8. $B_y(x)$: (1) $h = 13\,\mathrm{cm}$, (2) $h = 10\,\mathrm{cm}$, (3) $h = 9\,\mathrm{cm}$, (4) $h = 8\,\mathrm{cm}$ and (5) $h = 7\,\mathrm{cm}$.

Such bars (Fig. 6, Table 2) were made of steel (similar to the steel of the yoke in magnetic characteristics) on the pole of the SP-94 spectrometric magnet for the EXCHARM-II experiment in the 5H channel of the U-70 accelerator, IHEP, Protvino. Fig. 7 shows graphs for the basic version (without bars) and two versions with ferromagnet bars for the median plane, $y = 0$. It can be verified that a significant increase (for the component $B_y$) in the size of the area with a homogeneous field was observed inside the inter-pole 13 cm air gap.

## 6. SP-94 magnet, LHE, JINR

Fig. 8 shows the results of numerical modeling of the magnetic field distribution values for inter-pole air gaps of $h = 7, 8, 9, 10$ and 13 cm. The full current for half of the coil is $I = 0.73 \times 10^5$ A, generating a field of 1 T for the 13-cm air gap. The dependence functions constructed allow computation of the field for other configurations similar to those presented.

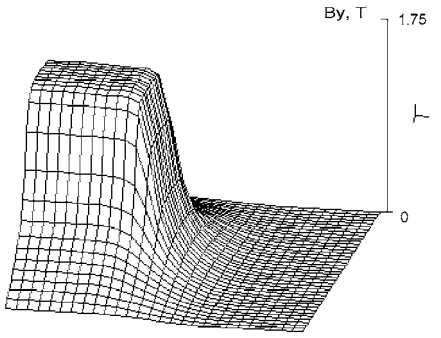



Fig. 11. 3D distribution of the magnetic field $B_y(x, 0, z)$.

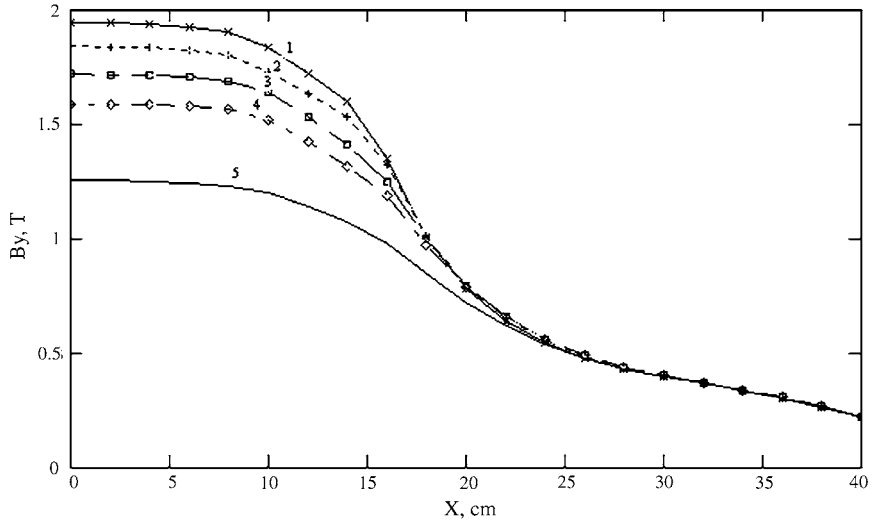



Fig. 9. Distribution of the magnetic field for various air gaps: (1) $h = 7$ cm, (2) $h = 8$ cm, (3) $h = 9$ cm, (4) $h = 10$ cm, and (5) $h = 13$ cm.

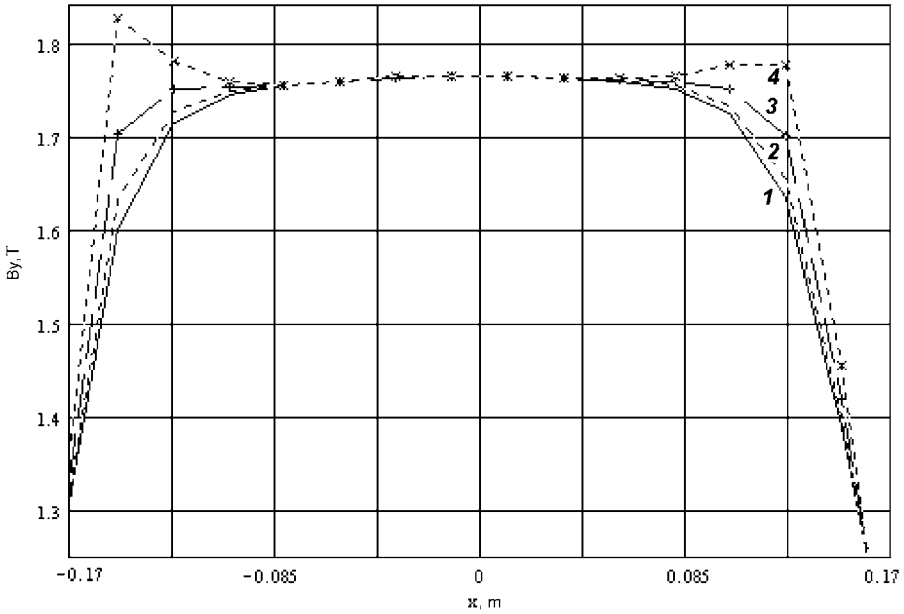



Fig. 10. Dependence of $B_y$ on $x$ for various values of $y$: (1) $y = 0$ cm, (2) $y = 0.9$ cm, (3) $y = 1.8$ cm and (4) $y = 2.7$ cm.

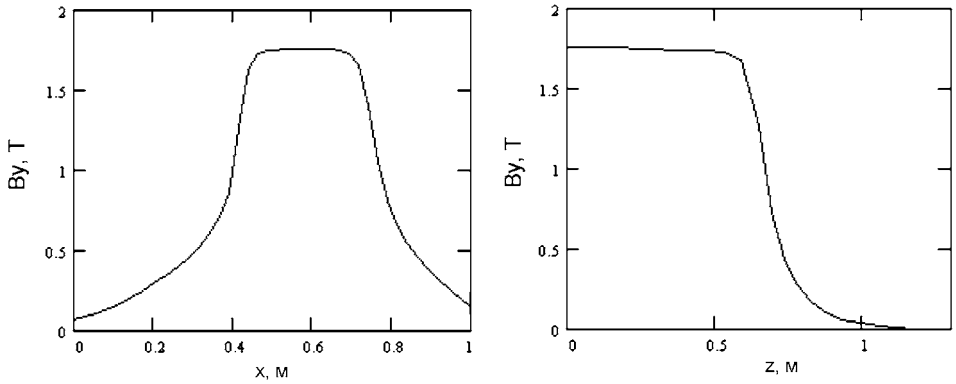



Fig. 12. The dependence of the field $B_y(x,0,0)$ on the $x$-axis and the field $B_y(0,0,z)$ on the $z$-axis.

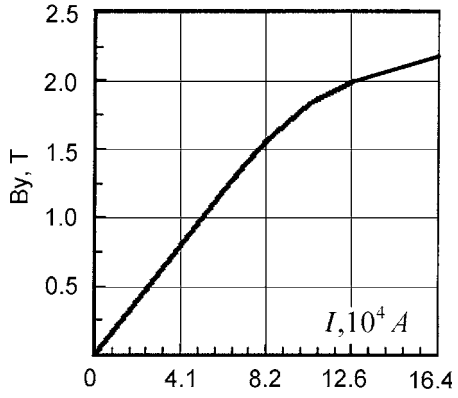



Fig. 13. The curve $B = B(I)$ for the magnet configuration.

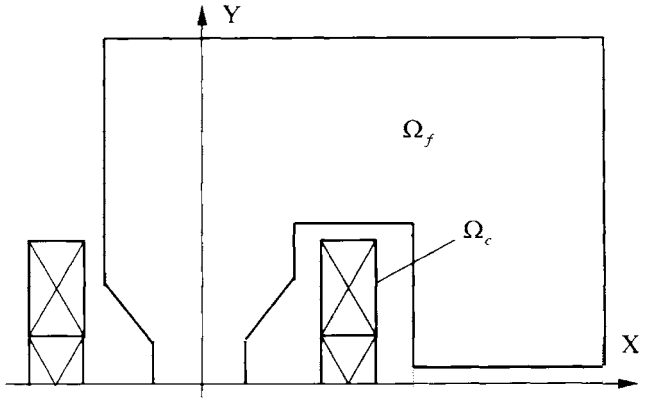



Fig. 14. The magnet with three parts to the current.

The nominal working current in one loop of the coil is 635 A, while the total current in the half coil is $I = 1.016 \times 10^5$ A. Such a current creates a field distribution for this configuration of the magnet with air gaps $h = 7, 8, 9, 10$ and 13 cm, which is shown in Fig. 9.

Numerical results for 3D calculation ($h = 9$ cm, pole width $s = 30$ cm) are shown in Fig. 10. The dependence of $B_y$ on $x$ for $z = 0$ and for various values of $y$ ($y = 0, 0.9, 1.8$ and 2.7 cm) is evident.

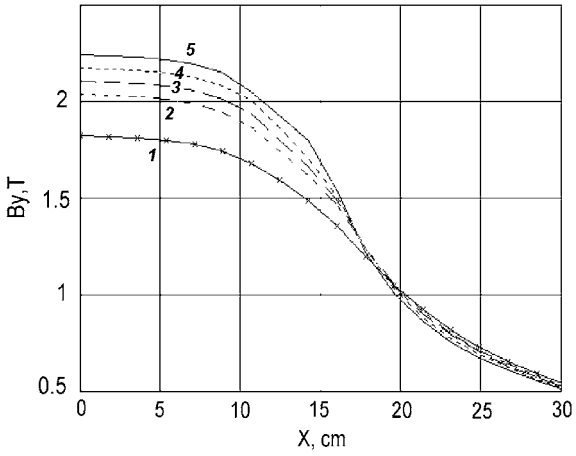



Fig. 15. Distribution of the component of the field $B_y$ along the $x$-axis for (1) $h = 13$ cm, (2) $h = 10$ cm, (3) $h = 9$ cm, (4) $h = 8$ cm and (5) $h = 7$ cm.

An inter-pole air gap of $h = 9$ cm with pole width $s = 30$ cm is now the working version of the SP-94 magnet at LHE, JINR, for the "Delta-Sigma" experiment.

The nominal field is $B_y(0,0,0) = 1.751$ T at the center. A 3D distribution of the magnetic field $B_y(x,0,z)$ for this working version with nominal current $I = 1.016 \times 10^5$ A is given in Fig. 11.

The dependence of $B_y = B_y(x,0,0)$ along the $x$-axis and the dependence of $B_y = B_y(0,0,z)$ along the $z$-axis are shown in Fig. 12.

The value of $\int_{-L}^{L} B_y(0,0,z)\,dz$ is equal to 2.4 T × m, where $B_y(0,0,0) = 1.751$ T, and $L = 1.5$ m is the size (along the $z$-axis) of the area where calculation of the magnetic field was carried out. The curve $B = B(I)$ of the excitation current is shown in Fig. 13 for this magnet configuration.

## 7. SP-94 magnet with three coils

Further investigations were directed toward solving the following inverse problem in magnetostatics. We were interested to find a magnet configuration that would provide a larger value of the working magnetic field inside the aperture, as well as a larger value of the field integral

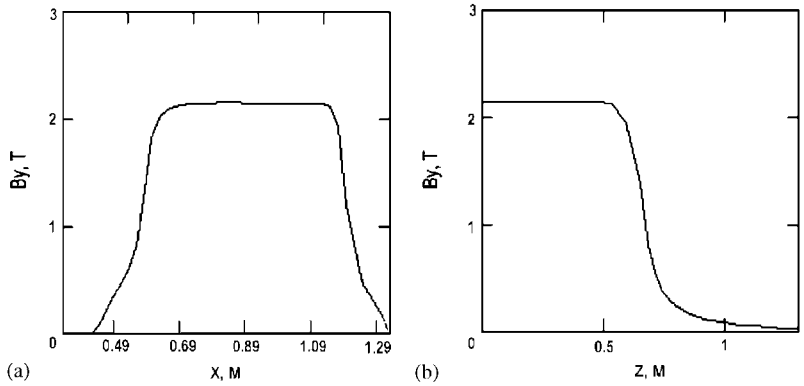



Fig. 16. (a) The distribution of $B_y$ along the $x$-axis for $y = z = 0$ and $h = 9$ cm; (b) the distribution of $B_y$ along the $z$-axis for $x = y = 0$ and $h = 9$ cm.

and possibly with a greater pole width. A technological possibility exists involving placement of an extra third coil, exactly the same size as, for example, the first (upper) or second (lower) coil in the basic version of the SP-94 magnet. The height of the ferromagnet yoke increases by 4 cm in this case.

The nominal current in one loop would, naturally, be kept the same, at 635 A. However, the total current increases by 1.5-fold for such a coil. The configuration suggested (half a section of the XOY plane) for the magnet is given in Fig. 14.

The results shown in Figs. 15 and 16 were obtained by numerical modeling. Fig. 15 demonstrates the field distribution of $B_y$ along the $x$-axis for $h = 7, 8, 9, 10$ and 13 cm. Fig. 16a shows the distribution of $B_y$ along the $x$-axis for $y = z = 0$ and $h = 9$ cm, while Fig. 16b shows the distribution of $B_y$ along the $z$-axis for $x = y = 0$ and $h = 9$ cm.

The value of $\int_{-L}^{L} B_y(0, 0, z)\,dz$ is equal to $3.019\,\mathrm{T} \times \mathrm{m}$. Here $B_y(0, 0, 0) = 2.21$ T, and $L = 1.5$ m is the size (along the $z$-axis) of the area where the field calculation was made. Thus, the value of the field integral for this configuration increases 1.3-fold compared to its value for the configuration of the magnet version ($h = 9$ cm) with two coils.

## 8. Main conclusions

1. Using numerical modeling, two- and three-dimensional distributions of the magnetic field were obtained for some modifications of the SP-94 spectrometric magnet, which is widely used in various nuclear physics experiments.
2. The optimal configuration was been found for the SP-94 magnet, which is well suited to implementation for different experiments.

## References

[1] E.P. Zhidkov, S. Lima, R.V. Polyakova, F. Fernandes Nodarse, I.P. Yudin, The Complex of Programs for the Modeling of Magnetic Systems, P11-93-256, Dubna, 1993.
[2] R.F. Holsinger, C. Iselin, POISCR-r604, User Guide, CERN, SPS/EMA, 1982.
[3] R. Brun, et al., PAW, CERN Program Library Entry Q121, Geneva, 1989.
[4] I.P. Yudin, E.E. Perepelkin, 3D field calculations of magnets by the two scalar potential method, ICAP 2004, St. Petersburg, 2004, Report NP-084.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 346–350
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# COSY INFINITY Version 9

Kyoko Makino*, Martin Berz

*Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*

Available online 6 December 2005

**Abstract**

In this paper, we review the features in the newly released version of COSY INFINITY, which currently has a base of more than 1000 registered users, focusing on the topics which are new and some topics which became available after the first release of the previous versions 8 and 8.1. The recent main enhancements of the code are devoted to reliability and efficiency of the computation, to verified integration, and to rigorous global optimization. There are various data types available in COSY INFINITY to support these goals, and the paper also reviews the feature and usage of those data types.


*Keywords:* COSY INFINITY; Verified computation; Differential algebra; Taylor model; Interval arithmetic

## 1. Introduction

Since the previous versions 8 and 8.1, the main enhancements of the code COSY INFINITY were focused on computational reliability and speed. Interval arithmetic, which serves as a stepping stone to guaranteed computations, is performed with fully verified software rounding in COSY INFINITY. The method is connected to the Differential Algebraic (DA) technique to provide a further efficient method of reliable computation, the Taylor model method, by utilizing the capability of high order computation of the DA technique. Both of these methods benefit from enhanced sparsity support including the ability to treat different variables to different orders.

Besides such basics to support the quality of the computation, there are several particle optical elements, analyzing tools and beam physics concepts newly added in the code since the last version. In particular now there are various ways to treat the dynamics of particle beams traveling in matter.

Through the changes of versions, emphasis has been placed on backwards compatibility and portability of the code to different platforms as well as transparent portability to four separate language platforms, namely F77, F90, C and C++.

## 2. Verified computations

### 2.1. Data types

COSY INFINITY supports various data types, starting from RE (double precision REal number), DA (Differential Algebra vector), and the GR (GRaphics) data type. Table 1 lists all the data types supported in the standard version of COSY INFINITY Version 9. Since the code is object oriented, new data types and the associated operations can be easily added and removed [1]. As some of the readers may have noticed, the Ordered Interval (OI) and Ordered interval Vector (OV) data types are not supported anymore. Besides the STring (ST), LOgical (LO) and GRaphics (GR) data types, all others are numerical computation objects based on double precision real numbers. If any higher precision computation environment like quadruple precision computation mode is available, all these numerical computation objects can be straightforwardly ported to the higher precision mode.

A REal number (RE) data type object corresponds to a double precision number; a CoMplex number (CM) data type object occupies two double precision numbers for the real and imaginary parts. A real number VEctor (VE) data type object consists of several double precision numbers in

*Corresponding author.
*E-mail addresses:* makino@msu.edu (K. Makino), berz@msu.edu (M. Berz).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.109

Table 1
Data types supported in COSY INFINITY

| | |
|---|---|
| RE | Real Number |
| ST | String |
| LO | Logical |
| DA | Differential Algebra Vector |
| VE | Real Number Vector |
| CM | Complex Number |
| IN | Interval |
| IV | Interval Vector |
| GR | Graphics |
| CD | Complex Differential Algebra Vector |
| TM | Taylor model (Remainder-enhanced Differential Algebra Object) |

a vector form, in a similar fashion as a one dimensional double precision array. The VE data type object is created by concatenating two RE data type objects using the operator "&", and more components can be added by further concatenating RE data type objects and/or VE data type objects; its advantage lies in enhanced performance in vectorizable operations.

A DA data type object carries all the non-negligible coefficients of the differential algebra ${}_nD_v$ with order $n$ and dimensionality $v$ [2], where vanishing coefficients and coefficients below the cutoff threshold $\varepsilon_c$ in magnitude are not retained. The value of the cutoff threshold $\varepsilon_c$ can be specified via the COSY intrinsic procedure DAEPS, its default is set to $10^{-20}$. The sparsity handling is particularly important for beam physics applications where usually half of coefficients vanish due to symmetry, and we have been striving to realize and keep superb performance of DA data type objects by the efficient sparsity handling algorithms [3]. A Complex Differential algebra vector (CD) data type object turns DA data type objects to be complex by carrying a set of DA coefficients for the real part and a set for the imaginary part. DA and CD data type objects truncate small coefficients below cutoff threshold, on the other hand, the TM (Taylor model, Remainder-enhanced Differential algebra object) data type objects keep errors associated to order truncation, cutoff and round-off in the data structure. It is worth noting that among various numerical algorithms in the code, those based on the DA-fixed point theorem [2] are particularly powerful, by achieving an $n$th order DA solution in at most $(n+1)$ iterations of the DA-fixed point operation.

The newly implemented feature of weighted order computation enables to carry different variables $x_i$ to different orders $w_i$, which can account for the fact that certain variables are more important than others. This fact is particularly helpful for the problem of integration of transfer maps for time-dependent systems with or without verification, where the length of the time step typically significantly exceeds the range of initial conditions, i.e. beam coordinates. This is achieved by simply "seeding" original variables as $x_i^{w_i}$ instead of $x_i$. In all subsequent operations, only multiples of $w_i$ appear as powers of $x_i$. Using the feature, optimal reduction of speed can be achieved by sparsity.

### 2.2. Data types for verified computations

The other data types INterval (IN), Interval Vector (IV) and TM are objects for verified computation. An IN data type object consists of two double precision numbers; one for the lower bound and one for the upper bound of an interval, and an IV data type object describes a vector with interval components. The concept of the IV data type is similar to that of the VE data type, except for that each component is an interval consisting of two double precision numbers. Interval methods for numerical computations express a set of numbers by an interval, and through various computational operations, interval arithmetic rigorously keeps all the possible outcomes in the resulting interval. (Refer to, for example [4] as well as references therein, and many more.) Floating point rounding errors are unavoidably associated with numerical computations, thus a correct implementation of interval methods requires proper handling of these rounding errors. The interval library in COSY INFINITY, which is the base for the verified computations for the data types IN, IV and TM, supports the directed rounding to assure the verification [5]. The COSY ASCII output of intervals is further rounded outward to avoid confusion caused by system dependent output truncation. To facilitate interval-related applications, some utility procedures are newly available. Particularly, INTSEC, INTUNI and INTINC are used for intersection, union and inclusion check, respectively.

The data type TM represents Taylor model objects. The DA data type objects truncate Taylor power series at order $n$, but the TM data type objects keep the contribution from the Taylor remainder term in an interval, the so-called remainder bound interval. The bulk amount of the functional dependency is kept in the polynomial part that has the same data structure as the DA data type objects. Thus, this data type is also called the (R)emainder-enhanced (D)ifferential algebra type. The remainder bound interval part is used conveniently to absorb other errors like floating point rounding errors and coefficient cutoff errors. This allows us to use floating point numbers for polynomial coefficients, while assuring verified computations.

The utilization of floating point number coefficients has numerous benefits. First, a Taylor model implementation can utilize a big part of the DA library code by mere DA subroutine calls. This helps to reduce the code implementation and maintenance effort as well as the size of the code, but it requires special care to properly handle floating point rounding errors and cutoff errors. Secondly, as summarized in an exhaustive paper on the method of Taylor models [6], this idea is one of key points to enable the suppression of the dependency problem. This is one source of overestimation often observed in interval based computations, and it limits the applicability of many verified

computations. Thirdly, the concept of DA-fixed point theorem can be extended straightforwardly to various Taylor model algorithms, allowing them to achieve the same level of computational efficiency to that of DA fixed point based algorithms. One superb example of this is the Taylor model algorithm for verified ODE (ordinary differential equation) integrations [7].

The arithmetic on Taylor models is introduced for binary operations and intrinsic functions. The implementation of Taylor models in COSY INFINITY utilizes the COSY DA and the COSY interval libraries optimally so that the efficiencies achieved for the DA and IN data types can be carried over to the TM data type. Refer to [6] for the theoretical background and implementation details. Furthermore, a binary output/input capability is newly added to COSY INFINITY to avoid any error growth associated to Taylor model file input and output. The binary I/O capability is supported for the RE and TM data types.

The Taylor model method provides sharp estimate while guaranteeing the result, and even only low-order Taylor models often perform better than sophisticated methods like the centered form and the mean value form [6,4]. Details can be found in Ref. [6], which also summarizes various Taylor model based algorithms and discusses some practical problems. The various Taylor model operations and intrinsics have been independently analyzed for rigor based on IEEE floating point standards [8] and subjected to extensive and challenging execution-based testing [9,10].

### 2.3. COSY-VI and COSY-GO

Verified integration of ODEs and global optimization require efficient computational methods with verification, and the method of Taylor models can be applied effectively. Based on the Taylor model implementation in COSY INFINITY, packages for those important applications are now available for release.

COSY-VI is the COSY Taylor model package for verified integration of ODEs, and the package offers various state of the art Taylor model algorithms for the task. The package also can be used for other types of problems like differential algebraic equations by reducing them to the form of ODEs [11]. Besides the general concern of controlling the dependency problem in verification problems, verified integrations of multidimensional ODEs exhibit a severe asymptotic overestimation problem of geometric nature, called the wrapping effect [4]. The striving for suppressing the wrapping effect has as long a history as the computer implemented interval method itself. The naive concept of Taylor models with multivariate polynomials allows not only the high-order Taylor expansion in time $t$, but also the high-order expansion in space variables $\vec{x}$. This feature enables a solution set at each integration time step to be enclosed by an $n$th order Taylor model, i.e. the set is approximated by an $n$th order polynomial while the approximation error is kept in a small remainder bound interval. Combined with preconditioning techniques, this approach much reduces the devastating wrapping effect. The COSY-VI package is further equipped with higher level algorithms like the method of shrink wrapping and various types of blunting for tighter control of error growth, details about which can be found in an exhaustive paper on the Taylor model ODE integrations [7]. In summary, the key features and algorithms of COSY-VI are

- High-order expansion not only in time but also in transversal variables.
- Capability of weighted order computation, allowing to suppress the expansion order in transversal variables.
- Shrink wrapping algorithm including blunting to control ill-conditioned cases.
- Pre-conditioning algorithms based on the Curvilinear, QR decomposition, and blunting pre-conditioners.
- Resulting data is available in various levels including graphics output.

COSY-GO is the COSY Taylor model package for global optimization. Different from intervals, Taylor models carry the information on local slope and convexity in the data structure. Utilizing the readily available information, various range bounding algorithms have been developed. The linear dominated bounder LDB provides fast multidimensional Taylor model range bounding, utilizing the linear part as a guideline on range enclosing and reducing the corresponding domain area. The quadratic dominated bounder QDB provides a thorough quadratic bounding of a multidimensional Taylor model by carrying out the convexity tests of the quadratic part. A $v$ dimensional box has $3^v$ surfaces, consisting of $2^v$ one dimensional corner points, various higher dimensional surfaces, and finally the $v$ dimensional box interior. Thus, a complete examination of stationary points for a $v$ dimensional quadratic polynomial requires $3^v - 2^v$ tests if conducted naively, which becomes impractical quickly as $v$ increases. The QDB bounder reduces the required efforts by utilizing the LDB and the efficient surface list handling, making high dimensional problems practically solvable. To facilitate the task of quadratic bounding for global optimization, a limited purpose quadratic bounder, the fast quadratic bounder QFB, is more practical. QFB is designed for a multidimensional Taylor model whose quadratic part is convex, which is characteristic of the most crucial bounding task, namely that of a Taylor model in the proximity of a local minimizer. This enables to eliminate the pure quadratic terms from the bounding task. The COSY-GO package is equipped with those state of the art Taylor model range bounding algorithms [12]. Since those quadratic bounders solve an infamous problem in global optimizations with verified methods, the so called cluster effect, COSY-GO makes various global optimization problems practically solvable, among them the long-term stability estimate of storage rings using Normal Form

analysis, which is the original motivation of the development of Taylor models. The core features and algorithms of COSY-GO can be summarized as follows:

- List management of boxes not yet determined to not contain the global minimizer. Loading a new box. Discarding a box with range above the current threshold value. Splitting a box with range not above the threshold value for further analysis. Storing a box smaller than the specified size.
- Application of a series of bounding schemes, starting from mere interval arithmetic to naive Taylor model bounding, LDB, then QFB. A higher bounding scheme is executed only if all the lower schemes fail.
- Update of the threshold cutoff value via various schemes. It includes upper bound estimates of the local minimum by corresponding bounding schemes, the mid point estimate, global estimates based on local behavior of function using gradient line search and convex quadratic form.
- Box size reduction using LDB.
- Resulting data is available in various levels including graphics output.

## 3. Particle optical elements and analyzing tools

COSY INFINITY offers various methods for particle trackings via transfer maps that relate the initial condition $\vec{z}_i$ to the final condition $\vec{z}_f$ via $\vec{z}_f = \mathscr{M}(\vec{z}_i)$. In the code, transfer map $\mathscr{M}$ is represented by the DA data type. For the purpose of gaining speed in computations, the VE data type is employed to represent the particle coordinates $\vec{z}$ in the tracking algorithm. Tracking can be performed not only in regular particle coordinates that provide intuitive understanding of the dynamics, but also in normal form coordinates that provide a means of quantitative analysis of the dynamics. When studying the dynamic aperture, it is important to be able to utilize appropriate symplectification. In the current version of COSY INFINITY, mere particle tracking without symplectification is performed by specifying the tracking mode TY $= -21$ for the command TR [13]. The symplectic tracking modes with generating functions of types $F_1$ to $F_4$ require the user to find the optimal generating function by trial and error for each problem. A new approach, the EXPO (The EXtended POincare generating function type), employs the optimal generating function for symplectification [14,15]. COSY INFINITY offers this method to promote the easy usage of symplectic trackings, and the feature is performed by specifying the tracking mode TY $= 0$. Another minor but useful tool is the command TRT, which allows the user to incorporate bookkeeping information in a tracking picture produced by the command TR. It is also possible to mark a specific particle in tracking pictures by coloring the particle via the command SR. Due to the additional memory consumption, this feature is turned off by default. To activate the feature, a few lines in the file cosy.fox have to be altered, which can be easily identified by searching a string "color."

Some of modern particle optical devices have large acceptance, for example those to be used for the various muon accelerator scenarios, where the difficulty and expense of cooling require the ability to manipulate a beam of unusually large emittance. This naturally has led to the usage of COSY INFINITY for such systems, because of the necessity of high-order nonlinear computations. As a result, several new algorithms and tools have been developed. This includes an extensive collection of solenoidal elements, efficient propagation of beams with tremendously large emittance, and treatment of dynamics of particle beams traveling through matter while experiencing scatterings. Refer to [16,17], for details.

## 4. Standard features and supported languages

With the rapid expansion of computer techniques in recent years, it is not a simple task to maintain a scientific computation code like COSY INFINITY with numerous users and a variety of computer environments, to adjust to newly emerging techniques and the disappearance of others. To efficiently confront this situation, we have strived to keep backward compatibility and portability of the code COSY INFINITY as much as possible in order to protect users from additional effort due to sudden code changes based on syntax modification.

There are some items worthwhile to mention in this paper about the current official distributions at the COSY web site cosy.pa.msu.edu. For the interactive graphics output purpose, the PGPLOT graphics library has been stable in the last years, and thus we keep the PGPLOT graphics drivers in COSY as the standard interactive graphics package [1]. On the other hand, the GKS graphics library is quickly becoming obsolete, so we demoted the GKS graphics drivers. The GKS drivers are merely commented in the code, so they are still easily available for the user. We keep the VGA graphics drivers for Lahey Fortran and the graPHIGS graphics drivers in the same commented form. The long swing between MicroSoft Windows PCs and Linux/UNIX for the COSY Fortran77 sources seems to be settled into "UNIX" version. It is because the COSY "PC" version was meant mostly for Lahey Fortran, and popular Fortran compilers lately available for MicroSoft Windows PCs are compatible with the COSY "UNIX" version.

Finally, for the increasing population and demand of non-Fortran77 languages, COSY INFINITY provides interface packages for Fortran90 and C++ to enhance portability. All the data types and the associated operations, functions and intrinsic procedures in COSY INFINITY are accessible via the interface packages as C++ classes for the C++ user, and as Fortran90 modules for the F90 user. These COSY interface C++ classes and Fortran90 modules outperform independent attempts of

creating DA packages in C++ and Fortran90. Refer to [1,5] for details on the interface packages.

## Acknowledgements

This work was supported by the US Department of Energy, the National Science Foundation, the German National Merit Foundation, an Alfred P. Sloan Fellowship, and the Illinois Consortium for Accelerator Research.

## References

[1] M. Berz, J. Hoefkens, K. Makino, COSY INFINITY Version 8.1—programming manual, Technical Report MSUHEP-20703, Department of Physics and Astronomy, Michigan State University, East Lansing, MI, 2002, see also http://cosy.pa.msu.edu.
[2] M. Berz, Modern Map Methods in Particle Beam Physics, Academic Press, San Diego, 1999, also available at http://bt.pa.msu.edu/pub.
[3] M. Berz, Forward algorithms for high orders and many variables, in: A. Griewank, G.F. Corliss (Eds.), Automatic Differentiation of Algorithms: Theory, Implementation and Application, SIAM, Philadelphia, 1991.
[4] R.E. Moore, Methods and Applications of Interval Analysis, SIAM, Philadelphia, 1979.
[5] J. Hoefkens, Verified methods for differential algebraic equations, Ph.D. Thesis, Michigan State University, East Lansing, MI, USA, 2001.
[6] K. Makino, M. Berz, Int. J. Pure Appl. Math. 6 (3) (2003) 239 available at http://bt.pa.msu.edu/pub.
[7] K. Makino, M. Berz, Suppression of the wrapping effect by Taylor model based validated integrators, submitted. Also MSUHEP-40910, available at http://bt.pa.msu.edu/pub.
[8] N. Revol, K. Makino, M. Berz, J. Logic Algebraic Program. 64 (1) (2004) 135–154.
[9] G.F. Corliss, J. Yu, Interval testing strategies applied to COSY's interval and taylor model arithmetic, in: R. Alt, et al. (Ed.), Numerical Software with Result Verification, Lecture Notes in Computer Science, vol. 2991, Springer, Berlin, 2004, pp. 91–106.
[10] G.F. Corliss, Testing COSY's INSRF, Technical Report MSUHEP 31228, Michigan State University, 2003.
[11] J. Hoefkens, M. Berz, K. Makino, Verified high-order integration of DAEs and higher-order ODEs, in: W. Kraemer, J.W.v. Gudenburg (Eds.), Scientific Computing, Validated Numerics, Interval Methods, Kluwer, Boston, 2001, pp. 281–292.
[12] M. Berz, K. Makino, Y.K. Kim, Long-term stability of the Tevatron by verified global optimization, Nucl. Instr. and Meth. A, in press, doi:10.1016/j.nima.2005.11.035.
[13] M. Berz, K. Makino, COSY INFINITY Version 8.1—user's guide and reference manual, Technical Report MSUHEP-20704, Department of Physics and Astronomy, Michigan State University, East Lansing, MI, 2002, see also http://cosy.pa.msu.edu.
[14] B. Erdélyi, M. Berz, Phys. Rev. Lett. 87 (11) (2001) 114302.
[15] B. Erdélyi, M. Berz, Int. J. Pure Appl. Math. 11 (3) (2004) 241 available at http://bt.pa.msu.edu/pub.
[16] K. Makino, M. Berz, C.J. Johnstone, D. Errede, Nucl. Instr. and Meth. A 519 (2004) 162.
[17] M. Berz, K. Makino, C.J. Johnstone, Propagation of a large-emittance muon beam through a straight, quadrupole-based precooling channel, in: NuFact03, AIP Conference Proceedings, 2004.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 351–355
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Recent advances of strong–strong beam–beam simulation

Ji Qiang[a,*], Miguel A. Furman[a], Robert D. Ryne[a], Wolfram Fischer[b], Kazuhito Ohmi[c]

[a]Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[b]Brookhaven National Laboratory, Upton, New York 11973, USA
[c]High Energy Accelerator Research Organization (KEK), 1-1 Oho, Tsukuba, 305-0801, Japan

Available online 28 November 2005

**Abstract**

In this paper, we report on recent advances in strong–strong beam–beam simulation. Numerical methods used in the calculation of the beam–beam forces are reviewed. A new computational method to solve the Poisson equation on nonuniform grid is presented. This method reduces the computational cost by a half compared with the standard FFT based method on uniform grid. It also appears to be more accurate than the standard method for a colliding beam with low transverse aspect ratio. In applications, we present the study of coherent modes with multi-bunch, multi-collision beam–beam interactions at RHIC. We also present the strong–strong simulation of the luminosity evolution at KEKB with and without finite crossing angle.


*PACS:* 29.27.Bd

*Keywords:* Beam–beam simulation

## 1. Introduction

The beam–beam interaction puts a strong limit on the luminosity of the high energy storage ring colliders. At the interaction points, the electromagnetic fields generated by one beam focus or defocus the opposite beam. This can cause beam blowup and a reduction of luminosity. An accurate simulation of the beam–beam interaction is needed to help optimize the luminosity in high energy colliders. In this case, the self-consistent strong–strong beam–beam simulation provides an invaluable tool for the study of the colliding beams.

In the strong–strong beam–beam simulation, a number of simulation particles (macroparticles) are used with the same charge-to-mass ratio as the real particles. Outside the interaction region, the macroparticles are transported through the simulated lattice using transfer maps associated with external elements, radiation damping, and quantum excitation. At the interaction point, the electromagnetic fields from the beams are calculated and applied to the particles of the opposing beam.

The soft Gaussian approximation is sometimes used to obtain the electromagnetic fields of the beams at the collision point [1–3]. While this approximation has the advantage of computational speed, it is not self-consistent because it assumes a Gaussian distribution for the macroparticles even when the actual distribution might differ substantially from the Gaussian shape. To take into account the effects of the beam distribution self-consistently, one has to solve the Poisson equation numerically during each collision for the actual macroparticle distribution at that instant. A number of methods have been used to solve the Poisson equation. The five-point finite difference method with Fourier analysis and cyclic reduction (FACR) has been used by Krishnagopal [4] and Cai et al. [5]. This method solves the Poisson equation efficiently with finite domain boundary conditions. For the open boundary conditions, which are appropriate in typical beam–beam simulations, the method requires finding an effective boundary condition on the problem boundary; this can be computationally expensive. In addition, this method is not efficient to handle the case with two widely

*Corresponding author. Tel.: +1 5104952608; fax: +1 5104952323.
*E-mail address:* jqiang@lbl.gov (J. Qiang).

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.045

separated beams, where the domain of the source particles (particle domain) and the domain of the electric field (field domain) are different. Another method based on the fast multipole expansion has been used by Herr et al. [6] to solve the Poisson equation. In this method, the computational cost scales linearly with the number of particles or with the number of total mesh points for the open boundary condition. The efficiency of this method is independent of the distribution of the source particles and the field domain, which makes it suitable to handle the situation with two separated beams. However, this method is an approximate algorithm in the sense that the accuracy of the expansion depends on the radius of convergence. The computational speed depends on the number of polynomials required in the multipole expansion.

The widely used method to solve the Poisson equation in beam–beam simulations is the Green function method with fast Fourier transform (FFT) on uniform grid. This method uses an FFT to calculate the cyclic summation on a doubled computational grid [7–11]. The computational cost scales as $N^2 \log(N)$, where $N$ is the number of grid points in one direction. By defining a new shifted integrated Green function, this method can handle the separated beams, and beams with large aspect ratio.

During the beam–beam interaction, when the bunch length is large compared with the beta function value or the beam–beam forces are strong, finite bunch length effects are not negligible. In this case, a multiple slice model has to be used. The computational cost scales as the square of the number of slices. For a hadron collider with small radiation damping, it is required to track the beams for many millions of turns to study the dynamics on the time scale of the lifetime of the beams. To study the beam–beam interaction fully self-consistently for both beams (i.e. a strong–strong formulation), and to include all the physical processes of long range off-centroid interactions, finite beam bunch length effects, and crossing angle collisions, requires computational resources far beyond the capability of current serial computers. A parallel beam–beam simulation code, Beam–Beam3D, with both weak–strong and strong–strong capabilities, that can simulate these physical processes accurately using high performance computers has been developed at Lawrence Berkeley National Laboratory [12]. In this paper, we present recent advances in the numerical method to calculate the beam–beam forces and in applications to the studies of beam–beam interactions at RHIC and KEKB.

The organization of the paper is as follows: The computational methods are described in Section 2. Applications to the studies of beam–beam interactions at RHIC and KEKB are given in Section 3. We summarize our results in Section 4.

## 2. Computational methods

In strong–strong beam–beam simulation, the electric fields generated by the opposite moving beam can be obtained from the solution of Poisson's equation. In Cartesian coordinate system, the solution of Poisson's equation can be written as

$$\phi(x,y) = \int G(x,\bar{x},y,\bar{y})\rho(\bar{x},\bar{y})\,\mathrm{d}\bar{x}\,\mathrm{d}\bar{y}, \tag{1}$$

where $G$ is the Green's function, $\rho$ is the charge density, and $(x,y)$ represent the coordinates in the plane perpendicular to the direction of motion of the beam. For the case of transverse open boundary conditions, the Green's function is given by

$$G(x,\bar{x},y,\bar{y}) = -\tfrac{1}{2}\ln((x-\bar{x})^2 + (y-\bar{y})^2). \tag{2}$$

Now consider a simulation of an open system where the computational domain containing the particles has a range of $(0, L_x)$ and $(0, L_y)$, and where each dimension has been discretized using $N_x$ and $N_y$ points, the electric potentials on the grid can be approximated as

$$\phi(x_i,y_j) = h_x h_y \sum_{i'=1}^{N_x}\sum_{j'=1}^{N_y} G(x_i - x_{i'}, y_i - y_{j'})\rho(x_{i'}, y_{j'}), \tag{3}$$

where $x_i = (i-1)h_x$ and $y_j = (j-1)h_y$. This convolution can be replaced by a cyclic convolution expression in a double-gridded computational domain. The cyclic convolution can be computed efficiently using an FFT as described by Hockney and Eastwood [13].

The method described above involves use of the FFT on a uniform computational grid. In high energy colliders, the colliding beams normally have a non-uniform transverse charge density distribution. A nonuniform grid may help resolve the charge density distribution more efficiently. Furthermore, using an FFT-based method on the nonuniform grid in cylindrical coordinates also reduces the computational cost of the Poisson solver by a factor of two, which could have an important effect in the strong–strong beam–beam simulation. This is because the Poisson solver to calculate the beam–beam forces is normally the major computational cost in the self-consistent strong–strong simulation. To use a nonuniform grid, we have transformed the charge density distribution from the Cartesian coordinates $(x,y)$ into a cylindrical coordinates $(r,\theta)$. Then, we define another transform between radial $r$ and a new variable $s$ as

$$s = \frac{1}{k_1}\log\left(\frac{r}{k_2}\right) \tag{4}$$

where the constants $k_1$ and $k_2$ control the scale and the rate of the function variation between $r$ and $s$. Using a uniform grid along $s$ will generate a nonuniform grid along $r$ since $\mathrm{d}r = k_1 k_2 r\,\mathrm{d}s$. For a uniform computational grid in $(s,\theta)$ coordinates, we can use the FFT based method to calculate the convolution for electric potential. A similar transform has been used in calculation of the gravitational potential in a disk galaxy system [14]. The new Green function in the

$(s,\theta)$ coordinate is

$$G(s,\theta) = -\tfrac{1}{2}\log(e^{2k_1 s} - 2e^{k_1 s}\cos(\theta) + 1). \quad (5)$$

In the $(s,\theta)$ coordinates, both the Green's function and the charge density distribution are periodic functions of $\theta$. Hence, we do not need to double the computational domain along $\theta$ to use the Hockey's algorithm. This reduces the computational cost and the storage by a factor of two compared with the standard FFT based Green method on uniform Cartesian coordinate.

As an example of the above algorithm, we have computed the radial electric field distribution generated by a round beam with a Gaussian density distribution. The left plot of Fig. 1 shows the analytical solution of the radial electric field $E_r$ as a function of radial distance $r$. The right plot of Fig. 1 shows the absolute error of $E_r$ as a function of $r$ using the nonuniform grid Green function method and the standard uniform grid Green function method. It can be seen that using the nonuniform grid Green function method, the numerical error of $E_r$ is about half of that using the uniform grid Green function method. In this example, the transverse aspect ratio of the colliding beam is one. This is true for most hadron collider where radiation damping is negligible. For electron–positron colliders such as KEKB and PEPII, the colliding beam can have a very large transverse aspect ratio. To test the applicability of the above algorithm, we have also calculated the electric field for a Gaussian charge density distribution with an aspect ratio of 30. The relative error of $E_x$ on the $x$ axis is given in Fig. 2 together with that calculated from using the integrated Green function method on uniform grid. Here, three relative errors of $E_x$ from using the nonuniform grid Green function method on a computational grid of $256\times512$, $512\times256$ and $1024\times512$ are given. The relative error from the integrated uniform Green function method uses a computational grid of $256\times256$. Since the nonuniform grid Green function does not need to double the computational domain in the $\theta$ direction, it has the same computational cost on a $256\times512$ grid as the uniform grid Green function method does on a $256\times256$ grid. It is seen that the integrated Green function method on a $256\times256$ uniform grid gives the least error. This suggests that the integrated Green function method might be more efficient for a beam with large aspect ratio.

## 3. Applications

In recent applications, we have studied the coherent modes of multi-bunch collisions at RHIC through a strong–strong beam–beam simulation [15]. Fig. 3 gives a schematic plot of two colliding beams at RHIC. Here each beam has three bunches. The six bunches couple with each other through collisions at four interaction points, IP2, IP6, IP8 and IP10. Table 1 gives a list of the physical parameters used in the simulation. All bunches are assumed to have the same physical parameters listed in the table. Fig. 4 shows power spectra of horizontal centroid motion of three bunches. There are only two distinct

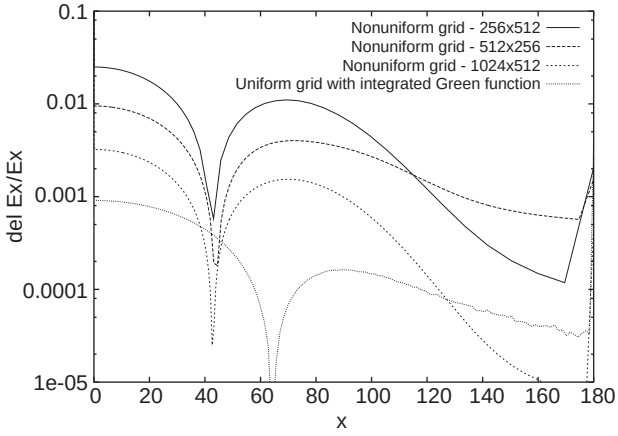



Fig. 2. Relative error of $E_x$ as a function of $x$ from using the nonuniform grid Green function method and from using the integrated uniform grid Green function method.

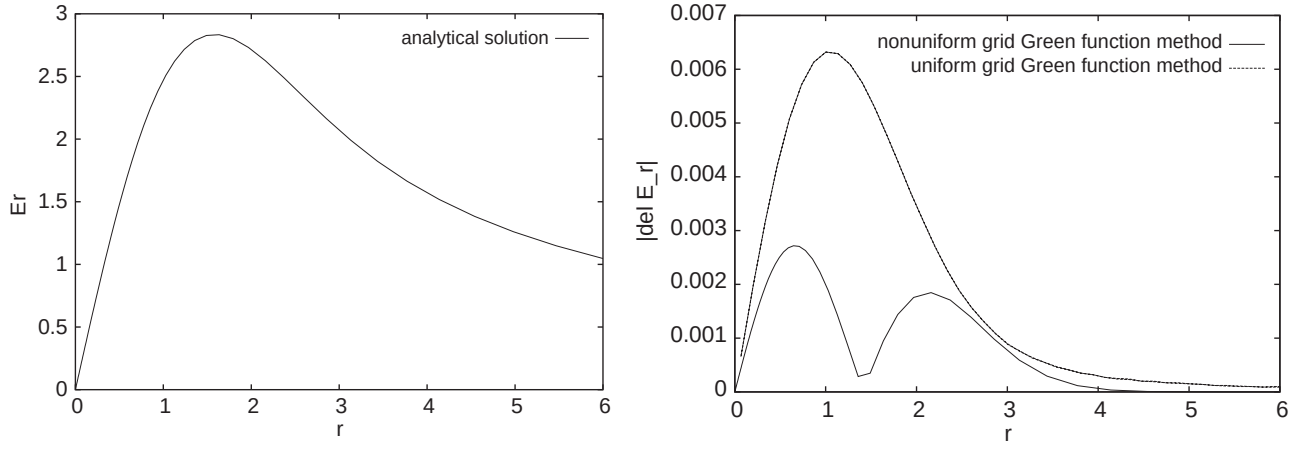



Fig. 1. Radial electric field $E_r$ and the absolute error of $E_r$ as a function of $r$ from the nonuniform grid Green function method and from the uniform grid Green function method.

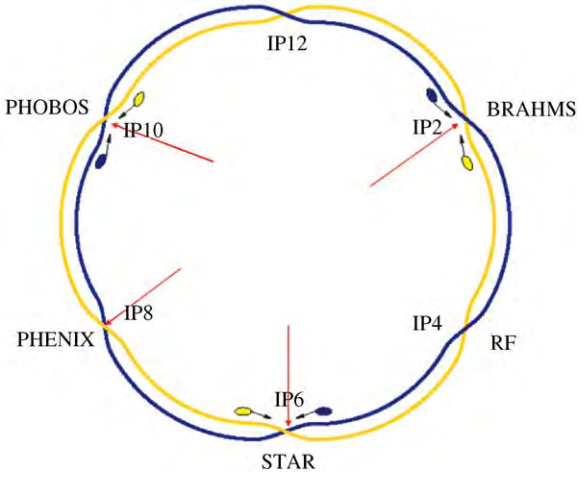



Fig. 3. A schematic plot of two colliding beams at RHIC.

Table 1
RHIC physical parameters for beam–beam simulations

| | |
|---|---|
| Beam energy (GeV) | 100 |
| Protons per bunch | $10.0 \times 10^{10}$ |
| $\beta^*$ (m) | 1.0 |
| RMS spot size at the IP (mm) | 0.176 |
| Betatron tunes $(\nu_x, \nu_y)$ | (0.22, 0.23) |
| Synchrotron tune $\nu_z$ | 3.7e − 4 |
| RMS bunch length (m) | 3.6 |
| Momentum spread | 1.6e − 3 |
| Beam–beam parameter $\xi$ | 0.00366 |

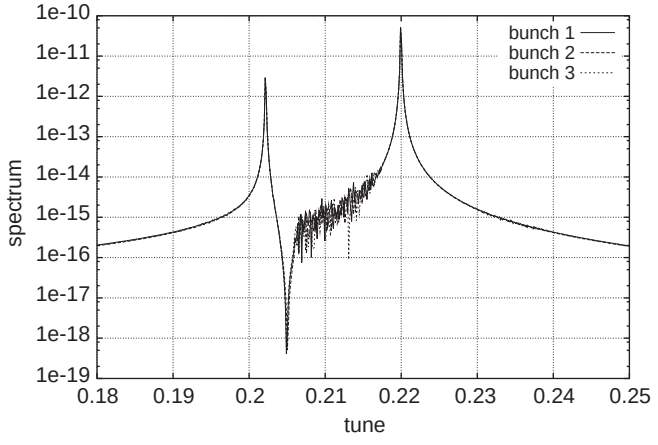



Fig. 4. Power spectra (arbitrary normalization) of the horizontal centroid motion of three bunches at RHIC.

eigenmodes, the $\pi$ mode (180 degree out of phase) and the $\sigma$ mode (in phase), which are observable in this example. The other four modes are degenerated and buried into the incoherent continuous spectra. The $\pi$ mode tune shift is $4.918\xi$ which is about of a factor of 4 times the single bunch $\pi$ tune shift $1.21\xi$. This is in agreement with the analytical calculation of Yokoya et al. [16]. The large tune shift of the $\pi$ mode due to the multi-bunch collisions presents a potential instability since it can not be damped out by the continuous spectra through the Landau damping. In above example, we have assumed that the two beams have the same parameters. In reality, the parameters of two rings can be controlled so that the two colliding beams have different tunes. Fig. 5 gives power spectra of horizontal centroid motion of three bunches with the horizontal tune of the second beam set as 0.2 while the first beam is set as 0.22. The two colliding beams lose the coherent motion and the dipole mode disappears into the continuous spectra.

In another application, we have studied the time evolution of luminosity at KEKB. The physical parameter used in the simulation is given in Table 2. Fig. 6 shows the time evolution of single bunch collision luminosity with (left plot) and without (right plot) a finite crossing angle. With 11 mrad of finite crossing angle, the luminosity has dropped by about a factor of two after about 2000 turns. This suggests that using a head-on collision (e.g. by using crab cavity at interaction point) will significantly improve the luminosity at KEKB.

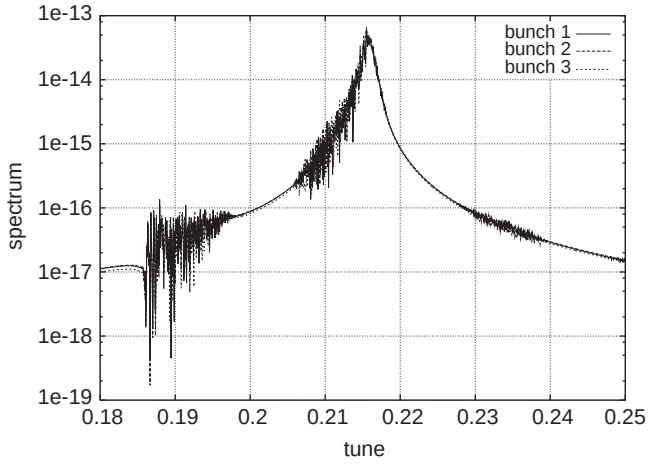



Fig. 5. Power spectra (arbitrary normalization) of the horizontal centroid motion of three bunches with different tune in each ring of RHIC.

Table 2
KEKB physical parameters for the beam–beam simulation

| | |
|---|---|
| $e^-/e^+$ beam energy (GeV) | 8.0/3.5 |
| $e^-/e^+$ per bunch | $4.375 \times 10^{10}/10.0 \times 10^{10}$ |
| $\beta^*$ (horizontal, vertical, longitudinal) (m) | (0.6, 0.007, 10.0) |
| Emittance (horizontal, vertical, longitudinal) (m-rad) | $(1.8 \times 10^{-18}, 1.8 \times 10^{-18}, 4.8 \times 10^{-6})$ |
| Betatron tunes $(\nu_x, \nu_y)$ | (0.5151, 0.5801) |
| Synchrotron tune $\nu_z$ | 0.016 |
| Damping time (horizontal, vertical, longitudinal) (/turn) | $(2.5 \times 10^{-4}, 2.5 \times 10^{-4}, 5.0 \times 10^{-4})$ |

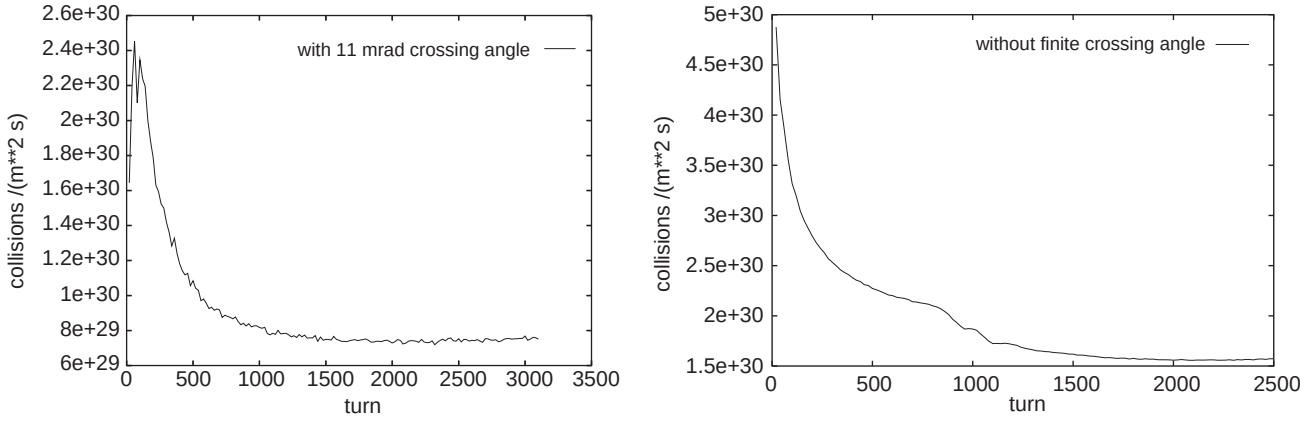



Fig. 6. Evolution of single bunch collision luminosity at KEKB with and without a finite crossing angle.

## 4. Summary

In this paper, we have reported on some recent advances of strong–strong beam–beam simulation. The new nonuniform grid Green function method for calculating the beam–beam forces has the advantage of better accuracy and less computational cost for low aspect ratio beam in hadron collider. The application to the study of the multi-bunch coherent modes at RHIC shows a much larger dipole mode tune shift than that of the single bunch collision. This mode can be removed with asymmetric tunes of two colliding beams. In the KEKB application, the collision with 11 mrad crossing angle shows a significant decrease of the luminosity compared with the head-on collision. This suggests that using a crab cavity to correct the crossing angle collision will improve the luminosity of the future machine operation.

## Acknowledgements

This research used resources of the National Energy Research Scientific Computing Center and the resources of the Center for Computational Sciences at OAK Ridge National Laboratory. Some of the computational work for this project was done on the LBL/NERSC Alvarez Cluster, an 80 node Pentium III Myrinet cluster. This work was supported by the US Department of Energy under Contract no. DE-AC03-76SF00098, and by a Scientific Discovery through Advanced Computing project, "Advanced Computing for 21st Century Accelerator Science and Technology," which is supported by the US DOE/SC Office of High Energy and Nuclear Physics and the Office of Advanced Scientific Computing Research.

## References

[1] K. Hirata, H. Moshammer, F. Ruggiero, Part. Accel. 40 (1993) 205.
[2] M.A. Furman, Beam–Beam Simulations with the Gaussian Code TRS, LBNL-42669, CBP Note 272, 1999.
[3] M.P. Zorzano, F. Zimmermann, Phys. Rev. Special Topics—Accel. Beams 3 (2000) 044401.
[4] S. Krishnagopal, Phys. Rev. Lett. 76 (1996) 235.
[5] Y. Cai, A.W. Chao, S.I. Tzenov, T. Tajima, Phys. Rev. Special Topics—Accel. Beams 4 (2000) 011001.
[6] W. Herr, M.P. Zorzano, F. Jones, Phys. Rev. Special Topics—Accel. Beams 4 (2001) 054402.
[7] E.B. Anderson, T.I. Banks, J.T. Rogers, in: Proceedings of the 1999 Particle Accelerator Conference, New York, 1999, pp. 1686.
[8] K. Ohmi, Phys. Rev. E 62 (2000) 7287.
[9] J. Shi, D. Yao, Phys. Rev. E 62 (2000) 1258.
[10] J. Qiang, M.A. Furman, R.D. Ryne, Phys. Rev. Special Topics—Accel. Beams 5 (2002) 104402.
[11] A. Kabel, in: Proceedings of the 2003 Particle Accelerator Conference, Portland, 2003, pp. 3545.
[12] J. Qiang, M.A. Furman, R.D. Ryne, J. Comp. Phys. 198 (2004) 278.
[13] R.W. Hockney, J.E. Eastwood, Computer Simulation Using Particles, McGraw-Hill, New York, 1985.
[14] J. Binney, S. Tremaine, Galactic Dynamics, Princeton University Press, Princeton, New Jersey, 1987.
[15] W. Fischer, L. Ahrens, M. Bai, M. Blaskiewicz, P. Cameron, R. Michnoff, F. Pilat, V. Ptitsyn, T. Sen, S. Tepikian, D. Trbojevic, M. Vogt, J. van Zeijts, Observation of coherent beam–beam modes in RHIC, BNL C-A/AP/75, 2002.
[16] K. Yokoya, H. Koiso, Part. Accel. 27 (1990) 181.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 356–359
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Bmad: A relativistic charged particle simulation library ☆

D. Sagan

*Wilson Laboratory, Cornell University, Ithaca, NY 14853, USA*

Available online 28 November 2005

**Abstract**

Bmad is a subroutine library for simulating relativistic charged particle beams in high-energy accelerators and storage rings. Bmad can be used to study both single and multi-particle beam dynamics using routines to track both particles and macroparticles. Bmad has various tracking algorithms including Runge–Kutta and symplectic (Lie algebraic) integration. Various effects such as wakefields, and radiation excitation and damping can be simulated. Bmad has been developed in a modular, object-oriented fashion to maximize flexibility. Interface routines allow Bmad to be called from C/C++ as well as Fortran programs. Bmad is well documented. Every routine is individually annotated, and there is an extensive manual.


*PACS:* 29.20.Dh; 29.27.−a

*Keywords:* Simulation; Particle accelerators; Relativistic beams

## 1. Overview

The Bmad [1] subroutine library for simulating relativistic charged particle beams was created to enable programmers to develop programs without the need to code from scratch commonly used functions such as lattice file parsing and particle tracking. Using a subroutine library such as Bmad cuts down on the time needed to develop programs and reduces programming errors.

Bmad has been developed at Cornell University's Laboratory for Elementary Particle Physics. The name Bmad is derived from the MAD [2] simulation program. Originally, the Bmad standard lattice input format was a subset of the MAD format, and the name Bmad was chosen as a shortening of the name "Baby MAD". As explained below, the Bmad lattice format now has numerous extensions that are not part of MAD.

Bmad, with some 400 routines, can do many things. Bmad can be used to study both single and multi-particle beam dynamics. It has routines to track both particles (represented by a phase-space position vector) and macroparticles (represented by a centroid phase-space vector with a $6 \times 6$ beam size sigma matrix). Bmad has various tracking algorithms including Runge–Kutta [3] and symplectic (Lie algebraic) integration. Bmad has routines for calculating transfer matrices, emittances, Twiss parameters, dispersion, coupling, etc. The elements that Bmad knows about include quadrupoles, RF cavities (both storage ring and LINAC accelerating types), solenoids, dipole bends, etc. In addition, elements can be defined to control the attributes of other elements. This can be used to simulate the "I-beams" which physically support components in the accelerator, or to easily simulate the action of control room "knobs".

Bmad, written in Fortran90, has been developed in a modular, object-oriented fashion to maximize flexibility. To facilitate interfacing Bmad with C and C++, Bmad has translation routines to convert Bmad structures to corresponding C/C++ structures. To facilitate working with Bmad in general, Bmad has been well documented [1]: every Bmad routine has header comments explaining how to use the routine, and there is an extensive manual covering nomenclature, physics, lattice input format and, programming.

☆ Work supported by the National Science Foundation.
*E-mail address:* dcs16@cornell.edu.

0168-9002/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2005.11.001

## 2. Lattice parsing

The Bmad library includes two lattice parser subroutines. One parser interfaces with the Extended Standard Input Format (XSIF) library developed at SLAC [4] so that XSIF lattice files can be read. The other parser implements the Bmad standard lattice format.

Aside from the standard MAD elements such as quadrupoles, bends, RF cavities, etc., the Bmad standard format includes wiggler [5], and Taylor map elements. Like the SLAC extension to MAD, the Bmad standard also implements LINAC accelerating cavity elements. Additionally, Bmad has a combination bend, solenoid, and quadrupole element with the solenoid field being able to vary linearly with longitudinal position

$$B_s = B_{s0} + \alpha s. \tag{1}$$

With such an element, a realistic solenoid with fringe fields can be easily simulated. All lattice elements can be individually positioned with respect to the reference orbit so, for example, the effect of lattice errors can easily be studied. Besides shifting elements with respect to the reference orbit, Bmad also has "patch" elements that can be used to shift the reference orbit itself.

The Bmad lattice format has the capability to easily superimpose elements on top of other elements as long as the combination can be represented as a valid Bmad element. A common example occurs when a quadrupole magnet is placed inside the field of a larger solenoid magnet. Thus, for example, a lattice file could contain the following:

```
Q: quad, l = 10
D: drift, l = 6
S: solenoid, l = 5, superimpose, &
  ref = Q, ref_end, offset = 1.4
lat: line = (Q, D)
use, lat
```

This places the center of S is 1.4 m away from the exit end of Q. The region where Q and S intersect will be treated as a combination solenoid/quadrupole element.

The Bmad lattice format also allows elements to be defined that control the attributes of other elements. An element, for example, could be defined to control the strengths of a string of quadrupoles, thus mimicking the effect of a power supply. In this way, control room "knobs" may easily be simulated. For example:

```
OV: overlay = {Q1, Q2/2.0, S[ks]/0.4}, k1 =
1.0.
```

This example defines a control element OV, called an overlay, that controls the k1 quadrupole strength of elements Q1 and Q2 and the ks solenoid strength of element S. The factors of 2.0 and 0.4 are scaling factors so that when, in this example, the value of OV is set to 1.0, the k1 strength of Q2 is 2.0, and the ks strength of S is 0.4. Control elements can control other control elements, so a hierarchy of controls may be established. Another type of control element is called an i_beam which simulates a support structure that orients the elements that are attached to it in space. For example:

```
IB: i_beam = {Q1, *, Q2}
IB[x_offset] = 0.012
IB[tilt] = 1e-4 * pi
```

would create an i_beam element that supports all elements between elements Q1 and Q2. Once IB is defined, it can be moved and tilted.

## 3. Tracking and transfer map calculations

Tracking and transfer map calculations are at the heart of most simulations, and different problems will have different requirements as to accuracy, speed, etc. To preserve flexibility, Bmad implements a number of different tracking and transfer map engines. Each lattice element can be separately assigned which engine to use. Furthermore, Bmad has routines to do tracking both in the forward and backwards directions.

The default "Bmad_standard" engine is implemented using thick element formulas which can be quickly evaluated. The drawback here is that the formulas are not necessarily symplectic. This is generally good for, for example, in lattice design where symplecticity is not an issue. The Polymorphic Tracking Code (PTC) library of Etienne Forest [6] has been interfaced to Bmad. PTC tracks using Lie algebraic techniques to integrate the appropriate Hamiltonians and this guarantees symplecticity. Differential Algebra (DA) is implemented in PTC with the result that Taylor maps of arbitrary order can be generated. PTC has a full suite of DA routines to manipulate Taylor maps, including map concatenation and normal form analysis. Alternatively, the second order transfer maps that are used in the MAD program are available for use. Furthermore, by writing the appropriate routines, custom calculations may be easily implemented. Runge–Kutta [3] and Boris [7] integrators are also available which allow tracking from a knowledge of the electric and magnetic fields within an element.

In addition to tracking with particles, macroparticles with wakefields have been implemented in Bmad. A macroparticle has a centroid which propagates like a particle, and a $6 \times 6$ sigma matrix which characterizes the size of the macroparticle. Since the full $6 \times 6$ sigma matrix is used, as opposed to only the transverse $4 \times 4$ part, macroparticles in Bmad can be propagated through bends.

## 4. C/C++ interface

Bmad is written in Fortran90 in an object-oriented fashion. To facilitate interfacing Bmad with routines written in C or C++, C structures with corresponding translation subroutines have been implemented to transfer

data between Fortran and C/C++. For example, a Fortran program with a call such as:

```
type (ele_struct) ele
call c_example (ele)
```

could be handled with a C++ routine like:

```
void c_example_(void* f_ele) {
 c_ele_struct c_ele;
 ele_struct_to_c_(f_ele, &c_ele);
 cout << "X_Beta: " << c_ele.x.beta <<
 endl;
 cout << "X_Eta: " << c_ele.x.eta << endl;
 c_ele_struct_deallocate (&c_ele);
}
```

Ele_struct is a Bmad structure that holds all the information about a single element (attribute values, tracking method used, etc.). The corresponding C structure is the c_ele_struct. The ele_struct_to_c routine transfers the information form an ele_struct to a c_ele_struct. The c_ele_struct_deallocate routine cleans up any allocated memory in c_ele.

## 5. Applications

The versatility of Bmad has led to its use in a number projects at Cornell. Bmad is the engine that drives the main software tools used for designing and simulating the Cornell Electron/positron Storage Ring CESR. Projects include dynamic aperture studies, lattice design, Beam-beam luminosity simulations, element misalignment studies, synchrotron radiation calculations, Baba scattering simulations, injection simulations, and measurement and correction of orbits, dispersion, coupling, beta, and betatron phase.

Bmad has also been used in conjunction with Cornell's Energy Recovery Linac (ERL) design effort and simulations of the Next Linear Collider and TESLA designs. Work is now in progress to use Bmad for lattice coupling simulation and correction at SLAC's PEP-II storage ring.

Fig. 1, which is a schematic diagram of the interaction region of the CESR storage ring, shows an example of how Bmad can be easily used. Two cryostats (only one is shown in the figure) holding four superconducting quadrupoles each are situated to either side of the interaction point. The quadrupoles come in skew/non-skew overlapping pairs and, to complicate matters, one pair of quadrupoles is partially inside the solenoidal field of the CLEO detector. To correct installation misalignments, each cryostat is movable via four eccentric cams. The standard CESR lattice file represents this situation by superimposing the superconducting quadrupoles on the lattice. The cams are represented by overlays that control the positions of the quadrupoles (it would be somewhat simpler to use an i_beam element here but the template for the standard CESR lattice was created before the i_beam element was implemented).

Since all the cam control information is in the lattice file, it is a simple and straightforward matter for any program to simulate cam motion. For example, the following snippet of code would calculate the change in orbit due to a movement of a cam by an amount `dr`:

```
call bmad_parser ('lattice_file_name', &
        lattice)
call closed_orbit_calc (lattice, &
        orbit_before, 4)
call element_locator (lattice, &
        'SC_CAM_4W', ix)
lattice%ele_(ix)%value(rho$) = &
        lattice%ele_(ix)%value(rho$) + dr
call control_bookkeeper (lattice, ix)
call closed_orbit_calc (lattice, &
        orbit_after, 4)
```

The call to bmad_parser reads in the lattice information and puts it into the `lattice` variable. The call to closed_orbit_calc calculates the closed orbit. Element_locator locates the cam in the element list contained within the `lattice` variable, and control_bookkeeper transfers the cam move to the appropriate elements that are used in tracking.

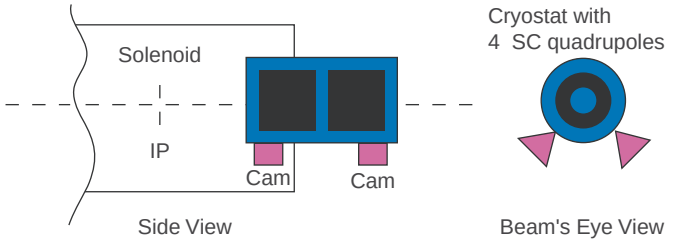



Fig. 1. Schematic diagram showing the positioning of the superconducting quadrupoles inside a cryostat near the interaction point of the CESR ring.

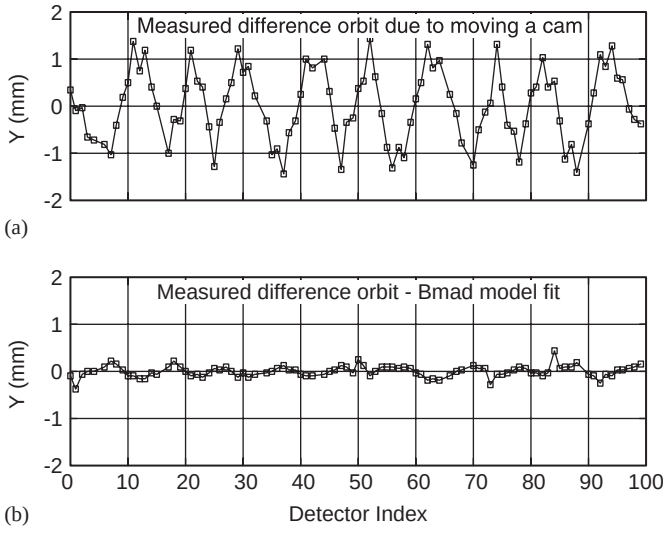



Fig. 2. (a) The measured change in orbit due to moving a single cam. (b) The measured change in orbit—the modeled change in orbit.

The results of a cam simulation are shown in Fig. 2. Fig. 2a shows the measured change in the vertical orbit due to a change in position of one of the cams. A program adjusted a simulated cam so that the simulated change in orbit best fit the measured orbit. A good fit to the data was achieved as shown in Fig. 2b which shows the difference between the measured and simulated orbit changes.

## 6. Summary

Bmad is a flexible, well documented, object-oriented environment for simulating accelerators and storage rings. It is not only useful in designing simulation programs but, since it is a software library, Bmad can be built into control room operation programs as well. Distributions with build scripts are available on OSF UNIX, Linux, and VMS. Bmad has also been compiled on Windows. The Bmad distribution and documentation may be obtained from Ref. [1]

## Acknowledgements

My thanks to David Rubin for his support, to Etienne Forest for use of his remarkable PTC/FPP library not to mention his patience in explaining everything to me, to Mark Palmer for all his work porting Bmad to different platforms, to Hans Grote for granting the adaptation of figures in the MAD manual for use in the Bmad manual, and to Richard Helms, Kim Moore, Jeremy Urban, Jeff Smith, and Mike Forster for their help.

## References

[1] For the Bmad distribution and documentation see: ⟨http://www.lepp.cornell.edu/~dcs/bmad⟩.
[2] H. Grote, F.C. Iselin, The MAD Program User's Reference Manual, Version 8.19, CERN/SL/90-13 (AP) (REV. 5) (1996). Can be obtained at: ⟨http://mad.home.cern.ch/mad⟩ (under MAD version 8).
[3] W. Press, B. Flannery, S. Teukolsky, W. Wetterling, Numerical Recipes in Fortran, the Art of Scientific Computing, second ed., Cambridge University Press, New York, 1992.
[4] P. Tenenbaum, LIBXSIF, a stand alone library for parsing the standard input format, in: Proceedings of 2001 Particular Accelerator Conference, 2001, pp. 3093–3095.
[5] D. Sagan, J. Crittenden, D. Rubin, A symplectic model for wigglers, Particular Accelerator Conference, 2003.
[6] E. Forest, Beam dynamics: a new attitude and framework, Harwood Academic Publishers, Amsterdam, 1998, see: ⟨http://acc-physics.kek.jp/forest/PTC/Introduction.html⟩.
[7] P.H. Stoltz, J.R. Cary, Efficiency of a Boris-like integration scheme with spatial stepping, Phys. Rev. Special Topics Accel. Beams 5 (2002) 094001.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 361–365
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Author index



doi:10.1016/S0168-9002(06)00134-3

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Nuclear Instruments and Methods in Physics Research A 558 (2006) 598–602
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH Section A
www.elsevier.com/locate/nima

# Subject index

ICAP-2004

Proceedings of the 8th International Computational

Accelerator Physics Conference

St. Petersburg, Russia, June 29–July 2, 2004



doi:10.1016/S0168-9002(06)00355-X

# NUCLEAR INSTRUMENTS AND METHODS IN PHYSICS RESEARCH

## Section A – Accelerators, Spectrometers, Detectors and Associated Equipment




**Aims and Scope**

Nuclear Instruments and Methods in Physics Research – Section A publishes papers on design, manufacturing and performance of scientific instruments with an emphasis on large scale facilities. This includes the development of particle accelerators, ion sources, beam transport systems and target arrangements as well as the use of secondary phenomena such as synchrotron radiation and free electron lasers. It also includes all types of instrumentation for the detection and spectrometry of radiations from high energy processes and nuclear decay, as well as instrumentation for experiments at nuclear reactors. Specialized electronics for nuclear and other types of spectrometry as well as computerization of measurements and control systems in this area also find their place. Theoretical as well as experimental papers are accepted.


**Publication Information**

Nuclear Instruments and Methods in Physics Research – Section A (ISSN 0168-9002). For 2006, *volumes* 555–568 (are) scheduled for publication. Subscription prices are available upon request from the Publisher or from the Regional Sales Office nearest you or from this journal's website (http://www.elsevier.com/locate/nima). Further information is available on this journal and other Elsevier products through Elsevier's website: (http://www.elsevier.com). Subscriptions are accepted on a prepaid basis only and are entered on a calendar year basis. Issues are sent by standard mail (surface within Europe, air delivery outside Europe). Priority rates are available upon request. Claims for missing issues should be made within six months of the date of dispatch.

*For orders, claims, and journal enquiries* (see backcover for manuscript enquiries): please contact the Customer Service Department at the Regional Sales Office nearest you:

Orlando: Elsevier, Customer Service Department, 6277 Sea Harbor Drive, Orlando, FL 32887-4800, USA; phone: (877) 8397126 or (800) 6542452 [toll free numbers for US customers]; (+1) (407) 3454020 or (+1) (407) 3454000 [customers outside US]; fax: (+1) (407) 3631354 or (+1) (407) 3639661; e-mail: usjcs@elsevier.com or elspcs@elsevier.com

Amsterdam: Elsevier, Customer Service Department, PO Box 211, 1000 AE Amsterdam, The Netherlands; phone: (+31) (20) 4853757; fax: (+31) (20) 4853432; e-mail: nlinfo-f@elsevier.com

Tokyo: Elsevier, Customer Service Department, 4F Higashi-Azabu 1-Chome Bldg, 1-9-15 Higashi-Azabu, Minato-ku, Tokyo 106-0044, Japan; phone: (+81) (3) 5561 5037; fax: (+81) (3) 5561 5047; e-mail: jp.info@elsevier.com

Singapore: Elsevier, Customer Service Department, 3 Killiney Road, #08-01 Winsland House I, Singapore 239519; phone: (+65) 63490222; fax: (+65) 67331510; e-mail: asiainfo@elsevier.com