International Journal of Pure and Applied Mathematics

Volume 11 No. 3 2004, 241-282

LOCAL THEORY AND APPLICATIONS OF EXTENDED GENERATING FUNCTIONS

Béla Erdélyi[1], Martin Berz[2 §]

[1]Fermi National Accelerator Laboratory
MS 220, P.O. Box 500, Batavia IL, 60510, USA

[2]Department of Physics and Astronomy
Michigan State University
East Lansing, MI, 48824, USA
e-mail: berz@msu.edu

Abstract: The time t maps of Hamiltonian flows are symplectic. The order n Taylor series approximation with respect to initial conditions of such a map is symplectic through terms of order n. Given an order n Hamiltonian symplectic map, there are a variety of procedures, called symplectification methods, which produce exactly symplectic maps with Taylor series that agree with the initial Taylor map through terms of order n.

Here we extend the generating function method of symplectification. To this end, we develop a general theory of generating functions of canonical transformations. It is shown that locally any symplectic map has uncountably many generating functions, each of which is associated with a conformal symplectic map. Within the subgroup of linear conformal symplectic maps, the available types can be organized into equivalence classes represented by symmetric matrices. Furthermore, equivalence of symplectification with and without factorization of the symplectic maps into linear and nonlinear parts is proved.

The method is illustrated with two examples; an anharmonic oscillator, and the dynamics in a proposed new particle accelerator, the so-called Neutrino Factory, which is known to exhibit a wide spectrum of nonlinear effects.

Received: October 22, 2003

[§]Correspondence author

AMS Subject Classification: 81S10, 58F05
Key Words: Hamiltonian flows, Hamiltonian symplectic map, symplectic map, symplectification, factorization of the symplectic maps, particle accelerator, Neutrino Factory, anharmonic oscillator

1. Introduction

A symplectic integration method is an integration method that preserves the symplectic structure at every time step. It is well-known that symplectic integration methods have favorable qualitative properties compared to non-symplectic ones when used for long-term integration. The good long-term behavior has been explained by favorable global error propagation (which is usually linear in the symplectic case, compared to generically quadratic in the non-symplectic case, and at least in certain cases stays bounded in the radial direction [31]), and the fact that the methods introduce only Hamiltonian perturbations of the original system; if the perturbations are small enough, according to the KAM Theorem, most invariant tori, and hence most of the geometric structure, survive [33]. Also, they have very good energy conservation properties. Although it is known that in general the symplectic structure and the energy cannot be conserved simultaneously by a numerical method for a Hamiltonian system [22], the Hamiltonian is preserved by a symplectic integration scheme up to a function of the accuracy of the integrator, up to exponentially long times [7]. Sometimes (as a function of time step and initial condition) quasi-periodic and bounded energy errors are observed that seemingly last forever. There are various implementations of symplectic integrators in the fields of molecular dynamics [27], celestial mechanics [44], non-equilibrium statistical mechanics [6, 36], beam physics, etc.

However, it is not clear geometrically what is the exact meaning of symplecticity. This is even more true for symplectic tracking with maps (symplectic integration, where the "timestep" is one turn around an accelerator, which can be kilometers long) as applied in the case of beam physics. On the one hand, the element by element symplectic integration (here the timestep is much smaller than in the above mentioned case, typically a few meters, or one magnetic element in the accelerator structure) usually is implemented in a second order (called the thin lens, kick, or leapfrog) approximation. While it is exactly symplectic, it is also slow, only second order accurate in the time step, and not applicable in the case of general, nonseparable Hamiltonians. On the other hand, to asses the long-term stability of particles in a periodic accelerator structure in a reasonable amount of time, it is customary to compute an approximation of the one turn map, and then track with the map. Unfortunately, by approximating the originally symplectic map, for example by truncation of the Taylor series of the true map, its symplecticity is lost. Tracking large numbers of turns with truncated Taylor maps thus can potentially give inaccurate results. But we can hope that by recovering the exact symplecticity "artificially" from the truncated map, the long term tracking with the map will restore the properties of the original system, and will speed up considerably the estimation of the region, where stable orbits exist (dynamic aperture in the beam physics jargon). As we shall see, in the Differential Algebraic Framework [8, 10] this can done to very high orders.

Therefore, tracking symplectically with high order maps is symplectic integration taken in its usual sense, but because of the use of integrators of very high order [10], it is usually more accurate in the time step and faster. It is also true that the speed is achieved at the expense of increasing the time step; here the time step is in fact one turn around the accelerator (using the arclength along the reference orbit as the independent variable). Sometimes it might be necessary to balance the length of the part of the system represented by a map with the required accuracy. This can be done by splitting the whole system into several pieces and representing each lump by a transfer map. Also, the map approach has the important advantage to be able to incorporate in the map effects that are otherwise very time consuming to compute, as, for example, fringe fields (stray, unwanted fields present at the magnet ends) [12, 46, 11].

The main step in tracking symplectically with maps is the symplectification of the truncated, order n symplectic, Taylor maps. Several methods have been developed to achieve the symplectification of maps. There are two main streams: one is based on factorization methods, and consist of Cremona symplectification [1], integrable polynomial factorization [34] and monomial factorization [23]; the other one is based on mixed variable generating function methods [9]. All methods provide valid symplectification schemes. However, the symplectified map depends on the specific method used. It was realized that the particular schemes applied often make considerable differences in the final results. This fact triggered the studies of optimal symplectification. For details concerning optimal Cremona symplectification see [1]. Here we extend the method of generating function symplectification to an exceedingly large class of generators. The optimality of generating function symplectification can be expressed in terms of Hofer metric [25]; the problem was studied in a previous paper [19].

The first mention of the possibility of symplectic integration using generating functions dates back to 1956 [37]. Later it was rediscovered by others; see for example [13, 14]. Specifically, in beam physics, symplectic tracking with maps based on generating functions was proposed in [9, 23, 17]. In particular, it has been shown that in the Differential Algebraic Framework it is straightforward to compute the order $n+1$ truncation of the generating function from the order n truncation of the one turn map to any order n [9, 10]. Symplectic tracking to order three was first implemented in the code *marylie* [16], and to arbitrary order it was first implemented in *cosy infinity* [28], among others. The possibility to estimate the SSC dynamic aperture with generating functions-based symplectic tracking with one turn maps has been considered in [45]. Another approach to generating functions and maps is based on fitted maps [38].

The generating function symplectification methods mentioned above use only the conventional $F_1, ..., F_4$ (in Goldstein notation) types of generating functions [24]. Recently, in [21, 20] a generating function based symplectic integration scheme has been developed. The authors of [21] show that actually there are infinitely many generating functions associated to a symplectic map. Their methods are based on [35], which is basically a linear algebra problem, and its local generalizations to the nonlinear case. The global theory on manifolds of the classical generating function theory has been developed in the 70's [39, 41, 42, 43]. We combine the two, and formulate the general theory of generating functions of canonical transformations, with an eye on usefulness for computation in the Differential Algebraic Framework [10] used in *cosy infinity* [28].

To be able to say which generating function is the best one, first requires a characterization of the various different types. This provides the motivation to develop the general theory of generating functions in Section 2. We show that locally there is an isomorphism between symplectic and gradient maps, and this leads to infinitely many generating function types for every symplectic map. In passing we note that there is a more general theory, which is based on transformation of the problem into a problem in symplectic geometry [18]. This approach also gives insight into various problems of locality versus globality of the generating functions. We mention that following Weinstein work, the geometric approach to generating functions in the physics literature appeared at various degrees of completeness, as, for example, in [2, 4, 29, 32]. However, for the purpose of this paper it is sufficient to develop only the local theory.

Also, we present some transformation properties of the generating functions in Section 3. The transformation rules allow us to form equivalence classes of generating functions in Section 4. Two types of generating function are equivalent if they produce exactly the same symplectified map when applied to a given order n symplectic Taylor map. Also, we present briefly in Appendix A how the conventional generators fit into this framework.

Sometimes it is preferred to factor out the linear part of the map to be symplectified, and apply the symplectification procedure to the nonlinear part only (in fact this part will have identity as linear part). We show in Appendix B that there is nothing to be gained by this approach if we use the appropriate types of generating functions. Appendix C describes briefly the implications of linear symplectic variable changes on the outcome of the symplectification process.

Section 5 gives some details about the implementation of the method in the code *cosy infinity*.

Finally, Section 6 is devoted to two examples, namely an anharmonic oscillator and a proposed lattice of the Neutrino Factory, a new type of accelerator for the future.

2. General Theory of Generating Functions of Canonical Transformations

In the classical mechanics literature, traditionally only the 4 Goldstein type of generating functions are well-known. However, it is easy to show that, for example, the identity transformation cannot be generated by the type 1 and 4 generating functions. On the other hand, this set can be easily extended to 2^{2n} generating functions. It is showed that for any symplectic map, at least one generating function from this set exists locally [10]. The common factor for this set is that they all depend on mixed coordinates, more specifically on n initial coordinates and momenta, and n final coordinates and momenta.

The first sign that in fact there are more generating functions is dating back to Poincaré, who used a different type of generating function, which not only is a mixed variable function in the sense discussed above, but also mixes (linearly) initial and final conditions in all the $2n$ variables [30]. Later this generating function reappeared in [40] and [21], where also a unifying approach to the theory of generating functions has been presented, from which resulted that there are infinitely many generating functions.

However, while the approach of [21] gives important computational insight, the general mathematical foundation of the theory is contained in the series of papers [39, 41, 42, 40]. On the other hand, the general theory lacks exactly the computational aspect. Our purpose is to give a rigorous account of the mathematical basis, and to cast the theory into a convenient computational tool within the framework of Differential Algebraic Methods. In order to keep the length of the paper within acceptable limits we present in detail only the local theory. The detailed account of the global theory will be published elsewhere.

Let us start by introducing a few notations. Every map is regarded as a column vector. Let

$$\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \tag{1}$$

be a diffeomorphism of a subset of $\mathbb{R}^{4n}$ onto its image, and let

$$\alpha^{-1} = \begin{pmatrix} \alpha^1 \\ \alpha^2 \end{pmatrix} \tag{2}$$

be its inverse. Notice that α_i and α^i, $i = 1, 2$, are the first $2n$ and second $2n$ components of α and α^{-1} respectively. This entails that $\alpha_i : U \subset \mathbb{R}^{4n} \to V \subset \mathbb{R}^{2n}$, and analogously for α^i. It is worthwhile to note that there is a geometric significance to the use of $\mathbb{R}^{4n}$. Both symplectic maps and functions under certain conditions can be given a geometric interpretation in the form of Lagrangian submanifolds of $\mathbb{R}^{4n}$ (Lagrangian submanifolds are $2n$ dimensional submanifolds of $4n$

dimensional symplectic manifolds on which the symplectic forms vanish identically [18]). Let

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{3}$$

be the $4n \times 4n$ Jacobian of α, split into $2n \times 2n$ blocks. Let

$$\tilde{J}_{4n} = \begin{pmatrix} J_{2n} & 0_{2n} \\ 0_{2n} & -J_{2n} \end{pmatrix}, \tag{4}$$

where

$$J_{2n} = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}, \tag{5}$$

and I_n is the unit matrix of appropriate dimension. A map α is called conformal symplectic if

$$(\mathrm{Jac}\,(\alpha))^T J_{4n}\, \mathrm{Jac}\,(\alpha) = \mu \tilde{J}_{4n}, \tag{6}$$

where μ is a non-zero real constant [5]. Also, we denote by $\mathcal{I}$ the identity map of appropriate dimension. A map $\mathcal{M}$ is called symplectic if its Jacobian M satisfies the symplectic condition [15], that is

$$M^T J M = J. \tag{7}$$

We always assume that the symplectic maps are origin preserving. We call a map gradient map if it has symmetric Jacobian N. It is well-known that gradient maps can be written as the gradient of a function (hence the name) [10], that is

$$N = \mathrm{Jac}\,(\nabla F)^T \tag{8}$$

(∇F is regarded as a row vector [3]). Strictly speaking, this is true over any simply connected domain, but we always assume that this is true. The function F is called the potential of the map.

The best way to formulate the main result of this section is a theorem.

Theorem 1. *Let $\mathcal{M}$ be a symplectic map. Then, for every point z there is a neighborhood of z such that $\mathcal{M}$ can be represented by functions F via the relation*

$$(\nabla F)^T = \left(\alpha_1 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right) \circ \left(\alpha_2 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right)^{-1}, \tag{9}$$

where α is any conformal symplectic map such that

$$\det\left(C\left(\mathcal{M}\left(z\right),z\right)\cdot Mz+D\left(\mathcal{M}\left(z\right),z\right)\right)\neq 0\,. \tag{10}$$

Conversely, let F be a twice continuously differentiable function. Then, the map $\mathcal{M}$, defined by

$$M=\left(NC-A\right)^{-1}\left(B-ND\right)\,, \tag{11}$$

is symplectic. The matrices A,B,C,D,M, and N are defined above.

Definition 2. The function F is called the generating function of type α of $\mathcal{M}$, and denoted $F_{\alpha,\mathcal{M}}$.

The theorem says that, once the generator type is fixed, locally there is a one-to-one correspondence between symplectic maps and scalar functions, which are unique up to an additive constant. The constant can be normalized to zero without loss of generality. Due to the fact that there exist uncountably many maps of the form (6), we can conclude that to each symplectic map one can construct infinitely many generating function types.

Remark 3. It can be shown that an alternate way to compute $\mathcal{M}$ from F is by inverting (9); it gives [18]

$$\mathcal{M}=\left(\alpha^1\circ\begin{pmatrix}(\nabla F)^T\\ \mathcal{I}\end{pmatrix}\right)\circ\left(\alpha^2\circ\begin{pmatrix}(\nabla F)^T\\ \mathcal{I}\end{pmatrix}\right)^{-1}. \tag{12}$$

In fact, (12) holds if and only if (9) holds.

We note that (9) and (12) cannot be simplified, due to the fact that the entries in the equations have different dimensions. For instance, $\alpha_i,\alpha^i:\mathbb{R}^{4n}\to\mathbb{R}^{2n}$ and $\begin{pmatrix}\mathcal{M}\\ \mathcal{I}\end{pmatrix},\begin{pmatrix}(\nabla F)^T\\ \mathcal{I}\end{pmatrix}:\mathbb{R}^{2n}\to\mathbb{R}^{4n}$.

Proof of Theorem 1. We notice that, by the implicit function theorem, the proof can be reduced to the linear case. In particular, the linearization of (9) at some point reads

$$N=\left(AM+B\right)\left(CM+D\right)^{-1}. \tag{13}$$

Here all the entries in the equation are matrices. Therefore, by the implicit function theorem, if (9) is well defined at some point, i.e. $\det\left(CM+D\right)\neq 0$, then it also holds in a neighborhood of that point. Therefore, the proof is complete if we prove the following lemma.

Lemma 4. *Let $A, B, C, D \in \mathbb{R}^{2n\times 2n}$ be such that*

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^T \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \mu \begin{pmatrix} J & 0 \\ 0 & -J \end{pmatrix} \tag{14}$$

holds. Let $M \in \mathbb{R}^{2n\times 2n}$ be given. If A, B, C, D is chosen such that

$$\det (CM + D) \neq 0, \tag{15}$$

and if N is defined as

$$N = (AM + B)(CM + D)^{-1}, \tag{16}$$

which is equivalent to

$$M = (NC - A)^{-1}(B - ND), \tag{17}$$

then the following are equivalent:
1) M is symplectic, i.e. $M^T JM = J$,
2) N is symmetric, i.e. $N^T = N$.

In the proof of the lemma we will need the following proposition.

Proposition 5. *Let* $\det (CM + D) \neq 0$ *and N defined as in (16). Then* $\det (NC - A) \neq 0$.

Proof. Denote

$$\alpha = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{18}$$

Taking the determinants on both sides of (14) if follows that $\det(\alpha) \neq 0$. Thus, denote its inverse by

$$\alpha^{-1} = \begin{pmatrix} \bar{A} & \bar{B} \\ \bar{C} & \bar{D} \end{pmatrix}. \tag{19}$$

Then, if we expand the relations $\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = I$, we obtain

$$A\bar{A} + B\bar{C} = \bar{A}A + \bar{B}C = C\bar{B} + D\bar{D} = \bar{C}B + \bar{D}D = I, \tag{20}$$
$$A\bar{B} + B\bar{D} = \bar{A}B + \bar{B}D = C\bar{A} + D\bar{C} = \bar{C}A + \bar{D}C = 0. \tag{21}$$

First we compute

$$
\begin{aligned}
&\left(\bar{C}N+\bar{D}\right)(CM+D)\\
&=\left[\bar{C}(AM+B)(CM+D)^{-1}+\bar{D}\right](CM+D) && (22)\\
&=\bar{C}(AM+B)+\bar{D}(CM+D) && (23)\\
&=\left(\bar{C}A+\bar{D}C\right)M+\left(\bar{C}B+\bar{D}D\right) && (24)\\
&=I. && (25)
\end{aligned}
$$

Taking determinants on both sides we obtain that

$$\det\left(\bar{C}N+\bar{D}\right)\neq 0. \tag{26}$$

Next consider the identity

$$\begin{pmatrix} I & 0 \\ \bar{C} & I \end{pmatrix}\begin{pmatrix} I & -N \\ 0 & \bar{C}N+\bar{D} \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A-NC & B-ND \\ 0 & I \end{pmatrix}. \tag{27}$$

Taking determinants on both sides yet again, we obtain that

$$\det\left(\bar{C}N+\bar{D}\right)\cdot\det(\alpha)=\det(A-NC). \tag{28}$$

But $\det(\alpha)\neq 0$, hence

$$\det\left(\bar{C}N+\bar{D}\right)\neq 0\Rightarrow\det(A-NC)\neq 0. \tag{29}$$

Combining (26) and (29) we arrive at

$$\det(CM+D)\neq 0\Rightarrow\det(NC-A)\neq 0, \tag{30}$$

and the proposition is proved. □

Proof of Lemma 4. We can proceed to prove the lemma. First, we show that from (14) together with N being symmetric follows that M is symplectic. From (16) we can deduce that

$$N^T=\left((CM+D)^{-1}\right)^T(AM+B)^T. \tag{31}$$

Using the assumption that $N^T=N$, we get that

$$\left((CM+D)^{-1}\right)^T(AM+B)^T=(AM+B)(CM+D)^{-1}. \tag{32}$$

Recall that for any regular matrix X, we have $\left(X^{-1}\right)^T = \left(X^T\right)^{-1}$. To remove the inverses, (53) can be rewritten as

$$\left(M^T C^T + D^T\right)(AM + B) = \left(M^T A^T + B^T\right)(CM + D). \tag{33}$$

Performing the operations and regrouping of terms gives

$$M^T \left(C^T A - A^T C\right) M + M^T \left(C^T B - A^T D\right) + \left(D^T A - B^t C\right) M + \left(D^T B - B^T D\right) = 0. \tag{34}$$

From the expansion of (14) it follows that

$$A^T C - C^T A = \mu J,\; B^T D - D^T B = -\mu J, \tag{35}$$

$$A^T C - C^T A = 0,\; B^T C - D^T C = 0. \tag{36}$$

This entails that (34) reduces to

$$M^T J M = J, \tag{37}$$

that is M is symplectic.

To complete the proof now we need to show that (14) together with M being symplectic implies that N is symmetric. First we notice that, according to the above proposition, (16) always can be solved for M to give

$$M = (NC - A)^{-1}(B - ND). \tag{38}$$

Therefore,

$$M^T = (B - ND)^T \left((NC - A)^{-1}\right)^T, \tag{39}$$

$$M^{-1} = (B - ND)^{-1}(NC - A). \tag{40}$$

Also, from the symplectic condition $M^T J M = J$ it follows that $J M^T = M^{-1} J$. Inserting (39) and (40) in this equation, it gives

$$J (B - ND)^T \left((NC - A)^{-1}\right)^T = (B - ND)^{-1}(NC - A) J, \tag{41}$$

which can be expressed as

$$(B - ND) J \left(B^T - D^T N^T\right) = (NC - A) J \left(C^T N^T - A^T\right). \tag{42}$$

Rearrangement of terms gives

$$N\left(CJC^T - DJD^T\right)N^T - N\left(CJA^T - DJB^T\right) \\ - \left(AJC^T - BJD^T\right)N^T + \left(AJA^T - BJB^T\right) = 0. \quad (43)$$

Next, we need to manipulate (14), which can be written in the compact form $\alpha^T J\alpha = \mu\tilde{J}$. It is equivalent to $\alpha^T J = \mu\alpha^{-1}\tilde{J}$. Transposition gives $J\alpha = \mu\alpha^{-T}\tilde{J}$, where we used that $J^T = -J$ and $\tilde{J}^T = -\tilde{J}$. Also, from $J^{-1} = -J$ and $\tilde{J}^{-1} = -\tilde{J}$ it finally follows that

$$\alpha\tilde{J}\alpha^T = \mu J. \quad (44)$$

This relation expanded reads

$$AJA^T - BJB^T = 0,\ CJC^T - DJD^T = 0, \quad (45)$$

$$AJC^T - BJD^T = \mu I,\ CJA^T - DJB^T = -\mu I. \quad (46)$$

As the last step, inserting them in (43) results that

$$N^T = N. \quad (47)$$

This completes the proof. □

Theorem 1 has a simple, intuitive interpretation. It provides a way to construct infinitely many generating function types to any given symplectic map. The various types are parametrized by the group of conformal symplectic maps. For the existence of a certain type of generator, $\det(CM + D) \neq 0$ must hold. Conversely, given any function and a conformal symplectic map, it provides a method for generation of symplectic maps.

Once we have the pool of types of generating functions to choose from, we can reduce the complexity of the optimal symplectification problem by noticing that the generating functions can be organized into equivalence classes. First, we need some transformation properties of the generating functions.

3. Transformation Properties of Generating Functions

If we look at how the generating functions transform under modifications of α and/or $\mathcal{M}$, we obtain a set of rules which we call transformation properties. These properties are based on the fact that if α is a conformal symplectic map such that

$$(\mathrm{Jac}\,(\alpha))^T J_{4n}\,\mathrm{Jac}\,(\alpha) = \mu \tilde{J}_{4n}, \tag{48}$$

then for any β and γ such that

$$(\mathrm{Jac}\,(\beta))^T J_{4n}\,\mathrm{Jac}\,(\beta) = J_{4n}, \tag{49}$$

$$(\mathrm{Jac}\,(\gamma))^T \tilde{J}_{4n}\,\mathrm{Jac}\,(\gamma) = \tilde{J}_{4n}, \tag{50}$$

the map $\beta \circ \alpha \circ \gamma$ is also a valid conformal symplectic map, that is, it follows from (49) and (50), and repeated application of the chain rule that $\beta\circ\alpha\circ\gamma$ satisfies (48). Therefore, it gives another type of generating function.

We begin with studying what happens to the generating function $F_{\alpha,\mathcal{M}}$ under the transformation $\alpha_1 \mapsto \lambda\alpha_1$, for some non-zero real λ. This affects only the conformality factor μ, of α, which becomes $\lambda\mu$. Slight rearrangement of (9) gives

$$\left((\nabla F_{\alpha,\mathcal{M}})^T \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0. \tag{51}$$

Then, we also have

$$\left(\left(\nabla F_{\begin{pmatrix} \lambda\alpha_1 \\ \alpha_2 \end{pmatrix},\mathcal{M}}\right)^T \circ \alpha_2 - \lambda\alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0, \tag{52}$$

which is equivalent to

$$\left(\left(\nabla \left(\lambda^{-1} F_{\begin{pmatrix} \lambda\alpha_1 \\ \alpha_2 \end{pmatrix},\mathcal{M}}\right)\right)^T \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0. \tag{53}$$

Comparing (51) with (53) we see that

$$\nabla F_{\alpha,\mathcal{M}} = \nabla \left(\lambda^{-1} F_{\begin{pmatrix} \lambda\alpha_1 \\ \alpha_2 \end{pmatrix},\mathcal{M}}\right), \tag{54}$$

that is

$$F_{\begin{pmatrix} \lambda\alpha_1 \\ \alpha_2 \end{pmatrix},\mathcal{M}} = \lambda F_{\alpha,\mathcal{M}} + c, \tag{55}$$

for some arbitrary constant c.

Next, we study what happens if we change the symplectic map, for example, by $\mathcal{M} \mapsto \mathcal{M} \circ \mathcal{A}$, for some symplectic map $\mathcal{A}$. From (9) we have

$$\left((\nabla F_{\alpha,\mathcal{M}\circ\mathcal{A}})^T \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \circ \mathcal{A} \\ \mathcal{I} \end{pmatrix} = 0\,, \tag{56}$$

$$\left((\nabla F_{\alpha,\mathcal{M}\circ\mathcal{A}})^T \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{A}^{-1} \end{pmatrix} = 0\,, \tag{57}$$

$$\left((\nabla F_{\alpha,\mathcal{M}\circ\mathcal{A}})^T \circ (\alpha_2 \circ T_{\mathcal{A}}) - (\alpha_1 \circ T_{\mathcal{A}})\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0\,, \tag{58}$$

where $T_{\mathcal{A}}$ is defined by $T_{\mathcal{A}}(\hat{z}, z) = \left(\hat{z}, \mathcal{A}^{-1}(z)\right)$. Equation (58) can also be interpreted as

$$\left((\nabla F_{\alpha\circ T_{\mathcal{A}},\mathcal{M}})^T \circ (\alpha_2 \circ T_{\mathcal{A}}) - (\alpha_1 \circ T_{\mathcal{A}})\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0, \tag{59}$$

from where we conclude

$$F_{\alpha,\mathcal{M}\circ\mathcal{A}} = F_{\alpha\circ T_{\mathcal{A}},\mathcal{M}} + c. \tag{60}$$

In the same manner, the left action of another symplectomorphism on the map, i.e. $\mathcal{M} \mapsto \mathcal{K} \circ \mathcal{M}$ leads to

$$\left((\nabla F_{\alpha,\mathcal{K}\circ\mathcal{M}})^T \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{K} \circ \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0. \tag{61}$$

Define $T_{\mathcal{K}}(\hat{z}, z) = (\mathcal{K}(\hat{z}), z)$. Then,

$$F_{\alpha,\mathcal{K}\circ\mathcal{M}} = F_{\alpha\circ T_{\mathcal{K}},\mathcal{M}} + c. \tag{62}$$

We are also interested what happens when we change the coordinates in the generating function, $F \mapsto F \circ \mathcal{L}$, by a diffeomorphism $\mathcal{L}$ (here not necessarily a symplectomorphism); we have

$$\left((\nabla (F \circ \mathcal{L}))^T \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0\,, \tag{63}$$

$$(\mathrm{Jac}(\mathcal{L}))^T \cdot \left((\nabla F)^T \circ \mathcal{L} \circ \alpha_2 - \alpha_1\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0\,, \tag{64}$$

$$\left(\nabla F_{T_{\mathcal{L}}\circ\alpha,\mathcal{M}}\right)^T \circ (\mathcal{L} \circ \alpha_2) - \left((\mathrm{Jac}(\mathcal{L}))^{-T} \cdot \alpha_1\right)\Big) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix} = 0\,, \tag{65}$$

where we defined $T_{\mathcal{L}}(\hat{z}, z) = \Big((\mathrm{Jac}(\mathcal{L}))^{-T} \cdot \hat{z}, \mathcal{L}(z)\Big)$. Hence,

$$\nabla \left(F_{T_{\mathcal{L}} \circ \alpha, \mathcal{M}} \circ \mathcal{L}\right) = \nabla F_{\alpha, \mathcal{M}}, \tag{66}$$

that is

$$F_{T_{\mathcal{L}} \circ \alpha, \mathcal{M}} = F_{\alpha, \mathcal{M}} \circ \mathcal{L}^{-1} + c. \tag{67}$$

We will use these transformation rules in the following sections.

4. Equivalence Classes of Generating Functions

We call two types of generating function equivalent if both types generate exactly the same symplectified map when applied to a truncated, order n symplectic, Taylor map. As we showed, the different types are parametrized by conformal symplectic maps. In this section we show that all the types generated by linear α, and which exist at least locally for a given symplectic map, can be organized into equivalence classes characterized by symmetric matrices.

Assume that for a given symplectic map $\mathcal{M}$ there exists a generating function of type α given by

$$(\nabla F)^T = \left(\alpha_1 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right) \circ \left(\alpha_2 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right)^{-1}, \tag{68}$$

and

$$\mathrm{Jac}(\alpha) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{69}$$

From (14) it readily follows that

$$(\mathrm{Jac}(\alpha))^{-1} = \begin{pmatrix} JC^T & -JA^T \\ -JD^T & JB^T \end{pmatrix}. \tag{70}$$

First of all, it is straightforward to see that one can always change the conformality factor to $\mu = 1$ using the transformation rule (55), by choosing $\lambda = \mu^{-1}$. From (12) and (70) it easily follows that we get the same symplectified map in both cases. Therefore, the conformality factor does not introduce any flexibility into the symplectification process.

Hence we can always assume that $\mu = 1$, which is the most convenient value from the numerical implementation point of view [18].

Denote the linear part of $\mathcal{M}$ by M. Then the generating function of the same type that generates the linear part M is given by

$$(\nabla F_0)^T = \left(\alpha_1 \circ \begin{pmatrix} M \\ \mathcal{I} \end{pmatrix}\right) \circ \left(\alpha_2 \circ \begin{pmatrix} M \\ \mathcal{I} \end{pmatrix}\right)^{-1}. \tag{71}$$

Subtraction of (71) from (68) gives

$$(\nabla (F - F_0))^T = \left[\left(\alpha_1 - (\nabla F_0)^T \circ \alpha_2\right) \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right] \circ \left(\alpha_2 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right)^{-1} \tag{72}$$

$$= \left(\bar{\alpha}_1 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right) \circ \left(\alpha_2 \circ \begin{pmatrix} \mathcal{M} \\ \mathcal{I} \end{pmatrix}\right)^{-1} = (\nabla G)^T, \tag{73}$$

where we used the notations $G = F - F_0$ and $\bar{\alpha}_1 = \alpha_1 - (\nabla F_0)^T \circ \alpha_2$. Define

$$\beta(\hat{w}, w) = \begin{pmatrix} \hat{w} - (\nabla F_0)^T (w) \\ w \end{pmatrix}. \tag{74}$$

Obviously, being a kick, i.e. changing only one component, β is a symplectic map for any function F_0. Clearly, with $\bar{\alpha}_2 = \alpha_2$ we have $\bar{\alpha} = \beta \circ \alpha$. Therefore, according to the transformation properties of the previous section, G is a valid generating function of type $\bar{\alpha}$. If we denote $N = \text{Jac}(\nabla F_0)^T = (AM + B)(CM + D)^{-1}$, the Jacobian of $\bar{\alpha}$ is given by

$$\text{Jac}(\bar{\alpha}) = \begin{pmatrix} A - NC & B - ND \\ C & D \end{pmatrix}, \tag{75}$$

and its inverse by

$$(\text{Jac}(\bar{\alpha}))^{-1} = \begin{pmatrix} JC^T & -J\left(A^T - C^T N\right) \\ -JD^T & J\left(B^T - D^T N\right) \end{pmatrix}. \tag{76}$$

Notice that N is actually the Hessian of a function, and hence symmetric, i.e. $N^T = N$.

Here we have to make an important observation. The symplectification procedure consists of starting with $\mathcal{M}_n$ and an a priori fixed α, and computing F_n using (9). Then (12) gives an exactly symplectic map, which we call the symplectified map. Unfortunately, on a computer (9) has to be represented by implicit equations and solved by fixed point iterations, but formally the Taylor expansion of the symplectified map (12) will be $\mathcal{M}_n$ up to order n. The point to be emphasized is that one needs an a priori fixed α that is exactly symplectic (not only up to order n) for the procedure to work. However, in general it is not easy to construct exactly symplectic polynomial maps of degree at most n. Even in the case that one constructs such a map, in general there is no reason to believe that $(\nabla F_0)^T$ as given by (71) will be a polynomial map of degree at most n. Thus, in this case the exact symplecticity of $\bar{\alpha}$ will be spoiled. Therefore, we are constrained to consider equivalence classes of the types of generating functions associated with the subgroup of linear conformal symplectic maps.

To this end, we can compare the two symplectified maps, that is the map obtained from F_n and α, and the map obtained from G_n and $\bar{\alpha}$. Notice that if α is linear, $(\nabla F_0)^T$ and hence $\bar{\alpha}$ are also linear. Then for the Jacobians of the symplectified maps we obtain from (12)

$$\begin{aligned}\mathrm{Jac}\,(\mathcal{M}_{F_n,\alpha}) &= \left[\begin{pmatrix} JC^T & -JA^T \end{pmatrix}\begin{pmatrix} \mathrm{Jac}\,(\nabla F_n)^T \\ I \end{pmatrix}\right] \\ &\quad\cdot\left[\begin{pmatrix} -JD^T & JB^T \end{pmatrix}\begin{pmatrix} \mathrm{Jac}\,(\nabla F_n)^T \\ I \end{pmatrix}\right]^{-1} \\ &= \left(JC^T\cdot\mathrm{Jac}\,(\nabla F_n)^T - JA^T\right)\cdot\left(-JD^T\cdot\mathrm{Jac}\,(\nabla F_n)^T + JB^T\right)^{-1},\end{aligned} \tag{77}$$

$$
\begin{aligned}
&\operatorname{Jac}\left(\mathcal{M}_{G_n,\bar{\alpha}}\right) \\
&= \left[\begin{pmatrix} JC^T & -J\left(A^T - C^T N\right) \end{pmatrix} \begin{pmatrix} \operatorname{Jac}\left(\nabla F_n\right)^T - N \\ I \end{pmatrix}\right] && (78) \\
&\cdot \left[\begin{pmatrix} -JD^T & J\left(B^T - D^T N\right) \end{pmatrix} \begin{pmatrix} \operatorname{Jac}\left(\nabla F_n\right)^T - N \\ I \end{pmatrix}\right]^{-1} && (79) \\
&= \left(JC^T \cdot \operatorname{Jac}\left(\nabla F_n\right)^T - JC^T N - JA^T + JC^T N\right) && (80) \\
&\cdot \left(-JD^T \cdot \operatorname{Jac}\left(\nabla F_n\right)^T + JD^T N + JB^T - JD^T N\right)^{-1} && (81) \\
&= \left(JC^T \cdot \operatorname{Jac}\left(\nabla F_n\right)^T - JA^T\right) \cdot \left(-JD^T \cdot \operatorname{Jac}\left(\nabla F_n\right)^T + JB^T\right)^{-1}. && (82)
\end{aligned}
$$

Since the maps are assumed to be origin preserving, we can conclude that

$$\mathcal{M}_{F_n,\alpha} = \mathcal{M}_{G_n,\bar{\alpha}}. \tag{83}$$

Thus we get the same symplectified map regardless of using F_n of type α, or G_n of type $\bar{\alpha}$. So why is G_n interesting? It is interesting because of the following property: if we denote the Jacobian of $\bar{\alpha}$ by

$$\begin{pmatrix} \bar{A} & \bar{B} \\ C & D \end{pmatrix}, \tag{84}$$

from (75) we observe that

$$
\begin{aligned}
\bar{A}M + \bar{B} &= (A - NC)\, M + (B - ND) && (85) \\
&= (AM + B) - N\,(CM + D) && (86) \\
&= (AM + B) - (AM + B)\,(CM + D)^{-1}\,(CM + D) && (87) \\
&= 0. && (88)
\end{aligned}
$$

Therefore we need to consider only the types that satisfy $\bar{A}M + \bar{B} = 0$, in addition to the usual constraints imposed by (14).

However, it is possible to further reduce the equivalence classes. We will use the transformation rule (67) with linear $\mathcal{L}$. Denoting $\operatorname{Jac}(\mathcal{L}) = L$ and $\tilde{\alpha} = T_{\mathcal{L}} \circ \bar{\alpha}$ we obtain

$$\operatorname{Jac}(\tilde{\alpha}) = \begin{pmatrix} \left(L^{-1}\right)^T \bar{A} & \left(L^{-1}\right)^T \bar{B} \\ LC & LD \end{pmatrix}. \tag{89}$$

We choose $L = (CM + D)^{-1}$. After writing out explicitly the constraints contained in (14), a straightforward calculation shows that

$$(CM + D)^{-1} = -M^{-1} J \bar{A}^T, \tag{90}$$

$$(CM + D)^T = -M^T J \bar{A}^{-1}. \tag{91}$$

This entails that

$$\mathrm{Jac}\,(\tilde{\alpha}) = \begin{pmatrix} -M^T J & J \\ -M^{-1} J \bar{A}^T C & -M^{-1} J \bar{A}^T D \end{pmatrix}, \tag{92}$$

and

$$(\mathrm{Jac}\,(\tilde{\alpha}))^{-1} = \begin{pmatrix} JC^T \bar{A} M J & M \\ -JD^T \bar{A} M J & I \end{pmatrix}. \tag{93}$$

As mentioned above, from (67) we obtain that $G_{\bar{\alpha},\mathcal{M}} = F_{\tilde{\alpha},\mathcal{M}} \circ \mathcal{L}$. We drop the subscript as there is no danger of confusion. Since $\mathcal{L}$ is linear, we also infer that

$$G_n = F_n \circ \mathcal{L}, \tag{94}$$

and as a consequence

$$(\nabla G_n)^T = L^T \cdot (\nabla F_n)^T \circ \mathcal{L}, \tag{95}$$

or

$$(\nabla F_n)^T = \left(L^{-1}\right)^T \cdot (\nabla G_n)^T \circ \mathcal{L}^{-1}. \tag{96}$$

We are now in position to compare the Jacobians of the two symplectified maps, and obtain

$$\begin{aligned} &\mathrm{Jac}\,(\mathcal{M}_{G_n,\bar{\alpha}}) \\ &\quad = \left[\begin{pmatrix} JC^T & -J\bar{A}^T \end{pmatrix} \begin{pmatrix} \mathrm{Jac}\,(\nabla G_n)^T \\ I \end{pmatrix}\right] \\ &\quad \cdot \left[\begin{pmatrix} -JD^T & -JM^T\bar{A}^T \end{pmatrix} \begin{pmatrix} \mathrm{Jac}\,(\nabla G_n)^T \\ I \end{pmatrix}\right]^{-1} \\ &\quad = \left(JC^T \cdot \mathrm{Jac}\,(\nabla G_n)^T - J\bar{A}^T\right) \\ &\quad\quad \cdot \left(-JD^T \cdot \mathrm{Jac}\,(\nabla G_n)^T - JM^T\bar{A}^T\right)^{-1}, \end{aligned} \tag{97}$$

where we used $\bar{B} = -\bar{A}M$ in the second line. Similarly, a somewhat cumbersome calculation shows that

$$\begin{aligned}
&\mathrm{Jac}\left(\mathcal{M}_{F_n,\tilde{\alpha}}\right) \\
&= \left[\begin{pmatrix} JC^T\bar{A}J\left(M^T\right)^{-1} & M \end{pmatrix}\begin{pmatrix} \left(L^{-1}\right)^T \cdot \mathrm{Jac}\left(\nabla G_n\right)^T \cdot L^{-1} \\ I \end{pmatrix}\right] \\
&\quad \cdot \left[\begin{pmatrix} -JD^T\bar{A}MJ & I \end{pmatrix}\begin{pmatrix} \left(L^{-1}\right)^T \cdot \mathrm{Jac}\left(\nabla G_n\right)^T \cdot L^{-1} \\ I \end{pmatrix}\right]^{-1} \\
&\quad = \left(JC^TAJ\left(M^T\right)^{-1}M^TJ\bar{A}^{-1} \cdot \mathrm{Jac}\left(\nabla G_n\right)^T - MM^{-1}J\bar{A}^T\right) \\
&\qquad \cdot \left(JD^TAMJM^TJ\bar{A}^{-1} \cdot \mathrm{Jac}\left(\nabla G_n\right)^T - M^{-1}J\bar{A}^T\right)^{-1} \\
&\qquad = \left(JC^T \cdot \mathrm{Jac}\left(\nabla G_n\right)^T - J\bar{A}^T\right) \\
&\qquad\quad \cdot \left(-JD^T \cdot \mathrm{Jac}\left(\nabla G_n\right)^T - JM^T\bar{A}^T\right)^{-1}, \qquad (98)
\end{aligned}$$

where we used $MJM^T = J$. Hence, we obtained again that

$$\mathcal{M}_{G_n,\bar{\alpha}} = \mathcal{M}_{F_n,\tilde{\alpha}}, \tag{99}$$

and after combining (83) and (99) we finally arrive at

$$\mathcal{M}_{F_n,\alpha} = \mathcal{M}_{F_n,\tilde{\alpha}}. \tag{100}$$

Thus, the symplectified map obtained from a truncated generating function of type α (linear) agrees with the symplectified map obtained from type $\tilde{\alpha}$. Denote

$$\mathrm{Jac}\left(\tilde{\alpha}\right) = \begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{pmatrix}. \tag{101}$$

Notice that the property from the first step of the reduction, that is

$$\begin{aligned}
\tilde{A}M + \tilde{B} &= -M^TJM + J \qquad (102) \\
&= -J + J = 0\,, \qquad (103)
\end{aligned}$$

is preserved, and in addition it has another very nice property, namely

$$\begin{aligned}
\tilde{C}M + \tilde{D} &= -M^{-1}J\bar{A}^T\left(CM + D\right) \qquad (104) \\
&= -M^{-1}J\bar{A}^T\left(\bar{A}^T\right)^{-1}JM = I, \qquad (105)
\end{aligned}$$

where we used (90) in the second line.

Therefore, every generating function type associated with linear maps, which exists at least locally for a given symplectic map, is equivalent for symplectification purposes with another type associated with

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{106}$$

such that the following relations hold:

$$AM + B = 0, \tag{107}$$
$$CM + D = I. \tag{108}$$

Equations (107) and (108) have to be imposed in addition to the usual constraints derived from (14), that is

$$AJA^T - BJB^T = 0,\ CJC^T - DJD^T = 0,\ DJB^T - CJA^T = I. \tag{109}$$

These five conditions restrict very much the pool of independent generator types. From (107) and (108) we obtain $B = -AM$ and $D = I - CM$ respectively, which inserted in (109) gives

$$A = -JM^{-1}, \tag{110}$$

and re-inserted in (107) gives

$$B = -AM = -\left(-JM^{-1}\right) M = J. \tag{111}$$

The first condition in (109) is automatically satisfied if we impose (107) and (108). Inserting $D = I - CM$ in the second relation of (109) we obtain

$$CMJ - (CMJ)^T = J. \tag{112}$$

We put

$$CM = \frac{1}{2}\left(I + JS\right). \tag{113}$$

This is always possible for some matrix S. Insertion of (113) in (112) gives that S must be symmetric, i.e.

$$S^T = S, \tag{114}$$

but can otherwise be arbitrary. Thus we obtained

$$C = \frac{1}{2}(I + JS) M^{-1}, \tag{115}$$

$$D = \frac{1}{2}(I - JS). \tag{116}$$

Therefore, every generator type belongs to an equivalence class $[S]$ associated with

$$\begin{pmatrix} -JM^{-1} & J \\ \frac{1}{2}(I + JS) M^{-1} & \frac{1}{2}(I - JS) \end{pmatrix}, \tag{117}$$

and represented by the symmetric matrix S.

Given an arbitrary type of generating function, how do we know which equivalence class it belongs to? We saw that

$$\tilde{C} = (CM + D)^{-1} C, \quad \tilde{D} = (CM + D)^{-1} D, \tag{118}$$

and similarly

$$\tilde{C} = \frac{1}{2}(I + JS) M^{-1}, \quad \tilde{D} = \frac{1}{2}(I - JS). \tag{119}$$

We can express $\tilde{C}M - \tilde{D}$ from the first two and second two relations respectively, obtaining

$$(CM + D)^{-1}(CM - D) = JS, \tag{120}$$

or equivalently

$$S = -J(CM + D)^{-1}(CM - D). \tag{121}$$

To remind ourselves, equivalence means that generating functions from the same equivalence class will produce indistinguishable results if used to symplectify a given order n symplectic map. Thus we just proved the following theorem.

Theorem 6. *Every generating function type associated with a linear conformal symplectic map that exists at least locally for a given symplectic map belongs to an equivalence class represented by a symmetric matrix. An arbitrary type of generator, associated with a linear* α *satisfying conditions (109) and (15), belongs to a class associated with (117), and characterized by the symmetric matrix given by (121).*

In Appendix A we show how the conventional generating function types fit into this framework.

Symplectification can be performed on the nonlinear part only by first factoring out the linear part, or on the whole map. The next question that arises naturally is whether there is something to be gained if one first factors out the linear part of the symplectic map to be symplectified. We address this problem in Appendix B.

Combining results of Appendix B, we show in Appendix C that not even a linear symplectic change of variables can provide additional freedom in the symplectification process.

5. Implementation

In this section we describe the implementation of symplectification in *cosy infinity*. The method starts with $\mathcal{M}_n$ given, and some arbitrary initial condition z. Utilizing equation (9) with α given by (117) we obtain the truncated α-generating function F_{n+1}. The arbitrary symmetric matrix S must be specified, fixing the type of generator utilized. All the necessary operations of map composition, map inversion, differentiation and integration are readily available in *cosy*. Then, notice that (9) can be expressed as

$$\hat{z} - M \cdot z = M \cdot J \cdot (\nabla F_n)^T \left(C \cdot (\hat{z} - M \cdot z) + z\right), \tag{122}$$

where we denoted $\hat{z} = \mathcal{M}_{[S]}(z)$, $\mathcal{M}_{[S]}$ representing formally the symplectified map, and

$$C = \frac{1}{2}(I + JS)M^{-1}. \tag{123}$$

To avoid as much as possible any problems with cancellation of digits, we denote

$$w = \hat{z} - M \cdot z, \tag{124}$$

which leads to

$$w = M \cdot J \cdot (\nabla F_n)^T \left(C \cdot w + z\right). \tag{125}$$

This can be solved by a fixed point iteration for w, and gives the final result by

$$\hat{z} = w + M \cdot z. \tag{126}$$

The orbit of z is then computed by iteration of the procedure (in the next step we take $\hat{z}$ as the initial condition, etc.).

Writing (125) as $w = f(w)$, we observe that to be able to solve (125) by a fixed point iteration, the right hand side must be contracting for a fixed z, i.e. is guaranteed to succeed if

$$|f(w_2) - f(w_1)| \leq q \cdot |w_2 - w_1|, \tag{127}$$

for some $q < 1$. To have a good chance of contractivity over an extended region, the first and second order derivatives of the generating functions must be small. From our experience, in general the fixed point iteration converges in the region where the generating function is defined.

Of course, (125) can be expressed as $f(w) - w = 0$, and solved for w by Newton method. We noticed that the results not only are sometimes dependent on the generating function type employed, but also on the numerics, that is the particular numerical method used to solve the implicit equations. Of course, if we start with an exactly order n symplectic map, and the convergence to the solution of the implicit equations is achieved over the tracking region for both methods, then the resulting pictures are identical. However, in practice the fixed point iteration works in a more stable manner. It is faster than Newton method when many particles are tracked simultaneously, and, in the vast majority of cases studied, its domain of convergence is larger. For maps of practical interest, Newton method often does not converge close to the dynamic aperture. Moreover, if the symplectification starts with truncated maps that are not exactly order n symplectic, the results depend on the way the truncated generating functions are computed. It seems that the general theory provides a good order n symplectification scheme. More about the performance of the algorithm will be presented in the next section, where we turn our attention to examples.

6. Examples

We illustrate the performance of the symplectification methods developed in this paper with two examples. In two dimensions we can easily generate symplectic maps to high orders. We can assume that these high order truncated maps are approximating the exact maps so well as to be considered numerically symplectic over a sizable phase space region. Then we can compute their generating functions, truncated at some modest order (say 7), and use them to generate exactly symplectic maps according to the above symplectification procedure. Finally, we can compare the various maps obtained this way. We will present some typical cases using various types of generating functions.

We study two examples: an anharmonic oscillator that has been studied previously in the symplectification literature [23, 1], and a lattice of the proposed Neutrino Factory [26]. Although we track the muons for their lifetime (only 1000 turns), it is still an interesting case due to the wide array of nonlinear effects which the lattice exhibits [12, 46, 11].

6.1. An Anharmonic Oscillator

We consider the $2D$ anharmonic oscillator described by the Hamiltonian

$$H = \frac{1}{2}\left(p^2 + q^2\right) - \frac{1}{4}q^4, \tag{128}$$

which has been studied previously in [23] to compare various symplectification methods, and in [1] to study optimal Cremona symplectification. To make the comparison easier we follow the same guidelines, and present the performance of our method by symplectifying the order 7 Taylor map of the time 1 map of the flow of (128). We track for 1000 turns and use as initial conditions

$$\begin{cases} q = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9,\ 0.95,\ 0,99,\ 1.0\,, \\ p = 0\,. \end{cases} \tag{129}$$

In Figure 1 we present the order 19 Taylor map for comparison purposes. We applied 8 different generating function symplectifications to the 7-th order Taylor map, which is shown in Figure 2. We displayed the results for the conventional F_1 through F_4 generators in Figure 3. Notice that F_1 is the best conventional generator for this example. Next we tracked with generators based on random symmetric matrices. In general, if the elements of S are chosen in $S_{ij} \in [-1, 1]$ we obtain better results than if we increased the norm of S. For example see Figure 4, and Figure 5 for another random type with $S_{ij} \in [-10, 10]$. These figures represent typical results. While one might say that F_1 is acceptable for estimating the dynamic aperture, we can show that it is not the optimal type. For now let us present two more generator types that give better results.

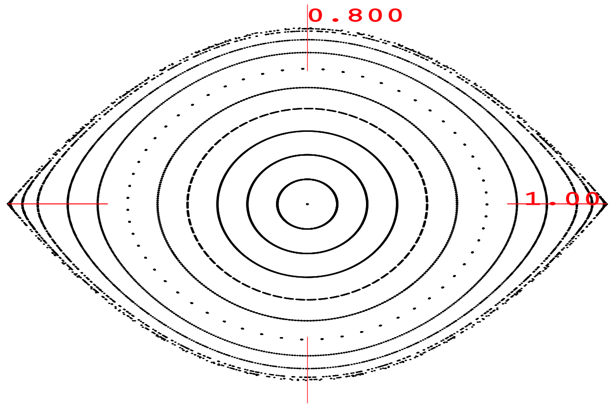

Figure 1: 1000 turn tracking of the anharmonic oscillator with the 19-th order Taylor map

These are the generators associated to $S = 0$, and to

$$S_b = \begin{pmatrix} 0.1 & 0 \\ 0 & -0.4 \end{pmatrix}. \tag{130}$$

The results are displayed in Figure 6. Comparing the various symplectic trackings with the order 19 Taylor map, we see that apparently the generator based on S_b is the best one, followed closely by the type associated to $S = 0$. Notice that the separatrix is very well reproduced. Also, the tunes are predicted accurately over a large phase space region. Therefore, at least for this example, order 7 it seems to be enough to estimate the dynamic aperture, if we use the best type of generating function symplectification.

6.2. A Neutrino Factory Lattice

Previous work exposed a variety of nonlinear effects in the lattice described in [26], of the proposed Neutrino Factory. Nonlinearities are due to the so-called kinematic effect, fringe fields, small circumference and large aperture [12, 46, 11]. The muons lifetime is less than 1000 turns. In spite of such a short tracking time, it is still interesting to see how the generating function symplectification method works in a case of practical interest, where nonlinearities play an important role. We computed order 8 maps of several realizations of the Neutrino Factory. Here we

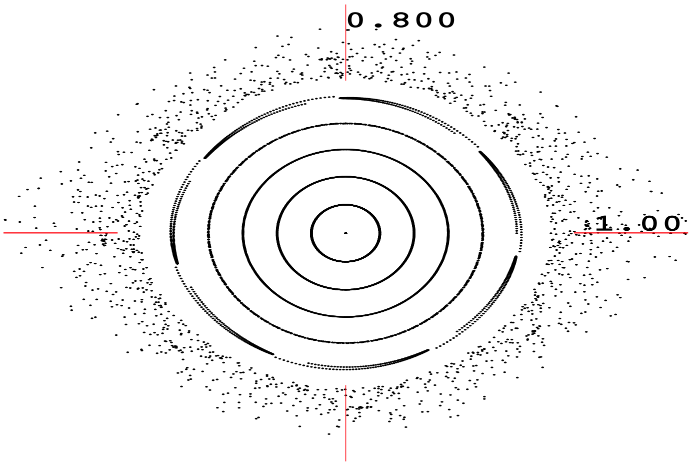

Figure 2: 1000 turn tracking of the anharmonic oscillator with the 7-th order Taylor map

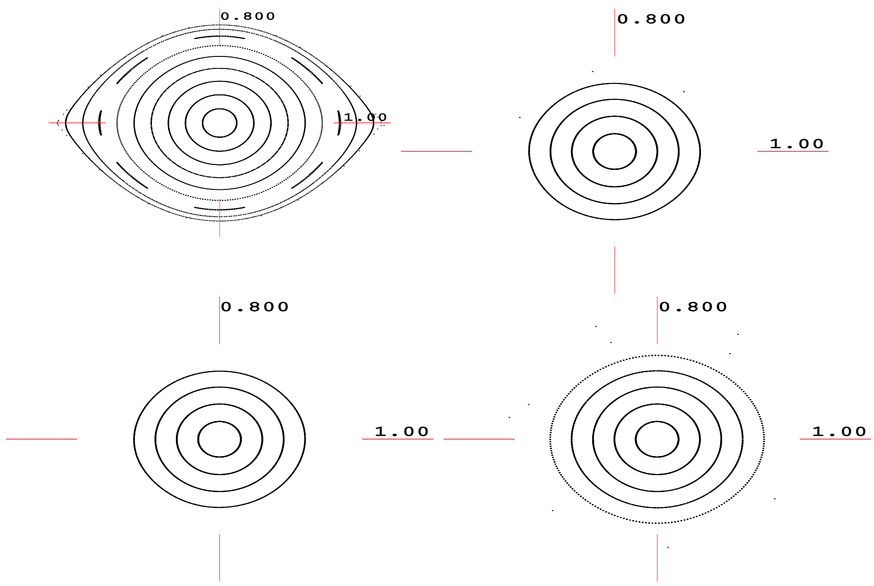

Figure 3: 1000 turn symplectic tracking of the anharmonic oscillator with the conventional generating functions (F_1 through F_4)

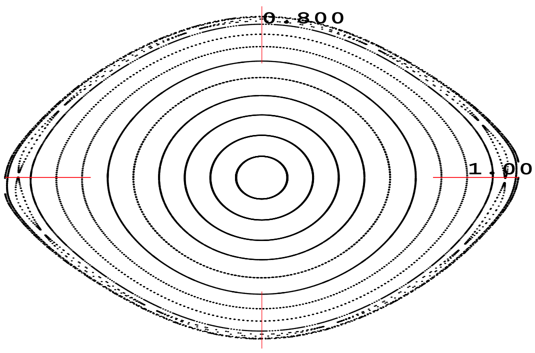

Figure 4: Typical 1000 turn symplectic tracking of the anharmonic oscillator with a generator type obtained from random symmetric matrices with entries in $[-1, 1]$

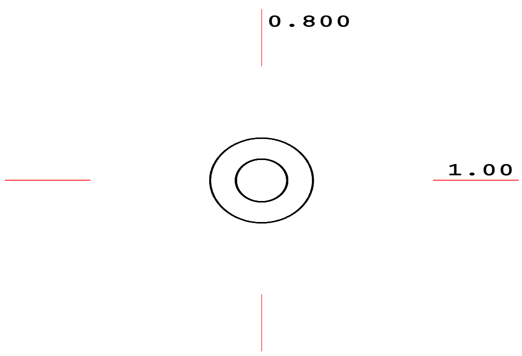

Figure 5: Typical 1000 turn symplectic tracking of the anharmonic oscillator with a generator type obtained from random symmetric matrices with entries in $[-10, 10]$

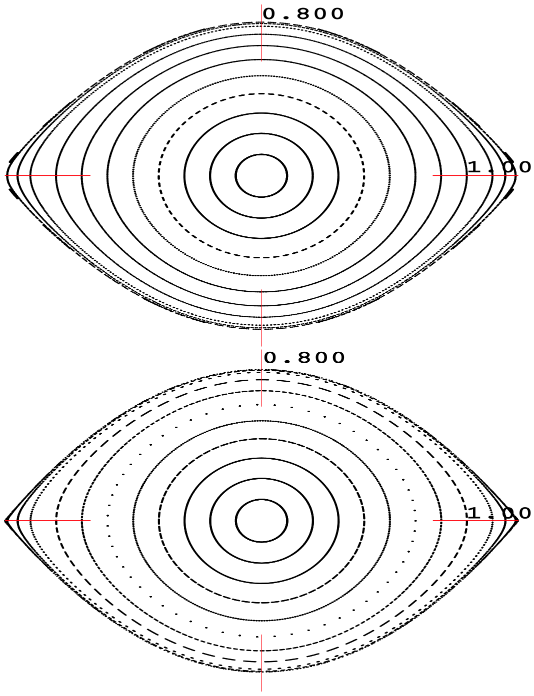

Figure 6: 1000 turn symplectic tracking of the anharmonic oscillator with the two best generator types: $S = 0$ on the left, and S_b on the right

use a weakly nonlinear one for which the 8-th order Taylor map looks quite accurate, and has a clearly defined 7-th order resonance structure. In the following we present the tracking pictures obtained from order 8 Taylor map tracking and symplectic tracking using different generator types. The Taylor map tracking is presented in Figure 7 and the symplectic trackings with conventional types in Figure 8. Contrary to the previous example, now F_4 gives a result that more closely resembles the truncated Taylor map. The other types give poor results. On the other hand, the type associated to $S = 0$ gives excellent results again, as seen in Figure 9. We can see that the resonance structure is preserved by the symplectified map, and we get a somewhat bigger dynamic aperture.

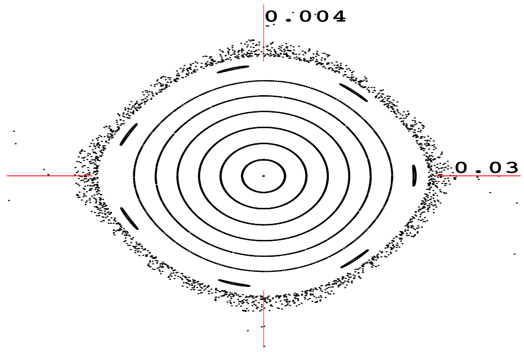

Figure 7: 1000 turn tracking of a lattice of the proposed Neutrino Factory with the 8-th order Taylor map

7. Conclusions

We developed the general theory of generating functions of canonical transformations. Using a modified definition of the generating function, we showed that there are many more generating functions than commonly known. The set of generating functions turned out to be very degenerate from the symplectification point of view. However, it was possible to reduce the pool of generating functions to equivalence classes associated with linear conformal symplectic maps. The remaining independent types were characterized by symmetric matrices. Also, we proved that by choosing appropriate types of generators, there is no advantage in factoring out linear parts and symplectifying only nonlinear parts, or first subjecting the map to be symplectified to a linear symplectic coordinate change. We illustrated the performance of this symplectification method by two examples. We showed that different generator types often give significantly different long term behavior of the symplectified maps. This fact points out the necessity for optimal generating function symplectification studies, which was solved in a very general setting based on Hofer metric [19].

Acknowledgments

This work was supported in part by the US Department of Energy Grant No. DE-FG02-95ER40931 and an Alfred P. Sloan Foundation Fellowship.

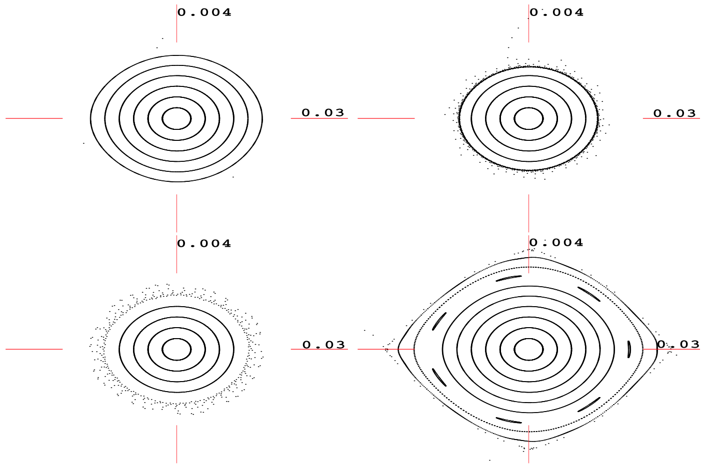

Figure 8: 1000 turn symplectic tracking of a lattice of the proposed Neutrino Factory with the conventional generating functions (F_1 through F_4)

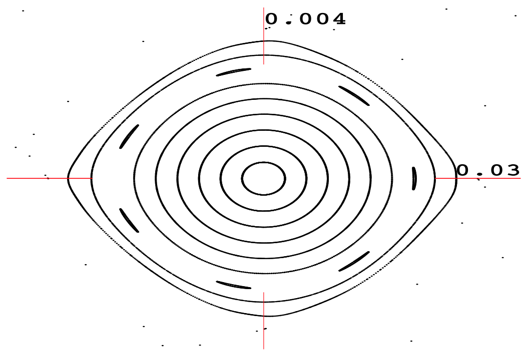

Figure 9: 1000 turn symplectic tracking of a lattice of the proposed Neutrino Factory with the generator type associated to $S = 0$

References

[1] D. Abell, Ph.D. Thesis, University of Maryland (1995).

[2] R. Abraham, J. Marsden, *Foundations of Mechanics*, Addison-Wesley, Reading, Second edition (1978).

[3] R. Abraham, J.E. Marsden, T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Springer-Verlag, New York (1988).

[4] J. Amiet, P. Huguenin, *Helv. Phys. Acta*, **53** (1980), 377.

[5] A. Banyaga, *The Structure of Classical Diffeomorphism Groups*, Dordrecht, Boston, Kluwer Academic (1997).

[6] G. Benettin, F. Fasso, *Appl. Numer. Math.*, **29**, No. 1 (1999), 73.

[7] G. Benettin, A. Giorgilli, *J. Stat. Phys.*, **74**, No. 5-6 (1994), 1117.

[8] M. Berz, *Particle Accelerators*, **24** (1989), 109.

[9] M. Berz, In: *Nonlinear Problems in Future Particle Accelerators*, World Scientific, Singapore, (1991), 288.

[10] M. Berz, *Modern Map Methods in Particle Beam Physics*, Academic Press, San Diego (1999).

[11] M. Berz, B. Erdelyi, K. Makino, *Phys. Rev. ST-Accel. Beams*, **3** (2000).

[12] M. Berz, K. Makino, B. Erdelyi, In: *Proceedings of HEMC'99, Workshop on Muon Colliders at Highest Energies*, Montauk, Long Island (1999).

[13] P. Channell, *Technical Report*, Internal Report AT-6, ATN-83-9, Los Alamos National Laboratory, Unpublished.

[14] P. Channell, J. Scovel, *Nonlinearity*, **3** (1990), 231.

[15] A.J. Dragt, In: *1981 Fermilab Summer School*, AIP Conference Proceedings, Volume **87**, New York (1982).

[16] A.J. Dragt, L.M. Healy, F. Neri, R. Ryne, *IEEE Transactions on Nuclear Science*, **NS-3,5** (1985), 2311.

[17] A. Dragt, D. Abell, *Fields Institute Communications*, **10** (1996), 59.

[18] B. Erdelyi, Ph.D. Thesis in preparation.

[19] B. Erdelyi, M. Berz, *Phys. Rev. Lett.*, **87**, No. 11 (2001), 114302.

[20] K. Feng, *J. Comp. Math.*, **16**, No. 6 (1998), 481.

[21] K. Feng, H. Wu, M. Quin, D. Wang, *J. Comp. Math.* **7**, No. 1 (1989), 71.

[22] Z. Ge, J. Marsden, *Phys. Lett. A*, **133** (1988), 134.

[23] I. Gjaja, A. Dragt, D. Abell, *IOP Conf. Ser.*, **131** (1993), 173.

[24] H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, MA (1980).

[25] H. Hofer, *Proc. Royal Soc. Edinburgh, Sect. A*, **115** (1990), 25.

[26] C. Johnstone, In: *Proceedings of HEMC'99, Workshop on Muon Colliders at Highest Energies*, Montauk, Long Island (1999).

[27] B. Leimkuhler, S. Reich, R.D. Skeel, In: *IMA Summer Program: Molecular Biology*, (Ed-s: K. Schulten, J. Mesirov), Springer-Verlag, New York (1994).

[28] K. Makino, M. Berz, *Nuclear Instruments and Methods*, **A427** (1999), 338.

[29] A. Ozorio de Almeida, *Proc. R. Soc. Lond. A*, **431** (1990), 403.

[30] H. Poincare, *New Methods of Celestial Mechanics*, Volume **3**, Chapter XXVIII, American Institute of Physics, New York (1893/1993).

[31] J. Sanz-Serna, M. Calvo, *Numerical Hamiltonian Problems*, Chapman and Hall, London (1994).

[32] M. Sewell, I. Roulstone, *Phil. Trans. R. Soc. Lond. A*, **345** (1993), 577.

[33] Z. Shang, *Numer. Math.*, **83** (1999), 477.

[34] J. Shi, P. Suwannakoon, *Phys. Rev. E*, **58**, No. 6 (1998), 7868.

[35] C. Siegel, *Amer. J. Math.*, **65**, No. 1 (1943).

[36] M. Suzuki, *Physica A*, **205** (1994), 65.

[37] R. deVogelaere, *Technical Report*, No. 4, Department of Mathematics, Univ. Notre Dame, Unpublished.

[38] R. Warnock, J. Ellison, *Appl. Numer. Math.*, **29** (1999), 89.

[39] A. Weinstein, *Advances in Math.*, **6** (1971), 329.

[40] A. Weinstein, *Invent. Math.*, **16** (1972), 203.

[41] A. Weinstein, *Ann. of Math.*, **98**, No. 2 (1973), 377.

[42] A. Weinstein, *Invent. Math.*, **20** (1973), 47.

[43] A. Weinstein, *Lectures on Symplectic Manifolds*, CBMS Reg. Conf. Series in Math., **29**, AMS, Providence, RI (1977).

[44] J. Wisdom, M. Holman, J. Touma, *Fields Institute Communications*, **10** (1996), 217.

[45] Y. Yan, P. Channell, M. Syphers, *Technical Report*, SSC Laboratory Report SSCL-157, Unpublished.

[46] F. Zimmermann, et al, *Technical Report* No. SL-2000-011 AP and NUFACT-NOTE 21, CERN, Unpublished.

A. Application to the Conventional Generating Function Types

In this appendix we present briefly how the conventional generating function types fit into the general theory framework. Because of (121), without loss of generality we can assume that the symplectic maps have identity as linear part. In canonical coordinates $(\vec{q}, \vec{p})$ an origin preserving symplectic map acts as

$$\mathcal{M}\begin{pmatrix} \vec{q} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{Q} \\ \vec{P} \end{pmatrix}. \tag{131}$$

The conventional type 1 (F_1) is determined by

$$(\nabla F_1)^T \begin{pmatrix} \vec{q} \\ \vec{Q} \end{pmatrix} = \begin{pmatrix} \vec{p} \\ -\vec{P} \end{pmatrix}. \tag{132}$$

It is straightforward to show that this type is associated with a linear α, namely (split into $n \times n$ blocks)

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} 0 & 0 & 0 & I \\ 0 & -I & 0 & 0 \\ 0 & 0 & I & 0 \\ I & 0 & 0 & 0 \end{pmatrix}. \tag{133}$$

However, it does not satisfy condition (15) for the local existence of the generating function, because $\det(C \cdot I + D) = 0$. Therefore, it does not belong to any equivalence class for symplectic maps having identity as linear part.

On the other hand, the second conventional type (F_2), determined by

$$(\nabla F_2)^T \begin{pmatrix} \vec{q} \\ \vec{P} \end{pmatrix} = \begin{pmatrix} \vec{p} \\ \vec{Q} \end{pmatrix}, \tag{134}$$

is associated with

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} 0 & 0 & 0 & I \\ I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \end{pmatrix}. \tag{135}$$

The transversality condition is satisfied, and can be easily checked from (121) that it belongs to the class represented by

$$S_2 = -\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}. \tag{136}$$

Similarly, the conventional type three (F_3), determined by

$$(\nabla F_3)^T \begin{pmatrix} \vec{p} \\ \vec{Q} \end{pmatrix} = \begin{pmatrix} -\vec{q} \\ -\vec{P} \end{pmatrix}, \tag{137}$$

is associated with

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} 0 & 0 & -I & 0 \\ 0 & -I & 0 & 0 \\ 0 & 0 & 0 & I \\ I & 0 & 0 & 0 \end{pmatrix}, \tag{138}$$

and belongs to the class

$$S_3 = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \tag{139}$$

differing only by a sign from the type F_2

$$S_3 = -S_2. \tag{140}$$

Finally, the conventional type four (F_4) is determined by

$$(\nabla F_4)^T \begin{pmatrix} \vec{p} \\ \vec{P} \end{pmatrix} = \begin{pmatrix} -\vec{q} \\ \vec{Q} \end{pmatrix}, \tag{141}$$

and is associated with

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} 0 & 0 & -I & 0 \\ I & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ 0 & I & 0 & 0 \end{pmatrix}, \tag{142}$$

The transversality condition being violated, it does not belong to any class.

Thus, we recovered from the general theory the well-known facts that F_1 and F_4 cannot, while F_2 and F_3 can be used to represent at least locally symplectic maps having identity as linear part. Also, we identified the equivalence classes which F_2 and F_3 belong to. The only difference if the symplectic maps do not have identity as linear part is that we obtain different symmetric matrices, and hence classes, which also can be computed using (121).

B. Equivalence of Symplectification Procedures with and without Linear Part

We write the symplectic map to be symplectified as

$$\mathcal{M} = M + H, \tag{143}$$

where M is the linear part, and H the higher order terms. We can distinguish three symplectification procedures: symplectify $\mathcal{M}_n$ directly, symplectify $\mathcal{M}_{L,n}$ obtained from

$$\mathcal{M}_L = I + M^{-1} \circ H, \tag{144}$$

or symplectify $\mathcal{M}_{R,n}$ obtained from

$$\mathcal{M}_R = I + H \circ M^{-1}. \tag{145}$$

In the latter two cases we first factored out the linear part from the left and right respectively. The relations among the maps are the following:

$$\mathcal{M} = M \circ \mathcal{M}_L, \tag{146}$$

$$\mathcal{M} = \mathcal{M}_R \circ M. \tag{147}$$

The question is whether these relations continue to hold for the symplectified versions of $\mathcal{M}_n$, $\mathcal{M}_{L,n}$, and $\mathcal{M}_{R,n}$. Suppose we symplectify the maps using a generator of type α with

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{148}$$

The local existence conditions are

$$\det\,(CM + D) \neq 0 \tag{149}$$

in the first case, and

$$\det\,(C + D) \neq 0 \tag{150}$$

in the second and third case. Being linear, we use M for the Jacobian of M too. Equations (149) and (150) are not compatible in general. Thus in general not every type of generating function exists in all three cases. The right question to ask is the following. Suppose one uses some type of generator to symplectify a given map, using the approach of one of the three cases. Then are there other types of generators which produce the same symplectification for the other two cases? In other words, we would like to find the appropriate type of generators such that relations (146) and (147) hold for the symplectified maps.

To this end, any generator for the first case is associated to one of the following:

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} -JM^{-1} & J \\ \frac{1}{2}\,(I + JS)\,M^{-1} & \frac{1}{2}\,(I - JS) \end{pmatrix}. \tag{151}$$

Its inverse is given by

$$(\mathrm{Jac}\,(\alpha))^{-1} = \begin{pmatrix} \frac{1}{2}MJ\,(I - SJ) & M \\ -\frac{1}{2}J\,(I + SJ) & I \end{pmatrix}. \tag{152}$$

Denoting the generator by $F_{[S]}$, $\mathrm{Jac}\left(\nabla F_{[S]}\right)^T =_n N_{[S]}$, and the symplectification of $\mathcal{M}_n$ by $\mathcal{M}_{[S]}$ we obtain

$$\begin{aligned}\mathrm{Jac}\left(\mathcal{M}_{[S]}\right) &= \left[\begin{pmatrix} \frac{1}{2}MJ\left(I-SJ\right) & M \end{pmatrix}\begin{pmatrix} N_{[S]} \\ I \end{pmatrix}\right] \\ &\quad \cdot \left[\begin{pmatrix} -\frac{1}{2}J\left(I+SJ\right) & I \end{pmatrix}\begin{pmatrix} N_{[S]} \\ I \end{pmatrix}\right]^{-1} \\ &= M\cdot\left(\frac{1}{2}J\left(I-SJ\right)\cdot N_{[S]}+I\right)\cdot\left(-\frac{1}{2}J\left(I+SJ\right)\cdot N_{[S]}+I\right)^{-1}.\end{aligned} \tag{153}$$

Now we turn our attention to the second case. Here the possible generators belong to one of the following classes:

$$\mathrm{Jac}\left(\beta\right) = \begin{pmatrix} -J & J \\ \frac{1}{2}\left(I+J\bar{S}\right) & \frac{1}{2}\left(I-J\bar{S}\right) \end{pmatrix}. \tag{154}$$

Clearly its inverse is

$$\left(\mathrm{Jac}\left(\beta\right)\right)^{-1} = \begin{pmatrix} \frac{1}{2}J\left(I-\bar{S}J\right) & I \\ -\frac{1}{2}J\left(I+\bar{S}J\right) & I \end{pmatrix}. \tag{155}$$

Again, denoting the symplectified version of $\mathcal{M}_{L,n}$ by $\mathcal{M}_{L[\bar{S}]}$ we obtain

$$\begin{aligned}\mathrm{Jac}\left(\mathcal{M}_{L[\bar{S}]}\right) &= \left[\begin{pmatrix} \frac{1}{2}J\left(I-\bar{S}J\right) & I \end{pmatrix}\begin{pmatrix} N_{[\bar{S}]} \\ I \end{pmatrix}\right] \\ &\quad \cdot \left[\begin{pmatrix} -\frac{1}{2}J\left(I+\bar{S}J\right) & I \end{pmatrix}\begin{pmatrix} N_{[\bar{S}]} \\ I \end{pmatrix}\right]^{-1} \\ &= \left(\frac{1}{2}J\left(I-\bar{S}J\right)\cdot N_{[\bar{S}]}+I\right)\cdot\left(-\frac{1}{2}J\left(I+\bar{S}J\right)\cdot N_{[\bar{S}]}+I\right)^{-1},\end{aligned} \tag{156}$$

where we used the notation $\mathrm{Jac}\left(\nabla F_{[\bar{S}]}\right)^T =_n N_{[\bar{S}]}$. Next, we use the transformation property (62) with $\mathcal{K} = M^{-1}$. It follows that

$$F_{\beta,\mathcal{M}_L} = F_{\beta\circ T_{\mathcal{K}},\mathcal{M}}. \tag{157}$$

Also notice that $\beta\circ T_{\mathcal{K}} = \alpha$ if and only if $\bar{S} = S$. In this case

$$N_{[\bar{S}]} = N_{[S]}. \tag{158}$$

Comparing equations (153) and (156), and using (158) we can conclude that

$$\mathcal{M}_{[S]} = M \circ \mathcal{M}_{L[\bar{S}]}, \tag{159}$$

if and only if

$$\bar{S} = S. \tag{160}$$

This proves that the symplectified version of (146) is equation (159), and holds only if (160) is satisfied.

We can proceed to the third case and follow the same route. To symplectify $\mathcal{M}_{R,n}$ we choose a generator type from the pool

$$\mathrm{Jac}\,(\gamma) = \begin{pmatrix} -J & J \\ \frac{1}{2}\left(I + J\tilde{S}\right) & \frac{1}{2}\left(I - J\tilde{S}\right) \end{pmatrix}, \tag{161}$$

with inverse

$$(\mathrm{Jac}\,(\gamma))^{-1} = \begin{pmatrix} \frac{1}{2}J\left(I - \tilde{S}J\right) & I \\ -\frac{1}{2}J\left(I + \tilde{S}J\right) & I \end{pmatrix}. \tag{162}$$

If we denote the symplectification of $\mathcal{M}_{R,n}$ by $\mathcal{M}_{R[\tilde{S}]}$ we get

$$\begin{aligned} \mathrm{Jac}\left(\mathcal{M}_{R[\tilde{S}]}\right) &= \left[\begin{pmatrix} \frac{1}{2}J\left(I - \tilde{S}J\right) & I \end{pmatrix}\begin{pmatrix} N_{[\tilde{S}]} \\ I \end{pmatrix}\right] \\ &\quad \cdot \left[\begin{pmatrix} -\frac{1}{2}J\left(I + \tilde{S}J\right) & I \end{pmatrix}\begin{pmatrix} N_{[\tilde{S}]} \\ I \end{pmatrix}\right]^{-1} \\ &= \left(\frac{1}{2}J\left(I - \tilde{S}J\right) \cdot N_{[\tilde{S}]} + I\right) \cdot \left(-\frac{1}{2}J\left(I + \tilde{S}J\right) \cdot N_{[\tilde{S}]} + I\right)^{-1}, \end{aligned} \tag{163}$$

using the notation $\mathrm{Jac}\left(\nabla F_{[\tilde{S}]}\right)^T =_n N_{[\tilde{S}]}$. Now using the transformation rule (60) with $\mathcal{A} = M^{-1}$ we obtain

$$F_{\gamma,\mathcal{M}_R} = F_{\gamma\circ T_{\mathcal{A}},\mathcal{M}}. \tag{164}$$

A straightforward calculation gives that

$$\mathrm{Jac}\,(\gamma \circ T_{\mathcal{A}}) = \begin{pmatrix} -J & JM \\ \frac{1}{2}\left(I + J\tilde{S}\right) & \frac{1}{2}\left(I - J\tilde{S}\right)M \end{pmatrix}, \tag{165}$$

with inverse

$$(\mathrm{Jac}\,(\gamma \circ T_{\mathcal{A}}))^{-1} = \begin{pmatrix} \frac{1}{2}J\left(I - \tilde{S}J\right) & I \\ -\frac{1}{2}M^{-1}J\left(I + \tilde{S}J\right) & M^{-1} \end{pmatrix}. \tag{166}$$

Then, equation (163) can be expressed as

$$\mathrm{Jac}\left(\mathcal{M}_{R[\tilde{S}]}\right) = X \cdot M^{-1}, \tag{167}$$

where we introduced the notation

$$X = \left(\frac{1}{2}J\left(I - \tilde{S}J\right) \cdot \nabla F_{\gamma \circ T_{\mathcal{A}}, \mathcal{M}} + I\right) \cdot \left(-\frac{1}{2}M^{-1}J\left(I + \tilde{S}J\right) \cdot \nabla F_{\gamma \circ T_{\mathcal{A}}, \mathcal{M}} + M^{-1}\right)^{-1}. \tag{168}$$

But as one can see from (166) this is nothing else than the Jacobian of the symplectified map obtained from $\mathcal{M}_n$ and generator of type (165). As shown in Section 4 this generator type for $\mathcal{M}_n$ belongs to the equivalence class represented by the symmetric matrix calculated using formula (121). A short calculation gives the result,

$$S = M^T \tilde{S} M. \tag{169}$$

Combining this result with equations (167) and (168) we can conclude that

$$\mathcal{M}_{[S]} = \mathcal{M}_{L[\tilde{S}]} \circ M. \tag{170}$$

This proves that the symplectified version of (147) is equation (170), and holds only if (169) is satisfied. Therefore, the main result of this appendix can be formulated as the following theorem.

Theorem 7. *The symplectified version of (146), i.e. equation (159), holds if and only if (160) is satisfied, and the symplectified version of (147), i.e. equation (170), holds if and only if (169) is satisfied.*

The main point we learned from this appendix is that from the optimal symplectification point of view there is no difference which way one proceeds. Once we obtained the best type of generator for one case, the best generators for the other cases automatically follow from (160) and (169). Therefore, there is nothing to be gained by factoring out linear parts and symplectifying the nonlinear parts only. Moreover, the first case (without factorization) is the most efficient when implemented numerically on a computer.

C. Equivalence in the Case of Symplectic Maps Conjugated by Linear Symplectic Maps

Combining the left and right factorizations in linear and nonlinear parts just discussed, we can address the special case of the linear symplectic change of variables. From equations (146) and (147) we can infer that

$$\mathcal{M}_R = M \circ \mathcal{M}_L \circ M^{-1}, \tag{171}$$

and from (159) and (170) that

$$\mathcal{M}_{R[\tilde{S}]} = M \circ \mathcal{M}_{L[\bar{S}]} \circ M^{-1}, \tag{172}$$

if

$$\bar{S} = M^T \tilde{S} M. \tag{173}$$

The two maps are conjugated by a linear symplectic transformation. However, this case is very special, since both $\mathcal{M}_R$ and $\mathcal{M}_L$ are obtained from the same map $\mathcal{M}$. We could relax the conditions, and ask if any two symplectic maps are conjugated by an arbitrary linear symplectic map, then the same holds true for their symplecified counterparts. That is, suppose that $\mathcal{M}$ and $\mathcal{N}$ are symplectic maps such that

$$\mathcal{N} = \mathcal{K} \circ \mathcal{M} \circ \mathcal{K}^{-1}, \tag{174}$$

for some linear symplectic $\mathcal{K}$. The possible types of generating functions are associated with

$$\mathrm{Jac}\,(\alpha) = \begin{pmatrix} -JKM^{-1}K^{-1} & J \\ \frac{1}{2}\,(I + JS)\,KM^{-1}K^{-1} & \frac{1}{2}\,(I - JS) \end{pmatrix}, \tag{175}$$

for $\mathcal{N}_n$, and

$$\mathrm{Jac}\,(\beta) = \begin{pmatrix} -JM^{-1} & J \\ \frac{1}{2}\left(I + J\bar{S}\right)M^{-1} & \frac{1}{2}\left(I - J\bar{S}\right) \end{pmatrix}, \tag{176}$$

for $\mathcal{M}_n$. As before, we denoted $\mathcal{M} = M + H$ and $\mathrm{Jac}\,(\mathcal{K}) = K$. The transformation rule to be used here is

$$F_{\alpha,\mathcal{N}} = F_{\alpha \circ T_{\mathcal{K}},\mathcal{M}}, \tag{177}$$

where

$$\mathrm{Jac}\,(T_{\mathcal{K}}) = \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix}. \tag{178}$$

Following the now well established procedures of Appendix B we proved that also

$$\mathcal{N}_{[S]} = \mathcal{K} \circ \mathcal{M}_{[\bar{S}]} \circ \mathcal{K}^{-1} \tag{179}$$

holds for the symplectified maps if

$$\bar{S} = K^T S K. \tag{180}$$

The details are left to the reader. Therefore, we gain no additional freedom in the symplectification process if we first subject the symplectic map to a linear symplectic variable change.