Lie Perturbation Theory

8. The Geometry and the Perturbing Equations

As is usual in perturbation theory, we begin with some integrable problem, with some Hamiltonian H_0, which we may think of in terms of flows, $\varphi_{_H}$, with parameter t, generated by the vector field $\widetilde{\Delta}_0$; and also the complete problem, with Hamiltonian H, **which is also assumed integrable**, generated by the yet-to-be-determined vector field $d/dt = \widetilde{\Delta}$:

$$H = H_0 + \epsilon H_1 \ , \qquad \text{action-angle variables for} \quad \begin{cases} H_0 \text{ defined as } \xi = (\phi_0, J_0) \ , \\ H \text{ defined as } \eta = (\phi, J) = \eta(\xi, t; \epsilon) \ . \end{cases} \tag{8.1a}$$

Since, for each value of ϵ, the variables ξ^α and the variables η^μ are canonical variables on the phasespace, there is a canonical transformation between them; i.e., as ϵ varies, we have a continuous, 1-parameter family of canonical transformations. Therefore, we may say that

$$\widetilde{\Delta} = \frac{d}{dt} = \frac{dq^a}{dt}\partial_{q^a} + \frac{dp_b}{dt}\partial_{p_b} + \frac{\partial}{\partial t} = \frac{\partial H}{\partial p_a}\partial_{q^a} - \frac{\partial H}{\partial q^a}\partial_{p_b} + \frac{\partial}{\partial t} \equiv \widetilde{X}_{_H} + \frac{\partial}{\partial t} = \{\bullet, H\} + \frac{\partial}{\partial t} \tag{8.1b}$$

is a vector field, with parameter t, in the tangent bundle to T^*Q, generated by the Hamiltonian H. Movement along the flow of this vector field corresponds to movement from some initial condition of the system, say ζ, equal to what one may call the initial condition, to some later condition, $\xi = \xi(\zeta, t)$, at time, t. Referring to this flow as the mapping ϕ_h, we may consider this mapping, for some value of t, as giving us $\xi = \xi(\zeta, t) = \phi_h(\zeta)$, where on the right-hand side the parameter value along the flow, i.e., t, has simply been suppressed (for ease of notation).

In the same way, we may say that there is a vector field in the tangent bundle, $d/d\epsilon$, with parameter ϵ, and movement from the initial value of $\epsilon = 0$ to some larger value of ϵ constitutes a "motion" from the unperturbed system to the perturbed one, with that value of ϵ as the coupling between them. Analogous to the Hamiltonian, we will take the **Lie generating function** of this motion as G, i.e., the infinitesimal generator of this family of canonical transformations, and realize that the vector $d/d\epsilon$ is just the same as the Hamiltonian vector field made from G; i.e., $\widetilde{X}_G = d/d\epsilon$ is analogous to $\widetilde{\Delta} = \widetilde{X}_H$. Therefore it is also true that we may write

$$\frac{d}{d\epsilon} = \frac{\partial G}{\partial p_a}\partial_{q^a} - \frac{\partial G}{\partial q^a}\partial_{p_b} + \frac{\partial}{\partial \epsilon} \equiv \widetilde{X}_{_G} + \frac{\partial}{\partial \epsilon} = \{\bullet, G\} + \frac{\partial}{\partial \epsilon} \ . \tag{8.2}$$

Beginning at some particular point in phasespace corresponding to the state of the system at, for instance, the initial time, and so the point with coordinates ζ in phasespace, we may move along a flow in the ϵ-direction, i.e., along the curves with tangent vector given by $\widetilde{X}_{_G}$, a parameter value ϵ away from 0, and come to the situation for the perturbed system, at that same time. Therefore, in an analogous way to the time development, we may suppose that we again begin at the initial time, and move forward some parameter value ϵ along this flow, which we denote by ϕ_g. We could write $\eta(\xi, 0; \epsilon) = \phi_g \zeta$ or, more generally, $\eta(\xi, t; \epsilon) = \phi_g \xi(\zeta, t)$. We conceive of this flow as a mapping of an initial neighborhood, $U_0 \in T^*Q$ to a final neighborhood, $U_\epsilon \in T^*Q$, and label the flow by $\phi_g : U_0 \to U_\epsilon$. The map ϕ_g maps points of the manifold into other points on the manifold. We, however, are dealing with functions defined over the manifold; indeed, even the way we label those points, via their coordinates, are in fact functions of those points. Therefore, we need now to have a way to "drag" those functions along when we follow along these curves; the standard way to do this is via a so-called *pullback of a map between manifolds*: Given a map between manifolds, or, as in this case, between different neighborhoods of the same manifold, we may define a map between the functions over those manifolds. We label those neighborhoods as U_0 and U_ϵ, and the space of (smooth) functions over those neighborhoods by $\mathcal{F}_0$ and $\mathcal{F}_\epsilon$:

$$\begin{aligned} &\phi_g : U_0 \rightarrow U_\epsilon \quad \text{i.e., for } \ P \in U_0\,, \quad \phi_g(P) = Q \in U_\epsilon\ ; \\ \Longrightarrow \quad &\phi_g^* : \mathcal{F}_\epsilon \rightarrow \mathcal{F}_0 \quad \text{i.e., for } \ f \in \mathcal{F}_\epsilon\,,\ (\phi_g^* f)(\xi) \equiv f[\phi_g(\xi)] = (f \circ \phi_g)(\xi) = f(\eta)\ , \end{aligned} \tag{8.3a}$$

where the $\circ$ denotes the convolution of functions. However, since our maps are canonical transformations, and therefore surely invertible, we may also involve the inverse flow:

$$\begin{aligned} \phi_g^{-1} : U_\epsilon \rightarrow U_0\,, \quad \text{i.e., } \ \xi \equiv \phi_g^{-1}(\eta)\ ; \\ \text{for } \ j \in \mathcal{F}_0\,,\ [(\phi_g^{-1})^* j](\eta) \ = (j \circ \phi_g^{-1})(\eta) = j(\xi)\ . \end{aligned} \tag{8.3b}$$

It should be noted that our ϕ_g^* is the operator that José refers to as $\mathcal{U}^{-1}$, and the same as the one that Lichtenberg refers to as T; as well, therefore, José's ϕ^G, which maps points with coordinates ξ into points with coordinates η, is the same as our ϕ_g. The equations just above have the same content as José's Eqs. (6.120-1).

We therefore now try to consider how these things vary as we vary ϵ. In particular, let us consider the following equation picked out of those above, which is the same as Jose's Eq. (6.121), and then differentiate it with respect to ϵ, assuming that we have a function $F = F(\eta, t; \epsilon)$ that depends on ϵ explicitly and also through η:

$$\begin{aligned} &\text{for } \ F \in \mathcal{F}_\epsilon\,, \quad F(\eta) = (\phi_g^* F)(\xi) \\ &\qquad \Longrightarrow \quad \frac{dF(\eta)}{d\epsilon} = [\frac{\partial \phi_g^*}{\partial \epsilon} F](\xi) + [\phi_g^* \frac{\partial F}{\partial \epsilon}](\xi) \ = \ (\phi_g^{-1})^* [\frac{\partial \phi_g^*}{\partial \epsilon} F](\eta) + \frac{\partial F}{\partial \epsilon}(\eta)\ , \end{aligned} \tag{8.4a}$$

where the last equality is required by the desire to express both sides of the equation as functions of the same variables, i.e., to define both sides as functions on the same neighborhood of the manifold. (To get the first term after the last equal sign we pulled back our composite function again.)

However, since G is the generating function for the flows that have ϵ as parameter, we may apply Eq. (8.2) to our function $F = F(\eta; \epsilon)$, remembering that it is also an explicit function of ϵ, but suppressing that in the notation, we have just the standard form involving Poisson brackets:

$$\frac{d}{d\epsilon} F(\eta) = \{F(\eta), G\} + \frac{\partial F}{\partial \epsilon}(\eta)\ . \tag{8.4b}$$

Comparing Eqs. (8.4a) and (8.4b) we obtain a very interesting and useful equation relating the derivative of the induced maps with the Poisson bracket taken with the generating function:

$$(\phi_g^{-1})^* [\frac{\partial \phi_g^*}{\partial \epsilon} F](\eta) = \{F(\eta), G\} = \widetilde{X}_G F(\eta) = \mathcal{L}_{\widetilde{X}_G} F(\eta)\ . \tag{8.5a}$$

Here the action of $\widetilde{X}_G$ on $F \in \mathcal{F}_\epsilon$ has also been written as the (equivalent) action of the Lie derivative in the direction of $\widetilde{X}_G$ on F. When acting on functions these two are equivalent; however, the Lie derivative is a more general object that is allowed to act on the elements of any sort of space of tensor fields; this will be useful in more advanced situations, and is a language that is incredibly common in this area of inquiry.

Factoring out the arbitrary function F, and the variables on which it depends, which live in U_ϵ, we have a very useful equality between our important operators, valid as an operator equation to act in any of the associated tensor spaces associated with U_ϵ:

$$\text{for } \; G \in \mathcal{F}_\epsilon \; \text{ and } \; \widetilde{X}_G \in \mathcal{T}_\epsilon \quad (\phi_g^{-1})^* \frac{\partial \phi_g^*}{\partial \epsilon} = \widetilde{X}_G = \mathcal{L}_{\widetilde{X}_G} = -\frac{\partial (\phi_g^{-1})^*}{\partial \epsilon} \phi_g^* \,, \tag{8.5b}$$

where the last equality follows because the flow forward followed by the flow backward simply gives the identity map. It is worth pointing out that this reasonably lengthy derivation has given us a very reasonable and expected answer, were we to have put in rather more time into worrying with the underlying differential geometry of the problem. By re-writing this equation in the more direct form, we may formally integrate the equation—in the case when G does not depend on ϵ—to obtain a different way to visualize our pullback mapping:

$$\frac{\partial \phi_g^*}{\partial \epsilon} = \phi_g^* \mathcal{L}_{\widetilde{X}_G} \quad \Longrightarrow \quad \phi_g^* = e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} \equiv \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} (\mathcal{L}_{\widetilde{X}_G})^n \,. \tag{8.5c}$$

Such a presentation is often referred to as a **Lie series**, and is a fairly standard way to implement explicit algebraic calculations of presentations of such mappings. One can also note that the 1-parameter flow, of the points on the manifold themselves, along the vector field $\widetilde{X}_G$, a parameter distance ϵ is often expressed in the form

$$\phi_g = e^{\epsilon \widetilde{X}_G} \;\Longrightarrow\; \frac{\partial \phi_g}{\partial \epsilon} = +\widetilde{X}_G \, e^{\epsilon \widetilde{X}_G} = \widetilde{X}_G \phi_g \,, \tag{8.5d}$$

so that our earlier equation, Eq. (8.5b), is simply the pullback of this standard form for the behavior of the generators of 1-parameter mappings, and the replacement of the vector field itself by the Lie derivative in that direction allows us to apply the formula on any tensor space, since, also, our manifold mapping is invertible. Lastly, it is worth noting that José's $\mathcal{P}_G$ would seem to me to be just the same as our $\widetilde{X}_G$, which we are now writing, sometimes, in the form of the Lie derivative in that direction. However, our equation Eq. (8.5b) is not his Eq. (6.125), which distinction we will try to understand below. On the other this operator is in fact equal to Lichtenberg's operator $-L_w$; i.e., $\widetilde{X}_G f = \{f, G\} = -\{G, f\} \equiv -L_G f$, since his w is the same as our G. The reason for the difference in sign will also be discussed below.

One reason for the common use of Lie series is the fact that they agree so very well with the properties of the symplectic form, i.e., with the Poisson bracket:

$$\begin{aligned} \mathcal{L}_{\widetilde{X}_G} f \; &\equiv \{f, G\} \,, \\ \Longrightarrow \quad \mathcal{L}_{\widetilde{X}_G} \{f, j\} &\equiv \{\{f, j\}, G\} = \{f, \{j, G\}\} + \{j, \{G, f\}\} = \{f, \mathcal{L}_{\widetilde{X}_G} j\} + \{\mathcal{L}_{\widetilde{X}_G} f, j\} \,, \\ \left[\mathcal{L}_{\widetilde{X}_G}, \mathcal{L}_{\widetilde{X}_W}\right] f &= \{\{f, W\}, G\} \; - \{\{f, G\}, W\} = -\{f, \{G, W\}\} = \; -\mathcal{L}_{\widetilde{X}_{\{G,W\}}} \,, \end{aligned} \tag{8.6a}$$

where the Jacobi identity for the Poisson bracket has been used in each of the last two lines. The second equality tells us that the Lie derivative simply moves inside a Poisson bracket and acts like a derivation there on its parts! Because of this, again when G is independent of ϵ, we may exponentiate this relation, using Eq. (8.5c), to obtain the action of the flow on the Poisson bracket:

$$\phi_g^* \{f, j\} = \{\phi_g^* f, \, j\} + \{f, \, \phi_g^* j\} \,. \tag{8.6b}$$

This says that motion along the ϵ-flow preserves the Poisson bracket; therefore, in particular when we use the ϵ-flow (generated by G), or the presentation via the Lie series, on a set of canonical coordinates, the set of coordinates remains canonical. [When one uses $d/d\epsilon$ instead of $\mathcal{L}_{\widetilde{X}_G}$, i.e., when the generating function G depends explicitly on ϵ, this virtue of the equations is still true, although the corresponding formulae become rather more complicated in appearance. In that case the general method is often referred to as **Lie transforms**, rather than Lie series, and will be discussed somewhat later, following Deprit.]

On the other hand, the last line of Eqs. (8.6a) is somewhat distracting, because of the perhaps-unexpected minus sign it contains. Because of this some people use a different definition for the Lie derivative with respect to a Hamiltonian vector field. I will refer to this definition using Lichtenberg's form,

$$L_G \equiv \{G, \bullet\} \quad \Longrightarrow \quad L_G f \equiv \{G, f\} = -\{f, G\} \ , \\ \Longrightarrow \quad \left[L_G, L_W\right] = +L_{\{G,W\}} \ . \tag{8.7a}$$

As we can see the definition of L_G leads to a "more natural" relationship between the commutator of two of the Lie derivatives and the Poisson bracket of the two functions. In fact, we can see that when we have a space of functions that are provided with a skew "product" we can always arrange for that space of functions to actually be a Lie algebra. In this case of course the skew product is given by the symplectic form; therefore, we may say that the space of functions over a symplectic manifold is always a Lie algebra. with the product as the Poisson bracket, and the mapping L, which sends such functions into the tangent bundle is a realization of that algebra in terms of vector fields, while the mapping $\mathcal{L}$ is an *anti-realization* in terms of vector fields. Since all this is true, and since we have spent considerable time discussing how to pullback functions, it seems now very useful to also see how this realization is changed under our mapping ϕ_g.

We begin with the situation already described above in Eqs. (8.5b): we have $G \in \mathcal{F}_\epsilon$ and $\phi_g : U_0 \;\rightarrow\; U_\epsilon$, i.e., $\phi_g(\xi) = \eta$. This allowed us to consider the pullback $\phi_g^* : \mathcal{F}_\epsilon \;\rightarrow\; \mathcal{F}_0$. If we now consider an arbitrary tangent vector field, $\widetilde{Y} \in \mathcal{T}_0$, we may define the (associated) "push-forward" of it to the tangent vector field, $\phi_{g*} : \mathcal{T}_0 \;\rightarrow\; \mathcal{T}_\epsilon$. We do this by insisting that when the new tangent vector acts on a function there, evaluated at a point there, it should give the same result as if we had, instead, pulled back the function to the place where the old tangent vector lives, and evaluated it there:

$$\text{for} \;\; Y \in \mathcal{T}_0 \ , \;\; \text{and} \;\; f \in \mathcal{F}_\epsilon \ , \\ \{(\phi_{g*}\widetilde{Y})[f]\}(\eta) \equiv \{\widetilde{Y}[\phi_g^* f]\}(\phi_g^{-1}\eta) = (\phi_g^{-1})^*\{\widetilde{Y}[\phi_g^* f]\}(\eta) \ , \\ \text{or} \quad \phi_{g*}\widetilde{Y} = (\phi_g^{-1})^*\widetilde{Y}\,\phi_g^* \;\in\; \mathcal{T}_\epsilon \ , \tag{8.7b}$$

where the mappings involved in the expression involving $\widetilde{Y}$, on the right-hand side of the last equation, all make sense because of the action that tangent vectors are supposed to have, on functions, and the expression on the left-hand side of the equation is simply the definition involving the (induced) mapping, ϕ_{g*}, which maps tangent vector fields as brought along by the initial mapping, ϕ_g, of the manifolds.

A different way to express the same thing, but perhaps more useful in the present circumstances, is to note that since $\phi_{g*}\widetilde{Y}$ in the expression above is an element of the tangent space over U_ϵ, we may think of it as the special case when it is the Hamiltonian vector field for some $f \in \mathcal{F}_\epsilon$. In that case we have

$$\text{choose} \;\; (\phi_{g*})\widetilde{Y} = \widetilde{X}_f \quad \Longrightarrow \quad \widetilde{X}_f = (\phi_g^{-1})^*\widetilde{X}_{\phi_g^* f}\,\phi_g^* \quad \text{or} \;\; \widetilde{X}_{\phi_g^* f} = \phi_g^*\widetilde{X}_f(\phi_g^{-1})^* \ . \tag{8.7c}$$

We may now, for instance, consider the application of this formula to the Hamiltonian vector field generated by $G \in \mathcal{F}_\epsilon$, as described in Eqs. (8.5b), using the inverse of our original flow:

$$
\begin{aligned}
\phi_g^{-1} : \; U_\epsilon \;\to\; U_0 \;, \qquad & \mathcal{T}_\epsilon \ni \widetilde{X}_G = \quad (\phi_g^{-1})^* \frac{\partial \phi_g^*}{\partial \epsilon} \\
\Longrightarrow \quad (\phi_g^{-1})_*(\widetilde{X}_G) = \phi_g^* \left\{ (\phi_g^{-1})^* \frac{\partial \phi_g^*}{\partial \epsilon} \right\} (\phi_g^{-1})^* \; & = \frac{\partial \phi_g^*}{\partial \epsilon} (\phi_g^{-1})^* \; \in \; \mathcal{T}_0 \; .
\end{aligned}
\tag{8.8a}
$$

It is to be noted that this is José's Eq. (6.125), telling us that his form is not "wrong," but, rather, simply written in a different tangent space! We should also simply append here the form for this version of the Hamiltonian vector field as written using the derivative of the inverse mapping:

$$
\begin{aligned}
& \phi_g^* (\phi_g^{-1})^* = \text{identity} \; : \; \mathcal{F}_0 \; \to \; \mathcal{F}_0 \; , \\
\Longrightarrow \quad & \mathcal{T}_0 \ni \; (\phi_g^{-1})_*(\widetilde{X}_G) = \frac{\partial \phi_g^*}{\partial \epsilon} (\phi_g^{-1})^* = -\phi_g^* \frac{\partial (\phi_g^{-1})^*}{\partial \epsilon} \; .
\end{aligned}
\tag{8.8b}
$$

At this point we have established the essential differential geometry, and must now apply it to some physics. We first note that since the coordinates $\eta = \eta(\xi, t; \epsilon)$ are simply functions on the manifold, where H is the Hamiltonian we may certainly write

$$
\frac{d}{dt}\eta = \{\eta, H\} + \frac{\partial \eta}{\partial t} \; . \tag{8.9a}
$$

Here we want to keep good track of what the partial derivative with respect to t means. We of course know that $\eta = \eta(\xi, t; \epsilon)$. In this equation the Poisson bracket takes care of the dependence of the ξ on t, were we to be following a particular motion; the $\partial/\partial t$ takes care of explicit dependence on t, still keeping ϵ fixed. A good way to think about this is to re-write $\eta = \phi ?$

On the other hand, were we to resolve the perturbed dynamical problem, the Hamiltonian would be $\overline{H}$ and the η and t would be the independent variables for the problem, so that we could also write

$$
\frac{d}{dt}\eta = \{\eta, \overline{H}\} \; . \tag{8.9b}
$$

In principle we may then simply subtract the two forms for the derivative of η along the dynamical motion, i.e., the vector field denoted by d/dt, and resolve it for the partial derivative:

$$
\frac{\partial \eta}{\partial t} = \{\eta, \overline{H} - H\} \equiv \{\eta, R\} \; . \tag{8.10}
$$

However, this is a somewhat illegitimate thing to do since $\overline{H}$ is defined in the neighborhood U_ϵ, i.e., is to be conceived of as depending on those points with coordinates η, while H is defined in the neighborhood U_0, i.e., depends on the coordinates ξ. We know how the ordinary working physicist would treat this problem. Working with explicit functions of coordinates, she wold simply write, for instance, $H = H(\xi)$, resolve the defining equations $\eta = \eta(\xi, t; \epsilon)$ for ξ and insert them into the arguments of H, still referring to that H by the same symbol, although it really isn't the same function anymore. Since the guiding idea behind our current strategies is to consider things in terms of functions, rather than values of functions at particular points (or coordinates), then we must be more careful, or precise. Using the format already presented in the second half of Eqs. (8.3b), the H that appears in our equation, Eq. (8.6a), must be a function of $\eta = \phi_g(\xi)$; however, the H that we would usually begin with would really be a function of ξ, i.e., $H \in \mathcal{F}_0$. This means that the quantity

H that should appear in our equation must actually be the induced form of H that depends on the variables η, i.e., it is really the function $(\phi_g^{-1})^*H$, which satisfies $[(\phi_g^{-1})^*H](\eta) = H(\phi_g^-1\eta) = H(\xi)$. This requires that we repeat our last equation, writing it correctly, with all parts defined in the same neighborhood:

$$\frac{\partial\eta}{\partial t} = \{\eta, R\} = \widetilde{X}_R\eta\ , \quad R \equiv \overline{H} - (\phi_g^{-1})^*H\ . \tag{8.10$'$}$$

Our next job is to think of this equation in terms of the mapping ϕ_g^*, which gives us $\eta = \phi_g^*\xi$. The total derivative of η with respect to t, as written above involves both how the ξ change with time, as given by the Poisson bracket, and, separately, the partial derivative of with respect to t which is what we have now; therefore, if we insert this form for η into the left-hand side of the equation, we can ignore any time dependence of the ξ. Secondly, since $\{\eta, R\}$ is a function, we may always write it as $\phi_g^*\{\xi, R\} = \phi_g^*\widetilde{X}_R\xi$. Then, since we now have ξ on both sides of the equation, which is an arbitrary point, we may remove it, obtaining the desired re-write of Eq. (8.10$'$), i.e., the functional relationship for the time-dependence of ϕ_g^*:

$$\frac{\partial\phi_g^*}{\partial t} = \phi_g^*\widetilde{X}_R\ . \tag{8.11}$$

At this point we know how the desired mapping, ϕ_g^*, varies as we, separately, vary its dependence on time, via Eq. (8.11), or on the parameter ϵ, via Eq. (8.5c). We must, however, still insist that the two distinct variations are consistent. This requirement will provide us with a differential equation which must be satisfied by R. The solution of that equation will, eventually, provide a relationship between the desired quantities K and G, which is the beginning of the Deprit approach to this problem. The consistency equation is simply the equality of mixed partial derivatives:

$$\begin{aligned} \frac{\partial}{\partial\epsilon}\frac{\partial\phi_g^*}{\partial t} &= \frac{\partial}{\partial t}\frac{\partial\phi_g^*}{\partial\epsilon} \\ \Longrightarrow \quad \frac{\partial\phi_g^*}{\partial\epsilon}\widetilde{X}_R + \phi_g^*\widetilde{X}_{\partial R/\partial\epsilon} &= \frac{\partial\phi_g^*}{\partial t}\widetilde{X}_G + \phi_g^*\widetilde{X}_{\partial G/\partial t} \\ \Longrightarrow \quad \phi_g^*\Big\{\widetilde{X}_G\widetilde{X}_R + \widetilde{X}_{\partial R/\partial\epsilon} &= \widetilde{X}_R\widetilde{X}_G + \widetilde{X}_{\partial G/\partial t}\Big\}\ , \end{aligned} \tag{8.12a}$$

where we have re-inserted the original equations for the first derivatives of ϕ_g^* and noted that the derivative of the Hamiltonian vector field, of some function, just means the Hamiltonian vector field made from the derivative of that function. At this point we have the commutator of two Hamiltonian vector fields, which we have already in Eqs. (8.6a). We insert that information, and note that ϕ_g^* is surely not zero, so that if something it acts on vanishes then that something must itself already vanish. Therefore, we must have that

$$\widetilde{X}_{\partial G/\partial t} - \widetilde{X}_{\partial R/\partial\epsilon} + \widetilde{X}_{\{G,R\}} = 0\ . \tag{8.12b}$$

Next, we note that Hamiltonian vector fields are additive, relative to their associated functions. Then lastly we note that if the Hamiltonian vector field with respect to some function vanishes, then that function must be a constant. Therefore the argument above must be some constant. However, surely R is not so well defined as we cannot absorb a constant into it without changing its definition. Therefore we obtain the following differential equation, as promised:

$$\frac{\partial G}{\partial t} = \frac{\partial R}{\partial\epsilon} + \{R, G\}\ . \tag{8.13a}$$

Before going on to put this into standard form, let us note that there is in fact a "formal" way to solve this equation, for R, assuming that we actually know the inhomogeneous part of the equation, $\frac{\partial G}{\partial t}$. Recall Eq. (8.5b), which can be written to define $\partial(\phi_g^{-1})^*/\partial\epsilon = -\widetilde{X}_G(\phi_g^{-1})^*$. Then define

$$S \equiv (\phi_g^{-1})^* \Big(\phi_{g_0}^* R(\epsilon_0) + \int_{\epsilon_0}^{\epsilon} d\epsilon'\, \phi_{g'}^* \frac{\partial G(\epsilon')}{\partial t}\Big) \equiv (\phi_g^{-1})^* N \; , \tag{8.13b}$$

where $g' \equiv G(\epsilon')$ and $g_0 \equiv G(\epsilon_0)$. We now claim that this S is the solution for R. We show this by inserting it into the equation as follows:

$$\frac{\partial S}{\partial \epsilon} = \frac{\partial(\phi_g^{-1})^*}{\partial\epsilon} N + (\phi_g^{-1})^* \phi_g^* \frac{\partial G}{\partial t} = -\widetilde{X}_G S + \frac{\partial G}{\partial t} \; , \tag{8.13c}$$

which is what is needed in order for this to be the desired solution for R. Agreeing that $R(0) = 0$, i.e., that K and $(\phi_g^{-1})^* H$ are the same when evaluated at $\epsilon = 0$, we may use this to write an explicit form for the new Hamiltonian, assuming that we already knew G:

$$K(\eta, t; \epsilon) = (\phi_g^{-1})^* \left\{ H(\xi, t; \epsilon) + \int_0^{\epsilon} d\epsilon'\, \phi_g^* \frac{\partial G}{\partial t} \right\} \; . \tag{8.13d}$$

We may now continue toward the desired equations that we will use to find, with yet more efforts concerning average values, to create perturbation series that we can try to resolve. We begin by inserting the definition of R into our earlier equation:

$$\frac{\partial G}{\partial t} = \frac{\partial \overline{H}}{\partial \epsilon} - \frac{\partial(\phi_g^{-1})^*}{\partial\epsilon} H - (\phi_g^{-1})^* \frac{\partial H}{\partial \epsilon} + \widetilde{X}_G(\overline{H} - (\phi_g^{-1})^* H) \; . \tag{8.14a}$$

At this point we use Eq. (8.5b), noting that when operating on functions the Lie derivative in direction $\widetilde{X}_G$ is the same as the action of that vector field itself on the function:

$$\widetilde{X}_G(\phi_g^{-1})^* H = -\frac{\partial(\phi_g^{-1})^*}{\partial\epsilon} \phi_g^* (\phi_g^{-1})^* H = -\frac{\partial(\phi_g^{-1})^*}{\partial\epsilon} H \; , \tag{8.14b}$$

so that this term cancels the second term in Eq. (8.14a), where we re-write the equation again just because it is indeed our fundamental equation which we will use to determine the desired quantities, $\overline{H}$ and G, via series expansions:

$$\frac{\partial G}{\partial t} = \frac{\partial \overline{H}}{\partial \epsilon} + \widetilde{X}_G \overline{H} - (\phi_g^{-1})^* \frac{\partial H}{\partial \epsilon} \; . \tag{8.15}$$

This is Eq. (2.5.27) of Lichtenberg, but is noticeably different from Eq. (6.128) of José. (It also agrees with all the other references I can find.)

We know we are going to expand all unknown quantities in infinite series in ϵ, and compare quantities term by term, when inserted into this equation. However, as we need the quantities ϕ_g^* and $\widetilde{X}_G$ in the above equation, and both of them are expressions determined by the generating function G, we first determine the relationship between their series expansions. Since we do not know G, we first expand it in an infinite series, where the "name" for the index of the expansion coefficients will turn out to be quite useful later on, in that it will cause quite plausible and easily remembered relationships:

$$G \equiv \sum_{n=0}^{\infty} \epsilon^n \, G_{n+1} \implies \widetilde{X}_G = \sum_{n=0} \epsilon^n \, \widetilde{X}_{G_{n+1}} \; . \tag{8.16a}$$

At this point I note that at first it bothered me that the coefficient of ϵ^0 was not something simple, since after all no perturbation has yet occurred. On the other hand, it does now seem clear that $\widetilde{X}_{G_1}$, the coefficient of ϵ^0, is just showing you the direction, in the phasespace, that the problem will be going when one allows ϵ to vary away from zero; i.e., it is like $d/d\epsilon$ evaluated at $\epsilon = 0$. I believe we will also see something relevant in that regard when I compare G to the associated generating function of type, for this situation, which will happen later.

The next step is to expand the flow mapping as well; the one we need is of course $(\phi_g^{-1})^*$:

$$(\phi_g^{-1})^* \equiv \sum_{n=0}^{\infty} \epsilon^n \, \Phi_n \; , \tag{8.16b}$$

which is normalized at the $\epsilon = 0$ level by setting $\Phi_0 = \mathbf{1}$, the identity map. As we know that there is a relationship between the flow and the Hamiltonian vector field for the associated generator, we now use that relationship to eliminate the need to make a separate expansion for the vector field, but of course taking account of the fact that linearity arranges for the relationship already noted in Eq. (8.16a). Inserting those two expansions into the relationship between them given by Eq. (8.5c), we have the following:

$$-\sum_{m=0}^{\infty} \epsilon^m \, \widetilde{X}_{G_{m+1}} \sum_{k=0}^{\infty} \epsilon^k \, \Phi_k = -\widetilde{X}_G \, (\phi_g^{-1})^* = \frac{\partial (\phi_g^{-1})^*}{\partial \epsilon} = \frac{\partial}{\partial \epsilon} \sum_{n=0}^{\infty} \epsilon^n \, \Phi_n = \sum_{n=1}^{\infty} n \epsilon^{n-1} \, \Phi_n \; . \tag{8.17}$$

To resolve the coefficient of every distinct power of ϵ on both sides of this equation, we first rewrite the double sum, in terms of $n \equiv k + m$ and k:

$$\sum_{n=1}^{\infty} n \epsilon^{n-1} \, \Phi_n = \sum_{m=0}^{\infty} \epsilon^m \, \widetilde{X}_{G_{m+1}} \sum_{k=0}^{\infty} \epsilon^k \, \Phi_k = \sum_{n=0}^{\infty} \epsilon^n \sum_{k=0}^{n} \widetilde{X}_{G_{n+1-k}} \, \Phi_k \; . \tag{8.18}$$

We may then compare both sides of this equation and determine a *recursion relation* for the coefficients Φ_n, beginning with the normalization condition already stated:

$$\Phi_n = -\frac{1}{n} \sum_{k=0}^{n-1} \widetilde{X}_{G_{n-k}} \Phi_k \; , \quad n > 0 \; ;$$
$$\implies \begin{cases} \Phi_0 = \mathbf{I} \; , \\ \Phi_1 = -\widetilde{X}_{G_1} \; , \\ \Phi_2 = -\frac{1}{2}(\widetilde{X}_{G_2} + \widetilde{X}_{G_1} \Phi_1) = \frac{1}{2}(-\widetilde{X}_{G_2} + (\widetilde{X}_{G_1})^2) \; , \\ \Phi_3 = -\frac{1}{3}(\widetilde{X}_{G_3} + \widetilde{X}_{G_2} \Phi_1 + \widetilde{X}_{G_1} \Phi_2) = \frac{1}{6}[-2\widetilde{X}_{G_3} + 2\widetilde{X}_{G_2} \widetilde{X}_{G_1} + \widetilde{X}_{G_1} \widetilde{X}_{G_2} - (\widetilde{X}_{G_1})^3] \\ \Phi_4 = \frac{1}{24}\Big[-6\widetilde{X}_{G_4} + 6\widetilde{X}_{G_3} \widetilde{X}_{G_1} + 6\widetilde{X}_{G_1} \widetilde{X}_{G_3} + 3(\widetilde{X}_{G_2})^2 - 3\widetilde{X}_{G_2} (\widetilde{X}_{G_1})^2 \\ \qquad\qquad -2\widetilde{X}_{G_1} \widetilde{X}_{G_2} \widetilde{X}_{G_1} - (\widetilde{X}_{G_1})^2 \widetilde{X}_{G_2} + (\widetilde{X}_{G_1})^4 \Big] \; , \end{cases} \tag{8.19}$$

where it does makes sense to have powers of vector fields, since it simply means the repeated action on functions.

Now expanding both H and $\overline{H}$ also in power series,

$$H = \sum_{n=0}^{\infty} \epsilon^n H_n \ , \qquad \overline{H} = \sum_{n=0}^{\infty} \epsilon^n \overline{H}_n \quad , \tag{8.20}$$

where we simply go ahead and maintain higher orders in ϵ in our original Hamiltonian as well, which doesn't particularly hurt anything even though we didn't begin with that form. Then we may insert it all back into the defining equation, Eq. (8.15), which gives us

$$\frac{\partial G_n}{\partial t} = n\overline{H}_n + \sum_{m=0}^{n-1} \widetilde{X}_{G_{n-m}} \overline{H}_m - \sum_{m=1}^{n} m\Phi_{n-m} H_m \ , \quad n = 1, 2, 3, \ \ldots \ . \tag{8.21}$$

Although this form will surely do, one may simplify it just a bit further by first noting that we know for sure that when $\epsilon = 0$ we have simply $\overline{H}_0 = H_0$, and writing out explicitly the terms in the two sums that do not have the same limits of summation. For the first sum, the first term is just $\widetilde{X}_{G_n} \overline{H}_0 = \widetilde{X}_{G_n} H_0 = \{H_0, G_n\}$, which we will move to the other side of the equation; while the last term in the second sum is $n\Phi_0 H_n = nH_n$, so that we may put this term together with the first term on the right-hand side of the equation. This results in the basic generating formulae:

$$\begin{aligned}
0 &= \overline{H}_0 - H_0 \ , \\
\frac{\partial G_1}{\partial t} + \{G_1, H_0\} &= \overline{H}_1 - H_1 \ , \\
\frac{\partial G_n}{\partial t} + \{G_n, H_0\} &= n(\overline{H}_n - H_n) + \sum_{m=1}^{n-1} \left\{ \widetilde{X}_{G_{n-m}} \overline{H}_m - m\Phi_{n-m} H_m \right\} \ , \ n = 2, 3, 4, \ \ldots \ .
\end{aligned} \tag{8.22}$$

If we now insert the values calculated above, in Eqs. (8.19), for ϕ_ℓ, we may write out higher-order expressions easily:

$$\begin{aligned}
\frac{\partial G_2}{\partial t} + \{G_2, H_0\} &= 2(\overline{H}_2 - H_2) + \widetilde{X}_{G_1}(\overline{H}_1 + H_1) \ , \\
\frac{\partial G_3}{\partial t} + \{G_3, H_0\} &= 3(\overline{H}_3 - H_3) + \widetilde{X}_{G_1}(\overline{H}_2 + 2H_2) + \widetilde{X}_{G_2}(\overline{H}_1 + \tfrac{1}{2}H_1) - \tfrac{1}{2}(\widetilde{X}_{G_1})^2 H_1 \ , \\
\frac{\partial G_4}{\partial t} + \{G_4, H_0\} &= 4(\overline{H}_4 - H_4) + \widetilde{X}_{G_1}(\overline{H}_3 + 3H_3) + \widetilde{X}_{G_2}(\overline{H}_2 + H_2) + \widetilde{X}_{G_3}(\overline{H}_1 + \tfrac{1}{3}H_1) \\
&\quad - (\widetilde{X}_{G_1})^2 H_2 - \tfrac{1}{6}[2\widetilde{X}_{G_2}\widetilde{X}_{G_1} + \widetilde{X}_{G_1}\widetilde{X}_{G_2} - (\widetilde{X}_{G_1})^3] H_1 \ , \\
\frac{\partial G_5}{\partial t} + \{G_5, H_0\} &= 5(\overline{H}_5 - H_5) + \widetilde{X}_{G_4}(\overline{H}_1 + \tfrac{1}{4}H_1) + \widetilde{X}_{G_3}(\overline{H}_2 + \tfrac{2}{3}H_2) + \widetilde{X}_{G_2}(\overline{H}_3 + \tfrac{3}{2}H_3) \\
&\quad + \widetilde{X}_{G_1}(\overline{H}_4 + 4H_4) - \tfrac{3}{2}(\widetilde{X}_{G_1})^2 H_3 - \tfrac{1}{3}[2\widetilde{X}_{G_2}\widetilde{X}_{G_1} + \widetilde{X}_{G_1}\widetilde{X}_{G_2} - (\widetilde{X}_{G_1})^3] H_2 \\
&\quad - \tfrac{1}{24}\Big[6\widetilde{X}_{G_3}\widetilde{X}_{G_1} + 6\widetilde{X}_{G_1}\widetilde{X}_{G_3} + 3(\widetilde{X}_{G_2})^2 - 3\widetilde{X}_{G_2}(\widetilde{X}_{G_1})^2 \\
&\qquad - 2\widetilde{X}_{G_1}\widetilde{X}_{G_2}\widetilde{X}_{G_1} - (\widetilde{X}_{G_1})^2\widetilde{X}_{G_2} + (\widetilde{X}_{G_1})^4\Big] H_1 \ , \\
&\text{recall that} \quad \widetilde{X}_{G_m} H_n \equiv \{H_n, G_m\} \ .
\end{aligned} \tag{8.23}$$

What I have written above agrees with the expressions in Lichtenberg, Eqs. (2.5.31).

It should be clear that the procedure to write down the equations to whatever order one wants is very straightforward and clear, even though surely it is true that the equations become longer as one goes to higher orders. On the other hand, just as in the canonical perturbation theory, when there are more than one independent variable then there will be problems with vanishing denominators at certain frequencies. In addition, the series does not usually converge, but, rather, is an asymptotic series only!

An even more serious difficulty with the equations is of course that we do not know either G_n nor $\overline{H}_n$. This is not a new difficulty, since we had the same problem with canonical perturbation theory and with Lindstedt's earlier version of the same generic questions. In all of these the solution basically amounted to ensuring that there are no secular terms in the results, since we insert, by hand, the additional requirement that the system is periodic. This was done by using an averaging scheme. We will do that, again, for the examples below. What is the basic "extra" ingredient for systems when we do not know that it will be forever periodic, I do not know.

I want now to discuss a little bit more how one might create ϕ_g from knowledge of G or, more precisely, from knowledge of $\widetilde{X}_G$. It is plausible that these very explicit calculations, borrowed from Deprit's original paper, may make some of the discussions above somewhat clearer, as well as show how they can be done in principle, at the least. In the first instance, as already stated without much proof, we go through the case when G is independent of ϵ. We begin by noting "obvious" properties of the Hamiltonian vector field, here written as $\mathcal{L}_{\widetilde{X}_G}$, for generality. If f and g are two appropriate functions, and α and β are constants, then the linearity and derivation properties are obvious:

$$\begin{aligned} \mathcal{L}_{\widetilde{X}_G}(\alpha f + \beta g) &= \alpha \mathcal{L}_{\widetilde{X}_G} f + \beta \mathcal{L}_{\widetilde{X}_G} g \ , \\ \mathcal{L}_{\widetilde{X}_G}(fg) &= f \mathcal{L}_{\widetilde{X}_G} g + (\mathcal{L}_{\widetilde{X}_G} f) g \ , \end{aligned} \tag{8.24}$$

On the other hand the Jacobi identity of Poisson brackets is needed to show the following two desirable features, where here both G and F are functions:

$$\begin{aligned} \mathcal{L}_{\widetilde{X}_G}\{f,g\} &= \{f, \mathcal{L}_{\widetilde{X}_G} g\} + \{\mathcal{L}_{\widetilde{X}_G} f, g\} \ , \\ [\mathcal{L}_{\widetilde{X}_G}, \mathcal{L}_{\widetilde{X}_F}] &= \mathcal{L}_{\widetilde{X}_{\{F,G\}}} \ . \end{aligned} \tag{8.25}$$

Defining the zero-th power of the Lie derivative as the identity, as one would expect, we may consider the extension of these derivational forms to n-th iterations, using the standard Leibnitz rule for n-th derivatives of products:

$$\begin{aligned} (\mathcal{L}_{\widetilde{X}_G})^n (fg) &= \sum_{m=0}^{n} \binom{n}{m} [(\mathcal{L}_{\widetilde{X}_G})^m f]\, (\mathcal{L}_{\widetilde{X}_G})^{n-m} g \ , \\ (\mathcal{L}_{\widetilde{X}_G})^n \{f,g\} &= \sum_{m=0}^{n} \binom{n}{m} \left\{ (\mathcal{L}_{\widetilde{X}_G})^m f, (\mathcal{L}_{\widetilde{X}_G})^{n-m} g \right\} \ . \end{aligned} \tag{8.26}$$

Now, we also assert that it is then straightforward to sum infinitely many of these to show the following relations, where the infinite sums in question do not act at all like derivations:

$$\begin{aligned} e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} (fg) &= [e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} f]\, e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} g \ , \\ e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} \{f,g\} &= \{e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} f, e^{\epsilon \mathcal{L}_{\widetilde{X}_G}} g\} \ . \end{aligned} \tag{8.27}$$

We prove this by expanding out the products on the right-hand side, beginning with the first one:

$$
\begin{aligned}
[e^{\epsilon\mathcal{L}_{\widetilde{X}_G}} f]\, e^{\epsilon\mathcal{L}_{\widetilde{X}_G}} g &= \sum_{k=0}^{\infty}\frac{\epsilon^k}{k!}(\mathcal{L}_{\widetilde{X}_G})^k f \sum_{m=0}^{\infty}\frac{\epsilon^m}{m!}(\mathcal{L}_{\widetilde{X}_G})^m g \\
&= \sum_{n=0}^{\infty}\epsilon^n \sum_{m=0}^{n}\frac{1}{m!}\frac{1}{(n-m)!}(\mathcal{L}_{\widetilde{X}_G})^{n-m} f(\mathcal{L}_{\widetilde{X}_G})^m g \qquad (8.27a)\\
&= \sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}\sum_{m=0}^{n}\binom{n}{m}(\mathcal{L}_{\widetilde{X}_G})^{n-m} f(\mathcal{L}_{\widetilde{X}_G})^m g = \sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}(\mathcal{L}_{\widetilde{X}_G})^n (fg) = e^{\epsilon\mathcal{L}_{\widetilde{X}_G}}(fg)\ .
\end{aligned}
$$

One can see that the algebra is not noticeably different for the second one, so it is not displayed here. However, the import of that one is the the transformation established by the following, where $G = G(\xi)$, is canonical:

$$
\eta(\xi,\epsilon) \equiv e^{\epsilon\mathcal{L}_{\widetilde{X}_G}}\xi \quad \text{is a canonical transformation.}
$$

(This of course because $e^{\epsilon\mathcal{L}_{\widetilde{X}_G}}$ acting on the constant matrix $\gamma^{\alpha\beta}$ just leaves it invariant.) Notice that with this transformation we may write the following since ξ is independent of ϵ:

$$
\begin{aligned}
\frac{\partial\eta}{\partial\epsilon} &= \sum_{n=0}^{\infty} n\frac{\epsilon^{n-1}}{n!}(\mathcal{L}_{\widetilde{X}_G})^n\xi = \sum_{m=0}^{\infty}\frac{\epsilon^m}{m!}(\mathcal{L}_{\widetilde{X}_G})^{m+1}\xi = \mathcal{L}_{\widetilde{X}_G}\,\eta\ , \\
\Longrightarrow \qquad \frac{\partial e^{\epsilon\mathcal{L}_{\widetilde{X}_G}}}{\partial\epsilon}\xi &= \mathcal{L}_{\widetilde{X}_G}\, e^{\epsilon\mathcal{L}_{\widetilde{X}_G}}\xi\ .
\end{aligned} \tag{8.28}
$$

We use this insight to consider an arbitrary function $f \in \mathcal{F}_\eta$, that does not depend explicitly on ϵ, and a mapping that sends it back to $\mathcal{F}_\xi$ as usual:

$$
(\phi_g^* f)(\xi;\epsilon) \equiv f[\eta(\xi;\epsilon)] \equiv f[\phi_g^\epsilon(\xi)]\ , \tag{8.29}
$$

where the superscript ϵ is put as an additional argument on the mapping only to remind us specifically as to where the value of ϵ comes from; i.e., it is the parameter value along which we have pushed ξ, which by itself does not depend on ϵ. We then may show that this mapping of functions is indeed generated specifically as the exponential we have been considering. First we consider the following, where we use Eq. (8.28) toward the end of the calculations:

$$
\begin{aligned}
\mathcal{L}_{\widetilde{X}_G}\phi_g^* f &= \{(\phi_g^* f)(\xi;\epsilon), G) = \gamma^{\alpha\beta}\frac{\partial\phi_g^* f}{\partial\xi^\alpha}\frac{\partial G}{\partial\xi^\beta} = \frac{\partial f}{\partial\eta^\mu}\gamma^{\alpha\beta}\frac{\partial\eta^\mu}{\partial\xi^\alpha}\frac{\partial G}{\partial\xi^\beta}\ , \\
&= \frac{\partial f}{\partial\eta^\mu}\{\eta^\mu, G\} = \frac{\partial f}{\partial\eta^\mu}\mathcal{L}_{\widetilde{X}_G}\eta^\mu = \frac{\partial f}{\partial\eta^\mu}\frac{\partial\eta^\mu}{\partial\epsilon} = \frac{d\,\phi_g^* f}{d\epsilon}\ .
\end{aligned} \tag{8.30}
$$

This is obviously true again when iterated an arbitrary number of times on both sides, so that we may write the following where we expand out in a Taylor series in ϵ:

$$
\phi_g^* f(\xi;\epsilon) = \sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}\frac{d^n\,\phi_g^* f}{d\epsilon^n}\Big|_{\epsilon=0} = \sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}(\mathcal{L}_{\widetilde{X}_G})^n(\phi_g^* f)(\xi;\epsilon=0) = e^{\epsilon\mathcal{L}_{\widetilde{X}_G}} f(\xi)\ . \tag{8.31}
$$

Comparing this with the definition of the function that has been pulled back, we may conclude that

$$
e^{\epsilon\mathcal{L}_{\widetilde{X}_G}} f(\xi) = (\phi_g^* f)(\xi;\epsilon) = f[e^{\epsilon\mathcal{L}_{\widetilde{X}_G}}\xi] = f[\eta(\xi;\epsilon)]\ . \tag{8.32}
$$

An interesting "sidelight" of this is the following statement that follows because $\{G,G\}=0$:

$$G[\eta(\xi;\epsilon)] = (\phi_g^* G)(\xi)\ . \tag{8.32a}$$

We may now generalize this discussion to the case where our function f does depend also on ϵ, but presumably in some analytic way so that we may write out a Taylor series for it:

$$f(\eta,\epsilon) = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} f_n(\eta)\ , \quad f_n(\eta) \equiv \left(\frac{\partial}{\partial\epsilon}\right)^n f(\eta,\epsilon)\Big|_{\epsilon=0}\ . \tag{8.33a}$$

We may now treat each of these functions $f_n(\eta)$ in the same way as we just did for functions that did not depend on ϵ explicitly:

$$\phi_g^* f_n[\eta(\xi;\epsilon)] = e^{\epsilon\mathcal{L}_{\widetilde{X}_G}} f_n(\xi) = \sum_{m=0}^{\infty} \frac{\epsilon^m}{m!} (\mathcal{L}_{\widetilde{X}_G})^m f_n(\xi)\ , \tag{8.33b}$$

which allows us to insert this into the previous equation to obtain a double sum, which can again be re-written as a sum over powers of ϵ:

$$\begin{aligned}(\phi_g^* f)(\xi,\epsilon) &= f[(\phi_g^*\eta)(\xi;\epsilon);\epsilon] = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \sum_{m=0}^{\infty} \frac{\epsilon^m}{m!} (\mathcal{L}_{\widetilde{X}_G})^m f_n(\xi) \\ &= \sum_{\ell=0}^{\infty} \frac{\epsilon^\ell}{\ell!} \sum_{m=0}^{\ell} \binom{\ell}{m} (\mathcal{L}_{\widetilde{X}_G})^m f_{\ell-m}(\xi)\ .\end{aligned} \tag{8.33c}$$

Having now looked at the case where both G and f were independent of ϵ, and then the (more complicated) case where f is allowed to depend on ϵ, Deprit then continues to the completely general case where both depend explicitly on ϵ, which is the one that is usually really wanted. In that case, also, he establishes the results as an infinite series, although in this case it is a doubly-infinite series. He explains it in some detail; however, it may really be considered as an algorithmic scheme to accomplish the calculations in question, that we have already described, more simply, as his scheme of infinitely many equations, in powers of ϵ. His approach is quite useful if one intends to program a computer to proceed all by itself in the process of solving the equations to an arbitrary, finite order.

9. Examples

It is probably useful to first consider the incredibly simple example put forward by Cary, to get perhaps a better feel for some of the abstract geometry. We suppose given in advance the Hamiltonian and also the generating function, G, in terms of a 1-dimensional problem, in coordinates (q,p) and desire to go to (Q,P), functions of (q,p), t, and also ϵ:

$$H = p^2/2m \qquad G = -\epsilon t p^2 \quad \Longrightarrow \quad \widetilde{X}_G = -2\epsilon t p \frac{\partial}{\partial q}\ ,\ . \tag{9.1}$$

This is of course only some illustration of the geometry, rather than a real problem, since any real problem one would have to find G; nonetheless, let's go quickly through the motions. We first want to find the mapping to $\eta = \eta(\xi, t; \epsilon)$, which we do by considering

$$\frac{\partial \eta}{\partial \epsilon} = \{\eta, G\} \;\Rightarrow\; \begin{cases} \dfrac{\partial Q}{\partial \epsilon} = \{-2\epsilon t p (\partial q/\partial q)\}_{|q=Q,p=P} = -2\epsilon t P \; , \\ \dfrac{\partial P}{\partial \epsilon} = \{-2\epsilon t p (\partial p/\partial q)\}_{|q=Q,p=P} = 0 \; , \end{cases} \tag{9.2a}$$

where we observe that we calculated the Poisson bracket in the ξ basis, and then substituted.

We first integrate the equation for P and insert it into that one for Q, which gives us

$$Q = q - \epsilon^2 t p \; , \quad P = p \; . \tag{9.2b}$$

As well, it is clear that if $f \in \mathcal{F}_\epsilon$ so that $f = f(Q, P, t, \epsilon)$, then

$$(\phi_g^* f)(q, p, t, \epsilon) = f[\phi_g(\eta)] = f[Q(q, p, t, \epsilon), P(q, p, t, \epsilon), , \epsilon] = f[q - \epsilon^2 t p, p, t, \epsilon] \; . \tag{9.2c}$$

On the other hand, we can now write

$$\frac{\partial \phi_g^*}{\partial \epsilon} = -(\phi_g^*)\, 2\epsilon t p \partial_q \quad \Longleftrightarrow \quad \frac{d\phi_g^*}{\phi_g^*} = -2 t p \partial_q \epsilon \, d\epsilon \quad \Longrightarrow \quad \phi_g^* = e^{-\epsilon^2 t p \partial_q} \; , \tag{9.2d}$$

so that this is just the (standard Taylor series expression for the) shift operator displayed in Eq. (9.2c), as expected. Notice, on the other hand, that it is not simply the Lie series for the vector field $\widetilde{X}_G$, which it should not have been since this G does depend on ϵ explicitly.

Now let us consider the continuing example of the (1-dimensional) pendulum. Following Lichtenberg, in §2.5, we rewrite that Hamiltonian, but now writing ϵ in such a way that it will simply be replaced by $+1$ when we want to revert to the original problem. (Recall that earlier in our own calculation it was last used to symbolize $1/6$):

$$\begin{aligned} H(\phi, J) &= H_0 + \epsilon\, H_1 + \epsilon^2\, H_2 = \omega_0 J - \epsilon \frac{G\, J^2}{6} \sin^4 \phi + \epsilon^2 \frac{G^2 J^3}{90\omega_0} \sin^6 \phi \\ &= \omega_0 J - \epsilon \tfrac{1}{48} G\, J^2 [3 - 4\cos 2\phi + \cos 4\phi] + \epsilon^2 \tfrac{1}{2880} \tfrac{G^2}{\omega_0} J^3 [10 - 15\cos 2\phi + 6\cos 4\phi - \cos 6\phi] \; . \end{aligned} \tag{9.3}$$

We next consider the fact that the "3" in the equation for H_1 will generate a secular term, or a non-zero average term, and therefore insist that $\overline{H}_1$ be found from it, to cancel it out:

$$\overline{H}_1 = \langle H_1 \rangle = -\tfrac{1}{16} G \overline{J}^2 \; , \tag{9.4a}$$

where we have written $\overline{J}$ instead of J, since $\overline{H}_1$ should of course be given as a function of $\overline{J}$, and the two are equal to lowest order in ϵ. Therefore, as before, we may write the complete new Hamiltonian to this order, and the associated frequency:

$$\overline{H} = \omega_0 \overline{J} - \epsilon \tfrac{1}{16} G \overline{J}^2 + O(\epsilon^2) \quad \Longrightarrow \quad \omega = \omega_0 - \epsilon \tfrac{1}{8} G \overline{J} + O(\epsilon^2) \; . \tag{9.4b}$$

Hamilton's equations, written below to this order. They may then be integrated, to this order, to determine the time-dependence of the new variables:

$$\begin{aligned} \frac{d}{dt}\overline{J} = -\frac{\partial\overline{H}}{\partial\overline{\phi}} = 0 \ , \qquad \frac{d}{dt}\overline{\phi} = +\frac{\partial\overline{H}}{\partial\overline{J}} = \omega_0 + \epsilon\tfrac{3}{4}J \ , \\ \overline{J} = J \ , \qquad \overline{\phi} = \phi + (\omega_0 + \epsilon\tfrac{3}{4}J)t \ . \end{aligned} \tag{9.4c}$$

We then write down the appropriate portion of Eq. (8.19), noting that since H does not depend on time, there is no reason that G should depend on time either, and also that H_0, being exactly integrable, does not depend on ϕ. We may then integrate for G_1:

$$\begin{aligned} \frac{\partial G_1}{\partial t} + \{G_1, H_0\} &= \overline{H}_1 - H_1 \ , \\ \Longrightarrow \quad \frac{\partial G_1}{\partial \phi}\frac{\partial H_0}{\partial J} + \frac{\partial G_1}{\partial J}\frac{\partial H_0}{\partial \phi} &= \tfrac{1}{48}GJ^2[-4\cos 2\phi + \cos 4\phi] \ , \\ \Longrightarrow \quad \omega_0\frac{\partial G_1}{\partial \phi} &= \tfrac{1}{48}GJ^2[-4\cos 2\phi + \cos 4\phi] \ , \\ \Longrightarrow \quad G_1 = \frac{GJ^2}{192\omega_0}[\sin 4\phi - 8\sin 2\phi] &= -\frac{GJ^2}{48\omega_0}\cos\phi\,[3\sin\phi + 2\sin^3\phi] \ . \end{aligned} \tag{9.5}$$

With this information we may calculate the inverse transformation to this order. We first recall exactly how this should work for this situation:

$$\begin{aligned} [(\phi_g^{-1})^* j](\eta, t; \epsilon) &= j[\xi(\eta, t; \epsilon)] \ ; \\ \text{choose } j \text{ as the identity function:} \quad \xi &= (\phi_g^{-1})^*\eta \ . \end{aligned} \tag{9.6a}$$

We now want to write this out to just the first order that we currently have. We use

$$\phi_g^{-1}* = \mathbf{I} + \epsilon\Phi_1 + O(\epsilon^2) = \mathbf{I} - e\widetilde{X}_{G_1} \ . \tag{9.6b}$$

However, we have G_1 in terms of the original variables, ξ, instead of the new variables. There are two different ways we can look at this problem. The first approach is that the action of $\widetilde{X}_{G_1}$ is through a Poisson bracket, and we know that the Poisson bracket is canonical so that we may calculate it either way. Following this approach we could calculate

$$\epsilon\widetilde{X}_{G_1}\eta = \epsilon\{\eta, G_1\}^\eta = \epsilon\{\eta, G_1\}^\xi = \epsilon\frac{\partial G_1}{\partial J}\frac{\partial \eta}{\partial \phi} - \epsilon\frac{\partial G_1}{\partial \phi}\frac{\partial \eta}{\partial J} \ , \tag{9.6c}$$

and then take note of the fact that since all this is multiplied by ϵ we only need to insert η to one lower order; however, that order is the original level at which we began, so that we may simply replace η by ξ, and continue. This would give us

$$\begin{aligned} \epsilon\widetilde{X}_{G_1}\overline{\phi} &= \epsilon\frac{\partial G_1}{\partial J} = \epsilon\tfrac{1}{96}\frac{GJ}{\omega_0}(\sin 4\phi - 8\sin 2\phi) \ , \\ \epsilon\widetilde{X}_{G_1}\overline{J} &= -\epsilon\frac{\partial G_1}{\partial \phi} = -\epsilon\tfrac{1}{48}\frac{GJ^2}{\omega_0}(\cos 4\phi - 4\cos 2\phi) \ . \end{aligned} \tag{9.6d}$$

However, since we really were supposed to have η's, we then simply replace the ξ's we have with η's, correct to this order.

On the other hand, the other approach is to begin by calculating the Poisson bracket in a more obvious way, in terms of the η variables, so that it is necessary to re-write the currently given form of $G_1 = G_1(\xi)$ in the form $[(\phi_g^{-1})^* G_1][\xi(\eta, t; \epsilon)]$. On the other hand, since, again, we are doing this to one lower order than the intended result, since this term is multiplied by ϵ, this simply amounts to re-writing G_1 by replacing the J and ϕ by $\overline{J}$ and $\overline{\phi}$, respectively. It should be clear that these two methods give exactly the same results, namely

$$\phi = \overline{\phi} - \epsilon \tfrac{1}{96} \frac{G\overline{J}}{\omega_0} (\sin 4\overline{\phi} - 8 \sin 2\overline{\phi}) \, , \qquad J = \overline{J} + \epsilon \tfrac{1}{48} \frac{G\overline{J}^2}{\omega_0} (\cos 4\overline{\phi} - 4 \cos 2\overline{\phi}) \, . \tag{9.7}$$

By inserting the time-development of the perturbed coordinates, for some arbitrary real motion of the system, correct to this order, given in Eqs. (9.4c), we then have the explicit time dependence, over the same motion, correct (only) to this order, for the original coordinates. **Do notice** that the time dependence of these original coordinates is now very much more complicated than it would have been without the perturbing terms, because the Φ's that appear in the relations above have now a linear time dependence, appearing inside the trigonometric functions!

One could then insert this back into that equation for G_2 to determine that; however, I didn't do it. Nonetheless, there are some interesting things that one can do yet: We take the equation for G_2 and average both sides, to determine $\overline{H}_2$:

$$\overline{H}_2 = \langle H_2 \rangle + \tfrac{1}{2} \langle \{G_1, H_1 - \langle H_1 \rangle\} \rangle = \frac{G^2 J^3}{\omega_0} \left[\frac{1}{288} - \frac{1}{2} \frac{17}{1152} \right] = - \frac{G^2 J^3}{256\omega_0} = - \frac{G^2 \overline{J}^3}{256\omega_0} \, . \tag{9.8}$$

From this the Hamiltonian and frequency to second order are

$$\overline{H} = \omega_0 \overline{J} - \epsilon \tfrac{1}{16} G\overline{J}^2 - \epsilon^2 \frac{G^2 \overline{J}^3}{256\omega_0} + O(\epsilon^3) \quad \Longrightarrow \quad \omega = \omega_0 - \epsilon \frac{G\overline{J}}{8} - \epsilon^2 \frac{3G^2 \overline{J}^2}{256\omega_0} + O(\epsilon^3) \, . \tag{9.9}$$

At this point I want this frequency, however, in terms of $\overline{H}$, the true, perturbed energy, at least to this order. Therefore, we must first resolve the equation above for $\overline{J}$ as a function of $\overline{H}/\omega_0$, and insert it into the frequency equation. As the appearance of $\overline{J}$ in the form for ω is always multiplied by a non-zero power of ϵ, and as we only have things correct to second order, we need only resolve this equation correct to first-order in ϵ, which is

$$\overline{J} = \frac{\overline{H}}{\omega_0} + \epsilon \frac{G}{16\omega_0} \left(\frac{\overline{H}}{\omega_0} \right)^2 + O^2(\epsilon) \, , \tag{9.10a}$$

which then gives us

$$\omega = \omega_0 - \epsilon \frac{G}{8} \frac{\overline{H}}{\omega_0} - \epsilon^2 \frac{5G}{256\omega_0} \left(\frac{\overline{H}}{\omega_0} \right)^2 + O(\epsilon^3) \, . \tag{9.10b}$$

To compare this with the exact answer, we first re-write it as a fraction, inserting the fact that $\omega_0 = \sqrt{FG}$:

$$\frac{\omega}{\omega_0} = 1 - \frac{\epsilon}{8} \frac{\overline{H}}{F} - \frac{5\epsilon^2}{256} \left(\frac{\overline{H}}{F} \right)^2 + O(\epsilon^3) \, , \tag{9.10c}$$

a function only of $\overline{H}/F$, where we maintain that $\overline{H}$ is the energy to this order.

We now retreat to the exact pendulum equation, discussed much earlier, and recall that for energies less than the "turn-over energy" at the top we need $E < F$, and

$$\frac{\omega}{\omega_0} = \frac{\pi/2}{\mathcal{K}(k)} \ , \tag{9.10d}$$

where here $\mathcal{K}(k)$ is the elliptic integral of the first kind and k, its argument, is determined by $k^2 = \frac{1}{2}(1 + E/F)$. However, we also recall that the argument was put together so that the stable equilibrium position for the pendulum, i.e., hanging "straight down," was such that $E = -F$, so that $k = 0$, as expected, so that the elliptic function would just be a trigonometric function, and ω would just be ω_0. Therefore, if we want to compare in some sensible way, we should actually use $1 + E/F = (E + F)/F \equiv E_{\rm normal}/F$ as our variable, since this definition of $E_{\rm normal}$ is simply zero when it is hanging straight down, and therefore is a "small" parameter when no longer simply hanging straight down at rest. Taking this ratio as $\beta \equiv 1 + E/F \equiv E_{\rm normal}/F$, then $k^2 = \beta/2$ and we may take the straight-forward Taylor series expansion for the elliptic integral, as a function of k^2. This gives us

$$\begin{gathered}\frac{\omega}{\omega_0} = \frac{\pi/2}{\mathcal{K}(k)} = 1 - \tfrac{1}{4}k^2 - \tfrac{5}{64}k^4 + O(k^6) = 1 - \tfrac{1}{8}\beta - \tfrac{5}{256}\beta^2 + O(\beta^3) \ , \\ \beta \equiv E_{\rm normal}/F \ , \ \ E_{\rm normal} = 0 \ \ \text{"when hanging straight down,"}\end{gathered} \tag{9.10e}$$

in perfect agreement with our calculation to second order, as determined above.

As an "amusement," the paper by Koseloff, mentioned below in the references, gives the perturbed Hamiltonian to very high order, the first 11 terms of which I copy here. However, I note that his is normalized to be as simple as possible, so that he has set $G = 1 = F$, i.e., $m = 1 = g\ell$, which means that $\omega_0 = 1$:

$$\begin{aligned} &\text{with} \quad F = 1 = G = \omega_0 \ , \\ \overline{H} = \overline{J} &- \frac{\epsilon}{2^4}\overline{J}^2 - \frac{\epsilon^2}{2^8}\overline{J}^3 - \frac{5\epsilon^3}{2^{13}}\overline{J}^4 - \frac{33\epsilon^4}{2^{18}}\overline{J}^5 - \frac{63\epsilon^5}{2^{21}}\overline{J}^6 - \frac{527\epsilon^6}{2^{26}}\overline{J}^7 - \frac{9387\epsilon^7}{2^{32}}\overline{J}^8 \\ &- \frac{175045\epsilon^8}{2^{38}}\overline{J}^9 - \frac{422565\epsilon^9}{2^{41}}\overline{J}^{10} - \frac{4194753\epsilon^{10}}{2^{46}}\overline{J}^{11} + O(\epsilon^{11}) \ . \end{aligned} \tag{9.10f}$$

References:

1. D. Boccaletti and G. Pucacco, *Theory of Orbits*, Vols. 1 and 2, Springer, Berlin, 1996. ISBN 3-540-58963-5.

2. John R. Cary, "Lie Transform Perturbation Theory for Hamiltonian Systems," Phys. Rep. **79**, 129-159 (1981).

3. John R. Cary, "Time dependent canonical transformations and the symmetry-equals-invariant theorem," J. Math. Phys. **18**, 2432-2435 (1977).

4. André Deprit, "Canonical Transformations Depending on a Small Parameter," Celestial Mechanics **1**, 12-30 (1969).

5. R.L. Dewar, "Renormalized canonical perturbation theory for stochastic propagators," J. Phys. A **9**, 2043-2057 (1976).

6. J. Draft and J.M. Finn, "Lie Series and Invariant Functions for Analytic Symplectic Maps," J. Math. Phys. **17**, 2215-2227 (1976).

7. Francesco Fassó, "On a Relation Among Lie Series," Celestial Mechanics and Dynamical Astronomy, **46**, 113-118 (1989).

8. Francesco Fassò, "Lie series method for vector fields and Hamiltonian perturbation theory," Zeitschrift für angewandte Mathematik und Physik **41**, 843-864 (1990).

9. P.-V. Koseleff, "Comparison between Deprit and Dragt-Finn Perturbation Methods," Celestial Mechanics, **58**, p. 17-36 (1994).

10. J. Sánchez Mondragón and K.B. Wolf, *Lie Methods in Optics*, Lecture Notes in Physics, v. 250, Springer, Berlin, 1986. (Procedings of the CIFMO-CIO workshop in León, Mexico, Jan., 1985.) ISBN 3-540-16471-5. Good articles on Lie series, and on the Dragt-Finn approach. Also nice summaries of the structure of the groups ***Sp***(n,m).

11. Kiran G. Sonnad and John R. Cary, "A Lie Transform Perturbation Scheme for Hamiltonian Averaging in Self Consistent Systems," Proc. of the 2003 Particle Accelerator Conference, IEEE, p. 1536-1538 (2003).