Hamiltonian Perturbation Theory

Daniel W. Crews

March 16, 2023

1 Lie transformation theory

These notes outline the basic theory of the operators involved in Lie transformation techniques for perturbation theory. Subsequent sections develop the techniques themselves. The theory of Lie series in a previous note "Theory of vector fields, flows, and Lie theory" is extended, and an expansion of a Lie operator in powers of ϵ is developed for use in the subsequent sections. It's recommended to understand the section "Hamiltonian Theory" in those notes before reading these notes.

1.1 Properties of the Lie transformation

(Reference: Boccaletti, "Theory of Orbits") Consider a Hamiltonian system in canonical coordinates, and let L_S denote the Hamiltonian vector field of a generating function S

$$L_S = \{\cdot, S\} = \frac{\partial S}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial S}{\partial q^i}\frac{\partial}{\partial p_i}. \tag{1}$$

The operator L_S is also the Lie derivative with respect to S, which acts on functions f on the phase space. If S were the Hamiltonian, it would generate the equations of motion. Here it generates a perturbation as a "continuous change of variables". Further, as a differential operator it has a Leibniz rule,

$$L_S(fg) = \{fg, S\} = g\{f, S\} + f\{g, S\} = gL_S f + fL_S g \tag{2}$$

and therefore follows the general Leibniz rule

$$L_S^n(fg) = \sum_{k=0}^{n} \binom{n}{k} L_S^{(n-k)} f L_S^k g. \tag{3}$$

Now also one can build up the Lie series of the operator,

$$\exp(\epsilon L_S) f = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} L_S^n f \tag{4}$$

which satisfies the properties of linearity and "commutation",

$$\exp(\epsilon L_S)(\alpha f + \beta g) = \alpha \exp(\epsilon L_S) f + \beta \exp(\epsilon L_S) g, \tag{5}$$
$$\exp(\epsilon L_S)(fg) = (\exp(\epsilon L_S) f)(\exp(\epsilon L_S) g). \tag{6}$$

The first property is straight-forward. The proof of the second is important for what follows[1],

$$\exp(\epsilon L_S)(fg) = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} L_S^n(fg), \tag{7}$$

$$= \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \sum_{k=0}^{n} \binom{n}{k} L_S^{(n-k)} f L_S^k g, \tag{8}$$

Now let $l = n - k$, so $\frac{1}{n!}\binom{n}{k} = \frac{1}{(n-k)!k!} = \frac{1}{l!k!}$. Then k and l run independently to infinity, so

$$\exp(\epsilon L_S)(fg) = \sum_{l=0}^{\infty} \frac{\epsilon^l}{l!} L_S^l f \sum_{k=0}^{\infty} \frac{\epsilon^k}{k!} L_S^k g, \tag{9}$$

$$= (\exp(\epsilon L_S) f)(\exp(\epsilon L_S) g). \tag{10}$$

By this standard method, the commutation result will hold for any finite product of functions $\prod f_i$. In fact, because any analytic function $f(z)$ can be expanded in power series at a point z of the phase space, this demonstrates a more general "commutation theorem"

$$\exp(\epsilon L_S) f(z_1, z_2, \cdots, z_{2n}) = f(\exp(\epsilon L_S) z_1, \exp(\epsilon L_S) z_2, \cdots, \exp(\epsilon L_S) z_{2n}). \tag{11}$$

As a corollary the Lie series $\exp(\epsilon L_S)$ produces canonical transformations

$$\exp(\epsilon L_S)\{f, g\} = \{\exp(\epsilon L_S) f, \exp(\epsilon L_S) g\} \tag{12}$$

by preserving the Poisson bracket of two analytic functions.

1.2 Nonautonomous Lie transformation

Now consider a more general operator which includes ϵ-dependent change,

$$\Delta_S = \frac{\partial}{\partial \epsilon} + L_S \equiv \partial_\epsilon + L_S \tag{13}$$

and let $S = S(\zeta, \epsilon)$ for ζ a point in the phase space. If S generated the equations of motion (with respect to the parameter ϵ), this operator would be the total time derivative. In the context of perturbation theory, one considers a continuous deformation of the initial problem (the "initial conditions" $z(\epsilon = 0) = \zeta$) and then looks at the behavior for infinitesimal ϵ.

As a differential operator, it is also linear and has a Leibniz rule,

$$\Delta_S^n(fg) = \sum_{k=0}^{n} \binom{n}{k} \Delta_S^{(n-k)} f \Delta_S^k g. \tag{14}$$

Consider the n'th rate of change of a function f under this operator near $\epsilon = 0$,

$$f^{(n)}(\zeta, 0) = [\Delta^n_{S(\zeta,\epsilon)} f(\zeta, \epsilon)]|_{\epsilon=0} \tag{15}$$

and then build up the corresponding Lie-like series as,

$$Tf \equiv \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} f^{(n)}(\zeta, 0). \tag{16}$$

[1] Gröbner, "Contributions to the method of Lie Series" (1967)

This is called *Deprit's Lie transformation operator*. It was introduced by A. Deprit, a researcher at Boeing laboratories, in 1968 to study perturbation problems in celestial mechanics. Today it can be recognized as the Lie series for the Lie derivative of the function at $\epsilon = 0$ along the flow of the "time-dependent" vector field of some generating function S. It can't be easily represented as an exponential because the system is nonautonomous.

The goal of perturbation theory is to choose S to transform a Hamiltonian function $f = H$ evolving under the flow of this perturbation to a normal form, or as close as possible. Like any perturbation theory, it will encounter resonances and blow up without applying averaging methods. While equivalent to the classical Poincaré-von Zeipel method (which perturbs the canonical Hamilton-Jacobi equation), it allows analysis at higher orders as there is a clear pattern in the perturbation equations, unlike in the classical method.

Soon after Deprit's publication a similar approach was proposed by Dragt and Finn which requires fewer Poisson bracket iterations at higher order and seems to be used more often in the plasma physics literature. Deprit's method is presented because it is based on one transformation and does not use extended phase space. In practice the Dragt-Finn approach is often easier to use but requires a few more concepts to be understood.

1.3 Expansion of the perturbation operator

Let us expand the Lie transformation L_S in powers of ϵ, supposing an expansion

$$S = \sum_{n=0}^{\infty} \epsilon^n S_{n+1}. \tag{17}$$

In relation to the initial conditions $z(0) = \zeta$, let $z = T(\epsilon)\zeta$. Then the equation of motion is

$$\frac{\partial z}{\partial \epsilon} = \{z, S(z, \epsilon)\} \tag{18}$$

which itself evolves with ϵ. That is, as ϵ varies the generating function changes, and the vector field $L_S = \{\cdot, S\}$ evolves. Now on the one hand,

$$\frac{\partial z}{\partial \epsilon} = L_S z \quad \Longrightarrow \quad \frac{\partial}{\partial \epsilon} T(\epsilon)\zeta = L_S T(\epsilon)\zeta \tag{19}$$

so as a formal equation for the operator, the problem and its solution are

$$\frac{\partial T}{\partial \epsilon} = L_{S(z,\epsilon)} T \quad \Longrightarrow \quad T = \exp\left(\int_0^{\epsilon} d\epsilon' L_{S(z,\epsilon')}\right) \tag{20}$$

as an exponential map, but this is not so useful because it is an integral over the *unknown* generating function $S(z, \epsilon)$ at the evolved points z at time ϵ. So instead, consider

$$\begin{aligned} \frac{\partial z}{\partial \epsilon} &= \{z, S(z, \epsilon)\} && (21) \\ &= \{T\zeta, S(T\zeta, \epsilon)\} = \{T\zeta, TS(\zeta, \epsilon)\} && (22) \\ &= T\{\zeta, S(\zeta, \epsilon)\} && (23) \\ &= T L_{S(\zeta,\epsilon)} \zeta && (24) \end{aligned}$$

so that the operator differential equation is in terms of the *supposed* expansion $S(\zeta, \epsilon)$,

$$\frac{\partial T}{\partial \epsilon} = T L_{S(\zeta,\epsilon)} \tag{25}$$

with initial condition $T(0) = 1$ (a near-identity transformation). First noting that

$$L_{S(\zeta,\epsilon)}\zeta = \{\zeta, S(\zeta,\epsilon)\} = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!}\{\zeta, S_{n+1}\} \equiv \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} L_{n+1}\zeta, \tag{26}$$

where $L_{n+1} \equiv L_{S_{n+1}}$, and also seeking an expansion of T in ϵ,

$$T = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} T_n \tag{27}$$

then substitution into Eqn. 25 gives,

$$\frac{d}{d\epsilon}\sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} T_n = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \frac{\epsilon^n \epsilon^m}{n!m!} T_n L_{m+1}, \tag{28}$$

$$\implies \sum_{n=1}^{\infty} \frac{\epsilon^{n-1}}{(n-1)!} T_n = \sum_{l=0}^{\infty} \frac{\epsilon^l}{l!} \sum_{n=0}^{l} \binom{l}{n} T_n L_{l+1-n}, \tag{29}$$

$$\implies \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} T_{n+1} = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \sum_{m=0}^{n} \binom{n}{m} T_m L_{n+1-m}, \tag{30}$$

after reversing the order of summation and relabelling $l \to n \to m$. Then the relation

$$T_n = \sum_{m=0}^{n-1} \binom{n-1}{m} T_m L_{n-m}, \quad n \geqslant 1 \tag{31}$$

holds for the expansion coefficients of the T operator, with $T_0 = 1$. The first few terms are,

$$T_0 = 1, \tag{32}$$

$$T_1 = L_1, \tag{33}$$

$$T_2 = L_2 + L_1^2, \tag{34}$$

$$T_3 = L_3 + L_2 L_1 + 2L_1 L_2 + L_1^3, \tag{35}$$

$$T_4 = L_4 + 3L_1 L_3 + 3L_2^2 + 3L_1^2 L_2 + L_3 L_1 + L_2 L_1^2 + 2L_1 L_2 L_1 + L_1^4, \tag{36}$$

$$\vdots \quad \vdots \tag{37}$$

The same procedure may be repeated for the inverse transformation. To find its differential equation, differentiate (noting that $\frac{d\zeta}{d\epsilon} = 0$ as ζ represents the initial conditions),

$$T^{-1}T\zeta = \zeta \implies \frac{d}{d\epsilon}(T^{-1}T)\zeta = 0, \tag{38}$$

$$\implies \frac{dT^{-1}}{d\epsilon} T\zeta + T^{-1}\frac{dT}{d\epsilon}\zeta = 0, \tag{39}$$

$$\implies (\frac{dT^{-1}}{d\epsilon} T + \Delta_{S(\zeta,\epsilon)})\zeta = 0, \tag{40}$$

$$\implies (\frac{dT^{-1}}{d\epsilon} T + L_S)\zeta = 0, \tag{41}$$

$$\implies \frac{dT^{-1}}{d\epsilon} = -L_S T^{-1}, \tag{42}$$

so seeking an expansion

$$T^{-1} = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} T_n^{-1} \tag{43}$$

then the very same procedure demonstrates the recurrence relation (with $T_0^{-1} = 1$),

$$T_n^{-1} = -\sum_{m=0}^{n-1} \binom{n-1}{m} L_{n-m} T_m^{-1}. \tag{44}$$

The first few expansion coefficients are

$$T_0^{-1} = 1, \tag{45}$$
$$T_1^{-1} = -L_1, \tag{46}$$
$$T_2^{-1} = -L_2 + L_1^2, \tag{47}$$
$$T_3^{-1} = -L_3 + 2L_2L_1 + L_1L_2 - L_1^3, \tag{48}$$
$$T_4^{-1} = -L_4 + 3L_3L_1 + 3L_2^2 - 3L_2L_1^2 + L_1L_3 - 2L_1L_2L_1 - L_1^2L_2 + L_1^4, \tag{49}$$
$$\vdots \quad \vdots \tag{50}$$

The numerical coefficients depend on whether factorials are used in the denominator of the assumed power series $f = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} f_n$ or not. The choice is conventional. Now while the expansion for the Lie transformation operator T is determined, exactly how to find the terms S_n in the generating function is not clear. The next section investigates this.

2 Deprit's method

This section describes how Deprit's method is used to analyze a Hamiltonian perturbation problem using the older methods found in Cary's report and the book by Lichtenberg and Lieberman. Deprit's original analysis is quite complicated. Dewar found a simpler (but rather opaque) method using operator notation [Cary, L& L]. Hopefully the reader will appreciate how much simpler the Dragt-Finn method in extended phase space is compared to the approach of this section. The complication arises as the system is non-autonomous. Now, the perturbation evolution operator $T(\epsilon)$ satisfies the differential equation

$$\frac{\partial T}{\partial \epsilon} = T L_S \tag{51}$$

and therefore any inhomogeneous Hamiltonian evolution equation in ϵ for unknown f

$$\frac{\partial f}{\partial \epsilon} + L_S f = g \tag{52}$$

can be solved using T, as

$$T\frac{\partial f}{\partial \epsilon} + T L_S f = Tg \tag{53}$$
$$\implies T\frac{\partial f}{\partial \epsilon} + \frac{\partial T}{\partial \epsilon} f = Tg \tag{54}$$
$$\implies \frac{\partial}{\partial \epsilon}(Tf) = Tg \tag{55}$$
$$\implies f(\epsilon) = T^{-1}(\epsilon) \int_0^{\epsilon} d\epsilon' T(\epsilon') g(\epsilon') + T^{-1}(\epsilon) f(0) \tag{56}$$

which is, in effect, a solution by method of characteristics, *i.e.* integration over a trajectory. Then one considers the transformation $z = z(\zeta, t, \epsilon)$ as discussed. If the original (unperturbed) Hamiltonian is $H = H(\zeta, t)$, then the original equations of motion are

$$\frac{d\zeta}{dt} = \{\zeta, H(\zeta, t)\}. \tag{57}$$

On the other hand, if the transformed Hamiltonian is denoted $K = K(z, t, \epsilon)$, then the Hamiltonian equation of motion in the transformed variables is

$$\frac{dz}{dt} = \{z, K(z, t, \epsilon)\}. \tag{58}$$

Now using the chain rule on $z = z(\zeta, t, \epsilon)$,

$$\frac{dz}{dt} = \frac{\partial z}{\partial t} + \frac{\partial z}{\partial \zeta}\frac{d\zeta}{dt} \tag{59}$$

$$= \frac{\partial z}{\partial t} + \{z, H(\zeta, t)\} \tag{60}$$

using Eqn. 57. Combining the equations of motion leads to

$$\frac{\partial z}{\partial t} = \{z, K(z, t, \epsilon) - H(\zeta, t)\} \equiv \{z, R(z, t, \epsilon)\}. \tag{61}$$

Here the Poisson brackets are with respect to the unperturbed variables. Then we seek R such that $K = H + R$ relates the perturbed and exact Hamiltonians. Such a relationship is a typical Hamilton-Jacobi one, so that R can be seen as the derivative of a generating function. This will be seen shortly. Then Eqns. 51 and 61 describe the two governing PDEs of the T operator,

$$\frac{\partial T}{\partial \epsilon} = TL_S, \tag{62}$$

$$\frac{\partial T}{\partial t} = TL_R. \tag{63}$$

Equating second partials of the above, one finds

$$TL_RL_S + TL_{S_t} = TL_SL_R + TL_{R_\epsilon} \tag{64}$$

and as this holds for all T, from the Jacobi identity $L_SL_R - L_RL_S = L_{\{R,S\}}$ we have

$$\frac{\partial R}{\partial \epsilon} + \{R, S\} = \frac{\partial S}{\partial t}. \tag{65}$$

This being an inhomogeneous equation of the form of Eqn. 52, its solution is

$$T(\epsilon)R(\epsilon) = \int_0^\epsilon d\epsilon' T(\epsilon')\frac{\partial S}{\partial t} \tag{66}$$

assuming that $R(\epsilon = 0) = 0$ (corresponding to $K(z(\zeta, t, 0), t, 0) = H(\zeta, t)$). Now

$$T(\epsilon)R(\epsilon) = T(\epsilon)K(z, t, \epsilon) - T(\epsilon)H(\zeta, t) = T(\epsilon)K(z, t, \epsilon) - H(z, t). \tag{67}$$

and therefore

$$K(z, t, \epsilon) = T^{-1}(\epsilon)H(z, t) + T^{-1}(\epsilon)\int_0^\epsilon d\epsilon' T(\epsilon')\frac{\partial S}{\partial t} \tag{68}$$

is the desired result. This result is closely related to the classic Hamilton-Jacobi equation for a mixed-variables generating function. The methods are equivalent if the Lie generating function $S = S(z, t, \epsilon)$ is related to the canonical generating function F_2 by

$$S(q, p, t, \epsilon) = -\frac{\partial F_2(q, \bar{p}, t, \epsilon)}{\partial \epsilon}\Big|_{\bar{p}=p}. \tag{69}$$

This equivalence was apparently first noted by Shniad (1970) and found again by Dewar (1978). Noting that

$$\frac{\partial}{\partial \epsilon}(TK) = T\frac{\partial K}{\partial \epsilon} + TL_S K = T\frac{dK}{d\epsilon} \tag{70}$$

then the differential form of Eqn. 68 is

$$TL_S K + T\partial_\epsilon K = \partial_\epsilon H + T\partial_t S \tag{71}$$

or equivalently,

$$(\partial_\epsilon + L_S)K = T^{-1}H + \partial_t S. \tag{72}$$

Equation 72 is the form of the Hamilton-Jacobi equation to which the perturbation is applied. Making Deprit's choice of expansion in the perturbing parameter,

$$K = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} K_n, \tag{73}$$

$$S = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} S_{n+1}, \tag{74}$$

$$H = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} H_n, \tag{75}$$

$$T^{-1} = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} T_n^{-1}, \tag{76}$$

we have,

$$\partial_\epsilon K = \sum_{n=1}^{\infty} \frac{\epsilon^{n-1}}{(n-1)!} K_n = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} K_{n+1}, \tag{77}$$

$$L_S K = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} L_{n+1} \sum_{m=0}^{\infty} \frac{\epsilon^m}{m!} K_m = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \sum_{m=0}^{n} \binom{n}{m} L_{m+1} K_{n-m}, \tag{78}$$

$$T^{-1}\partial_\epsilon H = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} T_n^{-1} \sum_{m=0}^{\infty} \frac{\epsilon^m}{m!} H_{m+1} = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \sum_{m=0}^{n} \binom{n}{m} T_m^{-1} H_{n+1-m}, \tag{79}$$

$$\partial_t S = \sum_{n=0}^{\infty} \frac{\epsilon^n}{n!} \frac{\partial S_{n+1}}{\partial t}. \tag{80}$$

Then substitution into Eqn. 72 yields, by equating like-powers of ϵ,

$$K_{n+1} + \sum_{m=0}^{n} \binom{n}{m} L_{m+1} K_{n-m} = \sum_{m=0}^{\infty} \binom{n}{m} T_m^{-1} H_{n+1-m} + \partial_t S_{n+1} \tag{81}$$

To simplify, pull out the highest order from the first sum and the lowest order from the second, with $T_0^{-1} = 1$,

$$K_{n+1} + L_{n+1}K_0 + \sum_{m=0}^{n-1} \binom{n}{m} L_{m+1}K_{n-m} = H_{n+1} + \sum_{m=1}^{n} \binom{n}{m} T_m^{-1} H_{n+1-m} + \partial_t S_{n+1} \tag{82}$$

and note that $K_0 = H_0$, so that the two terms

$$\partial_t S_{n+1} - L_{n+1}H_0 = (\partial_t + L_{H_0})S_{n+1} = \frac{dS_{n+1}}{dt} \tag{83}$$

combine as the total derivative of S_{n+1} along the unperturbed trajectories. Then shuffling the index $n+1 \to n$, a Hamilton-Jacobi like equation emerges to each order,

$$\frac{dS_n}{dt} = K_n - H_n + \sum_{m=1}^{n-1} \Big[\binom{n-1}{m-1} L_m K_{n-m} - \binom{n-1}{m} T_m^{-1} H_{n-m} \Big]. \tag{84}$$

This is the most useful result. It finally justifies choosing the expansion of S as beginning at S_1. It must be noted that all arguments are evaluated at the *transformed* variables. It is sometimes stated that this is a result of "operations on functions, rather than on variables". The first few equations for the generating function are

$$K_0 = H_0, \tag{85}$$
$$\frac{dS_1}{dt} = K_1 - H_1, \tag{86}$$
$$\frac{dS_2}{dt} = K_2 - H_2 + \{H_1 + K_1, S_1\}, \tag{87}$$
$$\frac{dS_3}{dt} = H_3 - K_3 + \{K_2 + 2H_2, S_1\} + \{2K_1 + H_1, S_2\} - \{\{H_1, S_1\}, S_1\}, \tag{88}$$
$$\vdots \qquad \vdots \tag{89}$$

Also, the transformation relating the canonical variables has already been found, $z = T\zeta$. The near-identity transformation evolving the canonical variables with $\zeta = (q, p)$ and $z = (\bar{q}, \bar{p})$ is

$$\bar{q}_0 = q, \tag{90}$$
$$\bar{q}_1 = \frac{\partial S_1}{\partial p}, \tag{91}$$
$$\bar{q}_2 = \frac{\partial S_2}{\partial p} + \{\frac{\partial S_1}{\partial p}, S_1\}, \tag{92}$$
$$\vdots \qquad \vdots \tag{93}$$
$$\bar{p}_0 = p, \tag{94}$$
$$\bar{p}_1 = -\frac{\partial S_1}{\partial q}, \tag{95}$$
$$\bar{p}_2 = -\frac{\partial S_2}{\partial q} - \{\frac{\partial S_1}{\partial q}, S_1\}, \tag{96}$$
$$\vdots \qquad \vdots \tag{97}$$

2.1 Example: analyzing the pendulum by Deprit's Lie series

(Ref: Lichtenberg and Lieberman) A classic nonlinear problem is the simple pendulum,

$$H = \frac{1}{2}p^2 + 1 - \cos(q). \tag{98}$$

While power series solutions of $p(t)$, $q(t)$ can be developed (in fact $p(t)$ can be solved in terms of elliptic functions), one can also treat the problem perturbatively. Expanding the potential, one has

$$1 - \cos(q) = \frac{q^2}{2} - \epsilon\frac{q^4}{4!} + \epsilon^2\frac{q^6}{6!} + \cdots \tag{99}$$

after inserting an ordering parameter ϵ. Then

$$H = \frac{1}{2}p^2 + \frac{1}{2}q^2 - \epsilon\frac{q^4}{4!} + \epsilon^2\frac{q^6}{6!} + \cdots . \tag{100}$$

To zeroth order there is simple harmonic oscillation, followed by higher-order nonlinearities. The problem is best treated in action-angle coordinates $(q, p) \to (\theta, J)$. Using the generating function

$$F_1(q, \theta) = \frac{1}{2}q^2\frac{\cos\theta}{\sin\theta} \tag{101}$$

one has

$$p = \frac{\partial F_1}{\partial q} = q\frac{\cos\theta}{\sin\theta}, \tag{102}$$

$$J = -\frac{\partial F_1}{\partial \theta} = \frac{1}{2}q^2\frac{1}{\sin^2\theta}, \tag{103}$$

so that upon rearranging,

$$J = \frac{1}{2}q^2 + \frac{1}{2}p^2, \tag{104}$$

$$\epsilon^{n-1}\frac{q^{2n}}{(2n)!} = \epsilon^{n-1}\frac{2^n}{(2n)!}J^n\sin^{2n}\theta, \tag{105}$$

and therefore

$$H = J - \frac{\epsilon}{6}J^2\sin^4\theta + \frac{\epsilon^2}{90}J^3\sin^6\theta + \cdots \tag{106}$$

Then in terms of the exact Hamiltonian, the successive powers of ϵ are

$$H_0 = J, \tag{107}$$

$$H_1 = -\frac{J^2}{6}\sin^4\theta = -\frac{J^2}{48}(3 - 4\cos(2\theta) + \cos(4\theta)), \tag{108}$$

$$H_2 = 2\frac{J^3}{90}\sin^6\theta = \frac{J^3}{1440}(10 - \cos(6\theta) + 6\cos(4\theta) - 15\cos(2\theta)), \tag{109}$$

$$\vdots \tag{110}$$

The problem is time-independent, so we seek a Lie transform generating function $S = S(J, \theta, \epsilon)$. Then $\frac{dS_1}{dt} = \partial_t S_1 + \{S_1, H_0\} = \{S_1, J\} = \partial_\theta S_1$. Thus the first of the perturbation equations is

$$\frac{\partial S_1}{\partial \theta} = K_1 - H_1. \tag{111}$$

By the Fredholm alternative theorem, the solvability condition is that the RHS be orthogonal to the null-space of the adjoint operator. Since $(\partial_\theta)^* = -\partial_\theta$, the Fredholm alternative equation is

$$-\partial_\theta f = 0 \qquad \Longrightarrow \qquad f = f(J) + c \tag{112}$$

for any function $f(J)$ and constant c. And for the right-hand side to be orthogonal to this null-space,

$$\langle K_1 - H_1, f(J) + c\rangle = 0 \tag{113}$$
$$\Longrightarrow \int_0^{2\pi} (K_1 - H_1)(f(J) + c) d\theta = 0 \tag{114}$$
$$\Longrightarrow \frac{1}{2\pi}\int_0^{2\pi} K_1 d\theta = \frac{1}{2\pi}\int_0^{2\pi} H_1 d\theta \tag{115}$$
$$\Longrightarrow \langle K_1\rangle_\theta = \langle H_1\rangle_\theta \tag{116}$$

where $\langle\cdot\rangle_\theta$ is an angle-average. Then the solvability condition can be satisfied by setting

$$H_1 = \langle H_1\rangle + \widetilde{H}_1 \tag{117}$$

with $\langle\cdot\rangle_\theta$ the angle-average and $\widetilde{\cdot}$ the fluctuating part, and choosing

$$K_1 = \langle H_1\rangle_\theta = -\frac{J^2}{16} \tag{118}$$

because $\langle K_1\rangle_\theta = \langle\langle H_1\rangle_\theta\rangle_\theta = \langle H_1\rangle_\theta$. Then the first-order generating function follows by integration,

$$\frac{\partial S_1}{\partial\theta} = -\widetilde{H}_1 = \frac{J^2}{48}(\cos(4\theta) - 4\cos(2\theta)) \tag{119}$$
$$\Longrightarrow S_1 = \frac{J^2}{192}(\sin(4\theta) - 8\sin(2\theta)). \tag{120}$$

Now the second-order perturbation equation is

$$\frac{\partial S_2}{\partial\theta} = K_2 - H_2 + \{H_1 + K_1, S_1\}. \tag{121}$$

Similarly, the solvability condition here can only be satisfied by removing the oscillation-average

$$K_2 = \frac{1}{2\pi}\int_0^{2\pi} (H_2 - \{K_1 + H_1, S_1\}) d\theta. \tag{122}$$

Examining the Poisson bracket terms, the first is

$$\langle\{K_1, S_1\}\rangle_\theta = \langle\{\langle H_1\rangle_\theta, S_1\}\rangle_\theta = -\left\langle\frac{\partial H_1}{\partial J}\right\rangle_\theta \left\langle\frac{\partial S_1}{\partial\theta}\right\rangle_\theta = 0 \tag{123}$$

because $\langle S_1\rangle_\theta = 0$. In the first, since $\langle\{\langle H_1\rangle, S_1\}\rangle_\theta = 0$, then as

$$\{\widetilde{H}_1, S_1\} = \frac{\partial S_1}{\partial J}\frac{\partial^2 S_1}{\partial\theta^2} - \frac{1}{2}\frac{\partial \widetilde{H}_1^2}{\partial J} = \frac{J^3}{1152}(17 - 9\cos(2\theta) + \cos(6\theta)), \tag{124}$$

the average is simply the constant part of this. So,

$$K_2 = \langle H_2\rangle_\theta + \langle\{\widetilde{H}_1, S_1\}\rangle_\theta = \frac{10}{1440}J^3 - \frac{17}{1152}J^3 = -\frac{J^3}{128}. \tag{125}$$

Therefore the energy to second order is

$$K = K_0 + \epsilon K_1 + \frac{1}{2}\epsilon^2 K_2 + \mathcal{O}(\epsilon^3) \tag{126}$$

$$= \bar{J} - \epsilon\frac{\bar{J}^2}{16} - \epsilon^2\frac{\bar{J}^3}{256} + \mathcal{O}(\epsilon^3) \tag{127}$$

recalling that the entire calculation was really done in terms of the transformed variable ("operations on functions"). To first order, the new action and angle variables are

$$\bar{\theta} = \theta + \epsilon\frac{\partial S_1}{\partial J} + \mathcal{O}(\epsilon^2), \tag{128}$$

$$\bar{J} = J + \epsilon\widetilde{H}_1 + \mathcal{O}(\epsilon^2). \tag{129}$$

The result that the first-order correction in action is the fluctuating part of the first-order exact Hamiltonian is a significant theme and is returned to later in the context of ponderomotive force.

3 Dragt and Finn's method

Deprit's method formulates the problem of Lie transformation as a power series technique to solve, effectively, a system of nonautonomous differential equations. Close examination of Eqn. 31, the recurrence relation for Deprit's evolution operator T, shows that upon iteration the number of terms in the operator T_n is 2^{n-1}. Dragt and Finn (1978) proposed a method whose complexity grows much less quickly at higher order.

The Lie transformation operator T is to be formulated in terms of a succession of generating functions $g_1, g_2, \cdots$ and ordered transformations $\exp(\epsilon L_{g_1}), \exp(\epsilon^2 L_{g_2}), \cdots$ such that

$$T = \cdots \exp(\epsilon^3 L_{g_3}) \exp(\epsilon^2 L_{g_2}) \exp(\epsilon L_{g_1}) \tag{130}$$

so that each generating function g_n appears at n'th order. To illustrate, seek an expansion of T in the ordering parameter,

$$T = I + \sum_{n=1}^{\infty} \epsilon^n T_n \tag{131}$$

and expand each exponential in Eqn. 130 in power series,

$$T = \cdots \Big(\sum_{n=0}^{\infty} \frac{\epsilon^{3n}}{n!} L_{g_3}^n\Big)\Big(\sum_{m=0}^{\infty} \frac{\epsilon^{2m}}{m!} L_{g_2}^m\Big)\Big(\sum_{k=0}^{\infty} \frac{\epsilon^{k}}{k!} L_{g_1}^k\Big). \tag{132}$$

Equating the ordering expansion and the power series expansion, the first few terms are

$$T_1 = \frac{1}{1!}L_{g_1}, \tag{133}$$

$$T_2 = \frac{1}{1!}L_{g_2} + \frac{1}{2!}L_{g_1}^2, \tag{134}$$

$$T_3 = \frac{1}{1!}L_{g_3} + \frac{1}{1!1!}L_{g_2}L_{g_1} + \frac{1}{3!}L_{g_1}^3, \tag{135}$$

$$T_4 = \frac{1}{1!}L_{g_4} + \frac{1}{1!1!}L_{g_3}L_{g_1} + \frac{1}{2!}L_{g_2}^2 + \frac{1}{1!2!}L_{g_2}L_{g_1}^2 + \frac{1}{4!}L_{g_1}^4, \tag{136}$$

$$\vdots \quad \vdots \tag{137}$$

so that the pattern is apparently

$$T_n = \sum_{m_1+2m_2+\cdots+nm_n=n} \frac{1}{m_n!\ldots m_2!m_1!} L_{g_n}^{m_n} \cdots L_{g_2}^{m_2} L_{g_1}^{m_1} \tag{138}$$

and the number of terms in the n'th operator is $p(n)$, the partitions of n. While $p(n)$ grows quickly, it is always less than the Deprit series' size of 2^{n-1} as demonstrated through the Hardy-Ramanujan asymptotic formula for partitions, *i.e.* $\lim_{n\to\infty} p(n) \sim \frac{1}{4\sqrt{3}n} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$. In this sense the Dragt-Finn method is more direct at high order, as far fewer Poisson bracket iterations are required.

The correct form of the inverse transformation is

$$T^{-1} = \cdots \exp(-\epsilon^3 L_{g_3}) \exp(-\epsilon^2 L_{g_2}) \exp(-\epsilon L_{g_1}) \tag{139}$$

which apparently does not give $TT^{-1} = I$ with the given ordering, but that's because it is really a pull-back of the inverse transform, explained in an upcoming section. Then by power series,

$$T_n^{-1} = \sum_{m_1+2m_2+\cdots+nm_n=n} \frac{(-1)^{m_n+\cdots+m_2+m_1}}{m_n!\ldots m_2!m_1!} L_{g_n}^{m_n} \cdots L_{g_2}^{m_2} L_{g_1}^{m_1} \tag{140}$$

and the first few terms are

$$T_1^{-1} = -\frac{1}{1!} L_{g_1}, \tag{141}$$

$$T_2^{-1} = -\frac{1}{1!} L_{g_2} + \frac{1}{2!} L_{g_1}^2, \tag{142}$$

$$T_3^{-1} = -\frac{1}{1!} L_{g_3} + \frac{1}{1!1!} L_{g_2} L_{g_1} - \frac{1}{3!} L_{g_1}^3, \tag{143}$$

$$T_4^{-1} = -\frac{1}{1!} L_{g_4} + \frac{1}{2!} L_{g_2}^2 - \frac{1}{1!2!} L_{g_2} L_{g_1}^2 + \frac{1}{1!1} L_{g_3} L_{g_1} + \frac{1}{4!} L_{g_1}^4, \tag{144}$$

$$\vdots \quad \vdots \tag{145}$$

To apply the Dragt-Finn method to a time-dependent Hamiltonian system one must go to extended phase space so that the system is autonomous through parameterization. The perturbation analysis can then be conducted through direct calculation. The section following that introduces a more elegant method which perturbs the symplectic 1-form, reproducing the direct approach.

4 Extended phase space

All that is eternal is a simultaneous whole.

- *Proclus* ($\sim$ 460 CE) expressing the (incorrect?) determinist paradigm, translated by E.R. Dodds.

References: Variational Principles of Mechanics - Cornelius Lanczos
Lagrange found that the equations of motion follow from extremizing the action integral,

$$\mathcal{A} = \int_{t_0}^{t_1} \mathcal{L}(q^1, \cdots, q^n, \dot{q}^1, \cdots, \dot{q}^n) dt \tag{146}$$

where $\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is the Lagrangian function on the configuration space. Further, the momenta p_i canonically conjugate to the coordinate variables q^i are given by the partial derivatives

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^i}. \tag{147}$$

The Lagrangian system on the tangent bundle is taken to a Hamiltonian one on the cotangent bundle by a Legendre transform (*i.e.* one which carries a function on a space into its dual),

$$\mathcal{H} = \sum_{i=1}^{n} p_i \dot{q}^i - \mathcal{L} \tag{148}$$

where the Hamiltonian function $\mathcal{H}$ represents the total energy of the mechanical system. When $\mathcal{H}$ depends on time (when energy is not conserved) the equations of motion become nonautonomous, becoming more difficult to solve. In solving a system of nonautonomous differential equations one would parameterize time and return to the autonomous case with one extra variable. We do the same in Hamiltonian mechanics by going to the *extended phase space*. That is, consider time to be an extra mechanical coordinate by setting $q^{n+1} = t$ and parameterizing time by $t(\tau)$. Under this reparametrization, the action integral becomes (with $\mathcal{L}_x$ for "extended"),

$$\mathcal{A} = \int_{\tau_0}^{\tau_1} \mathcal{L}\Big(q^1, \cdots, q^n, \frac{(q^1)'}{(q^{n+1})'}, \cdots \frac{(q^n)'}{(q^{n+1})'}\Big)(q^{n+1})' d\tau \equiv \int_{\tau_0}^{\tau_1} \mathcal{L}_x d\tau \tag{149}$$

where $(q^i)' \equiv \frac{dq^i}{d\tau}$. The function $\mathcal{L}_x$ is homogeneous of degree 1 in the reparametrized velocities $\Big\{(q^1)', (q^2)', \cdots, (q^{n+1})'\Big\}$, that is,

$$\mathcal{L}_x(q^1, \cdots q^n, \alpha(q^1)', \cdots, \alpha(q^{n+1})') = \mathcal{L}\Big(q^1, \cdots, q^n, \frac{(\alpha q^1)'}{(\alpha q^{n+1})'}, \cdots \frac{(\alpha q^n)'}{(\alpha q^{n+1})'}\Big)(\alpha q^{n+1})' \tag{150}$$

$$= \alpha \mathcal{L}_x \tag{151}$$

and therefore by Euler's theorem on homogeneous functions, it can be represented by

$$\mathcal{L}_x = \sum_{i=1}^{n+1} \frac{\partial \mathcal{L}_x}{\partial (q^i)'}(q^i)' = \sum_{i=1}^{n+1} p_i (q^i)'. \tag{152}$$

Then upon Legendre transformation of the extended Lagrangian $\mathcal{L}_x$, one finds

$$\mathcal{H}_x = \sum_{i=1}^{n+1} p_i (q^i)' - \mathcal{L}_x = 0 \tag{153}$$

so that the extended Hamiltonian $\mathcal{H}_x$ is identically zero. This does not mean that the coordinates $q(\tau), p(\tau)$ are constants. Rather, the solution is restricted to the energy surface $\mathcal{H}_x = 0$ which still has non-zero partial derivatives. To find the momentum conjugate to time, examine Eqn. 148 under the change of variables,

$$\mathcal{H} = \sum_{i=1}^{n} p_i \frac{(q^i)'}{(q^{n+1})'} - \mathcal{L} \quad \Longrightarrow \quad \mathcal{H}(q^{n+1})' = \sum_{i=1}^{n} p_i (q^i)' - \mathcal{L}(q^{n+1})' \tag{154}$$

and so as the extended Hamiltonian can be written (with $\mathcal{L}(q^{n+1})' = \mathcal{L}_x$)

$$\mathcal{H}_x = \sum_{i=1}^{n} p_i (q^i)' - \mathcal{H}(q^{n+1})' - \mathcal{L}_x \tag{155}$$

the momentum canonically conjugate to time is the negative of the original Hamiltonian,

$$p_{n+1} \equiv p_t = -\mathcal{H}. \tag{156}$$

The condition that $\mathcal{H}_x = 0$ should be seen as an auxiliary condition to the equations of motion, and can be interpreted in a number of ways (corresponding to the various mechanical minimization principles [Lanczos]). A particularly useful choice is to identify the extended Hamiltonian with

$$\mathcal{H}_x = p_t + \mathcal{H}. \tag{157}$$

Under this choice, the "most advanced" parameterized canonical equations of motion,

$$\frac{dq^i}{d\tau} = \frac{\partial \mathcal{H}_x}{\partial p_i}, \tag{158}$$

$$\frac{dp_i}{d\tau} = -\frac{\partial \mathcal{H}_x}{\partial q^i}, \tag{159}$$

are written in the expected simple, parameterized form in the $(2n+2)$-dimensional phase space

$$\frac{dq^i}{d\tau} = \frac{\partial \mathcal{H}}{\partial p_i}, \quad i \leqslant n, \tag{160}$$

$$\frac{dp_i}{d\tau} = -\frac{\partial \mathcal{H}}{\partial q_i}, \quad i \leqslant n, \tag{161}$$

$$\frac{dt}{d\tau} = 1, \tag{162}$$

$$\frac{dp_t}{d\tau} = -\frac{\partial \mathcal{H}}{\partial t}. \tag{163}$$

The latter two equations appear to be trivial, but in a relativistic treatment they are not. Also, any parameterization of time can be chosen and this is only the simplest choice.

Now just as the fundamental 1-form γ is of primary importance in unextended phase space

$$\gamma = p_i dq^i - \mathcal{H} dt \tag{164}$$

by leading to the symplectic 2-form ω^2 (phase space measure) as the exterior derivative of γ,

$$\omega^2 = d\gamma = dp_i \wedge dq^i - d\mathcal{H} \wedge dt \tag{165}$$

and hence to the equations of motion [Littlejohn], so too is the the fundamental form on the extended phase space of primary interest [Brizard]. It is given by

$$\gamma_E = p_i dq^i + p_t dt - \mathcal{H}_x d\tau, \tag{166}$$

$$= p_i dq^i + p_t dt - (p_t + \mathcal{H}) d\tau. \tag{167}$$

The extended phase space fundamental 1-form is theoretically very useful as the perturbation theory can be applied directly, in the very same way as techniques for an autonomous system of ODEs can be applied to nonautonomous ones by parametrization.

In addition, the extended Poisson bracket of two functions f, g is given by

$$\{f, g\}_x = \sum_{i=1}^{n} \Big(\frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \Big) + \frac{\partial f}{\partial t}\frac{\partial g}{\partial p_t} - \frac{\partial f}{\partial p_t}\frac{\partial g}{\partial t} \tag{168}$$

so the extended Poisson bracket of a function f with the extended Hamiltonian $\mathcal{H}_x$ must be

$$\{f, \mathcal{H}_x\}_x = \{f, \mathcal{H}\} + \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathcal{H}}\frac{\partial \mathcal{H}}{\partial t} \tag{169}$$

where $\{\cdot, \cdot\}$ is the unextended Poisson bracket, giving the total time derivative if $\frac{\partial f}{\partial \mathcal{H}} = 0$.

4.1 Hamiltonian perturbation theory in extended phase space

The extended phase space autonomizes a time-dependent Hamiltonian system, so that the gymnastics leading to Eqn. 72 are not necessary. Evolve the initial system points $\zeta = (q^1, p_1, \cdots, t, p_t)$ by a Lie transformation $z = T\zeta$, *i.e*

$$z(q^1, p_1, \cdots, t, p_t, \epsilon) = T(\epsilon)\zeta(q^1, p_1, \cdots, t, p_t) \tag{170}$$

using the Dragt-Finn choice of transformation operator $T(\epsilon)$ with extended Poisson brackets $\{\cdot,\cdot\}_x$ in the Lie series for the generating functions g_n. That is, the L operators in T are $L_{g_n} = \{\cdot, g_n\}_x$. Since the Hamiltonian system transformed is autonomous and the transformation is canonical, the transformed Hamiltonian $\mathcal{K}_x$ is related to the original Hamiltonian $\mathcal{H}_x$ by

$$\mathcal{K}_x(z) = \mathcal{H}_x(\zeta) \tag{171}$$

so that using Gröbner's commutation theorem, $\mathcal{K}_x(z) = \mathcal{K}_x(T\zeta) = T\mathcal{K}_x(\zeta)$, one has simply

$$\mathcal{K}_x(\zeta) = T^{-1}(\epsilon)\mathcal{H}_x(\zeta). \tag{172}$$

Expand both the transformed and original Hamiltonians in power series,

$$\mathcal{K}_x = \sum_{n=0}^{\infty} \epsilon^n \mathcal{K}_n, \tag{173}$$

$$\mathcal{H}_x = p_t + \sum_{n=0}^{\infty} \epsilon^n \mathcal{H}_n, \tag{174}$$

choosing that the time-momentum p_t will appear only in the lowest-order part of the extended Hamiltonian, *i.e.* $\mathcal{H}_{x,0} = p_t + \mathcal{H}_0$. Now using the relations of Section 3 for the inverse transformation T^{-1}, one obtains the perturbation equations,

$$\sum_{n=0}^{\infty} \epsilon^n \mathcal{K}_n = \sum_{j=0}^{\infty} \epsilon^j T_j^{-1} \sum_{k=0}^{\infty} \epsilon^k \mathcal{H}_{x,k} \tag{175}$$

$$= \sum_{n=0}^{\infty} \epsilon^n \sum_{m=0}^{n} T_{n-m}^{-1} \mathcal{H}_{x,m} \tag{176}$$

$$\Longrightarrow \mathcal{K}_n = \sum_{m=0}^{n} T_{n-m}^{-1} \mathcal{H}_{x,m}. \tag{177}$$

For example, for $n = 1$ the relation is

$$\mathcal{K}_1 = T_1^{-1}\mathcal{H}_{x,0} + T^{-1}\mathcal{H}_1 \tag{178}$$

$$= -L_{g_1}\mathcal{H}_{x,0} + \mathcal{H}_1 \tag{179}$$

and the derivative term, acting on $\mathcal{H}_{x,0} = p_{n+1} + \mathcal{H}_0$, is

$$-L_{g_1}\mathcal{H}_{x,0} = -\{\mathcal{H}_{x,0}, g_1\}_x = \{g_1, \mathcal{H}_{x,0}\}_x = \{g_1, \mathcal{H}_0\} + \frac{\partial g_1}{\partial t} \tag{180}$$

by *choosing* $\frac{\partial g_1}{\partial \mathcal{H}_0} = 0$, so that the transformation leaves the time coordinate unperturbed (for this non-relativistic analysis). In this way, the first three perturbation equations are

$$\mathcal{K}_0 = \mathcal{H}_{0,x}, \tag{181}$$
$$\frac{\partial g_1}{\partial t} + \{g_1, \mathcal{H}_0\} = \mathcal{K}_1 - \mathcal{H}_1, \tag{182}$$
$$\frac{\partial g_2}{\partial t} + \{g_2, \mathcal{H}_0\} = \mathcal{K}_2 - \mathcal{H}_2 + \frac{1}{2} L_{g_1}(\mathcal{H}_1 + \mathcal{K}_1), \tag{183}$$
$$\frac{\partial g_3}{\partial t} + \{g_3, \mathcal{H}_0\} = \mathcal{K}_3 - \mathcal{H}_3 + L_{g_1}\mathcal{H}_2 + L_{g_2}\mathcal{K}_1 - \frac{1}{3} L_{g_1}^2 \mathcal{H}_1 - \frac{1}{6} L_{g_1}^2 \mathcal{K}_1, \tag{184}$$
$$\vdots \quad \vdots \tag{185}$$

found by substituting the previous results into each order, and so on. Observe that this result is quite similar to that of Deprit's series, but derived much more simply. Also at higher order much fewer terms appear (not appearing here as 3rd-order is not high enough to illustrate). Here the coefficients are different as well because, for convenience, the factorial factors are not used in the denominator of the power series expansions for $\mathcal{H}_n$, $\mathcal{K}_n$.

4.2 Example: return of the pendulum

To illustrate the Dragt-Finn method the pendulum example is repeated here. As it is an autonomous problem the time derivatives vanish from the perturbation hierarchy. The only main difference is that, since factorials are not assumed in the expansion denominators, the expansion of the Hamiltonian in action-angle coordinates has a factor $1/n!$, that is,

$$\mathcal{H}_0 = J, \tag{186}$$
$$\mathcal{H}_1 = -\frac{J^2}{6} \sin^4 \theta = -\frac{J^2}{48}(3 - 4\cos(2\theta) + \cos(4\theta)), \tag{187}$$
$$\mathcal{H}_2 = \frac{J^3}{90} \sin^6 \theta = \frac{J^3}{2880}(10 - \cos(6\theta) + 6\cos(4\theta) - 15\cos(2\theta)), \tag{188}$$
$$\vdots \quad \vdots \tag{189}$$

As before, the solvability condition on the first-order perturbation equation

$$\frac{\partial g_1}{\partial \theta} = \mathcal{K}_1 - \mathcal{H}_1 \tag{190}$$

is that $\mathcal{K}_1 = \langle H_1 \rangle_\theta = -\frac{J^2}{16}$, so

$$g_1 = \frac{J^2}{192}(\sin(4\theta) - 8\sin(2\theta)). \tag{191}$$

The second-order equation is

$$\frac{\partial g_2}{\partial \theta} = \mathcal{K}_2 - \mathcal{H}_2 + \frac{1}{2}\{\mathcal{H}_1 + \mathcal{K}_1, g_1\} \tag{192}$$

so that, as before,

$$\frac{1}{2}\{\mathcal{H}_1 + \mathcal{K}_1, g_1\} = \frac{1}{2}\{\widetilde{H}_1, g_1\} = \frac{J^3}{2304}(17 - 9\cos(2\theta) + \cos(6\theta)). \tag{193}$$

Now choosing K_2 so that the perturbation equation is solvable,

$$K_2 = \langle H_2 \rangle_\theta + \frac{1}{2}\langle\{\widetilde{H}_1, g_1\}\rangle_\theta = \frac{10J^3}{2880} - \frac{17J^3}{2304} = -\frac{J^3}{256} \tag{194}$$

the pendulum energy to second-order is

$$\mathcal{K} = \bar{J} - \epsilon\frac{\bar{J}^2}{16} - \epsilon^2\frac{\bar{J}^3}{256} + \mathcal{O}(\epsilon^3) \tag{195}$$

just as before. Further, the evolution of the variables follows in the very same manner with,

$$\bar{\theta} = T\theta = \theta + \epsilon\frac{\partial g_1}{\partial J} + \mathcal{O}(\epsilon^2), \tag{196}$$

$$\bar{J} = TJ = J + \epsilon\widetilde{H}_1 + \mathcal{O}(\epsilon^2). \tag{197}$$

5 Perturbation theory with differential forms

This section presents a more general and extensible take on Hamiltonian perturbation theory, beginning with why the general method works, presenting the method, and then demonstrating its equivalence with the findings of the last section. The initial observation is that the equations of motion are direct consequences of the symplectic form

$$\omega^2 = d\gamma = dp_i \wedge dq^i - d\mathcal{H} \wedge dt \tag{198}$$

so that the basic object to be perturbed is γ, not $\mathcal{H}$. Its component matrix ω_{ij}, where $\omega^2 = \omega_{ij} dx^i \wedge dx^j$, is given by the "Lagrange brackets" (dual to the Poisson brackets) of the coordinates. That is, it satisfies

$$J^{ij}\omega_{jk} = \delta^i_k \tag{199}$$

where $J^{ij} = \{z^i, z^j\}$ is the matrix of coordinate Poisson brackets, sometimes called the Poisson tensor. For this reason ω_{ij} is also called the Lagrange tensor. The Lagrange brackets have generally fallen out of use and so are not given. In canonical coordinates, this $(2n+1)\times(2n+1)$ matrix has the simple form

$$\omega_{ij} = \begin{bmatrix} 0 & -I & -\mathcal{H}_q \\ +I & 0 & -\mathcal{H}_p \\ \mathcal{H}_q & \mathcal{H}_p & 0 \end{bmatrix} \tag{200}$$

where $\mathcal{H}_i$ is partial derivative with respect to variable i. It posseses a null eigenvector

$$\begin{bmatrix} 0 & -I & -\mathcal{H}_q \\ +I & 0 & -\mathcal{H}_p \\ \mathcal{H}_q & \mathcal{H}_p & 0 \end{bmatrix} \begin{bmatrix} H_p \\ -H_q \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{201}$$

This null vector field generates a flow by its corresponding system of differential equations

$$\frac{dq}{ds} = \frac{\partial\mathcal{H}}{\partial p}, \tag{202}$$

$$\frac{dp}{ds} = -\frac{\partial\mathcal{H}}{\partial q}, \tag{203}$$

$$\frac{dt}{ds} = 1. \tag{204}$$

The significance of it being a null-field is that nowhere do the trajectories lie entirely in the (p, q)-plane for some t. In analogy with the kinematics of an ideal fluid and the generalized Stokes theorem, these trajectories can be considered as the vortex lines of the fundamental 1-form γ. So in a sense, it is really the fundamental form γ that generates the equations of motion through its derivative $\omega = d\gamma$ in the same way that electrodynamics may be understood through the dynamics of the potential. Now as γ is a potential, the symplectic form ω is invariant under gauge transformations

$$\bar{\gamma} \equiv \gamma + dS \quad \Longrightarrow \quad \bar{\omega} = d\bar{\gamma} = d\gamma = \omega \tag{205}$$

because $d(dS) = 0$. This is the basis of the theory of canonical transformations. For example, consider a gauge transformation which brings the transformed Hamiltonian to some other form. That is, choose $S = S(q, \bar{q}, t)$ so that $dS = S_q dq + S_{\bar{q}} d\bar{q} + S_t dt$, and so

$$\bar{\gamma} = \gamma - dS(q, \bar{q}, t) \tag{206}$$

$$\bar{p} d\bar{q} - K dt = p dq - H dt - dS(q, \bar{q}, t), \tag{207}$$

$$(\bar{p} + S_{\bar{q}}) d\bar{q} - K dt = (p - S_q) dq - (H + S_t) dt, \tag{208}$$

so that matching values of the basis 1-forms,

$$p = \frac{\partial S(q, \bar{q}, t)}{\partial q}, \tag{209}$$

$$\bar{p} = -\frac{\partial S(q, \bar{q}, t)}{\partial \bar{q}}, \tag{210}$$

$$K(\bar{q}, \bar{p}, t) = H(q, p, t) + \frac{\partial S(q, \bar{q}, t)}{\partial t}, \tag{211}$$

reproduces the classical canonical transformation with a mixed-variables generating function *through a gauge transformation of the symplectic potential.* This is a differential forms perspective on Jacobi's theory of canonical transformations, in which case a total differential leaves the action integral invariant [Lanczos]. As a 1-form, integrating γ to find the action on some curve is always possible, as differential forms are after all "whatever appear underneath an integral sign". So the two approaches are of course equivalent.

5.1 Perturbation theory

The gauge functions S must be chosen as mixed-variables to preserve the equations of motion in a simple way. However, if the transformation is near-identity then the gauge function can be in terms of the original variables only[2]. This fact is the basis of perturbation theory.

Generate a canonical transformation $\bar{z} = Tz$ by a Lie transformation operator T,

$$T = \cdots T_3 T_2 T_1 \tag{212}$$

in the framework of the Dragt-Finn method, as in Section 3. The transformation T induces a push-forward T_* on vector fields and a pull-back T^* on differential forms,

$$T_{n*} = \exp(-\epsilon^n L_n), \tag{213}$$

$$T_n^* = \exp(\epsilon^n L_n), \tag{214}$$

[2]Considering infinitesimal canonical transformations actually leads to understanding the evolutionary phase flow itself as a sequence of near-identity transformations. In fact this approach reproduces Hamilton's principal function S and the Hamilton-Jacobi equation $\frac{\partial S}{\partial t} + H = 0$. [Lanczos]

where the inverse pull-back is given by

$$T^{-1} = T_1^{-1}T_2^{-1}T_3^{-1}\cdots, \tag{215}$$

$$T^{*-1} = \cdots T_3^{*-1}T_2^{*-1}T_1^{*-1}. \tag{216}$$

As differential forms pull back under transformations, consider how the symplectic form $\omega = T^*\bar{\omega}$ transforms. In terms of transformed vector fields $\bar{X} = T_*X, \bar{Y} = T_*Y$,

$$\bar{\omega}(\bar{X},\bar{Y}) = \bar{\omega}(T_*X, T_*Y) = T^*\bar{\omega}(X,Y) = \omega(X,Y) \tag{217}$$

so the symplectic form, and the equations of motion, are preserved under the Lie transformation. Now since the exterior derivative commutes with the pull-back, we also have

$$\bar{\omega} = d\bar{\gamma} = d(T^{*-1}\gamma) = T^{*-1}d\gamma = T^{*-1}\omega \tag{218}$$

so it suffices to transform the symplectic potential alone. Addition of a gauge transformation produces a more abstract and elegant formulation of Lie transform perturbation theory,

$$\bar{\gamma} = T^{*-1}\gamma + dS. \tag{219}$$

Now the perturbation equations follow directly from considering power series expansions,

$$\bar{\gamma} = \sum_{n=0}^{\infty} \epsilon^n \bar{\gamma}_n, \tag{220}$$

$$\gamma = \sum_{n=0}^{\infty} \epsilon^n \gamma_n, \tag{221}$$

$$S = \sum_{n=0}^{\infty} \epsilon^n S_n, \tag{222}$$

$$T^{*-1} = \sum_{n=0}^{\infty} \epsilon^n T_n^{*-1} \tag{223}$$

where γ has a known expansion, the Dragt-Finn expansion follows from Section 3, and the gauge function S is to be chosen. Substitution yields the perturbation equations

$$\bar{\gamma}_0 = \gamma_0 + dS_0, \tag{224}$$

$$\bar{\gamma}_1 = \gamma_1 + L_{S_1}\gamma_0 + dS_1, \tag{225}$$

$$\bar{\gamma}_2 = \gamma_2 + L_{S_1}\gamma_1 + (\frac{1}{2}L_{S_1}^2 + L_{S_2})\gamma_0 + dS_2, \tag{226}$$

$$\vdots \qquad \vdots \tag{227}$$

5.2 Demonstration of equivalence with direct method

Now each Lie derivative $L_{G_1} \equiv L_1$ can be simplified using Cartan's formula,

$$L_G\gamma = \imath_G d\gamma + d(\imath_G\gamma). \tag{228}$$

The second term is an exact differential and so does not influence the transformation. Only the first term needs to be considered. With the vector field in components,

$$G = G^q\partial_q + G^p\partial_p + G^t\partial_t \tag{229}$$

that interior product is[3] (with $\gamma = pdq - Hdt \implies d\gamma = dp \wedge dq - dH \wedge dt$),

$$\imath_G(dp \wedge dq) = G^p dq - G^q dp \tag{230}$$

$$\imath_G(H_q dq \wedge dt) = H_q G^q dt - H_q G^t dq \tag{231}$$

$$\implies \imath_G d\gamma = (G^p - H_q G^t)dq + (-G^q - H_p G^t)dp - (H_q G^q + H_p G^p)dt \tag{232}$$

$$= G^p dq - G^q dp - (H_q G^q + H_p G^p)dt \tag{233}$$

having taken $G^t = 0$ so that time is unperturbed. Then choose for zeroth order

$$dS_0 = 0 \implies \bar{\gamma}_0 = \gamma_0 = pdq - H_0 dt. \tag{234}$$

The first interesting perturbation equation is first order (with $\gamma_1 = -H_1 dt$ the desired form)

$$\bar{\gamma}_1 = \gamma_1 + \imath_{G_1} d\gamma_0 + dS_1, \tag{235}$$

$$= [-H_1 dt] + [G_1^p dq - G_1^q dp - (H_{0q} G_1^q + H_{0p} G_1^p)dt] + S_{1q} dq + S_{1p} dq + S_{1t} dt \tag{236}$$

$$-K_1 dt = (G_1^p + S_{1q})dq + (S_{1p} - G_1^q)dp - (S_{1t} + H_{0q} G_1^q + H_{0p} G_1^p - H_1)dt. \tag{237}$$

Collecting terms, the coefficients of dq, dp identify the components of the vector field as $G_1^p = -S_{1q}$, $G_1^q = S_{1p}$, and therefore $H_{0q} G_1^q + H_{0p} G_1^p = H_{0q} S_{1p} - H_{0p} S_{1q} = \{H_0, S_1\}$. Then the first Lie transform perturbation equation is recovered

$$K_1 = H_1 + \frac{dS_1}{dt} \tag{238}$$

where $\frac{dS_1}{dt} = \frac{\partial S_1}{\partial t} + \{S_1, H_0\}$. This shows that the perturbation of the symplectic potential is transferred onto the Hamiltonian function, and helps to elucidate the connection between classic Hamilton-Jacobi equations for canonical transformations and the hieracrhy of perturbation equations for near-identity Lie transformations.

[3] Some authors use $\imath_G \gamma \equiv G \cdot \gamma$ as a more intuitive notation.