10. Normal Forms

Perturbation theory is one of the few ways that one can bridge the gap between the behavior of a real nonlinear system and its linear approximation. Because the theory of linear systems is so much simpler, investigators are tempted to fit the problem at hand to a linear model without proper justification. Such a linear model may lead to quantitative as well as qualitative errors. On the other hand, so little is known about the general behavior of a nonlinear system that some sort of approximation has to be made.

Many interesting problems can be formulated as a system of equations that depend on a small parameter ε with the property that when $\varepsilon = 0$ the system is linear, or at least integrable. This chapter develops a very powerful and general method for handling the formal aspects of perturbations of linear and integrable systems, and the next two chapters contain rigorous results that depend on these formal considerations.

10.1 Normal Form Theorems

In this section the main theorems about the normal form at an equilibrium and at a fixed point developed in this chapter are summaries without proof. Upon a first sitting a reader may want read this section, skip the details in the rest of the chapter, and go on to other topics.

10.1.1 Normal Form at an Equilibrium Point

Consider a Hamiltonian system of the form

$$H_{\#}(x) = \sum_{i=0}^{\infty} H_i(x). \tag{10.1}$$

In order to study this system we change coordinates so that the system in the new coordinates is simpler. The definition of simpler depends on the problem at hand. In this chapter we construct formal, symplectic, near-identity changes of variables $x = X(y) = y + \cdots$, such that in the new coordinates the Hamiltonian becomes

K.R. Meyer et al., *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem*, Applied Mathematical Sciences 90, DOI 10.1007/978-0-387-09724-4_10,

$$H^{\#}(y) = \sum_{i=0}^{\infty} H^i(y). \tag{10.2}$$

If the Hamiltonian $H^{\#}$ meets the criteria for being simple then the system is said to be in normal form. It is important to understand the implications of a formal transformation. Even though the original system (10.1) is a convergent series for x in some domain, the series expansion for the change of variables $X(y)$ will not converge in general. Thus the series (10.2) does not necessarily converge. The only way to obtain rigorous results based on this theory is to truncate the series expansion for X at some finite order to obtain a finite (hence convergent) series for X. In this case only the first few terms of $H^{\#}$ are in normal form. In general, if the series for X is truncated after the Nth term then the series for $H^{\#}$ will be convergent, but only the terms up to and including the Nth will be in normal form.

Various methods for transforming a system into normal form have been given because the middle of the nineteenth century, but we present the method of Lie transforms because of its great generality and simplicity. The simplicity of this method is the result of its recursive algorithmic definition which lends itself to easy computer implementation.

Our first example is the classical theorem on the normal form for a Hamiltonian system at a simple equilibrium point. Consider an analytic Hamiltonian, $H_{\#}$, which has an equilibrium point at the origin in $\mathbb{R}^{2n}$, and assume that the Hamiltonian is zero at the origin. Then $H_{\#}$ has a Taylor series expansion of the form (10.1) where H_i is a homogeneous polynomial in x of degree $i+2$; so, $H_0(x) = \frac{1}{2}x^T Sx$, where S is a $2n \times 2n$ real symmetric matrix, and $A = JS$ is a Hamiltonian matrix. The linearized equation about the critical point $x = 0$ is

$$\dot{x} = Ax = JSx = J\nabla H_0(x), \tag{10.3}$$

and the general solution of (10.3) is $\phi(t, \xi) = \exp(At)\xi$. A traditional analysis is to solve (10.3) by linear algebra techniques and then hope that the solutions of the nonlinear problem are not too dissimilar from the solutions of the linear equation. In many cases this hope is unjustifiable. The next best thing is to put the equations in normal form and to study the solutions of the normal form equations. This too has its pitfalls.

Theorem 10.1.1. *Let A be diagonalizable. Then there exists a formal, symplectic change of variables, $x = X(y) = y + \cdots$, which transforms the Hamiltonian* (10.1) *to* (10.2) *where H^i is a homogeneous polynomial of degree $i+2$ such that*

$$H^i(e^{At}y) \equiv H^i(y), \tag{10.4}$$

for all $i = 0, 1, \ldots,$ all $y \in \mathbb{R}^{2n}$, and all $t \in \mathbb{R}$.

For example consider a two degree of freedom system in the case when the matrix A is diagonalizable and has distinct pure imaginary eigenvalues

$\pm i\omega_1$, $\pm\omega_2$. In this case we may assume that after a symplectic change of variables the quadratic terms are

$$H_0(x) = \frac{\omega_1}{2}(x_1^2 + x_3^2) + \frac{\omega_2}{2}(x_2^2 + x_4^2) = \omega_1 I_1 + \omega_2 I_2, \tag{10.5}$$

where in the second form we use the action–angle coordinates

$$I_1 = \frac{1}{2}(x_1^2 + x_3^2), \quad I_2 = \frac{1}{2}(x_2^2 + x_4^2), \quad \phi_1 = \tan^{-1}\frac{x_3}{x_1}, \quad \phi_2 = \tan^{-1}\frac{x_4}{x_2}.$$

The linear equations (10.3) in action–angle coordinates become

$$\dot{I}_1 = 0, \quad \dot{I}_2 = 0, \quad \dot{\phi}_1 = -\omega_1, \quad \dot{\phi}_2 = -\omega_2.$$

The condition (10.4) requires the terms in the normal form to be constant on the solutions of the above equations. These equations have as solutions $I_1 = I_1^0$ and $I_2 = I_2^0$ where I_1^0 and I_2^0 are constants. $I_1 = I_1^0$ and $I_2 = I_2^0$ where $I_1^0 > 0$ and $I_2^0 > 0$ defines a 2-torus with angular coordinates ϕ_1 and ϕ_2. This type of flow on a torus was discussed in detail in Section 1.9.

There are two cases depending on whether the ratio ω_1/ω_2 is rational or irrational. In the case when the ratio is irrational the flow on the torus defined by the equations above is dense on the torus and so the only continuous functions defined on the torus are constants, therefore, the terms in the normal form will depend only on the action variables I_1, I_2. On the other hand, if the ratio is rational, say $\omega_1/\omega_2 = p/q$, then the terms in the normal form may contain a dependence on the single angle $\psi = q\phi_1 - p\phi_2$.

Thus: *If* H_0 *in* (10.1) *is of the form* (10.5) *then the normal form for the system is*

$$H^{\#} = \sum_{i=0}^{\infty} H^i(I_1, I_2)$$

when the ratio ω_1/ω_2 *is irrational, and*

$$H^{\#} = \sum_{i=0}^{\infty} H^i(I_1, I_2, q\phi_1 - p\phi_2)$$

when $\omega_1/\omega_2 = p/q$.

This covers the normal form at the equilibrium point $\mathcal{L}_4$ of the restricted 3-body problem when $0 < \mu < \mu_1$. A multitude of interesting stability and bifurcation results follow from simple inequalities on a finite number of terms in this normal form.

In the case where the matrix A is not diagonalizable the only change in the statement of Theorem 10.1.1 is that the condition (10.4) is replaced by

$$H^i(e^{A^T t}y) \equiv H^i(y),$$

where A^T is the transpose of A.

Consider a two degree of freedom Hamiltonian system at an equilibrium point when the exponents are $\pm i\omega$ with multiplicity two and the linearized system is not diagonalizable. The normal form for the quadratic part of such a Hamiltonian was given as

$$H_0 = \omega(x_2y_1 - x_1y_2) + \frac{\delta}{2}(x_1^2 + x_2^2),$$

where $\omega \neq 0$ and $\delta = \pm 1$. In this case

$$A = \begin{bmatrix} 0 & \omega & 0 & 0 \\ -\omega & 0 & 0 & 0 \\ -\delta & 0 & 0 & \omega \\ 0 & -\delta & -\omega & 0 \end{bmatrix}.$$

The normal form in this case depends on the four quantities

$$\begin{array}{ll} \Gamma_1 = x_2y_1 - x_1y_2, & \Gamma_2 = \frac{1}{2}(x_1^2 + x_2^2), \\ \Gamma_3 = \frac{1}{2}(y_1^2 + y_2^2), & \Gamma_4 = x_1y_1 + x_2y_2. \end{array}$$

Note that $\{\Gamma_1, \Gamma_2\} = 0$ and $\{\Gamma_1, \Gamma_3\} = 0$. The system is in Sokol'skii normal form if the higher-order terms depend on the two quantities Γ_1 and Γ_3, that is, the Hamiltonian is of the form

$$H^{\#} = \omega(x_2y_1 - x_1y_2) + \frac{\delta}{2}(x_1^2 + x_2^2) + \sum_{k=1}^{\infty} H_{2k}(x_2y_1 - x_1y_2, y_1^2 + y_2^2),$$

where H_{2k} is a polynomial of degree k in two variables. The first few terms of this normal form determine the nature of the stability and bifurcations at the equilibrium point $\mathcal{L}_4$ of the restricted problem when $\mu = \mu_1$.

10.1.2 Normal Form at a Fixed Point

The study of the stability and bifurcation of a periodic solution of a Hamiltonian system of two degrees of freedom can be reduced to the study of the Poincaré map in an energy level (i.e., level surface of the Hamiltonian). Sometimes the value of the Hamiltonian must be treated as a parameter.

Consider a diffeomorphism of the form

$$F_{\#} : N \subset \mathbb{R}^2 \to \mathbb{R}^2 : x \to f(x), \tag{10.6}$$

where N is a neighborhood of the origin in $\mathbb{R}^2$, and f is a smooth function such that

$$f(0) = 0, \qquad \det \frac{\partial f}{\partial x}(x) \equiv 1.$$

The origin is a fixed point for the diffeomorphism because $f(0) = 0$, and it is orientation-preserving and area-preserving because $\det \partial f/\partial x \equiv 1$. This map should be considered as the Poincaré map associated with a periodic solution of a two degree of freedom Hamiltonian system.

The linearization of this map about the origin is $x \to Ax$ where A is the 2×2 matrix $(\partial f/\partial x)(0)$. Because the determinant of A is 1 the product of its eigenvalues must be 1. The eigenvalues λ, λ^{-1} of A are called the multipliers of the fixed point. There are basically four cases:

1. Hyperbolic fixed point: multipliers real and $\lambda \neq \pm 1$
2. Elliptic fixed point: multipliers complex conjugates and $\lambda \neq \pm 1$
3. Shear fixed point: $\lambda = +1$, A not diagonalizable
4. Flip fixed point: $\lambda = -1$, A not diagonalizable

As before in order to study an area-preserving map we can change coordinates so that the map in the new coordinates is simpler. Here we consider a formal symplectic, near-identity change of variables $x = X(y) = y + \cdots$, such that in the new coordinates the map (10.6) becomes

$$F^{\#} : y \to g(y). \tag{10.7}$$

If the map $F^{\#}$ meets the criteria for being simple then the map is said to be in normal form. It is important to understand the implications of a formal transformation. Even though the original system (10.6) is a convergent series for x in some domain, the series expansion for the change of variables $X(y)$ will not converge in general. Thus the series (10.7) does not converge in general. The only way to obtain rigorous results based on this theory is to truncate the series expansion for X at some finite order to obtain a finite (hence convergent) series for X. In this case only the first few terms of $F^{\#}$ will be in normal form. In general, if the series for X is truncated after the Nth term then the series for $F^{\#}$ will be convergent, but only the terms up to and including the Nth will be in normal form.

Hyperbolic fixed point. In the hyperbolic case after a change of variables we may assume that

$$A = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix},$$

with $\lambda \neq \pm 1$ and real. The mapping (10.7) is in normal form with $F^{\#} : (u, v) \to (u', v')$ where $y = (u, v)$ with

$$u' = u\ell(uv)$$

$$v' = v\ell(uv)^{-1},$$

where ℓ is a formal series in one variable, $\ell(uv) = \lambda + \cdots$.

The map takes the hyperbolas $uv = \textit{constant}$ into themselves. The transformation to normal form actually converges by a classical theorem of Moser (1956).

Elliptic fixed point. In the elliptic case when λ is a complex number of unit modulus certain reality conditions must be met. Consider the case when

A has eigenvalues $\lambda^{\pm 1} = \exp(\pm\omega i) \neq \pm 1$; i.e., the origin is an elliptic fixed point. First assume that λ is not a root of unity. Change to action–angle variables (I, ϕ); The normal form in action–angle variables in this case is $F^{\#} : (I, \phi) \to (I', \phi')$ where

$$I' = I, \qquad \phi' = \phi + \ell(I),$$

where ℓ has a formal expansion $\ell(I) = -\omega + \beta I \cdots$. If a diffeomorphism is in this form with $\beta \neq 0$, then the origin is called a general elliptic point, or $F^{\#}$ is called a twist map. This map takes circles, $I = const$, into themselves and rotates each circle by an amount $\ell(I)$.

Now consider the case when the diffeomorphism has an elliptic fixed point whose multiplier is a root of unity. Let λ be a kth root of unity; so, $\lambda^k = 1$, $k > 2$, and $\lambda = \exp(h2\pi i/k)$, where h is an integer. The origin is called a k-resonance elliptic point in this case. The normal form in action–angle variables in this case is $F^{\#} : (I, \phi) \to (I', \phi')$ where

$$\begin{aligned} I' &= I + 2\alpha I^{k/2} \sin(k\phi) + \cdots, \\ \phi' &= \phi + (2\pi h/k) + \alpha I^{(k-2)/2} \cos(k\phi) + \beta I + \cdots. \end{aligned} \tag{10.8}$$

Shear fixed point. Consider the cases where the multiplier is $+1$. If $A = I$, the identity matrix, the system is so degenerate that there is no normal form in general. Otherwise, by a coordinate change we have

$$A = \begin{bmatrix} 1 & \pm 1 \\ 0 & 1 \end{bmatrix}. \tag{10.9}$$

The important terms of the normal form $F^{\#} : (u, v) \to (u', v')$ are

$$\begin{aligned} u' &= u \pm v - \cdots, \\ v' &= v - \beta u^2 + \cdots. \end{aligned} \tag{10.10}$$

The ellipsis may contain other quadratic terms and higher-order terms.

Flip periodic point. Now consider the case when A has eigenvalue -1. In this case the generic form for A is

$$A = \begin{bmatrix} -1 & \pm 1 \\ 0 & -1 \end{bmatrix}.$$

The quadratic terms can be eliminated and the important terms of the normal form $F^{\#} : (u, v) \to (u', v')$ are

$$\begin{aligned} u' &= -u - v + \cdots, \\ v' &= -v + \beta u^3 + \cdots. \end{aligned} \tag{10.11}$$

The ellipsis may contain other cubic terms and higher-order terms.

10.2 Forward Transformations

One of the most general methods of mathematics is to simplify a problem by a change of variables. The method of Lie transforms developed by Deprit (1969) and extended by Kamel (1970) and Henrard (1970b) is a general procedure to change variables in a system of equations that depend on a small parameter. Deprit's original method was for Hamiltonian systems only, but the extensions by Kamel and Henrard handle non-Hamiltonian equations. Only the Hamiltonian case is treated here.

10.2.1 Near-Identity Symplectic Change of Variables

The general idea of this method is to generate a symplectic change of variables depending on a small parameter as the general solution of a Hamiltonian system of differential equations; see Theorem 6.1.2. $X(\varepsilon, y)$ is said to be a near-identity symplectic change of variables (or transformation), if X is symplectic for each fixed ε and is of the form $X(\varepsilon, y) = y + O(\varepsilon)$; i.e., $X(0, y) = y$. Because $X(0, y) = y$, $\partial X(\varepsilon, y)/\partial y$ is nonsingular for small ε so by the inverse function theorem, the map $y \to X(\varepsilon, y)$ has a differentiable inverse for small ε. Both X and its inverse are symplectic for fixed ε.

Consider the nonautonomous Hamiltonian system

$$\frac{dx}{d\varepsilon} = J\nabla W(\varepsilon, x) \tag{10.12}$$

and the initial condition

$$x(0) = y, \tag{10.13}$$

where W is smooth. The basic theory of differential equations asserts that the general solution of this problem is a smooth function $X(\varepsilon, y)$ such that $X(0, y) \equiv y$, and by Theorem 6.1.2, the function X is symplectic for fixed ε. That is, the differential equation (10.12) and the initial condition (10.13) define a near-identity symplectic change of variables.

Conversely, let $X(\varepsilon, y)$ be a near-identity symplectic change of variables with inverse function $Y(\varepsilon, x)$ such that $X(\varepsilon, Y(\varepsilon, x)) \equiv x$ and $Y(\varepsilon, X(\varepsilon, y)) \equiv y$ where defined. Y is symplectic too. Differentiating $Y(\varepsilon, X(\varepsilon, y)) \equiv y$ with respect to ε yields

$$\frac{\partial Y}{\partial x}(\varepsilon, X(\varepsilon, y))\frac{\partial X}{\partial \varepsilon}(\varepsilon, y) + \frac{\partial Y}{\partial \varepsilon}(\varepsilon, X(\varepsilon, y)) \equiv 0$$

or

$$\frac{\partial X}{\partial \varepsilon}(\varepsilon, y) \equiv \left[\frac{\partial Y}{\partial x}(\varepsilon, X(\varepsilon, y))\right]^{-1} \frac{\partial Y}{\partial \varepsilon}(\varepsilon, X(\varepsilon, y)).$$

This means that $X(\varepsilon, y)$ is the general solution of

$$\frac{dx}{d\varepsilon} = U(\varepsilon, x), \qquad \text{where } U(\varepsilon, x) = \left[\frac{\partial Y}{\partial x}(\varepsilon, x)\right]^{-1} \frac{\partial Y}{\partial \varepsilon}(\varepsilon, x).$$

This equation is Hamiltonian so, there is a function $W(\varepsilon, x)$ such that $U(\varepsilon, x) = J\nabla W(\varepsilon, x)$. This proves the following.

Proposition 10.2.1. *$X(\varepsilon, y)$ is a near-identity symplectic change of variables if and only if it is the general solution of a Hamiltonian differential equation of the form* (10.12) *satisfying initial condition* (10.13)*.*

A Hamiltonian system of equations generates symplectic transformations directly, which is in contrast to the symplectic transformations given by the generating functions in Theorem 6.2.1, where the new and old variables are mixed.

10.2.2 The Forward Algorithm

Let $X(\varepsilon, y)$, $Y(\varepsilon, x)$, and $W(\varepsilon, x)$ be as above; so, $X(\varepsilon, y)$ is the solution of (10.12) satisfying (10.13). Think of $x = X(\varepsilon, y)$ as a change of variables $x \to y$ that depends on a parameter. Throughout this chapter, when we change variables, we do not change the parameter ε.

Let $H(\varepsilon, x)$ be a Hamiltonian and $G(\varepsilon, y) \equiv H(\varepsilon, X(\varepsilon, y))$; so, G is the Hamiltonian H in the new coordinates. We call G the Lie transform of H (generated by W). Sometimes H is denoted by H_* and G by H^*, and sometimes G is denoted by $\mathcal{L}(W)H$ to show that G is the Lie transform of H generated by W. Let the function $H = H_*$, $G = H^*$, and W all have series expansions in the small parameter ε. The forward algorithm of the method of Lie transforms is a recursive set of formulas that relate the terms in these various series expansions.

In particular let

$$H(\varepsilon, x) = H_*(\varepsilon, x) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) H_i^0(x), \tag{10.14}$$

$$G(\varepsilon, y) = H^*(\varepsilon, y) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) H_0^i(y), \tag{10.15}$$

$$W(\varepsilon, x) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) W_{i+1}(x). \tag{10.16}$$

The method of Lie transforms introduces a double indexed array $\{H_j^i\}$, $i, j = 0, 1, \dots$ which agrees with the definitions given in (10.14) and (10.15) when either i or j is zero. The other terms are intermediary terms introduced to facilitate the computation.

Theorem 10.2.1. *Using the notation given above, the functions $\{H_j^i\}$, $i = 1, 2, \dots$, $j = 0, 1, \dots$ satisfy the recursive identities*

$$H_j^i = H_{j+1}^{i-1} + \sum_{k=0}^{j} \binom{j}{k} \{H_{j-k}^{i-1}, W_{k+1}\}. \tag{10.17}$$

Remarks. The above formula contains the standard binomial coefficient

$$\binom{j}{k} = \frac{j!}{k!(j-k)!}.$$

Note that because the transformation generated by W is a near identity transformation, the first term in H_* and H^* is the same, namely H_0^0. Also note that the first term in the expansion for W starts with W_1. This convention imparts some nice properties to the formulas in (10.17). Each term in 10.17 has indices summing to $i+j$, and each term on the right-hand side has upper index $i-1$.

In order to construct the change of variables $X(\varepsilon, y)$, note that X is the transform of the identity function or $X(\varepsilon, y) = \mathcal{L}(W)(id)$, where $id(x) = x$.

The interdependence of the functions $\{H_j^i\}$ can easily be understood by considering the Lie triangle

$$\begin{array}{lll}
H_0^0 & & \\
\downarrow & & \\
H_1^0 \rightarrow & H_0^1 & \\
\downarrow & \downarrow & \\
H_2^0 \rightarrow & H_1^1 \rightarrow & H_0^2\,. \\
\downarrow & \downarrow & \downarrow
\end{array}$$

The coefficients of the expansion of the old function H_* are in the left column, and those of the new function H^* are on the diagonal. Formula (10.17) states that to calculate any element in the Lie triangle, you need the entries in the column one step to the left and up.

For example, to compute the series expansion for H^* through terms of order ε^2, you first compute H_0^1 by the formula

$$H_0^1 = H_1^0 + \{H_0^0, W_1\}, \tag{10.18}$$

which gives the term of order ε, and then you compute

$$H_1^1 = H_2^0 + \{H_1^0, W_1\} + \{H_0^0, W_2\},$$

$$H_0^2 = H_1^1 + \{H_0^1, W_1\}.$$

Then $H^*(\varepsilon, x) = H_0^0(x) + H_0^1(x)\varepsilon + H_0^2(x)(\varepsilon^2/2) + \cdots$.

Proof. (Theorem 10.2.1) Recall that $H^*(\varepsilon, y) = G(\varepsilon, y) = H(\varepsilon, X(\varepsilon, y))$, where $X(\varepsilon, y)$ is the general solution of (10.12). Define the differential operator $\mathcal{D} = \mathcal{D}_W$ by

$$\mathcal{D}F(\varepsilon, x) = \frac{\partial F}{\partial \varepsilon}(\varepsilon, x) + \{F, W\}(\varepsilon, x),$$

so that

$$\frac{d}{d\varepsilon}\left(F(\varepsilon, x)\Big|_{x=X(\varepsilon,y)}\right) = \mathcal{D}F(\varepsilon, x)\Big|_{x=X(\varepsilon,y)}.$$

Define new functions by $H^0 = H$, $H^i = \mathcal{D}H^{i-1}$, $i \geq 1$. Let these functions have series expansions

$$H^i(\varepsilon, x) = \sum_{k=0}^{\infty}\left(\frac{\varepsilon^k}{k!}\right) H^i_k(x)$$

so,

$$\begin{aligned} H^i(\varepsilon, x) &= \mathcal{D}\sum_{k=0}^{\infty}\left(\frac{\varepsilon^k}{k!}\right) H^{i-1}_k(x) \\ &= \sum_{k=1}^{\infty}\left(\frac{\varepsilon^{k-1}}{(k-1)!}\right) H^{i-1}_k(x) + \left\{\sum_{k=0}^{\infty}\left(\frac{\varepsilon^k}{k!}\right) H^{i-1}_k(x), \sum_{s=0}^{\infty} W_{s+1}\right\} \\ &= \sum_{j=0}^{\infty}\left(\frac{\varepsilon^j}{j!}\right)\left(H^{i-1}_{j+1} + \sum_{k=0}^{j}\binom{j}{k}\{H^{i-1}_{j-k}, W_{k+1}\}\right). \end{aligned}$$

So the functions H^i_j are related by (10.17). It remains to show that $H^* = G$ has the expansion (10.15). By the above and Taylor's theorem

$$\begin{aligned} G(\varepsilon, y) &= \sum_{n=0}^{\infty}\left(\frac{\varepsilon^n}{n!}\right)\frac{d^n}{d\varepsilon^n}G(\varepsilon, y)\Big|_{\varepsilon=0} \\ &= \sum_{n=0}^{\infty}\left(\frac{\varepsilon^n}{n!}\right)\frac{d^n}{d\varepsilon^n}\left(H(\varepsilon, x)\Big|_{x=X(\varepsilon,y)}\right)_{\varepsilon=0} \\ &= \sum_{n=0}^{\infty}\left(\frac{\varepsilon^n}{n!}\right)\left(\mathcal{D}^n H(\varepsilon, x)\Big|_{x=X(\varepsilon,y)}\right)_{\varepsilon=0} \\ &= \sum_{n=0}^{\infty}\left(\frac{\varepsilon^n}{n!}\right) H^n_0(y). \end{aligned}$$

10.2.3 The Remainder Function

Assume now that the Hamiltonian and hence the equations are time dependent; i.e., consider

$$\dot{x} = J\nabla H(\varepsilon, t, x), \tag{10.19}$$

where H has an expansion

$$H(\varepsilon, t, x) = H_*(\varepsilon, t, x) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) H_i^0(t, x). \tag{10.20}$$

Make a symplectic change of coordinates, $x = X(\varepsilon, t, y)$, which transforms (10.19) to the Hamiltonian differential equation

$$\dot{y} = J\nabla G(\varepsilon, t, y) + J\nabla R(\varepsilon, t, y) = J\nabla K(\varepsilon, t, y),$$

where $G(\varepsilon, t, y) = H^*(\varepsilon, t, y) = H(\varepsilon, t, X(\varepsilon, t, y))$ is the Lie transform of H, R is the remainder function, and $K = G + R$ is the new Hamiltonian. Let G, R, and K have series expansions of the form

$$G(\varepsilon, t, y) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) H_0^i(t, y), \qquad R(\varepsilon, t, y) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) R_0^i(t, y)$$

$$K(\varepsilon, t, y) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) K_0^i(t, y).$$

Let the symplectic change of variables $X(\varepsilon, t, y)$ be the general solution of the Hamiltonian system of equations

$$\frac{dx}{d\varepsilon} = J\nabla W(\varepsilon, t, x), \qquad x(0) = y,$$

where $W(\varepsilon, t, x)$ is a Hamiltonian function with a series expansion of the form

$$W(\varepsilon, t, x) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) W_{i+1}(t, x).$$

The variable t is simply a parameter, and so the function $G = H^*$ can be computed by formulas (10.17) in Theorem 10.2.1 using the Lie triangle as a guide. The remainder term R needs further consideration.

Theorem 10.2.2. *The remainder function is given by*

$$R(\varepsilon, t, y) = -\int_0^{\varepsilon} \mathcal{L}(W) \left(\frac{\partial W}{\partial t}\right) (s, t, y) ds. \tag{10.21}$$

Proof. Making the symplectic change of variable $x = X(\varepsilon, t, y)$ in (10.19) directly gives

$$\dot{y} = \left(\frac{\partial X}{\partial y}\right)^{-1} (\varepsilon, t, y) J\nabla_x H(\varepsilon, t, X(\varepsilon, t, y)) - \left(\frac{\partial X}{\partial y}\right)^{-1} (\varepsilon, t, y) \frac{\partial X}{\partial t}(\varepsilon, t, y).$$

By the discussion in Section 6.1 the first term on the right-hand side is $J\nabla G$, and so,

$$J\nabla R(\varepsilon,t,y) = -\left(\frac{\partial X}{\partial y}(\varepsilon,t,y)\right)^{-1}\frac{\partial X}{\partial t}(\varepsilon,t,y).$$

$A(\varepsilon) = \partial X(\varepsilon,t,y)/\partial y$ is the fundamental matrix solution of the variational equation; i.e., it is the matrix solution of

$$\frac{dA}{d\varepsilon} = \left(J\frac{\partial^2 W}{\partial x^2}(\varepsilon,t,X(\varepsilon,t,y))\right)A, \qquad A(0) = I.$$

Differentiating $\partial X(\varepsilon,t,y)/\partial\varepsilon = J\nabla W(\varepsilon,t,X(\varepsilon,t,y))$ with respect to t shows that $B(\varepsilon) = \partial X(\varepsilon,t,y)/\partial t$ satisfies

$$\frac{dB}{d\varepsilon} = \left(J\frac{\partial^2 W}{\partial x^2}(\varepsilon,t,X(\varepsilon,t,y))\right)B + J\frac{\partial^2 W}{\partial x\partial t}(\varepsilon,t,X(\varepsilon,t,y)).$$

Because $X(0,t,y) \equiv y$, $B(0) = 0$, and so, by the variation of constants formula,

$$B(\varepsilon) = \int_0^\varepsilon A(\varepsilon)A(s)^{-1}J\frac{\partial^2 W}{\partial x\partial t}(s,t,X(s,t,y))ds;$$

therefore,

$$\begin{aligned} J\nabla R(\varepsilon,t,y) &= -\left(\frac{\partial X}{\partial y}(\varepsilon,t,y)\right)^{-1}\frac{\partial X}{\partial t}(\varepsilon,t,y) = -A(\varepsilon)^{-1}B(\varepsilon) \\ &= -\int_0^\varepsilon A(s)^{-1}J\frac{\partial^2 W}{\partial x\partial t}(s,t,X(s,t,y))ds \\ &= -\int_0^\varepsilon JA(s)^T\frac{\partial^2 W}{\partial x\partial t}(s,t,X(s,t,y))ds \\ &= -J\frac{\partial}{\partial y}\int_0^\varepsilon \frac{\partial W}{\partial t}(s,t,X(s,t,y))ds \\ &= -J\frac{\partial}{\partial y}\int_0^\varepsilon \mathcal{L}(W)\left(\frac{\partial W}{\partial t}\right)(s,t,y)ds. \end{aligned}$$

In the above, the fact that A is symplectic is used to make the substitution $A^{-1}J = JA^T$.

Thus, to compute the remainder function, first compute the transform of $-\partial W/\partial t$, and then integrate it. That is, let $S_*(\varepsilon,t,x) = \sum(\varepsilon^i/i!)S_i^0(t,x)$, where $S_i^0(t,x) = -\partial W_{i-1}(t,x)/\partial t$. Compute the Lie transform of S_* by the previous algorithms to get $\mathcal{L}(W)(S) = S^*(\varepsilon,t,x) = \sum(\varepsilon^i/i!)S_0^i(t,x)$. Then $R_0^i = S_0^{i-1}$.

For example, to compute the series expansion for $K = G + R$, the new Hamiltonian, through terms of order ε^2, set $K_0^0 = H_0^0$, then compute K_0^1 by the formulas

$$H_0^1 = H_1^0 + \{H_0^0, W_1\}, \quad R_0^1 = -\frac{\partial W_1}{\partial t}, \quad K_0^1 = H_0^1 + R_0^1,$$

which gives the term of order ε, and then compute

$$H_1^1 = H_1^0 + \{H_1^0, W_1\} + \{H_0^0, W_2\}, \quad H_0^2 = H_1^1 + \{H_0^1, W_1\},$$

$$R_0^2 = -\frac{\partial W_2}{\partial t} - \left\{\frac{\partial W_1}{\partial t}, W_1\right\}, \qquad K_0^2 = H_0^2 + R_0^2.$$

Then $K^*(\varepsilon, x) = K_0^0(x) + \varepsilon K_0^1(x) + \frac{\varepsilon^2}{2} K_0^2(x) + \cdots$.

10.3 The Lie Transform Perturbation Algorithm

In many of the cases of interest, the Hamiltonian is given, and the change of variables is sought to simplify it. When the Hamiltonian, and hence the equations, are in sufficiently simple form, they are said to be in "normal form," an expression whose meaning is discussed in detail later.

10.3.1 Example: Duffing's Equation

In (6.15) the Hamiltonian of Duffing's equation was given as

$$H = \frac{1}{2}(q^2 + p^2) + \frac{\gamma}{4}q^4 \tag{10.22}$$

in rectangular coordinates, (q, p), and in action–angle variables, (I, ϕ), it was given as

$$H = I + \frac{\gamma}{8} I^2 (3 + 4\cos 2\phi + \cos 4\phi). \tag{10.23}$$

The Hamiltonian is analytic in rectangular coordinates, and so has the d'Alembert character. Consider γ as a small parameter by setting $\varepsilon = \gamma/8$; so, $H(\varepsilon, I, \phi) = H^*(\varepsilon, I, \phi) = H_0^0(I, \phi) + \varepsilon H_1^0(I, \phi)$, where

$$H_0^0 = I, \qquad H_1^0 = I^2(3 + 4\cos 2\phi + \cos 4\phi).$$

By formula (10.18),

$$H_0^1 = H_1^0 + \{H_0^0, W_1\};$$

so,

$$H_0^1 = I^2(3 + 4\cos 2\phi + \cos 4\phi) - \frac{\partial W_1}{\partial \phi}.$$

Choose W_1 so that H_0^1 contains as few terms as possible (one definition of normal form). For the transformation generated by W_1 to be analytic in rectangular coordinates, W must be a Poisson series with the d'Alembert character. Thus the simplest form for H_0^1 is

$$H_0^1 = 3I^2,$$

which is accomplished taking

$$W_1 = I^2(2\sin 2\phi + \frac{1}{4}\sin 4\phi).$$

With this W_1, the Hamiltonian in the new coordinates, (J, θ), would be

$$H_*(\varepsilon, J, \theta) = J + \frac{3\gamma}{8}J^2 + O(\gamma^2),$$

and the equations of motion would be

$$\dot{J} = O(\gamma^2), \qquad \dot{\theta} = -1 - \frac{3\gamma}{4}J + O(\gamma^2).$$

In these coordinates, up to terms $O(\gamma^2)$, the solutions move on circles $J =$ constant with uniform angular frequency $-1 - (3\gamma/4)J$.

Let us do this simple example again, but this time in complex coordinates $z = q + ip$, $\bar{z} = q - ip$. This change of variables is symplectic with multiplier $2i$; so, the Hamiltonian becomes

$$H(z, \bar{z}) = iz\bar{z} + \frac{\gamma i}{32}(z^4 + 4z^3\bar{z} + 6z^2\bar{z}^2 + 4z\bar{z}^3 + \bar{z}^4).$$

H is real in the rectangular coordinates (q, p), so H is conjugated by interchanging z and $\bar{z}$; i.e., $H(z, \bar{z}) = H(\bar{z}, z)$. This is the reality condition in these variables. Let $\varepsilon = \gamma/32$ and

$$H_0^0 = iz\bar{z}, \quad H_1^0 = i(z^4 + 4z^3\bar{z} + 6z^2\bar{z}^2 + 4z\bar{z}^3 + \bar{z}^4);$$

so Equation (10.18) becomes

$$H_0^1 = i(z^4 + 4z^3\bar{z} + 6z^2\bar{z}^2 + 4z\bar{z}^3 + \bar{z}^4) + \frac{1}{2}\left(z\frac{\partial W}{\partial z} - \bar{z}\frac{\partial W}{\partial \bar{z}}\right).$$

Try $W = az^\alpha\bar{z}^\beta$; then $(z\partial W/\partial z + \bar{z}\partial W/\partial \bar{z})/2 = a(\alpha - \beta)z^\alpha\bar{z}^\beta/2$; so, all the terms in H_1^0 can be eliminated except those with $\alpha = \beta$. That is, if we take

$$W = -i(z^4/2 + 4z^3\bar{z} - 4z\bar{z}^3 - \bar{z}^4/2),$$

then

$$H_* = H_0^0 + H_1^0 = iz\bar{z} + (3\gamma i/16)(z\bar{z})^2 + O(\varepsilon^2).$$

Notice that both W and H_* satisfy the reality condition and so are real functions in the original coordinates (q, p). The two methods of solving the problem (action–angle variables and complex variables) give the same results when written in rectangular coordinates.

10.3.2 The General Algorithm

The main Lie transform algorithm starts with a given Hamiltonian that depends on a small parameter ε, and constructs a change of variables so that the Hamiltonian in the new variables is simple. The algorithm is built around the following observation.

Consider the Hamiltonian $H_*(\varepsilon, x)$ with series expansion as given in Equation (10.14); so, all the H_i^0 are known. Assume that all the entries in the Lie triangle are known down to the Nth row; so, the H_j^i are known for $i+j \leq N$, and assume that the W_i are known for $i \leq N$. Let $L_j^i, i+j \leq N$, be computed from the same initial Hamiltonian, but with $U_1, \dots, U_N$ where $U_i = W_i$ for $i = 1, 2, \dots, N-1$ and $U_N = 0$. Then

$$\begin{aligned} H_j^i &= L_j^i \text{ for } i+j < N \\ H_j^i &= L_j^i + \{H_0^0, W_N\} \quad \text{ for } i+j = N. \end{aligned} \tag{10.24}$$

This is easily seen from the recursive formulas in Theorem 10.2.1. Recall the remark that the sum of all the indices must add to the row number; so, W_N does not affect the terms in the first $N-1$ rows. The second equation in (10.24) follows from a simple induction across the Nth row.

From this observation, the algorithm is as follows. Assume all the rows in the Lie triangle have been computed down to the $(N-1)$st row, that $W_1, \dots, W_{N-1}$ have been determined, and that the terms $H_0^1, \dots, H_0^{N-1}$ are in normal form; i.e., simple in some sense. Now it is time to compute W_N so that H_0^N is in normal form. To compute the Nth row do the following.

Step 1: Compute the Nth row from the formulas in Theorem 10.2.1 assuming that $W_N = 0$, and call these terms L_j^i, $i+j = N$.

Step 2: Solve the equation $H_0^N = L_0^N + \{H_0^0, W_N\}$ for W_N and H_0^N, so that H_0^N is in normal form or simple.

Step 3: Add $\{H_0^0, W_N\}$ to each term in the Nth row; i.e., calculate $H_j^i = L_j^i + \{H_0^0, W_N\}$ for all $i+j = N$.

Step 4: Repeat for the next row.

Of course the definition of normal form and simple depends on the equation $H_0^N = L_0^N + \{H_0^0, W_N\}$, which in turn depends on H_0^0. This equation is called the Lie equation or the homology equation.

10.3.3 The General Perturbation Theorem

The algorithm can be used to prove a general theorem that includes almost all applications. Use the notation of Section 10.2.

Theorem 10.3.1. *Let $\{\mathcal{P}_i\}_{i=0}^\infty$, $\{\mathcal{Q}_i\}_{i=1}^\infty$, and $\{\mathcal{R}_i\}_{i=1}^\infty$ be sequences of linear spaces of smooth functions defined on a common domain O in $\mathbb{R}^{2n}$ with the following properties.*

1. $\mathcal{Q}_i \subset \mathcal{P}_i$, $i = 1, 2, \ldots$.
2. $H_i^0 \in \mathcal{P}_i$, $i = 0, 1, 2, \ldots$.
3. $\{\mathcal{P}_i, \mathcal{R}_j\} \subset \mathcal{P}_{i+j}$, $i + j = 1, 2, \ldots$.
4. *for any* $D \in \mathcal{P}_i$, $i = 1, 2, \ldots$, *there exist* $B \in \mathcal{Q}_i$ *and* $C \in \mathcal{R}_i$ *such that*

$$B = D + \{H_0^0, C\}. \tag{10.25}$$

Then there exists a W *with a formal Hamiltonian of the form* (10.16) *with* $W_i \in \mathcal{R}_i, i = 1, 2, \ldots$, *which generates a near-identity symplectic change of variables* $x \to y$ *such that the Hamiltonian in the new variables has a series expansion given by* (10.15) *with* $H_0^i \in \mathcal{Q}_i, i = 1, 2, \ldots$.

Remarks. The Lie equation (10.25) is the heart of a perturbation problem. H_0^0 defines the unperturbed system when $\varepsilon = 0$, so it is supposed to be well understood. For example, it might be the harmonic oscillator or the 2-body problem. Equation (10.25) can be rewritten

$$B = D + \mathcal{F}(C)$$

where $\mathcal{F} = \{H_0^0, \cdot\}$ is a linear operator on functions. One must analyze this operator to determine in what linear spaces the equation (10.25) is solvable. Roughly speaking the Hamiltonian (10.14) starts with terms in the $\mathcal{P}$-spaces ($H_i^0 \in \mathcal{P}_i$), and the equation in normal form has terms in the $\mathcal{Q}$-space ($H_0^i \in \mathcal{Q}_i$). The $\mathcal{Q}$-spaces are smaller than the $\mathcal{P}$-spaces ($\mathcal{Q}_i \subset \mathcal{P}_i$). So the normal form is "simpler." The transformation is generated by a Hamiltonian differential equation with Hamiltonian W in the $\mathcal{R}$-spaces ($W_i \in \mathcal{R}_i$). D is an old term, B is a new term, and C is a generator.

Proof. Use induction on the rows of the Lie triangle.
Induction hypothesis I_n: Let $H_j^i \in \mathcal{P}_{i+j}$ for $0 \le i + j \le n$ and $W_i \in \mathcal{R}_i, H_0^i \in \mathcal{Q}_i$ for $1 \le i \le n$.

I_0 is true by assumption, and so assume I_{n-1}. By Equation (10.17)

$$H_{n-1}^1 = H_n^0 + \sum_{k=0}^{n-2} \binom{n-1}{k} \{H_{n-1-k}^0, W_{k+1}\} + \{H_0^0, W_n\}.$$

The last term is singled out because it is the only term that contains an element, W_n, which is not covered by the induction hypothesis or the hypothesis of the theorem. All the other terms are in $\mathcal{P}_n$ by I_{n-1} and (3). Thus

$$H_{n-1}^1 = E^1 + \{H_0^0, W_n\},$$

where $E^1 \in \mathcal{P}_n$ is known. A simple induction on the columns of the Lie triangle using (10.17) shows that

$$H_{n-s}^s = E^s + \{H_0^0, W_n\},$$

where $E^s \in \mathcal{P}_n$ for $s = 1, 2, \ldots, n$, and so

$$H_n^0 = E^n + \{H_0^0, W_n\}.$$

By (4), solve for $W_n \in \mathcal{R}_n$ and $H_0^n \in \mathcal{Q}_i$. Thus I_n is true.

The theorem given above is formal in the sense that the convergence of the various series is not discussed. In interesting cases the series diverge, but useful information can be obtained in the first few terms of the normal form. One can stop the process at any order N to obtain a W that is a polynomial in ε and so converges. From the proof given above, it is clear that the terms in the series for H^* up to order N are unaffected by the termination. Thus the more useful form of Theorem 10.3.1 is the following.

Corollary 10.3.1. *Let $N \geq 1$ be given, and let $\{\mathcal{P}_i\}_{i=0}^N, \{\mathcal{Q}_i\}_{i=1}^N$, and $\{\mathcal{R}_i\}_{i=1}^N$ be sequences of linear spaces of smooth functions defined on a common domain O in $\mathbb{R}^{2n}$ with the following properties.*

1. *$\mathcal{Q}_i \subset \mathcal{P}_i$, $i = 1, 2, \ldots, N$.*
2. *$H_i^0 \in \mathcal{P}_i$, $i = 0, 1, 2, \ldots, N$.*
3. *$\{\mathcal{P}_i, \mathcal{R}_j\} \subset \mathcal{P}_{i+j}$, $i + j = 1, 2, \ldots, N$.*
4. *For any $D \in \mathcal{P}_i$, $i = 1, 2, \ldots, N$, there existB $\in \mathcal{Q}_i$ and $C \in \mathcal{R}_i$ such that*

$$B = D + \{H_0^0, C\}. \tag{10.26}$$

Then there exists a polynomial W,

$$W(\varepsilon, x) = \sum_{i=0}^{N-1} \left(\frac{\varepsilon^i}{i!}\right) W_{i+1}(x), \tag{10.27}$$

with $W_i \in \mathcal{R}_i, i = 1, 2, \ldots, N$, such that the change of variables $x = X(\varepsilon, y)$ where $X(\varepsilon, y)$ is the general solution of $dx/d\varepsilon = J\nabla W(\varepsilon, x), x(0) = y$, transforms the convergent Hamiltonian

$$H(\varepsilon, x) = H_*(\varepsilon, x) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) H_i^0(x) \tag{10.28}$$

to the convergent Hamiltonian

$$G(\varepsilon, x) = H^*(\varepsilon, y) = \sum_{i=0}^{N} \left(\frac{\varepsilon^i}{i!}\right) H_0^i(y) + O(\varepsilon^{N+1}), \tag{10.29}$$

with $H_0^i \in \mathcal{Q}_i, i = 1, 2, \ldots, N$.

The nonautonomous case. In the nonautonomous case, the algorithm is slightly different. The remainder function, $\mathcal{R}(\varepsilon, t, y)$, is the indefinite integral of $S^*(\varepsilon, t, y)$, where $S^*(\varepsilon, t, y) = -\mathcal{L}(W)(\partial W/\partial t)(s, t, y)$, the Lie transform of $S_* = -\partial W/\partial t$. One constructs two Lie triangles, one for the Hamiltonian H and one for the function S. Because $\mathcal{R}$ is the indefinite integral of S^*, if you want the new Hamiltonian up to terms of order ε^N, then you need all the Lie triangle for H_* down to the Nth row, but only down to the $(N-1)$st for S. One simply works down the two triangles together, but with the S triangle one row behind.

Assume that all the entries in the Lie triangle for H are known down to the Nth row $(H_j^i, i+j \leq N)$ and that all the entries in the Lie triangle for S_* are known down to the $(N-1)$st row $(S_j^i, i+j \leq N-1)$ using the W_i for $i \leq N$. Let $G_j^i, i+j \leq N$, be computed from the same Hamiltonian; so, $G_i^0 = H_i^0$ for all i, but with $U_1, \ldots, U_N$, where $U_i = W_i$ for $i = 1, 2, \ldots, N-1$ and $U_N = 0$. Let Q_j^i be the terms in the Lie triangle for the remainder using the U_i's. Then

$$H_j^i = G_j^i \text{ for } i+j < N, \qquad S_j^i = Q_j^i \text{ for } i+j < N-1,$$

$$H_j^i = G_j^i + \{H_0^0, W_N\} \text{ for } i+j = N, \quad S_j^i = Q_j^i - \frac{\partial W_N}{\partial t} \text{ for } i+j = N-1. \tag{10.30}$$

This is easily seen from the recursive formulas in Theorem 10.2.1.

From this observation, the algorithm is as follows. Assume that all the rows in the Lie triangle for H have been computed down to the $(N-1)$st row, that all the rows in the Lie triangle for S_* have been computed down to the $(N-2)$nd row and that $W_1, \ldots, W_{N-1}$ have been determined, and that the $H_0^1, \ldots, H_0^{N-1}$ are in normal form. Now it is time to compute W_N so that H_0^N is in normal form.

Step 1: Compute the Nth row for H and the $(N-1)$st row for the remainder assuming that $W_N = 0$, and call these terms $G_j^i, i+j = N$, and $Q_j^i, i+j = N-1$, respectively.

Step 2: Solve the equation $H_0^N = G_0^N + \mathcal{Q}_0^{N-1} + \{H_0^0, W_N\} - \partial W_N/\partial t$ for W_N and H_0^N so that H_0^N is in normal form or simple.

Step 3: Add $\{H_0^0, W_N\}$ to each term in the Nth row for H, and add $\partial W_N/\partial t$ to each term in the $(N-1)$st row for S.

Step 4: Repeat.

The nonautonomous version of Theorem 10.3.1 is as follows.

Theorem 10.3.2. *Let $\{\mathcal{P}_i\}_{i=0}^\infty$, $\{\mathcal{Q}_i\}_{i=1}^\infty$, and $\{\mathcal{R}_i\}_{i=1}^\infty$ be sequences of linear spaces of smooth functions defined on a common domain O in $\mathbb{R}^1 \times \mathbb{R}^{2n}$. Let $\dot{\mathcal{R}}_i$ be the space of all derivatives of functions in $\mathcal{R}_i$. Assume the following:*

1. *$\mathcal{Q}_i \subset \mathcal{P}_i, i = 1, 2, \ldots$.*
2. *$H_i^0 \in \mathcal{P}_i, i = 0, 1, 2, \ldots$.*
3. *$\{\mathcal{P}_i, \mathcal{R}_j\} \subset \mathcal{P}_{i+j}$ and $\{\mathcal{P}_i, \dot{\mathcal{R}}_j\} \subset \mathcal{P}_{i+j}$. for $i+j = 1, 2, \ldots$.*
4. *For any $D \in \mathcal{P}_i, i = 1, 2, \ldots,$ there exists $B \in \mathcal{Q}_i$ and $C \in \mathcal{R}_i$ such that*

$$B = D + \{H_0^0, C\} - \frac{\partial C}{\partial t}. \tag{10.31}$$

Then there exists a W *with a formal Hamiltonian of the form* (10.16) *with* $W_i \in \mathcal{R}_i, i = 1, 2, \ldots,$ *that generates a near-identity symplectic change of variables* $x \to y$ *such that the Hamiltonian in the new variables has a series expansion given by* (10.15) *with* $H_0^i \in \mathcal{Q}_i, i = 1, 2, \ldots.$

Duffing's equation revisited. Consider the Hamiltonian (10.23) of Duffing's equation as written in action–angle variables. The operator $\{H_0^0, C\} = \partial C / \partial \phi$ is very simple to understand. Equation (10.31) becomes

$$B = D + \frac{\partial C}{\partial \phi}.$$

If D is a finite Poisson series with d'Alembert character, then by taking B to be the term of D that is independent of the angle ϕ and $C = \int (B - D) d\phi, B$ and C satisfy this equation. This leads us to the following definitions of the spaces.

Let $\mathcal{P}_i$ be the space of all finite Poisson series with d'Alembert character corresponding to homogeneous polynomials of degree $2i + 2$ in rectangular coordinates. So an element in $\mathcal{P}_i$ is of the form I^{i+1} times a finite Fourier series in ϕ. Let $\mathcal{Q}_i$ be the space of all polynomials of the form AI^{i+1}, where A is a constant. Let $\mathcal{R}_i$ be the subspace of $\mathcal{P}_i$ of Poisson series without a term independent of ϕ. So $\mathcal{P}_i = \mathcal{Q}_i \oplus \mathcal{R}_i$. Because the Poisson bracket of homogeneous polynomials of degree $2i + 2$ and degree $2j + 2$ is a polynomial of degree $2(i+j)+2$, and because symplectic changes of coordinates preserve Poisson brackets, we have $\{\mathcal{P}_i, \mathcal{R}_j\} \subset \mathcal{P}_{i+j}$. Thus by Corollary 10.3.1, there exists a formal, symplectic transformation that transforms the Hamiltonian of Duffing's equation into the form

$$H^*(\varepsilon, J) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!} \right) H_0^i(J)$$

and the equations of motion become

$$\dot{J} = 0, \qquad \dot{\phi} = -\frac{\partial H}{\partial \phi}(\varepsilon, J) = -\omega(\varepsilon, J).$$

Thus formally, the solutions move on circles with a uniform frequency $\omega(\varepsilon, J)$, which depends on ε and J. By the theorems of Poincaré (1885) and Rüssman (1959) the series converges in this simple case.

Uniqueness of normal forms: One of the important special cases where Theorem 10.3.1 applies is when the operator $\mathcal{F}_i = \{H_0^0, \cdot\} : \mathcal{P}_i \to \mathcal{P}_i$ is simple; i.e., when $\mathcal{P}_i = Q_i \oplus \mathcal{R}_i$, $Q_i = \text{kernel}\,(\mathcal{F}_i)$, and $\mathcal{R}_i = \text{range}\,(\mathcal{F}_i)$. In this case, the Lie equation (10.25) has a unique solution. This is not enough to assure uniqueness of the normal form. One needs one extra condition.

Theorem 10.3.3. *Let $\{\mathcal{P}_i\}_{i=0}^{\infty}$, be sequences of linear spaces of smooth functions defined on a common domain O in $\mathbb{R}^{2n}$. Let $\mathcal{F}_i = \{H_0^0, \cdot\} : \mathcal{P}_i \to \mathcal{P}_i$ be simple; so, $\mathcal{P}_i = \mathcal{Q}_i \oplus \mathcal{R}_i$, $\mathcal{Q}_i =$ kernel$(\mathcal{F}_i)$, $\mathcal{R}_i =$ range$(\mathcal{F}_i)$. Assume*

1. $H_i^0 \in \mathcal{P}_i,\ i = 0, 1, 2, \ldots.$
2. $\{\mathcal{P}_i, \mathcal{R}_j\} \subset \mathcal{P}_{i+j},\ i + j = 1, 2, \ldots.$

Then there exists a W with a formal expansion of the form (10.16) *with $W_i \in \mathcal{R}_i, i = 1, 2, \ldots$, such that W generates a near-identity symplectic change of variables $x \to y$ which transforms the Hamiltonian $H_*(\varepsilon, x)$ with the formal series expansion given in Equation* (10.14) *to the Hamiltonian $H^*(\varepsilon, y)$ with the formal series expansion given by Equation* (10.15) *with $H_0^i \in \mathcal{Q}_i, i = 1, 2, \ldots.$*

Moreover, if

$$\{\mathcal{Q}_i, \mathcal{Q}_j\} = 0, \quad i, j = 1, 2, \ldots,$$

then the terms in the normal form are unique.

Remark. All the obvious remarks about the time-dependent cases hold here also. The normal form is unique, but the transformation taking the equation need not be unique. Clearly this theorem applies to the Duffing example. We do not need this theorem in our development. See Liu (1985) for a proof or see the Problems section.

10.4 Normal Form at an Equilibrium

Consider an analytic Hamiltonian H that has an equilibrium point at the origin in $\mathbb{R}^{2n}$, and assume that the Hamiltonian is zero at the origin. Then H has a Taylor series expansion of the form

$$H(x) = H_{\#}(x) = \sum_{i=0}^{\infty} H_i(x), \tag{10.32}$$

where H_i is a homogeneous polynomial in x of degree $i + 2$; so, $H_0(x) = \frac{1}{2}x^T Sx$, where S is a $2n \times 2n$ real symmetric matrix, and $A = JS$ is a Hamiltonian matrix. The linearized equations about the critical point $x = 0$ are

$$\dot{x} = Ax = JSx = J\nabla H_0(x), \tag{10.33}$$

and the general solution of (10.33) is $\phi(t, \xi) = \exp(At)\xi$.

The classical case. The matrix A is simple if it has $2n$ linearly independent eigenvectors that may be real or complex. The matrix A being simple is equivalent to A being similar to a diagonal matrix by a real or complex similarity transformation. This is why A is sometimes said to be diagonalizable. The classical theorem on normal forms is as follows.

Theorem 10.4.1. *Let A be simple. Then there exists a formal symplectic change of variables,*

$$x = X(y) = y + \cdots , \tag{10.34}$$

that transforms the Hamiltonian (10.32) *to*

$$H^{\#}(y) = \sum_{i=0}^{\infty} H^i(y), \tag{10.35}$$

where H^i is a homogeneous polynomial of degree $i+2$ such that

$$H^i(e^{At}y) \equiv H^i(y), \tag{10.36}$$

for all $i = 0, 1, \ldots$, all $y \in \mathbb{R}^{2n}$, and all $t \in \mathbb{R}$.

Remark. Formula (10.36) is the classical definition of normal form for a Hamiltonian near an equilibrium point with a simple linear part. Formula (10.36) says that H^i is an integral for the linear system (10.33); so, by Theorem 1.3.1, (10.36) is equivalent to

$$\{H^i, H^0\} = 0 \tag{10.37}$$

for all i.

Proof. In order to study the solutions near the origin, scale the variables by $x \to \varepsilon x$. This is a symplectic transformation with multiplier ε^{-2}; so, the Hamiltonian becomes

$$H(\varepsilon, x) = H_*(\varepsilon, x) = \sum_{i=0}^{\infty} \left(\frac{\varepsilon^i}{i!}\right) H_i^0(x), \tag{10.38}$$

where $H_i^0 = i! H_i$. Because we are working formally, we can set $\varepsilon = 1$ at the end, or we can rescale by $x \to \varepsilon^{-1}$x.

Let $\mathcal{P}_i$ be the linear space of all real homogeneous polynomials of degree $i+2$; so, $H_i^0 \in \mathcal{P}_i$. Because A is simple, A has $2n$ linearly independent eigenvectors $s_1, \ldots, s_{2n}$ corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_{2n}$. The s_i are row eigenvectors; so, $s_i A = \lambda_i s_i$. Let $2r$ of the eigenvalues be complex, and number them so that $\lambda_i = \bar{\lambda}_{n+i}$ for $i = 1, \ldots, r$. Choose the eigenvectors so that $s_i = \bar{s}_{n+i}$ for $i = 1, \ldots, r$. The other eigenvalues and eigenvectors are real. Let $K \in \mathcal{P}_i$; so, K is a homogeneous polynomial of degree $i+2$. Because the s_i are independent, K may be written in the form

$$K = \sum \kappa_{m_1 m_2 \ldots m_{2n}} (s_1 x)^{m_1} (s_2 x)^{m_2} \cdots (s_{2n} x)^{m_{2n}}, \tag{10.39}$$

where the sum is over all $m_1 + \cdots + m_{2n} = i + 2$. So the monomials in

$$B = \{(s_1 x)^{m_1} (s_2 x)^{m_2} \cdots (s_{2n} x)^{m_{2n}} : m_1 + \cdots + m_{2n} = i + 2\} \tag{10.40}$$

span $\mathcal{P}_i$. It is also clear that they are independent; so, form a basis for $\mathcal{P}_i$. The coefficients in (10.39) may be complex but must satisfy the reality condition that interchanging the subscripts m_i and m_{n+i} for $i = 1, \ldots, r$ in the κ coefficients is the same as conjugation.

Now let $\mathcal{F} = \mathcal{F}_i : \mathcal{P}_i \to \mathcal{P}_i$ be the linear operator of Theorem 10.3.3 as it applies to Hamiltonian systems, that is, define $\mathcal{F}$ by $\mathcal{F}(G) = \{H_0^0, G\} = -(\partial G/\partial x)Ax$; so,

$$\begin{aligned}\mathcal{F}(&(s_1x)^{m_1}(s_2x)^{m_2}\cdots(s_{2n}x)^{m_{2n}})\\&= -(m_1\lambda_1 + \cdots + m_{2n}\lambda_{2n})(s_1x)^{m_1}(s_2x)^{m_2}\cdots(s_{2n}x)^{m_{2n}}.\end{aligned}$$

So the elements of B are eigenvectors of $\mathcal{F}$ and the eigenvalues are $-(m_1\lambda_1 + \cdots + m_{2n}\lambda_{2n})$, $m_1 + \cdots + m_{2n} = i + 2$. Thus we can define $\mathcal{F}$-invariant subspaces

$$\begin{aligned}\mathcal{K}_i = \operatorname{span}\{&(s_1x)^{m_1}(s_2x)^{m_2}\cdots(s_{2n}x)^{m_{2n}} : m_1 + \cdots + m_{2n} = i + 2,\\ &m_1\lambda_1 + \cdots + m_{2n}\lambda_{2n} = 0\},\end{aligned}$$

$$\begin{aligned}\mathcal{R}_i = \{&(s_1x)^{m_1}(s_2x)^{m_2}\cdots(s_{2n}x)^{m_{2n}} : m_1 + \cdots + m_{2n} = i + 2,\\ &m_1\lambda_1 + \cdots + m_{2n}\lambda_{2n} \neq 0\}.\end{aligned}$$

In summary, $\mathcal{K}_i = \text{kernel}\,(\mathcal{F})$, $\mathcal{R}_i = \text{range}\,(\mathcal{F})$, and $\mathcal{P}_i = \mathcal{K}_i \oplus \mathcal{R}_i$. Thus this classical theorem follows from the first part of Theorem 10.3.3 because we have shown that the operators $\mathcal{F}_i : \mathcal{P}_i \to \mathcal{P}_i$ are simple. However, the extra condition in Theorem 10.3.3 is not satisfied in general; so, the normal form may not be unique.

Birkhoff (1927) considered a special case of the above.

Corollary 10.4.1. *Assume that the quadratic part of* (10.32) *is of the form*

$$H_0(x) = \sum_{j=1}^{n} \lambda_j x_j x_{n+j}, \tag{10.41}$$

where the λ_js are independent over the integers; i.e., there is no nontrivial relation of the form

$$\sum_{i=1}^{n} k_j \lambda_j = 0, \tag{10.42}$$

where the k_j are integers. Then there exists a formal symplectic change of variables $x = X(y) = y + \cdots$ that transforms the Hamiltonian (10.32) *to the Hamiltonian* (10.35)*, where $H^j(y)$ is a homogeneous polynomial of degree $j + 1$ in the n products $y_1y_{n+1}, \ldots, y_ny_{2n}$. So, $H^\#(y_1, \ldots, y_{2n}) = H^\#(y_1y_{n+1}, \ldots, y_ny_{2n})$ where $H^\#$ is a function of n variables. Moreover, in this case, the normal form is unique.*

Remark. Formally the equations of motion for the system in normal form are

$$\dot{y}_j = y_j D_j H^{\#}(y_1 y_{n+1}, \ldots, y_n y_{2n}),$$

$$\dot{y}_{j+n} = -y_{j+n} D_j H^{\#}(y_1 y_{n+1}, \ldots, y_n y_{2n}).$$

Here D_j stands for the partial derivative with respect to the jth variable. In this form, the system of equations has n formal integrals in involution, $I_1 = y_1 y_{n+1}, \ldots, I_n = y_n y_{2n}$.

In the case when the $\lambda_j = i\omega_j$ are pure imaginary and the y_j are the complex coordinates discussed in Lemma 3.3.4, then we can switch to action–angle variables by $y_j = \sqrt{I_j/2} e^{i\phi_j}$, $y_{n+j} = \sqrt{I_j/2} e^{-i\phi_j}$. The Hamiltonian in normal form is a function of the action variables only; so, the Hamiltonian is $H^{\dagger}(I_1, \ldots, I_n)$, and the equations of motion are

$$\dot{I}_j = \frac{\partial H^{\dagger}}{\partial \phi_j} = 0, \qquad \dot{\phi}_j = -\frac{\partial H^{\dagger}}{\partial I_j} = \omega_j(I_1, \ldots, I_n).$$

Here $\omega_i(I_1, \ldots, I_n) = \pm\omega_i + \cdots$, and the sign is determined by the cases in Lemma 3.3.2. Setting the action variables equal to nonzero constants, $I_1 = c_1, \ldots, I_n = c_n$, defines an invariant set which is an n-torus with n angular coordinates $\phi_1, \ldots, \phi_n$. On each torus the angular frequencies $\omega_j(I_1, \ldots, I_n)$, are constant, and so, define a linear flow on the torus as discussed in Section 1.2. The frequencies vary from torus to torus in general.

Notation. For this proof, and subsequent discussions, some notation is useful. Let $\mathcal{Z} = \mathbb{Z}_+^{2n}$ denote the set of all $2n$-tuples of nonnegative integers; so, $k \in \mathcal{Z}$ means $k = (k_1, \ldots, k_{2n}), k_i \geq 0$, k_i an integer. Let $|k| = k_1 + \cdots + k_{2n}$. If $x \in \mathbb{R}^{2n}$ and $k \in \mathcal{Z}$, then define $x^k = x_1^{k_1} x_2^{k_2} \cdots x_{2n}^{k_{2n}}$.

Proof. The linear part is clearly simple. Let $H^i(y) = \sum h_k y^k$, where the sum is over $k \in \mathcal{Z}, |k| = i + 2$. The general solution of the linear system is $y_i = y_{i0} \exp(\lambda_i t), y_{i+n} = y_{i+n,0} \exp(-\lambda_i t)$ for $i = 1, \ldots, n$. Formula (10.36) implies that $\sum h_k \exp\, t\{(k_1 - k_{n+1})\lambda_1 + \cdots + (k_n - k_{2n})\lambda_n\} y^k$ is constant in t, and this implies that $\{(k_1 - k_{n+1})\lambda_1 + \cdots + (k_n - k_{2n})\lambda_n = 0$. But because the $\lambda_i' s$ are independent over the integers, this implies $k_1 = k_{n+1}, \ldots, k_n = k_{2n}$. That is, H^i is a function of the products $y_1 y_{n+1}, \ldots, y_n y_{2n}$ only.

By the remark above, the kernel consists of those functions that depend only on $I_1, \ldots, I_n$ and not on the angles in action–angle variables. Therefore, the extra condition of Theorem 10.3.3 holds, and the normal form is unique.

Remark. If the condition (10.42) only holds for $|k_1| + \cdots + |k_n| \leq N$, then the terms in the Hamiltonian up to the terms of order N can be put in normal form, and these terms are unique.

The general equilibria. In the 1970s, the question of the stability of the Lagrange triangular point $\mathcal{L}_4$ was studied intensely. For Hamiltonian systems, it is not enough to look at the linearized system alone, because the higher-order terms in the normalized equations can change the stability (see the discussion in Chapter 11). The matrix of the linearization of the equations at $\mathcal{L}_4$ when $\mu = \mu_1$ is not simple as was seen in Section 4.1. The normal form for this case, and other similar cases was carried out by the Russian school; see Sokol'skij (1978). First Kummer (1976,1978) and then Cushman, Deprit, and Mosak (1983), used group representation theory. Representation theory is very helpful in understanding the general case, but there are simpler ways to understand the basic ideas and examples. In Meyer (1984b) a theorem like Theorem 10.4.1 above was given for non-Hamiltonian systems but A was replaced by A^T in (10.36); so, the terms in the normal form are invariant under the flow $\exp(A^T t)$. A far better proof can be found in Elphick et al. (1987), which is what we present here.

The proof of Theorem 10.4.1 rested on the fact that for a simple matrix, A, the vector space $\mathbb{R}^{2n}$ is the direct sum of the range and kernel of A, and this held true for the operator $\mathcal{F} = \{H_0^0, \cdot\}$ defined on homogeneous polynomials as well. The method of Elphick et al. is based on the following simple lemma in linear algebra known as the Fredholm alternative and an inner product defined on homogeneous polynomials given after the lemma.

Lemma 10.4.1. *Let $\mathbb{V}$ be a finite-dimensional inner product space with inner product $(\cdot,\cdot)$. Let $A : \mathbb{V} \to \mathbb{V}$ be a linear transformation, and A^* its adjoint (so $(Ax, y) = (x, A^*y)$ for all $x, y \in \mathbb{V}$). Then $\mathbb{V} = R \oplus K^*$ where R is the range of A and K^* is the kernel of A^*.*

Proof. Let $x \in R$; so, there is a $u \in \mathbb{V}$ such that $Au = x$. Let $y \in K^*$; so, $A^*y = 0$. Because $0 = (u, 0) = (u, A^*y) = (Au, y) = (y, x)$, it follows that R and K^* are orthogonal subspaces. Let K be the kernel of A. In a finite dimensional space, $\dim \mathbb{V} = \dim R + \dim K$ and $\dim K = \dim K^*$. Because R and K^* are orthogonal, $\dim(R + K^*) = \dim R + \dim K^* = \dim \mathbb{V}$; so, $\mathbb{V} = R \oplus K^*$.

Let $\mathcal{P} = \mathcal{P}_j$ be the linear space of all homogeneous polynomials of degree j in $2n$ variables $x \in \mathbb{R}^{2n}$. So if $P \in \mathcal{P}$, then

$$P(x) = \sum_{|k|=j} p_k x^k = \sum_{|k|=j} p_{k_1 k_2 \ldots k_{2n}} x_1^{k_1} x_2^{k_2} \cdots x_{2n}^{k_{2n}}.$$

Define $P(\partial)$ to be the differential operator

$$P(\partial) = \sum_{|k|=j} p_k \frac{\partial^k}{\partial x^k},$$

where we have introduced the notation

$$\frac{\partial^k}{\partial x^k} = \frac{\partial^{k_1}}{\partial x_1^{k_1}} \frac{\partial^{k_2}}{\partial x_2^{k_2}} \cdots \frac{\partial^{k_{2n}}}{\partial x_{2n}^{k_{2n}}}.$$

Let $Q \in \mathcal{P}$, $Q(x) = \sum q_h x^h$ be another homogeneous polynomial, and define an inner product $< \cdot, \cdot >$ on $\mathcal{P}$ by

$$< P, Q >= P(\partial)Q(x).$$

To see that this is indeed an inner product, note that $\partial^k x^h / \partial x^k = 0$ if $k \neq h$ and $\partial^k x^h / \partial x^k = k! = k_1! k_2! \cdots k_{2n}!$ if $k = h$; so,

$$< P, Q >= \sum_{|k|=j} k! p_k q_k.$$

Let $A = JS$ be a Hamiltonian matrix where S is a symmetric matrix of the quadratic Hamiltonian H^0; so, $H^0(x) = \frac{1}{2} x^T S x$. From Theorem 10.3.1 and the proof of Theorem 10.4.1, the operator of importance is $\mathcal{F}(A) : \mathcal{P} \to \mathcal{P}$, where

$$\mathcal{F}(A)P = \{H_0^0, P\} = -\frac{\partial P}{\partial x} A x = \frac{d}{dt} P(e^{At} x)\Big|_{t=0}. \tag{10.43}$$

Lemma 10.4.2. *Let $A : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ be as above and A^T its transpose (so A^T is the adjoint of A with respect to the standard inner product in $\mathbb{R}^{2n}$). Then for all $P, Q \in \mathcal{P}$,*

$$< P(x), Q(Ax) >=< P(A^T x), Q(x) > \tag{10.44}$$

and

$$< P, \mathcal{F}(A)Q >=< \mathcal{F}(A^T)P, Q > . \tag{10.45}$$

That is, the adjoint of $\mathcal{F}(A)$ with respect to $< \cdot, \cdot >$ is $\mathcal{F}(A^T)$.

Proof. Equation (10.44) follows from (10.43) because (10.43) implies

$$< P(x), Q(e^{At} x) >=< P(e^{A^T t} x), Q(x) > .$$

Differentiating this last expression with respect to t and setting $t = 0$ gives (10.45).

Let $y = Ax$ (i.e., $y^i = \sum_j A^i_j x^j$) and $F(y) = F(Ax)$. Inasmuch as

$$\frac{\partial F(y)}{\partial x^j} = \sum_i \frac{\partial F(y)}{\partial y^i} \frac{\partial y^i}{\partial x^j} = \sum_i \frac{\partial F(y)}{\partial y^i} A^i_j,$$

it follows that $\partial / \partial x = A^T \partial / \partial y$.

$$< P(x), Q(Ax) >= P(\partial_x) Q(Ax) = P(A^T \partial_y) Q(y) =< P(A^T y), Q(y) > .$$

Theorem 10.4.2. *Let A be a Hamiltonian matrix. Then there exists a formal symplectic change of variables, $x = X(y) = y + \cdots$, that transforms the Hamiltonian* (10.32) *to*

$$H^{\#}(y) = \sum_{j=0}^{\infty} H^j(y), \tag{10.46}$$

where H^j is a homogeneous polynomial of degree $j + 2$ such that

$$H^j(e^{A^T t} y) \equiv H^j(y), \tag{10.47}$$

for all $j = 0, 1, \ldots,$ all $y \in \mathbb{R}^{2n}$, and all $t \in \mathbb{R}$.

Remark. Let $H_0^T(x) = H_T^0(x) = \frac{1}{2}x^T Rx$ be the quadratic Hamiltonian for the adjoint linear equation; so, $A^T = JR$. Then (10.47) is equivalent to

$$\{H^i, H_T^0\} = 0$$

for $j = 1, 2, \ldots$.

Proof. By Theorem 10.3.1, we must solve Equation (10.25) or $\mathcal{F}(A)C + D = B$, where $D \in P_j = \mathcal{P}$ is given, and $C \in Q_j = \mathcal{P}$, and $D \in Q_j =$ kernel $(\mathcal{F}(A^T))$. By Lemma 10.4.2, we can write $D = B - G$, where $B \in$ kernel$(\mathcal{F}(A^T))$; so, $\{B, H_T^0\} = 0$, and $G \in$ range $(\mathcal{F}(A))$; so, $G = \mathcal{F}(A)C$, $C \in \mathcal{P}$. With these choices, (10.30) is solved. Verification of the rest of the hypothesis in Theorem 10.3.1 is just as in the proof of Theorem 10.4.1.

Theorem 10.4.1 is a corollary of this theorem because when A is simple, it is diagonalizable, and so, its own adjoint. We proved Theorem 10.4.1 separately, because the proof is constructive.

Examples of normal forms in the nonsimple case. Consider the Hamiltonian system (10.32), where $n = 1$ and $x = (q, p)$. Let

$$H_0(q, p) = p^2/2, \qquad H_0^T(q, p) = -q^2/2,$$

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad A^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Because

$$\exp(A^T t) = \begin{bmatrix} 1 & 0 \\ 1 + t & 1 \end{bmatrix},$$

(10.47) implies that the higher-order terms in the normal form are independent of p, or $H^i = H^i(p, \cdot)$. Thus the Hamiltonian in normal form is $p^2/2 + G(q)$, which is the Hamiltonian for the second-order equation $\ddot{q} + g(q) = 0$, where $g(q) = \partial G(q)/\partial \text{q}$.

Now consider a Hamiltonian system with two degrees of freedom with a linearized system with repeated pure imaginary roots that are nonsimple. In Section 4.6, the normal form for the quadratic part of such a Hamiltonian was given as

$$H_0 = \omega(\xi_2\eta_1 - \xi_1\eta_2) + \frac{\delta}{2}(\xi_1^2 + \xi_2^2),$$

where $\omega \neq 0$ and $\delta = \pm 1$. The linearized equations are

$$\begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\eta}_1 \\ \dot{\eta}_2 \end{bmatrix} = \begin{bmatrix} 0 & \omega & 0 & 0 \\ -\omega & 0 & 0 & 0 \\ -\delta & 0 & 0 & \omega \\ 0 & -\delta & -\omega & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \eta_1 \\ \eta_2 \end{bmatrix}.$$

The transpose is defined by the Hamiltonian

$$H_0^T = \omega(\xi_2\eta_1 - \xi_1\eta_2) - \frac{\delta}{2}(\eta_1^2 + \eta_2^2),$$

and the transposed equations are

$$\begin{bmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\eta}_1 \\ \dot{\eta}_2 \end{bmatrix} = \begin{bmatrix} 0 & -\omega & -\delta & 0 \\ \omega & 0 & 0 & -\delta \\ 0 & 0 & 0 & -\omega \\ 0 & 0 & \omega & 0 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \eta_1 \\ \eta_2 \end{bmatrix}.$$

Sokol'skij (1978) suggested changing to polar coordinates (see Section 6.2) to make the transposed equations simple. That is, he changed coordinates by

$$\eta_1 = r\cos\theta, \qquad R = (\xi_1\eta_1 + \xi_2\eta_2)/r,$$

$$\eta_2 = r\sin\theta, \qquad \Theta = \eta_1\xi_2 - \eta_2\xi_1.$$

In these coordinates,

$$H_0^T = -\omega\Theta + \frac{\delta}{2}r^2, \qquad H_0 = \omega\Theta + \frac{\delta}{2}\left(R^2 + \frac{\Theta^2}{r^2}\right),$$

and the transposed equations are

$$\dot{r} = 0, \quad \dot{\theta} = \omega, \quad \dot{R} = \delta r, \quad \dot{\Theta} = 0.$$

Thus the higher order terms in the normal form are independent of θ and R and so depend only on $r^2 = \eta_1^2 + \eta_2^2$ and $\Theta = \eta_1\xi_2 - \eta_2\xi_1$.

Thus the theory of the normal form in this case depends on three qualities

$$\Gamma_1 = \xi_2\eta_1 - \xi_1\eta_1, \qquad \Gamma_2 = \frac{1}{2}(\xi_1^2 + \xi_2^2), \qquad \Gamma_3 = \frac{1}{2}(\eta_1^2 + \eta_2^2).$$

The Hamiltonian $H_0 = \omega\Gamma_1 + \Gamma_2$ and the higher-order terms in the normal form are functions of Γ_1 and Γ_3 only. This is known as Sokol'skij's normal form.

10.5 Normal Form at $\mathcal{L}_4$

Recall that in Section 4.1, we showed that the linearization of the restricted 3-body problem at the Lagrange triangular point $\mathcal{L}_4$ had two pairs of pure imaginary eigenvalues, $\pm i\omega_1, \pm i\omega_2$ when $0 < \mu < \mu_1 = \frac{1}{2}(1 - \sqrt{69}/9)$, and that there are symplectic coordinates so that the quadratic part of the Hamiltonian is

$$H_2 = \omega_1 I_1 - \omega_2 I_2,$$

where I_1, I_2, ϕ_1, ϕ_2 are action–angle variables.

Recall that in Section 8.5, we defined μ_r to be the value of μ for which $\omega_1/\omega_2 = r$, and that $0 \cdots < \mu_3 < \mu_2 < \mu_1$. When $0 < \mu < \mu_1$, and $\mu \neq \mu_2, \mu_3$ then by Corollary 10.4.1, the Hamiltonian of the restricted 3-body problem can be normalized through the fourth-order terms; so, the Hamiltonian becomes

$$H = \omega_1 I_1 - \omega_2 I_2 + \frac{1}{2}(AI_1^2 + 2BI_1I_2 + CI_2^2) + \cdots.$$

After six months of hand calculations, Deprit and Deprit-Bartholome computed:

$$A = \frac{1}{72}\omega_2^2 \frac{(81 - 696\omega_1^2 + 124\omega_1^4)}{(1 - 2\omega_1^2)^2(1 - 5\omega_1^2)},$$

$$B = -\frac{1}{6}\frac{\omega_1\omega_2(43 + 64\omega_1^2\omega_2^2)}{(1 - 2\omega_1^2)^2(1 - 5\omega_1^2)},$$

$$C(\omega_1, \omega_2) = A(\omega_2, \omega_1).$$

Meyer and Schmidt (1986) computed the normal form through terms of sixth-order by computer. The results are too lengthy to reproduce here. It did serve as an independent check of the calculations of Deprit and Deprit-Bartholome. In Section 4.1, the quadratic part of the Hamiltonian of the restricted 3-body problem at $\mathcal{L}_4$ for $\mu = \mu_1$ was brought into normal form by a linear symplectic change of coordinates. In these coordinates, the quadratic part of the Hamiltonian is of the form

$$H_0 = \omega(\xi_2\eta_1 - \xi_1\eta_2) + \frac{1}{2}(\xi_1^2 + \xi_1^2) = \omega\Gamma_1 + \Gamma_2,$$

where $\omega = \sqrt{2}/2$ and $\delta = +1$.

The normal form for the Hamiltonian of the restricted 3-body problem at $\mathcal{L}_4$ for $\mu = \mu_1$ is of the form

$$\begin{aligned} H &= \omega\Gamma_1 + \Gamma_2 + c\Gamma_1^2 + 2d\Gamma_1\Gamma_3 + 4e\Gamma_3^2 + \cdots \\ &= \omega(\xi_2\eta_1 - \xi_1\eta_2) + \tfrac{1}{2}(\xi_1^2 + \xi_1^2) \\ &\qquad + c(\xi_2\eta_1 - \xi_1\eta_2)^2 + d(\eta_1^2 + \eta_2^2)(\xi_2\eta_1 - \xi_1\eta_2) + e(\eta_1^2 + \eta_2^2)^2 + \cdots \end{aligned}$$

where c, d, e are constants. As another related problem, consider a quadratic Hamiltonian $Q(y, \varepsilon)$ that depends on a parameter ε, which for $\varepsilon = 0$ is H_0. That is, $Q(y, \varepsilon) = Q_0(y) + \varepsilon Q_1(y) + \cdots$, where $Q_0 = H_0$. Then this Hamiltonian can be brought into normal form to an order so that $Q_1, Q_2, \ldots$ depend only on Γ_1 and Γ_3. (See Schmidt (1990) for the calculations.)

The quadratic part of the Hamiltonian of the restricted 3- body problem at the Lagrange triangular point, $\mathcal{L}_4$, for values of the mass ratio parameter $\mu = \mu_1 + \varepsilon$ can be brought into normal form by a linear symplectic change of coordinates. The normal form up to order 4 looks like

$$\begin{aligned}
Q &= \omega\Gamma_1 + \Gamma_2 + \varepsilon\{a\Gamma_1 + b\Gamma_3\} + \cdots \\
&= \omega(\xi_2\eta_1 - \xi_1\eta_2) + \tfrac{1}{2}(\xi_1^2 + \xi_1^2) \\
&\qquad + \varepsilon\{a(\xi_2\eta_1 - \xi_1\eta_2) + \tfrac{1}{2}b(\eta_1 + \eta_2) + \cdots .
\end{aligned}$$

Schmidt (1990) calculated that

$$a = 3\sqrt{69}/16, \qquad b = 3\sqrt{69}/8.$$

10.6 Normal Forms for Periodic Systems

This section reduces the study of the normal forms for symplectomorphisms to the study of normal forms of periodic systems. Then as examples, the normal forms for symplectomorphisms of the plane are given in preparation for the study of generic bifurcations of fixed points given in Chapter 11.

The reduction. The study of a neighborhood of a periodic solution of an autonomous Hamiltonian system was reduced to the study of the Poincaré map in an energy surface by the discussion in Section 8.5. This Poincaré map is a symplectomorphism with a fixed point corresponding to the periodic orbit.

Let the origin be a fixed point for the symplectomorphism

$$\Psi(x) = \Gamma x + \psi(x), \tag{10.48}$$

where Γ is a $2n \times 2n$ symplectic matrix, and ψ is higher-order; i.e., $\psi(0) = \partial\psi(0)/\partial x = 0$. By Theorem 8.2.1 and the discussion following that theorem, if Γ has a logarithm, then (10.48) is the period map of a periodic Hamiltonian system. Because $\Psi^2(x) = \Gamma^2 x + \cdots$, and Γ^2 always has a logarithm, if Ψ is not a period map, then Ψ^2 is. Except for one example given at the end of this chapter, only the case when Γ has a real logarithm is treated here.

Given a periodic system, by the Floquet–Lyapunov theorem (see Theorem 3.4.2 and the discussion following it), there is a linear, symplectic, periodic change of variables that makes the linear part of Hamiltonian equations constant in t. Thus the study of symplectomorphisms near a fixed point is equivalent to studying a 2π-periodic Hamiltonian system of the form

$$H_\#(t, x) = \sum_{i=0}^{\infty} H_i(t, x), \tag{10.49}$$

where H_i is a homogeneous polynomial in x of degree $i + 2$ with 2π-periodic coefficients, and $H_0(t, x) = \frac{1}{2}x^T S x$ where S is a $2n \times 2n$ real constant symmetric matrix, and $A = JS$ is a constant, real, Hamiltonian matrix. The linearized equations about the critical point $x = 0$ are

$$\dot{x} = Ax = JSx = J\nabla H_0(x), \tag{10.50}$$

and the general solution of (10.50) is $\phi(t, \xi) = \exp(At)\xi$.

The general periodic case. Here, the generalization of the general normal form given in Section 10.4 is extended to periodic systems. As before, we consider the periodic system (10.49) but no longer assume that the linear system is simple. First let us consider the generalization of Theorem 10.4.2.

Consider the 2π-periodic equations

$$\dot{x} = A(t)x + f(t), \tag{10.51}$$

$$\dot{x} = A(t)x, \tag{10.52}$$

$$\dot{y} = -A(t)^T y. \tag{10.53}$$

Equation (10.52) is the homogeneous equation corresponding to the nonhomogeneous equation (10.51), and (10.53) is the adjoint equation of (10.52).

Lemma 10.6.1. *The nonhomogeneous equation* (10.51) *has a 2π-periodic solution $\phi(t)$ if and only if*

$$\int_0^{2\pi} y^T(s)f(s)ds = 0,$$

for all 2π-periodic solutions $y(t)$ of the adjoint equation (10.53)*.*

Proof. Let $x(t, x_0)$ be the solution of (10.51) with $x(0, x_0) = x_0$. Then

$$x(t, x_0) = X(t)x_0 + \int_0^t X(t)Y^T(s)f(s)ds,$$

where $X(t)$ and $Y(t)$ are the fundamental matrix solutions of (10.52) and (10.53), respectively; so, $X^{-1} = Y^T$. The solution is 2π-periodic if and only if $x(t, x_0) = x$, or

$$Bx_0^0 = g,$$

where

$$B = I - X(2\pi), \qquad g = \int_0^{2\pi} X(2\pi)Y^T(s)f(s)ds.$$

By Lemma 10.4.1, the linear equation $Bx_0 = g$ has a solution if and only if $v^T g = 0$ for all v with $B^T v = 0$. That is, there is a 2π-periodic solution if and only if

$$\int_0^{2\pi} v^T X(2\pi)Y^T(s)f(s)ds = 0 \quad \text{for all } v \text{ with } X(2\pi)^T v = v.$$

But if $X(2\pi)^T v = v$, then the integral above is $\int_0^{2\pi} v^T Y^T(s)f(s)ds = 0$. But $X(2\pi)^T v = v$ if and only if $Y(2\pi)v = v$ and if and only if $Y(s)v$ is a 2π-periodic solution of (10.53).

Consider the periodic Hamiltonian system (10.49). Scale by $x \to \varepsilon x$ as in the proof of Theorem 10.4.1, and use the same notation for the scaled Hamiltonian. By Theorem 10.3.2 we must define spaces $\mathcal{P}_i, \mathcal{Q}_i$, and $\mathcal{R}_i$ with $\mathcal{Q}_i \subset \mathcal{P}_i$, $H_i^0 \in \mathcal{P}_i$, $H_0^i \in \mathcal{Q}_i$, $W_i \in \mathcal{R}_i$. The Lie equation to be solved in this case is

$$E = D + \{H_0^0, C\} - \frac{\partial C}{\partial t},$$

where D is given in $\mathcal{P}_i$, and we are to find $E \in \mathcal{Q}_i$ and $C \in \mathcal{R}_i$.

Let B be the adjoint of A; i.e., the transpose in the real case. Define $K(x) = (1/2)x^T Rx$, where $B = JR$; so, K is the Hamiltonian of the adjoint linear system. Let $\mathcal{P}_i$ be the space of polynomials in x with coefficients that are smooth 2π-periodic functions of t. Let $\mathcal{F} = \{H_0^0, \cdot\} : \mathcal{P}_i \to \mathcal{P}_i$, and let $\mathcal{T} = \{K, \cdot\} : \mathcal{P}_i \to \mathcal{P}_i$. $\mathcal{T}$ is the adjoint of $\mathcal{F}$ if we use the metric defined by Elphick et al. that was used in Section 10.4. Therefore, given D, the Lie equation has a unique 2π-periodic solution, C, where E is a 2π-periodic solution of the homogeneous adjoint equation

$$0 = \{K, E\} + \frac{\partial E}{\partial t}. \tag{10.54}$$

Characterizing the 2π-periodic solutions of (10.54) defines the normal form. Expand the elements of $\mathcal{P}_i$ in Fourier series. Let $E = d(x)e^{imt}$, and substitute into (10.54) to get

$$0 = \{K, d\} + imd.$$

Thus one characterization of the normal form is in terms of the eigenvectors of $\mathcal{T} = \{K, \cdot\} : \mathcal{P}_i \to \mathcal{P}_i$. That is, $\mathcal{Q}_i$ has a basis of the form $\{d(x)e^{imt} : d$ is an eigenvector of $\mathcal{T}$ corresponding to the eigenvalue im. $\}$

Theorem 10.6.1. *Let $H^0(x) = H_0(x) = \frac{1}{2}x^T Sx$, where $A = JS$ is an arbitrary, constant Hamiltonian matrix, and let B be the adjoint of A. Then there exists a formal, symplectic, 2π-periodic change of variables $x = X(t, y) = y + \cdots$ which transforms the Hamiltonian* (10.49) *to the Hamiltonian system*

$$\dot{y} = J\nabla H^{\#}(t, y), \qquad H^{\#}(t, y) = \sum_{i=0}^{\infty} H^i(t, y), \tag{10.55}$$

where

$$\{H^i, K\} + \frac{\partial H^i}{\partial t} = 0 \quad \textit{for } i = 1, 2, 3, \ldots, \tag{10.56}$$

or equivalently,

$$H^i(t, e^{Bt}x) \equiv H^i(0, x) \quad \textit{for } i = 1, 2, 3, \ldots. \tag{10.57}$$

Corollary 10.6.1. *Let A be simple and have eigenvalues $\pm\lambda_1, \ldots, \pm\lambda_n$. Assume that $\lambda_1, \ldots, \lambda_n$ and i are independent over the integers; i.e. there is no relation of the form $k_1\lambda_1 + \cdots + k_n\lambda_n = mi$, where $k_1, \ldots, k_n$ and m are* *integers. Then there exists a formal, symplectic, 2π-periodic change of variables $x = X(t, y) = y + \cdots$ which transforms the Hamiltonian* (10.49) *to an autonomous Hamiltonian system*

$$\dot{y} = J\nabla H^{\#}(y), \quad H^{\#}(y) = \sum_{i=0}^{\infty} H^i(y), \tag{10.58}$$

where $H^0 = H_0$, and

$$\{H^i, H^0\} = 0, \tag{10.59}$$

or equivalently,

$$H^i(e^{At}y) \equiv H^i(y) \tag{10.60}$$

for all $i = 0, 1, 2, \ldots,\ y \in \mathbb{R}^{2n}, t \in \mathbb{R}$.

Proof. Let $A = B = \operatorname{diag}(\lambda_1, \ldots, \lambda_n, -\lambda_1, \ldots, -\lambda_n)$. A typical term in the normal form given by Theorem 10.6.1 is of the form $h(t, x) = h_k e^{imt} x^k$. Applying (10.57) to this term gives

$$h_k \exp\{im + (k_1 - k_{n+1})\lambda_1 + \cdots + (k_n - k_{2n})\lambda_n\}t = 0.$$

By the assumption on the independence, this can only hold if $m = 0, k_1 = k_{n+1}, \ldots, k_n = k_{2n}$. Thus the Hamiltonian is in the normal form of Birkhoff as described in Corollary 10.4.1.

Corollary 10.6.2. *Let Γ be simple and have a real logarithm. Then there exists a formal, near-identity, symplectic change of variables $x \to y$ such that in the new coordinates the symplectomorphism in* (10.48) *is of the form*

$$\Phi(y) = \Gamma y + \phi(y), \tag{10.61}$$

where

$$\phi(\Gamma y) \equiv \Gamma \phi(y) \quad \text{or} \quad \Phi(\Gamma y) \equiv \Gamma \Phi(y). \tag{10.62}$$

Proof. Let $\Gamma = \exp(2\pi A)$. Because Γ is simple, so is A, and therefore it can be taken as its own adjoint. Then by the reduction given above, the map (10.48) is the period map of a system of Hamiltonian differential equations. Assume that the symplectic change of coordinates has been made so that the Hamiltonian is in normal form, and let the equations in these coordinates be $\dot{y} = Ay + f(t, y)$. Condition (10.57) implies $f(t, e^{At}x) = e^{At} f(0, x)$, and this implies $f(t, \Gamma x) = \Gamma f(t, x)$. Let $\xi(t, \eta)$ be a solution of this equation with $\xi(0, \eta) = \eta$. Define $\zeta(t, \eta) = \Gamma \xi(t, \Gamma^{-1}\eta)$, so $\xi(0, \eta) = \zeta(0, \eta) = \eta$. $\dot{\xi} = \Gamma\{A\xi + f(t, \xi)\} = A\Gamma\xi + \Gamma f(t, \xi) = A\Gamma\xi + f(t, \Gamma\xi) = A\zeta + f(t, \zeta)$. By the uniqueness theorem for ordinary differential equations $\xi(t, \eta) = \zeta(t, \eta) = \Gamma\xi(t, \Gamma\eta)$; so, the period map satisfies (10.61).

General hyperbolic and elliptic points. Consider as examples the case when $n = 1$; so, Ψ in (10.48) is a symplectomorphism of the plane with a fixed point at the origin.

First, consider the case when Γ has eigenvalues μ, μ^{-1}, where $0 < \mu < 1$; i.e., the origin is a hyperbolic fixed point. By Lemma 3.3.7, there are symplectic coordinates, say x, so that

$$\Gamma = \begin{bmatrix} \mu & 0 \\ 0 & \mu^{-1} \end{bmatrix}.$$

Let $2\pi\alpha = \ln \mu$; so, $\Gamma = \exp(2\pi A)$ where

$$A = \begin{bmatrix} \alpha & 0 \\ 0 & -\alpha \end{bmatrix}.$$

By the discussion given above, the symplectomorphism Ψ is the period map of the 2π-periodic system (10.49) with $H_0(x) = \alpha x_1 x_2$. By Corollary 10.6.1, there is a formal, 2π-periodic, symplectic change of variables $x \to y$ that transforms (10.49) to the autonomous system (10.58) with (10.60) holding. The solution of the linear system is $y_1(t) = y_{10}e^{\alpha t}$, $y_2(t) = y_{20}e^{-\alpha t}$, therefore the condition (10.60) implies that the Hamiltonian (10.58) is a function of the product $y_1 y_2$ only. Let $H^{\#}(y) = K^{\#}(y_1 y_2) = \alpha y_1 y_2 + K(y_1 y_2)$. By the above discussion, the normal form for (10.48) is the time 2π-map of the autonomous system whose Hamiltonian is $K^{\#}$. The equations defined by $K^{\#}$ are

$$\dot{y}_1 = y_1(\alpha + k(y_1 y_2)),$$

$$\dot{y}_2 = -y_2(\alpha + k(y_1 y_2)),$$

where k is the derivative of K. These equations have $y_1 y_2$ as an integral, and so the equations are solvable, and the solution is

$$y_1(t) = y_{10} \exp(t(\alpha + k(y_1 y_2))),$$

$$y_2(t) = y_{20} \exp(-t(\alpha + k(y_1 y_2))).$$

Thus the normal form for (10.48) in this case is

$$\Psi(y) = \begin{bmatrix} y_1 g(y_1 y_2) \\ \\ y_2 g(y_1 y_2)^{-1} \end{bmatrix},$$

where g has a formal expansion $g(u) = \mu u + \cdots$. If a symplectomorphism is in this form, then the origin is called a general hyperbolic point. This map takes the hyperbolas $y_1 y_2 =$ constant into themselves. In this case, the transformation to normal form converges by a classical theorem of Moser (1956).

Next consider the case when A has eigenvalues $\lambda = \alpha + \beta i$, $\bar{\lambda} = \alpha - \beta i$, where $\alpha^2 + \beta^2 = 1, \beta \neq 0$; i.e., the origin is an elliptic fixed point. By Lemma 3.3.9 there are symplectic coordinates, say x, so that

$$\Gamma = \begin{bmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{bmatrix}.$$

Let $\Gamma = \exp(2\pi A)$, where

$$A = \begin{bmatrix} \omega i & 0 \\ 0 & -\omega i \end{bmatrix}.$$

Assume that ω is not an integer; that is, λ is not a root of unity. By the discussion given above, the symplectomorphism Ψ is the period map of the 2π-periodic system (10.49) with $H_0(x) = i\omega x_1 x_2$. By Corollary 10.6.1, there is a formal, 2π-periodic, symplectic change of variables, $x \to y$, which transforms (10.49) to the autonomous system (10.58) satisfying (10.60). Equation (10.60) implies that the Hamiltonian is a function of $y_1 y_2$ only. Let $H^{\#}(y) = K^{\#}(y_1 y_2) = i\omega y_1 y_2 + iK(y_1 y_2)$. By the above discussion, the normal form for (10.48) is the time 2π-map of the autonomous system whose Hamiltonian is $K^{\#}$. Change to action–angle variables (I, ϕ); so, the Hamiltonian becomes $H^{\#}(I, \phi) = K^{\#}(I) = \omega I + K(I)$. The equations defined by $K^{\#}$ are

$$\dot{I} = 0, \qquad \dot{\phi} = \omega - k(I)$$

where k is the derivative of K. These equations have I as an integral, and so the equations are solvable, and the solution is

$$I(t) = I_0, \qquad \phi(t) = \phi_0 + (-\omega + k(I_0))t.$$

Thus the normal form for (10.48) in action–angle variables in this case is

$$\Psi(I, \phi) = \begin{bmatrix} I \\ \phi + g(I) \end{bmatrix},$$

where g has a formal expansion $g(I) = -\omega + \beta I \cdots$. If a symplectomorphism is in this form with $\beta \neq 0$, then the origin is called a general elliptic point, or Ψ is called a twist map. This map takes circles into circles and rotates each circle by an amount $g(I)$.

Higher resonance in the planar case. Let us consider the case when $n = 1$, and the symplectomorphism Ψ has an elliptic fixed point whose multiplier is a root of unity. Theorem 10.6.1 and Corollary 10.6.2 apply as well.

Let Γ have eigenvalues $\lambda = \alpha + \beta i$, $\bar{\lambda} = \alpha - \beta i$, where λ is a kth root of unity; so, $\lambda^k = 1$, $k > 2$, and $\lambda = \exp(h 2\pi i/k)$, where h is an integer. The origin is called a k-resonance elliptic point in this case. By Lemma 3.3.9, there are symplectic coordinates, say x, so that

$$\Gamma = \begin{bmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{bmatrix}.$$

Let $\Gamma = \exp(2\pi A)$, where

$$A = \begin{bmatrix} (h/k)i & 0 \\ 0 & -(h/k)i \end{bmatrix}.$$

Because A is diagonal, it is its own adjoint. By the discussion given above the symplectomorphism Ψ is the period map of the 2π-periodic system (10.49) with $H_0(x) = (hi/k)(x_1x_2)$, where the reality condition is $\bar{x}_1 = x_2$. The normal form for the Hamiltonian is given by Theorem 10.6.1 above.

Let $h(t, x)$ be a typical term in the normal form expansion, so

$$h = e^{ist} x_1^{m_1} x_2^{m_2}.$$

The term h satisfies (10.57) if and only if

$$(h/k)(m_1 - m_2)i + si = 0;$$

so it is in the normal form if h is

$$(x_1x_2)^m \qquad \text{or} \qquad x_1^{m_1} x_2^{m_2} e^{-rit},$$

where $r = (m_1 - m_2)h/k$, and m, m_1, m_2, r are integers.

In action–angle coordinates (I, ϕ), $H^0(I, \phi) = (h/k)I$, and the solution of the linear system is $I = I_0$, $\phi = \phi_0 - (h/k)t$. Thus $H^{\#}(t, I, \phi)$ is a function of I and $(k\phi + ht)$; so, let $H^{\#}(t, I, \phi) = K^{\#}(I, k\phi + ht) = (h/k)I + K(I, k\phi + ht)$.

The lowest-order terms that contain t, the new terms, are $x_1^k e^{-hit}$ and $x_2^k e^{hit}$. In action–angle coordinates these terms are like $I^{k/2} \cos(k\phi + ht)$ and $I^{k/2} \sin(k\phi + ht)$. Thus the normalized Hamiltonian is a function of I and $(k\phi + ht)$ only, and it is of the form

$$\begin{aligned} H^{\#}(t, I, \phi) = {} & (h/k)I + aI^2 + bI^3 + \cdots + \\ & + I^{k/2}\{\alpha \cos(k\phi + ht) + \beta \sin(k\phi + ht)\} + \cdots . \end{aligned} \tag{10.63}$$

The equations of motion are

$$\begin{aligned} \dot{I} &= I^{k/2}\{-\alpha \sin(k\phi + ht) + \beta \cos(k\phi + ht)\} + \cdots, \\ \dot{\phi} &= -\frac{h}{k} - 2aI - \frac{k}{2} I^{(k-2)/2}\{\alpha \cos(k\phi + ht) + \beta \sin(k\phi + ht)\} + \cdots . \end{aligned} \tag{10.64}$$

By a rotation, $\phi \to \phi + \delta$; the first sin term can be absorbed into the cos term, so there is no loss in generality in assuming that $\beta = 0$ in (10.63) and (10.64). Henceforth, we assume this rotation has been made, and so, $\beta = 0$.

Note that in the $\dot{\phi}$ equation in (10.64), there are two nonlinear terms. When $k > 4$, the term that contains the angle is of higher-order in I, whereas

$k = 3$ it is lower-order. When $k = 4$, the two terms are both of order I^1. We show in later chapters on applications that the cases when $k = 3$ or 4 must be treated separately.

The 2π-map is then of the form

$$\begin{aligned} I &= I_0 - \alpha I_0^{k/2} \sin(k\phi_0) + \cdots, \\ \phi &= \phi_0 - (2\pi h/k) - 4\pi a I_0 + \alpha \pi k I^{(k-2)/2} \cos(k\phi_0) + \cdots. \end{aligned} \tag{10.65}$$

Normal forms when multipliers are ± 1. Consider the cases where the multiplier is $+1$ first. For this problem no trigonometric functions are used, therefore assume that the periodic systems are periodic with period 1. If Γ has the eigenvalue $+1$, then either Γ is the identity, and A is the zero matrix, or there are symplectic coordinates such that

$$\Gamma = \exp A = \begin{bmatrix} 1 & \pm 1 \\ 0 & 1 \end{bmatrix}, \quad \text{where } A = \begin{bmatrix} 0 & \pm 1 \\ 0 & 0 \end{bmatrix}. \tag{10.66}$$

In the first case, when $\Gamma = I$ and $A = 0$, Theorem 10.6.1 gives no information, and this is because the situation is highly degenerate and nongeneric.

Therefore, consider the case when Γ and A are as in (10.66) with the plus sign; so, the adjoint of A is B where

$$B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad \exp(Bt) = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}.$$

Let $x = (u, v)$. Condition (10.57) of Theorem 10.6.1 is $H^i(u, v + ut, t) \equiv H^i(u, v, 0)$. This condition and the fact that H^i must be periodic in t implies that $H^i(u, v, t) = K^i(u)$. Thus the normal form is

$$H^{\#}(t, u, v) = v^2/2 + K(u) = v^2/2 + \beta u^3/3 + \cdots \tag{10.67}$$

and the equations of motion are

$$\begin{aligned} \dot{u} &= v + \cdots, \\ \dot{v} &= -\frac{\partial K}{\partial u}(u) = -\beta u^2 + \cdots. \end{aligned} \tag{10.68}$$

The period map is not so easy to compute and is not so simple. Fortunately, in applications, the critical information occurs at a very low order. By using the Lie transform methods discussed in the Problem section one finds that the period map is $(u, v) \to (u', v')$ where

$$\begin{aligned} u' &= u + v - \frac{\beta}{12}(6u^2 + 4uv + v^2) + \cdots, \\ v' &= v - \frac{\beta}{3}(3u^2 + 3uv + v^2) + \cdots. \end{aligned} \tag{10.69}$$

Now consider the case when Γ has eigenvalue -1. Consider the case when

$$\Gamma = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

first because it has a real logarithm,

$$\Gamma = \exp 2\pi A, \quad A = \begin{bmatrix} 0 & 1/2 \\ -1/2 & 0 \end{bmatrix}.$$

This is almost the same as the higher-order resonance considered in the previous subsection. Corollary 10.6.2 implies that the normal form in this case is simply an odd function. That is, $\Phi(y) = -y + \phi(y)$ is in normal form when $\phi(-y) = -\phi(y)$.

Now consider the case when

$$\Gamma = \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix}.$$

We make two changes of coordinates to bring this case to normal form. First, instead of the usual uniform scaling, scale by $x_1 \to \varepsilon x_1$, $x_2 \to \varepsilon^2 x_2$ so that the map (10.48) becomes $\Psi(x) = -x + O(\varepsilon)$. This nonuniform scaling moves the off-diagonal term to the higher-order terms, and now the lead term is the same as discussed in the last paragraph. Thus there is a symplectic change of coordinates $z = R(x)$ such that in the new coordinates z, the map (10.48) is odd; i.e., $R \circ \Psi \circ R^{-1}(z) = \Xi(z) = \Gamma z + \cdots$ is odd.

Write

$$\Xi(z) = -\Lambda(z) = -\{\Omega z + \zeta(z)\}, \quad \text{where} \quad \Omega = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Now Ω is of the form discussed above, and so, there is a symplectic change of coordinates $y = S(z)$ which puts Λ in the normal form given by the time 1-map of a Hamiltonian system of the form (10.67), where now $K(u)$ is even. Because Λ is odd, the transformation S can be made odd also; see problems. Thus $S \circ \Lambda \circ S^{-1} = \Theta$ is in the normal form given by the time 1-map of a Hamiltonian system of the form

$$H^{\#}(t, u, v) = v^2/2 + K(u) = v^2/2 + \beta u^4/4 + \cdots. \tag{10.70}$$

Using the method discussed in the problems gives $\Theta : (u, v) \to (u', v')$, where

$$\begin{aligned} u' &= u + v - \frac{\beta}{20}(10u^3 + 10u^2 v + 5uv^2 + v^3) + \cdots, \\ v' &= v - \frac{\beta}{3}(4u^3 + 6u^2 v + 4uv^2 + v^3) + \cdots. \end{aligned} \tag{10.71}$$

Combining these changes of coordinates and using the fact that S is odd, it follows that $(S \circ R) \circ \Psi \circ (S \circ R)^{-1} = -\Theta$. That is, in the new coordinates, the map is just the negative of (10.71), or the normal form for the map is

$$
\begin{aligned}
u' &= -u - v + \frac{\beta}{20}(10u^3 + 10u^2 v + 5uv^2 + v^3) + \cdots, \\
v' &= -v + \frac{\beta}{3}(4u^3 + 6u^2 v + 4uv^2 + v^3) + \cdots.
\end{aligned}
\tag{10.72}
$$

Problems

1. a) The normal form for a Hamiltonian system with $H_0^0(q,p) = p^2/2$ is $H^*(q,p) = p^2/2 + Q(q)$. This normal form also appears in Section 10.6 when the case of multipliers equal to $+1$ is discussed. Carefully draw the phase portrait for the system with Hamiltonian $H(q,p) = p^2/2 + \beta q^3$ when $\beta = +1$ and -1.
   b) In Section 10.6 when the multiplier -1 is discussed the normal form is $H^*(q,p) = p^2/2 + Q(q)$ with Q even. Carefully draw the phase portrait for the system with Hamiltonian $H(q,p) = p^2/2 + \beta q^4$ when $\beta = +1$ and -1.
2. a) Compute the next term in the normal form of the unforced Duffing equation (10.22) by hand. Recall that H_0^0, H_1^0, H_0^1 and W_1 are given in Section (10.3). (Hint: To get the next term you do not have to compute all of H_1^1, H_0^2 and W_2. H_0^2 is the term which is independent of ϕ in $H_1^1 + \{H_1^0, W_1\}$. Show that $\{H_1^0, W_1\}$ has no term independent of ϕ. Now $H_1^1 = H_2^0 + \{H_1^0, W_1\} + \{H_0^0, W_2\}$, $H_0^2 = 0$, so you need to compute the term independent of ϕ in $\{H_1^0, W_1\}$.)
   b) Using Maple, Mathematica, etc., find the first four terms in the normal form for the unforced Duffing equation.
3. The Hamiltonian for Duffing's equation is of the form $(q^2 + p^2)/2 + P(p)$ where P is an even polynomial.
   a) Show that such a Hamiltonian in action–angle variables is a Poisson series with only cosine terms.
   b) Show that the Poisson bracket of two Poisson series, one of which is a cosine series and the other of which is a sine series, is always a cosine series.
   c) Let H_j^i and W_i be from the normalization of such a Hamiltonian with an even potential. Show that H_j^i can always be taken as a cosine series and W_i as a sine series. (Hint: Define the spaces $\mathcal{P}_i$, $\mathcal{Q}_i$, and $\mathcal{R}_i$ of Theorem 10.3.1.)
4. Consider a Hamiltonian differential equation of the form

$$\dot{x} = \varepsilon F_\#(\varepsilon, t, x) = \varepsilon F_1(t,x) + \varepsilon^2 F_2(t,x) + \cdots,$$

where F is T-periodic in t. Show that there is a formal symplectic series expansion $x = X(\varepsilon, t, y) = y + \cdots$ which is T-periodic in t and transforms the equation to the autonomous Hamiltonian system $\dot{y} = \varepsilon F^{\#}(y) = \varepsilon F^1(y) + \varepsilon^2 F^2(y) + \cdots$. Show that $F^1(y) = (1/T)\int_0^T F_1(\tau, y)d\tau$; i.e., F^1 is the average of F_1 over a period. This is called the method of averaging. (Hint: Use Theorem 10.6.1 and remember $F_0^0 = 0$.)

5. Use the notation of the previous problem. Show that if $F^1(\xi) = 0$ and $\partial F^1(\xi)/\partial x$ is nonsingular, then the equation $\dot{x} = \varepsilon F_{\#}(\varepsilon, t, x)$ has a T-periodic solution $\phi(t) = \xi + O(\varepsilon)$.
6. Analyze the forced Duffing's equation,

$$\ddot{x} + x = \varepsilon\{\delta x + \gamma x^3 + A\cos t\} = 0$$

in three different ways, and show that the seemingly different methods give the same intrinsic results. The parameter δ is called the detuning and is a measure of the difference between the natural frequency and the external forcing frequency. Remember that a one degree of freedom autonomous system has a phase portrait given by the level lines of the Hamiltonian.

   a) Write the system in action–angle coordinates, and compute the first term in the normal form, F_0^1, as was done for Duffing's equation. Analyze the truncated equation by drawing the level lines of the Hamiltonian. (See Section 9.2.)
   b) Write the system in complex coordinates and compute the first term in the normal form, F_0^1, as was done for Duffing's equation in Section 10.3. Analyze the equation.
   c) Make the "van der Pol" change of coordinates

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

   and then compute the first average of the equations via Problems 4 and 5. Analyze the equations. See McGehee and Meyer (1974).
7. Consider a Hamiltonian of two degrees of freedom of the form (10.32), $x \in \mathbb{R}^4$. Let $H_0(x)$ be the Hamiltonian of two harmonic oscillators. Change to action–angle variables $(I_1, I_2, \phi_1, \phi_2)$ and let $H_0 = \omega_1 I_1 + \omega_2 I_2$. Use Theorem 10.4.1 to show that the terms in the normal form are of the form $aI_1^{p/2}I_2^{q/2}\cos(r\phi_1 + s\phi_2)$ or $\mathrm{b}\mathrm{I}_1^{p/2}I_2^{q/2}\sin(r\phi_1 + s\phi_2)$, a and b constants, if and only if $r\omega_1 + s\omega_2 = 0$, and the terms have the d'Alembert character. See Henrard (1970b).
8. Consider a Hamiltonian $H(x)$ with general solution $\phi(t, \xi)$. Observe that the ith component of ϕ is the Lie transform of x_i; i.e., $\phi_i(t, \xi) = \mathcal{L}_H(x_i)(\xi)$, where ε is replaced by t.
   a) Show that $\phi_i(t, \xi) = \left[x_i + \{x_i, H\}t + \{\{x_i, H\}, H\}t^2/2 + \cdots\right]_{x=\xi}$.
   b) Using Maple, Mathematica, etc., write a simple function to compute the time 1 maps given in (10.69) and (10.71) (Make sure that you

compute the time series far enough to pick up all the quadratic and cubic terms in the initial conditions.)

9. Prove Theorem 10.3.3, the uniqueness theorem. (Hint: Show that if the normal form is not unique then there are two different Hamiltonians H and K which are both in normal form and a generating function W carrying one into the other. Show that the terms in the series expansion for W must lie in the kernel of $\mathcal{Q}_i$. Then show that this implies that $W \equiv 0$.)