SIAM J. MATH. ANAL.
Vol. 15, No. 5, September 1984

002

SCALING HAMILTONIAN SYSTEMS*

K. R. MEYER†

Abstract. This paper presents a detailed discussion of scaling techniques for Hamiltonian systems of equations. These scaling techniques are used to introduce small parameters into various systems of equations in order to simplify the proofs of the existence of periodic solutions. The discussion proceeds through a series of increasingly more complex examples taken from celestial mechanics. In particular, simple proofs are given for Lyapunov's center theorem, the continuation theorem of Hadjidemetriou, and several theorems on periodic solutions by the author.

1. Introduction. Perturbation analysts often argue over which general method is best—the methods of averaging, Lie transformations, two-timing, Lyapunov–Schmidt, etc., all have their strong advocates. However, no matter what perturbation technique is used, an important, fundamental and often overlooked question is the correct selection of the equations of the first approximation. In some cases it is so obvious what the first approximation is that there really is no choice, but in other cases the choice can drastically affect all subsequent analysis. In celestial mechanics the equations of the first approximation are called the main problem, and I shall use this term since it emphasizes the importance of these equations.

A historic example where the choice of the main problem had important consequences is found in lunar theory. Until the works of Hill were completely understood, researchers looking for a good approximate solution to the equations of celestial mechanics which described the motion of the moon used as their main problem two decoupled Kepler problems. The two Kepler problems were the equations of motion of the earth and moon about their combined center of mass, and the equations of motion describing the sun and the center of mass of the earth–moon system. Coupling terms were neglected in the main problem. Various perturbation techniques were used, but the approximate solutions failed to agree with the observational data over long periods of time. In a series of papers [5], Hill redefined the main problem of lunar theory by taking into account the fact that the motion of the moon is strongly affected by the sun. Hill's main problem took into account more terms, and as a result the perturbations were smaller and the series converged better numerically. In fact, for many years lunar ephemerides were computed from series developed by Brown, who used Hill's main problem. Even today searchers for more accurate lunar theories use Hill's main problem.

In this survey paper, I want to discuss a general procedure for deciding the correct definition of the main problem in various situations in celestial mechanics. The examples are taken from my own work and therefore consist mainly of problems of finding periodic solutions in Hamiltonian formalism.

The method present is certainly not new—in fact it is so old that I have no idea of when it originated. The method is also not obscure—in fact almost all perturbation techniques are based either explicitly or implicitly on this method. The method is simply that of scaling variables. Since the problems discussed here are written in Hamiltonian formalism, the scaling will be done so that the resulting equations are

*Received by the editors June 2, 1982, and in revised form June 23, 1983. This research was partially supported by the National Science Foundation under grant MCS-80-01851.

† Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio 45221.

again Hamiltonian and so the scaling is symplectic (canonical). Therefore I call the method symplectic scaling. There have been other discussions of scaling as a general procedure in applied mathematics; see for example [16].

Scaling is often presented as a triviality. Sometimes an author starts his discussion with a single statement like: "scale by $x \to \varepsilon x$ and $y \to \varepsilon y$" and then proceeds with page upon page of detailed calculation. Usually there is no discussion of why the equations were scaled nor whether this scaling is the best. In fact, I have seen many papers that could be greatly simplified if the author had used a different scaling (say $x \to \varepsilon x$ and $y \to \varepsilon^2 y$). The examples given below illustrate how to obtain the correct scaling for a particular problem in celestial mechanics.

2. Review of transformation theory. I shall deal exclusively with autonomous Hamiltonian systems. Even though this paper attempts to be reasonably self-contained, I assume that the reader has some background in differential equations and Hamiltonian mechanics. The excellent introductory book by Pollard [13] should be more than adequate. I shall not bog myself down with topological or smoothness questions, since the results given below are local in nature. All functions and vector fields will be assumed to be C^∞ on some open set in $\mathbb{R}^{2n}$ or even defined on all of $\mathbb{R}^{2n}$. Also vectors will be column vectors unless otherwise stated, but will be written as row vectors in the text for typographical reasons.

If $\phi: \mathbb{R}^l \to \mathbb{R}^k$ and $y = \phi(x)$ then $\partial\phi/\partial x$ or $\partial y/\partial x$ will denote the $k \times l$ Jacobian matrix. Thus if H: $\mathbb{R}^{2n} \to \mathbb{R}$, $x \in \mathbb{R}^{2n}$ then $\partial H/\partial x$ is a row vector. Define $\nabla_x H = \nabla H = (\partial H/\partial x)^T$ where the superscript T denotes the transpose.

An autonomous Hamiltonian system of n degrees of freedom in $\mathbb{R}^n$ is a system of ordinary differential equations of the form

$$\dot{x} = J \nabla_x H(x) \tag{2.1}$$

where $H: \mathbb{R}^{2n} \to \mathbb{R}^1$, $x \in \mathbb{R}^{2n}$, $\dot{} = d/dt$ and J is the $2n \times 2n$ constant matrix

$$J = \begin{pmatrix} 0 & I \\ -1 & 0 \end{pmatrix}$$

where 0 and I are the $n \times n$ zero and identity matrices. The independent variable t will be called time, the function H, the Hamiltonian and (2.1), the equations of motion. If $x = (q, p)$ where $q, p \in \mathbb{R}^n$ then the equations of motion take the classical form

$$\dot{q} = \frac{\partial H}{\partial p}, \qquad \dot{p} = -\frac{\partial H}{\partial q}. \tag{2.2}$$

Thus there is a well-defined prescription for obtaining the equations of motion from the Hamiltonian. This prescription is not invariant under all changes of variables. That is, if one changes variables in both the equations of motion and the Hamiltonian, then the new equations of motion may not be obtained from the new Hamiltonian by the prescription given in (2.1). Those changes of variables which preserve this prescription are known as symplectic or canonical. There is a vast literature on the subject of symplectic changes of variables, but fortunately only one basic fact will be needed for the subsequent discussion. The texts by Pollard [13], Wintner [17] and Abraham and Marsden [1] contain more details.

Consider the change of coordinates $x = \phi(y)$ for (2.1). These equations become

$$\dot{y} = P^{-1}(y) J \nabla_x H(\phi(y)) = P^{-1}(y) J \left\{ \frac{\partial H}{\partial x}(\phi(y)) \right\}^T \tag{2.3}$$

where $P(y)$ is the Jacobian $\partial\phi(y)/\partial x$. As noted above, these equations need not be in Hamiltonian form; i.e., the right-hand side is not of the form $J\nabla_y K(y)$ where $K:\mathbb{R}^{2n}\to\mathbb{R}$. However, if we assume that

$$J=\mu TJT^T \tag{2.4}$$

where μ is a nonzero constant, then

$$\dot{y}=T^{-1}J\left\{\frac{\partial H}{\partial x}\right\}^T=\mu JT^T\left\{\frac{\partial H}{\partial x}\right\}^T=\mu J\left\{\frac{\partial H}{\partial x}\frac{\partial\phi}{\partial y}\right\}^T=J\nabla_y\{\mu H(\phi(y))\},$$

or

$$\dot{y}=J\nabla_y K(y) \tag{2.5}$$

where

$$K(y)=\mu H(\phi(y)). \tag{2.6}$$

A change of variables $x=\phi(y)$ which satisfies (2.4) for all y and for some nonzero constant μ is called a symplectic transformation with multiplier μ. What was just shown is that these transformations preserve the Hamiltonian character of the equations and in particular transform (2.1) to (2.5).

If $\mu=1$ then the change of variables is simply called symplectic. Many elementary texts consider only this case, but the added generality of having a μ different from 1 is very important for scaling.

As an example, consider the problem of changing units in the N-body problem. Let $q_1,\cdots,q_N$ be the position vectors with respect to a Newtonian frame of N point masses moving in $\mathbb{R}^3$. Let $p_1,\cdots,p_N$ be the momentum vectors and $m_1,\cdots,m_N$ be masses of these point masses. Then the Hamiltonian for the N-body problem is

$$H=\sum_{i=1}^{N}\frac{\|p_i\|^2}{2m_i}-\sum_{1\le i<j\le N}\frac{km_im_j}{\|q_i-q_j\|} \tag{2.7}$$

where k is the universal gravitational constant. If $x=(q_1,\cdots,q_N,p_1,\cdots,p_N)$, then $x\in\mathbb{R}^{6N}$ and the equations of motion for the N-body problem are (2.1).

Scaling and changing units are essentially the same thing. Let's say for example that the quantities in this problem are all measured in the CGS system. Then $k=6.67\times 10^{-8}$. If we wish to change the unit of length, then we set $q_i=\alpha\bar{q}_i$, $p_i=\alpha\bar{p}_i$ where α is the conversion factor ($\alpha=100$ cm/m if the new unit of length is meters). This change of variables is symplectic with multiplier α^{-2}, and so the Hamiltonian becomes

$$H=\sum_{i=1}^{N}\frac{\|\bar{p}_i\|^2}{2m_i}-\sum_{1\le i<j\le N}\frac{k}{\alpha^3}\frac{m_im_j}{\|\bar{q}_i-\bar{q}_j\|}. \tag{2.8}$$

In this mixed system of units (MGS) the gravitational constant becomes $k/\alpha^3=6.67\times 10^{-14}$. In theoretical work it is convenient to use nonmetric units and to take $\alpha^3=k$ so that the gravitational constant is 1. We shall take $k=1$ in all subsequent discussions.

Since the bars over variables are not esthetic and are besides a lexicographical nuisance, it will be convenient to drop them in all subsequent discussions. The operation of first changing variables by $q_i=\alpha\bar{q}_i$, $p_i=\alpha\bar{p}_i$ and then dropping the bars is denoted by $q_i\to\alpha q_i$, $p_i\to\alpha p_i$. It should be carefully noted that this notation implies a change of variables and is only used to limit the proliferation of symbols.

Sometimes it is necessary to change the independent variable t also. If $t=\beta\tau$, where β is a constant, then (2.1) becomes

$$x'=J\nabla_x K(x) \tag{2.9}$$

where $'=d/d\tau$ and $K=\beta H$. Thus scaling time is equivalent to multiplying the Hamiltonian by a factor. In the scaling notation $t\to\beta t$ and $H\to\beta H$.

3. The noncritical case. Since the main application of scaling to be discussed in this paper is to establish the existence of periodic solutions, I shall summarize some of the known elementary results. Let $\phi(t,\xi,\lambda)$ be the solution of the Hamiltonian system

$$\dot{x}=J\nabla_x H(x,\lambda) \tag{3.1}$$

which satisfies $\phi(0,\xi,\lambda)=\xi$ where λ is a real parameter. Since (3.1) is autonomous, a necessary and sufficient condition for a particular solution $\phi(t,\xi_0,\lambda_0)$ to be T-periodic ($T>0$) is

$$\phi(T,\xi_0,\lambda_0)=\xi_0. \tag{3.2}$$

This is easily proved by observing that both $\phi(t,\xi_0,\lambda_0)$ and $\phi(t+T,\xi_0,\lambda_0)$ are both solutions of (3.1) and that (3.2) implies that both these solutions satisfy the same initial condition. Thus the uniqueness theorem for ordinary differential equations assures that the two solutions are identical.

The necessary and sufficient condition (3.2) is interesting since it shows that the existence of periodic solutions of a differential equation is equivalent to solution of a system of (nondifferential) equations. In theory, at least, only finite dimensional methods could be used to establish the existence of periodic solutions. This is certainly not the case when we are trying to establish the existence of almost periodic solutions, invariant manifolds, etc. For these problems infinite dimensional methods are essential.

One approach to solving (3.2) is the use of the implicit function theorem. If (T,ξ_0,λ_0) satisfies (3.2) then the implicit function theorem would give nearby solutions, provided the Jacobian matrix

$$\frac{\partial\phi}{\partial\xi}(T,\xi_0,\lambda_0)-I \tag{3.3}$$

were nonsingular or equivalently that the Jacobian matrix

$$\frac{\partial\phi}{\partial\xi}(T,\xi_0,\lambda_0) \tag{3.4}$$

did not have the eigenvalue $+1$. The eigenvalues of (3.4) are so important in the study of periodic solutions that they are named the characteristic multipliers (or simply multipliers) of the periodic solution. Unfortunately, $+1$ is always a multiplier of a periodic solution of an autonomous system. Even worse, since (3.2) admits H as a first integral, the algebraic multiplicity of $+1$ as a multiplier is greater than or equal to 2. In the class of nonautonomous periodic equations, the usual case is that a periodic solution does not have the multiplier $+1$, and so this is usually called the noncritical case. In the class of autonomous equations, the usual case is that a periodic solution has the characteristic multiplier $+1$ with multiplicity precisely $=+1$ [7], and so for autonomous systems this is the noncritical case. In the case of autonomous Hamiltonian systems, the usual case is that a periodic solution has the characteristic multiplier $+1$ with algebraic multiplicity precisely equal to 2, and so for such systems this is the usual case [15].

In each of the cases listed above we have defined a noncritical case for each choice of our universe of discourse. The reason I call these the noncritical cases is that for each of these definitions of the noncritical case there is a theorem which states that in the noncritical case a small perturbation within the universe of discourse causes a slight perturbation in the periodic solution. Moreover, each of these theorems admits an elementary proof based on the implicit function theorem. The most satisfying discussion of these theorems is contained in Poincaré [14], but a clear, elementary discussion in modern notation can be found in Deprit and Henrard [2].

In the autonomous Hamiltonian case the precise statement of the theorem alluded to above is:

THEOREM 3.1. *Let* $\lambda \in \mathbb{R}^k$, $k=0,1,2,\cdots,H:\mathbb{R}^{2n}\times\mathbb{R}^k\to\mathbb{R}^1$ *be smooth and let* $\phi(t,\xi,\lambda)$ *be the solution of* (3.1) *so that* $\phi(0,\xi,\lambda)=\xi$. *Assume that* (T,ξ_0,λ_0), $T>0$, *satisfies*

i) $$\phi(T,\xi_0,\lambda_0)=\xi_0$$

and

ii) $$\operatorname{rank}\left\{\frac{\partial\phi}{\partial\xi}(T,\xi_0,\lambda_0)-I\right\}=2n-2.$$

Then the periodic solution $\phi(t,\xi_0,\lambda_0)$ *is smoothly embedded in a* $(k+2)$*-parameter family of periodic solutions. That is, there are a neighborhood* 0 *of* λ_0 *in* $\mathbb{R}^k$, *a neighborhood P of* $(0,0)$ *in* $\mathbb{R}^2$, *and smooth maps* $\tau:P\times 0\to\mathbb{R}^n$ *and* $\xi:P\times 0\to\mathbb{R}^n$ *such that* $\tau(0,0,\lambda_0)=T$, $\xi(0,0,\lambda_0)=\xi_0$, *and* $\phi(t,\xi(\alpha,\beta,\lambda),\lambda)$ *is a* $\tau(\alpha,\beta,\lambda)$*-periodic solution of* (3.1) *where* $(\alpha,\beta)\in P$ *and* $\lambda\in 0$.

Note that even when the equation does not depend on a parameter (i.e. $k=0$), the periodic solution is still embedded in a 2-parameter family of periodic solutions. These two additional parameters can be chosen as the value of the integral H on the periodic solution and the time from a well-chosen epoch along the periodic solution. In this case these periodic solutions locally fill a cylinder in $\mathbb{R}^{2n}$; see [1, Fig. 8.2-1]. Again we refer the reader to [2] for a simple clean proof of this theorem.

The remainder of this section is devoted to illustrating the method of symplectic scaling as a tool for reducing a given system to one to which the above theorem applies. Consider first the famous Lyapunov center theorem by glancing at the proofs given in [4], [6], [8]. In this theorem we assume that the equation (3.1) has an equilibrium point, say at $x=0$, and then expand the right-hand side in a Taylor series to get

$$\dot{x}=Ax+f(x) \tag{3.5}$$

where $f(0)=0$, $\partial f(0)/\partial x=0$ and A is a $2n\times 2n$ constant matrix. The Hamiltonian becomes

$$H(x)=\tfrac{1}{2}x^TSx+K(x) \tag{3.6}$$

where S is the Hessian of H at $x=0$, $A=JS$, K and the first and second partials of K vanish at $x=0$. By setting $f=0$ in (3.5) we obtain the linearization of the equations about the equilibrium point—an obvious candidate for the main problem. We would expect or at least hope that the solutions of the linear system are nearly the same as the solutions of the full equation when x is small, but how can we demonstrate this connection? To obtain a measure of x being small, scale by $x\to\varepsilon x$. This scaling is symplectic of order ε^2 and so (3.5) becomes

$$\dot{x}=Ax+O(\varepsilon) \tag{3.7}$$

and the Hamiltonian becomes

$$H(x)=\tfrac{1}{2}x^TSx+O(\varepsilon). \tag{3.8}$$

When $\varepsilon=0$, (3.7) becomes linear and the general solution is $\phi(t,\xi,0)=(\exp At)\xi$. In order to apply the theorem above, this system must have a periodic solution. Therefore let the eigenvalues of A be $\lambda_1,\lambda_2,\cdots,\lambda_{2n}$ and assume $\lambda_1=+i\omega$, $\lambda_2=-i\omega$, $\omega>0$. Let η and $\bar{\eta}$ be the corresponding eigenvectors so $A\eta=i\omega\eta$ and $(\exp At)\eta=(\exp i\omega t)\eta$. Thus $\mathrm{Re}(\exp At)\eta=(\exp At)\xi_0$ is a $T=2\pi/\omega$ periodic solution. The Jacobian matrix (3.4) becomes in this case

$$\frac{\partial\phi}{\partial\xi}(T,\xi_0,0)=\left(\exp\frac{2\pi}{\omega}A\right), \tag{3.9}$$

which has eigenvalues $\exp\pm 2\pi i=1$, $\exp(\lambda_2 2\pi/\omega),\cdots,\exp(\lambda_{2n}2\pi/\omega)$. Thus for the second condition of the theorem to hold it must be assumed that

$$\frac{\lambda_k 2\pi}{\omega}\not\equiv 0 \bmod 2\pi i \quad \text{for } k=2,\cdots,2n, \tag{3.10}$$

or

$$\frac{\lambda_k}{\lambda_1} \text{ is not an integer} \quad \text{for } k=2,\cdots,2n. \tag{3.11}$$

If this condition applies, then (3.7) has a 3-parameter family of periodic solutions which are of the form $(\exp At)\xi_0+O(\varepsilon)$. The original equation (3.5) has a two-parameter family of periodic solutions of the form $\varepsilon(\exp At)\xi_0+O(\varepsilon^2)$. It may seem that we have proved that (3.5) has a 3-parameter family also, but this equation is independent of ε, and the theorem given above gives precisely a $k+2$ manifold of periodic solutions whose period is close to T. Thus one of the parameters is redundant. That proves Lyapunov's center theorem!

As a second example, consider the relationship between the full three-body problem and the restricted three-body problem. In the traditional derivation of the restricted three-body problem, one is asked to consider the motion of an infinitesimally small particle moving in the plane under the influence of the gravitational attraction of two finite particles which revolve around each other on a circular orbit of the Kepler problem. Although this description is picturesque, it hardly clarifies the relationship between the restricted three-body problem and the full problem. Consider the planar N-body problem where $N=2$ or 3 written in rotating coordinates [1]. The Hamiltonian is

$$H_N=\sum_{i=1}^{N}\frac{\|y_i\|^2}{2m_i}-x_i^TKy_i-\sum_{1\le i<j\le N}\frac{m_im_j}{\|x_i-x_j\|} \tag{3.12}$$

where m_i is the mass, x_i is the position and y_i is the momentum of the ith particle in a rotating coordinate system and $K=\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$. In order to consider the case when one particle is small, set $m_3=\varepsilon^\alpha$ where α is a positive number to be determined later and ε will be treated as a small parameter (we are not scaling at this point!). Making this substitution in (3.12) with $N=3$ and rewriting yields

$$H_3=\frac{\|y_3\|^2}{2\varepsilon^\alpha}-x_3^TKy_3-\sum_{i=1}^{2}\frac{\varepsilon^\alpha m_i}{\|x_i-x_3\|}+H_2. \tag{3.13}$$

Here the terms involving the third particle have been removed, leaving the Hamiltonian of the two-body problem as a remainder. Since ε is a small parameter which already measures the smallness of one mass, we should attempt to make ε also measure the deviation of the motion of the first two particles from a circular orbit. That is ε, or a power of ε, should measure not only the smallness of m_3, but also how close the first two particles come to a circular orbit. To accomplish this we must prepare the Hamiltonian H_2 so that one variable represents the deviation from a circular orbit. Actually part of this preparation has already been done, since in rotating coordinates a circular orbit appears as an equilibrium solutions. Let $Z=(x_1,x_2,y_1,y_2)$, so H_2 is a function of the 8-vector Z, and let $Z^*=(a_1,a_2,b_1,b_2)$ be a critical point of H_2, so $\nabla H_2(Z^*)=0$. (Later we shall give explicit values for the a's and b's, but for now it is enough to know that they exist.) By Taylor's theorem

$$H_2(Z)=H_2(Z^*)+\tfrac{1}{2}(Z-Z^*)S(Z-Z^*)+o\left(\|Z-Z^*\|^2\right) \tag{3.14}$$

where S is the Hessian of H_2 evaluated at Z^*. Since constant terms in the Hamiltonian drop out when the equations of motion are formed, we shall ignore $H_2(Z^*)$ by setting it to zero. If the motion of the first two particles is nearly circular, then $Z-Z^*$ should be small, so this suggests the scaling

$$Z-Z^*\to\varepsilon^{\beta}U \tag{3.15}$$

where U is a new variable and β is a positive number to be determined. So far we have implemented the assumptions that the third mass is small, that the deviation of the motion of the first two particles from a circular orbit is small, and that the smallness relationships is in the form of a power law. α and β have not been given yet, and so the precise relationship between the two small quantities is not yet established. This is the point at which symplectic scaling gives some guidance on how to proceed. Note first that (3.15) is a symplectic change of the U variables with multiplier $\varepsilon^{-2\beta}$; however, (3.15) is not a symplectic change of variables on the whole space since x_3 and y_3 have not been changed yet. The scaling (3.15) implies $x_1=a_1+O(\varepsilon^{\beta})$ and $x_2=a_2+O(\varepsilon^{\beta})$ where a_1 and a_2 are the constant vectors defined above, so x_1 and x_2 are order zero in ε. Since we are not interested in the case when x_3 is close to a_1 or a_2 (the collision problem) nor that when x_3 is large (the case of a comet), we shall take x_3 as order zero in ε also. Thus for a change of variables on the whole phase space to be symplectic, it is necessary that $y_3\to\varepsilon^{2\beta}\eta$. Thus we complete (3.15) with

$$x_3\to\xi,\qquad y_3\to\varepsilon^{2\beta}\eta. \tag{3.16}$$

Using (3.14), (3.15) and (3.16) in (3.13) yields

$$H_3=\varepsilon^{\alpha-2\beta}\frac{\|\eta\|^2}{2}-\xi^TK\eta-\varepsilon^{\alpha-2\beta}\sum_{i=1}^{2}\frac{m_i}{\|\xi-a_i\|}+\cdots+\frac{1}{2}U^TSU+\cdots. \tag{3.17}$$

In order to make the first and third terms in (3.17) just as important as the second and fourth, it is necessary to have $\alpha=2\beta$. Setting $\beta=1$, $\alpha=2$ gives a small integer solution of this relation. To summarize: if $m_3=\varepsilon^2$ then

$$x_3\to\xi,\quad y_3\to\varepsilon^2\eta,\quad Z-Z^*\to\varepsilon U \tag{3.18}$$

is a symplectic change of variables which reduces (3.13) to

$$H_3=\left\{\frac{\|\eta\|^2}{2}-\xi^T K\eta-\sum_{i=1}^{2}\frac{m_i}{\|\xi-a_i\|}\right\}+\frac{1}{2}U^T SU+O(\varepsilon). \tag{3.19}$$

The quantity in the braces above is the Hamiltonian of the restricted three-body problem if we take $m_1+m_2=1$, $m_1=\mu$, $m_2=1-\mu$, $a_1=(1-\mu,0)$ and $a_2=(-\mu,0)$. The quadratic term in U is simply the Hamiltonian of the linearization of the equations of motion of the two-body problem about the circular solution. For $\varepsilon=0$ the Hamiltonian H_3 is a sum of these two Hamiltonians, and so the equations of motion decouple. If $\xi=\phi(t)$, $\eta=\psi(t)$ is any solution of the restricted problem, then $\xi=\phi(t)$, $\eta=\psi(t)$, $U\equiv 0$ is a solution of the equations of motion defined by (3.18) with $\varepsilon=0$. Thus for bounded times, there are solutions of the full three-body problem of the form $\xi=\phi(t)+O(\varepsilon)$, $\eta=\psi(t)+O(\varepsilon)$ and $U=O(\varepsilon)$.

Looking at (3.18) we see that since y_3 is the momentum of the third particle and $m_3=\varepsilon^2$, the variable η is really the velocity of the third particle. Thus all the new quantities have been given physical meaning, and the relationship between the small quantities has been established.

The problem defined by the Hamiltonian (3.18) is still degenerate due to the fact that the original three-body problem admits symmetries and integrals. Specifically, the Hamiltonian H_3 is invariant under the full group of Euclidean motions of the plane and admits linear and angular momentum as integrals. Holding these integrals fixed and then identifying configurations which differ by a Euclidean motion only leads to a Hamiltonian on a reduced space. The details of this reduction are unimportant for the present discussion and are classical. It is enough to say that after this reduction is done the Hamiltonian (3.19) becomes

$$H_3=\left\{\frac{\|\eta\|^2}{2}-\xi^T K\eta-\sum_{i=1}^{2}\frac{m_i}{\|\xi-a_i\|}\right\}+\frac{1}{2}\{r^2+R^2\}+O(\varepsilon) \tag{3.20}$$

where r and R are scalar variables. See [9] for a complete discussion of this reduction. Thus if $\xi=\phi(t)$, $\eta=\psi(t)$ is a τ-periodic solution of the restricted problem with characteristic multipliers 1, 1, λ, λ^{-1}, then $\xi=\phi(t)$, $\eta=\psi(t)$, $r=R=0$ is a τ-periodic solution of the three-body problem defined by (3.20) with $\varepsilon=0$ with characteristic multipliers 1, 1, λ, λ^{-1}, $\exp\pm i\tau$. Thus if $\lambda\neq 1$ and $\tau\not\equiv 0 \bmod 2\pi$, this represents a nondegenerate τ-periodic solution of the three-body problem defined by (3.20) with $\varepsilon=0$.

Now the classical perturbation theorem applies to yield the theorem of Hadjidemetriou [3], namely, that any nondegenerate periodic solution of the restricted problem whose period is not a multiple of 2π can be continued into the full three-body problem.

There is another restricted three-body problem, known as Hill's lunar equations, which is derived under slightly different assumptions. The traditional description [5] of this equation is even more picturesque then the description of the restricted problem. One is asked to consider the motion of an infinitesimal body (the moon) which is attracted to a finite body (the earth) which is fixed at the origin of a rotating coordinate system. The coordinate system rotates so that the positive x-axis points to an infinite body (the sun) which is infinitely far away. The ratio of the two infinite quantities [sic] is taken so that the gravitational attraction of the sun on the moon is finite.

Briefly, I shall indicate how Hill's lunar equations can be derived from the three-body problem; for details see [12]. In this problem two masses m_1 and m_2 (the earth and moon) are small relative to the mass of the third (the sun). Also the distance between the earth and moon is small relative to the distance between their center of mass and the sun. The first assumption is easy to implement: simply set $m_1=\varepsilon^6\mu_1$, $m_2=\varepsilon^6\mu_2$ and $m_3=\mu_3$. (Here I have fixed the exponents since I already worked out what they should be.) In order to implement the second assumption, we must choose coordinates so that one variable represents the distance between the two bodies. A classical set of symplectic cordinates known as Jacobi coordinates has one coordinate which represents the distance between two of the bodies and so is the logical choice here. The Jacobi position vectors are u_0, the position of the center of mass of the triple; u_1, the position of particle 2 relative to particle 1; and u_2, the position of particle 3 relative to the center of mass of particles 1 and 2. The variables v_0, v_1 and v_2 are the corresponding momenta where v_0 is the total linear momentum of the system. Making the initial scaling

$$v_1\to\varepsilon^6 v_1, \qquad v_2\to\varepsilon^6 v_2 \tag{3.21}$$

as in the previous example and fixing the center of mass at the origin, $u_0=0$, and ignoring the total linear momentum v_0 leads to the following Hamiltonian for the full three-body problem:

$$\begin{aligned} &H_3=H'+H''+O(\varepsilon^6),\\ &H'=\frac{\|v_1\|^2}{2M_1}-u_1^TJv_1-\varepsilon^6\frac{\mu_0\mu_1}{\|u_1\|},\\ &H''=\frac{\|v_2\|^2}{2M_2}-u_2^TJv_2-\frac{\mu_1\mu_2}{\|u_2-\nu_0u_1\|}-\frac{\mu_0\mu_2}{\|u_0+\nu_1u_1\|}. \end{aligned} \tag{3.22}$$

Here M_1, M_2, ν_0, ν_1 are all positive constants defined in terms of the original classes. The only property needed here is $\nu_0+\nu_1=1$.

The Hamiltonian H' contains only u_1 and v_1, the variables of the earth–moon pair, whereas the Hamiltonian H'' contains cross terms. Since u_1 is to be taken as a small quantity later, rewrite H'' as

$$\begin{aligned} &H''=H^*+H^{**},\\ &H^*=\frac{\|v_2\|^2}{2M_2}-u_2^TJv_2-\frac{\mu_2(\mu_0+\mu_1)}{\|u_2\|},\\ &H^{**}=\frac{\mu_2(\mu_0+\mu_1)}{\|u_2\|}-\frac{\mu_1\mu_2}{\|u_2-\nu_0u_1\|}-\frac{\mu_0\mu_1}{\|u_0+\nu_1u_1\|}. \end{aligned} \tag{3.23}$$

Now H^* contains only u_2 and v_2, the variables describing the motion of the earth–moon pair about the sun. Since this motion is assumed to be nearly circular, we set $Z=(u_2,v_2)$ and $Z^*=(a,b)$ as before so that

$$H^*(Z)=H^*(Z^*)+\tfrac{1}{2}(Z-Z^*)S(Z-Z^*)+\cdots. \tag{3.24}$$

Now the full set of physical assumptions can be affected by the following scaling:

$$u_1\to\varepsilon^2u_1, \qquad v_1\to\varepsilon^2v_1, \tag{3.25}$$

$$Z-Z_0\to\varepsilon^2U. \tag{3.26}$$

Scaling (3.25) says that the distance between the earth and the moon is small, and scaling (3.26) says that the earth–moon system moves about the sun in a nearly circular orbit. This scaling is symplectic with multiplier ε^{-4} and greatly simplifies the problem. The Hamiltonian H^{**} is not completely ignorable, though, so it must be expanded in a series of Legendre polynomials, the details of which are not appropriate here.

The end result is that the Hamiltonian of the full three-body problem becomes, under these assumptions,

$$H_3=\left\{\frac{\|\eta\|^2}{2}-\xi^TJ\eta-\frac{1}{\|\xi\|}+\left(3\xi_1^2-\|\xi\|^2\right)\right\}+\frac{1}{2}U^TSU+O(\varepsilon^2) \tag{3.27}$$

where ξ and η are essentially u_1 and v_1 (the variables describing the motion of the earth–moon system) and U measures the deviation from a circular orbit of the motion of the earth–moon system around the sun. The quantity in the braces in (3.27) is the Hamiltonian for Hill's lunar equations, and the last expression in parenthesis comes from H^{**}. As with the restricted problem, it is easy to prove that any nondegenerate periodic solution of Hill's lunar equation whose period is not a multiple of 2π can be continued into the full three-body problem. See [12] for a complete account of this derivation, including the details of the expansion of H^{**} in Legendre polynomials.

4. The critical cases. Scaling is particularly useful in the critical cases, since it is usually not obvious which terms in the Taylor expansion of the Hamiltonian are important for the perturbation analysis. The correct scaling not only defines the main problem, but also orders all the terms according to the strength of their influence on the problem at hand. Obviously, since the critical case is the complement of a nice case, further subdivision is necessary. Also, there will always be a system which is so degenerate that it does not fall into any of the previously defined subcases. This section defines what I consider to be the first critical subcase for Hamiltonian systems. This subcase is defined as all systems which can be analyzed by Lemma 4.1. The only new tool necessary to prove this lemma is the variation of constants formula, and so the proof is not much more difficult than the proof of Theorem 3.1.

The lemma deals with a Hamiltonian system of the form

$$\dot z=\nabla H(z,\varepsilon)=Az+\varepsilon f(z,\varepsilon) \tag{4.1}$$

where $z\in\mathbb{R}^{2n}$, $\varepsilon\in\mathbb{R}$, A is a $2n\times 2n$ nonsingular matrix such that $\exp AT=I$ for some $T>0$, and f is a smooth function. Since $\exp AT=I$, all solutions of (4.1) when $\varepsilon=0$ are T-periodic and all their characteristic multipliers are $+1$. Thus when $\varepsilon=0$ the system fails to satisfy the hypotheses of the perturbation theorem of the previous section. In order to restrict the level of degeneracy of this system, some condition must be placed on the higher order terms represented by f.

Let β be a real parameter and define

$$B(\beta,\zeta)=\beta A\zeta+\int_0^T e^{-As}f(e^{As}\zeta,0)\,ds. \tag{4.2}$$

The function B (sometimes called the describing function) is defined entirely in terms of the known quantities A and f and does not depend on the unknown solutions of (4.1). The first critical subcase is defined by:

LEMMA 4.1. *If there exist smooth functions $\zeta(\alpha)$, $\beta(\alpha)$, where α is real and $\zeta(\alpha)\in\mathbb{R}^n$, $\beta(\alpha)\in\mathbb{R}$, such that*

i) $$B(\beta(\alpha),\zeta(\alpha))=0,$$

ii) $$\operatorname{rank}\left(\frac{\partial B}{\partial \beta}, \frac{\partial B}{\partial \zeta}\right)(\beta(\alpha), \zeta(\alpha)) = 2n-1$$

for $|\alpha| \le \alpha_0$, *then there exists a smooth 2-parameter family of periodic solutions of* (4.2), *denoted by* $\phi(t, \alpha, \varepsilon)$, *such that*

iii) $\phi(t,\alpha,\varepsilon)$ *is* $T(\alpha,\varepsilon)$ *periodic for* ε *small and* $|\alpha| \le \alpha_0$,

iv) $$\phi(t,\alpha,0) = (\exp At)\zeta(\alpha),$$

v) $$T(\alpha,\varepsilon) = T + \varepsilon\beta(\alpha) + O(\varepsilon^2).$$

The details of the proof can be found in [11]. The essential step in the proof is a simple calculation of the general solution of (4.2). Let $\psi(t, \zeta_0, \varepsilon)$ be the solution of (4.1) which satisfies $\psi(0, \zeta_0, \varepsilon) = \zeta_0$ and seek the periodic solutions whose period is $T + \varepsilon\beta$. From the variation of constants formula

$$\psi(T+\varepsilon\beta, \zeta_0, \varepsilon) = \zeta_0 + \varepsilon B(\beta, \zeta_0) + O(\varepsilon^2),$$

so the problem of finding an initial condition leading to a periodic solution is just solving

$$B(\beta, \zeta_0) + O(\varepsilon) = 0.$$

One simply applies the implicit function theorem to this system of $2n$ equations to solve $2n-1$ of the equations and then uses the integral H to show that the last equation is also satisfied.

As the first example of how scaling can be used to reduce a problem to a system where this lemma applies, consider the restricted three-body problem where the small mass is far from the primaries. The Hamiltonian of the restricted three-body problem is

$$H = \frac{\|\eta\|^2}{2} - \xi^T K \eta - \sum_1^2 \frac{m_i}{\|\xi - a_i\|} \tag{4.3}$$

where the notation is the same as in (3.19). The equations of motion are

$$\dot{\xi} = K\xi + u, \qquad \dot{\eta} = K\eta + \sum_1^2 \frac{m_i(a_i - \xi)}{\|a_i - \xi\|^3}. \tag{4.4}$$

In order to study this problem for large ξ, scale by $\xi \to \varepsilon^{-2}\xi$ and $\eta \to \varepsilon\eta$. This is a symplectic scaling with multiplier ε, so the Hamiltonian becomes

$$H = -\xi^T K \eta + \varepsilon^3 \left\{ \frac{\|\eta\|^2}{2} - \frac{1}{\|\xi\|} \right\} + O(\varepsilon^5), \tag{4.5}$$

and the equations of motion become

$$\dot{\xi} = K\xi + \varepsilon^3 \eta + O(\varepsilon^5), \qquad \dot{\eta} = K\eta + \varepsilon^3 \|\xi\|^{-3} \xi + O(\varepsilon^5). \tag{4.6}$$

To lowest order in ε, these equations are linear and the general solution is $\xi = (\exp Kt)\xi_0$, $\eta = (\exp Kt)\eta_0$. So if $z = (\xi, \eta)$, $A = \operatorname{diag}(K, K)$, the system (4.6) is of the form (4.1) with ε^3 replacing ε. Since $\exp Kt$ is the rotation matrix by an angle t for small ε, the solutions are nearly circular with periods near 2π. Since rotating coordinates are being used, this means that near infinity the infinitesimal body mainly feels the effect of the Coriolis and centrifugal forces, and in a fixed frame it would be nearly at rest. The coefficient of the ε^3-term is the Hamiltonian of the Kepler problem where the central body has mass 1 (we have assumed that the sum of the masses of the primaries is 1). This can be interpreted as meaning that the next most important force felt by the infinitesimal body is the attraction of a fixed body at the center of mass of the two primaries whose mass is equal to the sum of the masses of the two primaries.

The function B in (4.2) is easy to compute. Setting $\zeta=(\xi,\eta)$, $B=0$ becomes

$$\beta K\xi+2\pi\eta=0, \qquad \beta K\eta-2\pi\frac{\xi}{\|\xi\|^3}=0. \tag{4.7}$$

It is not difficult to analyze these equations and show that Lemma 4.1 applies. (The details are found in [10].) The main conclusion of this analysis is that the restricted three-body problem has two families of nearly circular orbits of large radius.

The last example illustrates the proper method of scaling when one encounters nonelementary divisors in a matrix. The restricted three-body problem always has two equilibrium points which are at the vertices of an equilateral triangle, one of whose sides is the line segment joining the two primaries. The linearized equations about this equilibrium point consist of two harmonic oscillators when the mass ratio is small, and form a complex saddle when the mass ratio is near $\frac{1}{2}$. There is one specific value of the mass ratio where the linearized equation has two equal pairs of imaginary eigenvalues and the Jordan canonical form for the coefficient matrix has off-diagonal elements. When restricting to symplectic similarity transformations, the canonical form for the linearized system is

$$\begin{pmatrix} \omega i & 1 & 0 & 0 \\ 0 & \omega i & 0 & 0 \\ 0 & 0 & -\omega i & 0 \\ 0 & 0 & -1 & -\omega i \end{pmatrix} \tag{4.8}$$

with certain reality conditions.

After some preparation the Hamiltonian is of the form

$$H=i\omega(z_1z_3+z_2z_4)+z_2z_3+\left(a_1z_1^2z_3^2+a_2z_1^2z_3z_4+a_3z_1^2z_4^2\right)+\cdots \tag{4.9}$$

where the z's are complex coordinates satisfying the reality conditions $\bar{z}_1=-z_4$, $\bar{z}_2=z_3$. The best scaling for this problem will push the off-diagonal terms in the matrix into the higher order terms. Introducing a small parameter ε and scaling by

$$z_1\to\varepsilon z_1, \quad z_2\to\varepsilon^2 z_2, \quad z_3\to\varepsilon^2 z_3, \quad z_4\to\varepsilon z_4 \tag{4.10}$$

accomplishes this task. The scaling in (4.10) is symplectic with multiplier ε^{-3}, and so the Hamiltonian becomes

$$H=i\omega(z_1z_3+z_2z_4)+\varepsilon\left(z_2z_3+a_3z_1^2z_4^2\right)+\cdots. \tag{4.11}$$

The equations of motion implied by (4.11) are in the form (4.1), and the function B in (4.2) is easy to compute and analyze. The proper scaling in this problem has simplified not only the zeroth order terms but the first order terms as well, and this greatly simplifies the analysis. Applying the lemma to this problem establishes the existence of two families of periodic solution which emanate from this equilibrium point for the restricted three-body problem. The reader is referred to [11] where this problem and a more interesting one are discussed in detail.

There are many other interesting problems where scaling can greatly ease the analysis. The examples given here were chosen to illustrate a variety of different situations where scaling can help, without getting us too deeply involved in the technical aspects of the problem. The main point of this survey is to demonstrate how the correct scaling is obtained when suddenly the equations are greatly simplified, and suddenly (but after the fact) it is obvious why you should use that scaling.

REFERENCES

[1] R. ABRAHAM AND J. MARSDEN, *Foundations of Mechanics*, 2nd ed., Benjamin/Cummings, Reading, MA, 1978.

[2] A. DEPRIT AND J. HENRARD, *A manifold of periodic solutions*, Adv. Astron. Astrophys., 6 (1968), pp. 12 ff .

[3] J. D. HADJIDEMETRIOU, *The continuation of periodic orbits from the restricted to the general three-body problem*, Celestial Mech., 12 (1975), pp. 155–174.

[4] J. K. HALE, *Ordinary Differential Equations*, Wiley-Interscience, New York, 1969.

[5] G. W. HILL, *Researches in the lunar theory*, Amer. J. Math., 1 (1878), pp. 5–26, 129–147, 245–260.

[6] A. KELLY, *On the Liapunov sub-center manifold*, J. Math. Anal. Appl., 18 (1967), pp. 472–478.

[7] I. KUPKA, *Contribution à la théorie des champs génériques*, Contribution to Diff. Eqs., 2 (1963), pp. 457–484.

[8] V. V. NEMYTSKII AND V. V. STEPANOV, *Qualitative Theory of Differential Equations*, Princeton Univ. Press, Princeton, NJ, 1960.

[9] K. R. MEYER, *Periodic solutions of the N-body problem*, J. Differential Equations, 39 (1981), pp. 2–38.

[10] ——, *Periodic orbits near infinity in the restricted N-body problem*, Celestial Mech., 23 (1981), pp. 69–81.

[11] K. R. MEYER AND D. S. SCHMIDT, *Periodic orbits near L_4 for mass ratios near the critical mass ratio of Routh*, Celestial Mech., 4 (1971), pp. 99–109.

[12] ——, *Hill's lunar equations and the three-body problem*, J. Differential Equations, 44 (1982), pp. 1–10.

[13] H. POLLARD, *Mathematical Introduction to Celestial Mechanics*, Prentice-Hall, Englewood Cliffs, NJ, 1966.

[14] H. POINCARE, *Les méthodes nouvelles de la mécanique celeste*, Gauthier-Villar, Paris, 1892.

[15] C. ROBINSON, *Generic properties of conservative systems* I, II, Amer. J. Math., 92, pp. 562–603, 897–906.

[16] L. SEGEL, *Simplification and scaling*, SIAM Rev., 14 (1972), pp. 547–571.

[17] A. WINTNER, *The Analytic Foundations of Celestial Mechanics*, Princeton Univ. Press, Princeton NJ, 1941.