# STABILITY OF NONLINEAR HAMILTONIAN MOTION FOR A FINITE BUT VERY LONG TIME ${ }^{(a)}$ 

Robert L. Warnock and Ronald D. Ruth<br>Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309


#### Abstract

By constructing action variables that are very nearly invariant in a region $\Omega$ of phase space, and by examining their residual variation, we set long-term bounds on any orbit starting in an open subregion of $\Omega$. A new and generally applicable method for constructing the required high-precision invariants is applied. The technique is illustrated for transverse oscillations in a circular accelerator, a case with $21 / 2$ degrees of freedom and strong nonlinearity.


Submitted to Physical Review Letters.

[^0]In applications of nonlinear Hamiltonian mechanics it is often interesting to establish theoretical limits on the motion over extremely long intervals of time. For instance, in the design of cyclic particle accelerators and storage rings one would like to predict stability of particle orbits over a beam storage time of several hours. During such a time, a particle makes a stupendous number of interactions with localized nonlinear magnetic fields. In the Superconducting Supercollider (SSC) the number will be around $10^{12}$, while the stored particles make about $10^{8}$ turns around the ring.

To approach such problems it is useful to consider invariant surfaces in phase space. The effective phase space has dimension $D=2 d+\tau$ for a system with $d$ degrees of freedom, where $\tau=0$ if the system is autonomous, and $\tau=1$ if the Hamiltonian $\boldsymbol{H}$ is a periodic function of the time. We emphasize the latter case, which is of interest for accelerators, and exclude nonperiodic time dependence of $\boldsymbol{H}$. For nearly integrable systems as studied in the Kolmogorov-Arnol'd-Moser (KAM) theory: a large set of invariant tori exists. If $D \leq 4$, an invariant torus divides the space into two disjoint regions. We then have stability for all time, since an orbit evolving from a point inside the torus must stay inside forever. If $D \geq 5$, as is the case in relevant models of storage rings, an invariant torus no longer divides the space so as to confine orbits. Furthermore, arbitrarily near an invariant torus there are initial conditions for orbits that visit regions of phase space far removed from that torus. Such orbits follow stochastic layers near resonance structures that form a connected web permeating phase space. This phenomenon is broadly referred to as Arnol'd Diffusion, following the demonstration of such an effect by Arnol'd in an example with $D=5$. Thus, in high-dimensional systems the existence of invariant surfaces has no direct bearing on stability of orbits in a laboratory experiment. We must approach the stability question in a different way, and attempt to show that
the drift of orbits along resonances is so slow as to be harmless under conditions of interest.

Mathematical results in this direction were obtained by Nekhoroshev, ${ }^{4,5}$ who studied a wide class of nearly integrable systems of arbitrary dimension. The Nekhoroshev Theorem states that an orbit will be confined to a specified, bounded region of phase space at least for a time $T$ that increases exponentially as the strength $\epsilon$ of the non-linear- perturbation tends to zero. Unfortunately, the Nekhoroshev Theorem has no direct practical application, since $\epsilon$ must be absurdly small to guarantee a stability time $T$ of suitable magnitude. This situation results from pessimistic estimates that are required in the rigorous analysis. The true time of stability is almost certainly underestimated by a huge margin.

We wish to show that an argument in the spirit of Nekhoroshev's proof, but quite different in technique, can be carried out numerically. Without a severe restriction on perturbation strength, we obtain bounds on the motion for suitably long time intervals. Owing to the finite nature of numerical analysis, the bounds are not mathematically rigorous, but in our view they are persuasive and much more reliable than conclusions based on the conventional method of "tracking". In the latter, one follows a few orbits by symplectic numerical integration of the equations of motion over a time interval that is usually much less than the desired time for stability. An extrapolation to claim stability on the desired interval is risky, since examples are frequent in which an orbit is apparently well-behaved for a long time, but finally displays instability.

We first describe the motion in terms of action-angle variables (I, $\boldsymbol{\Phi}$ ) of an underlying integrable system. The Hamiltonian will have the form

$$
\begin{equation*}
H(\mathbf{I}, \mathbf{\Phi}, \theta)=H_{o}(\mathbf{I})+V(\mathbf{I}, \mathbf{\Phi}, \theta) \tag{1}
\end{equation*}
$$

where bold-faced symbols denote d-dimensional vectors, and $\theta$ is the independent variable of Hamilton's equations, a monotonically increasing function of the time. The perturbation $V$ is $2 \pi$-periodic in $\mathbf{\Phi}$ and $\theta$. The following discussion requires only minor modifications if $\boldsymbol{V}$ is independent of $\theta$. Our treatment makes no direct reference to the Hamiltonian, being based on the time evolution map $\boldsymbol{M}$ for $N_{o}$ periods of $V$ :

$$
\begin{equation*}
M: \quad(\mathbf{I}, \mathbf{\Phi})_{\theta=0} \mapsto(\mathbf{I}, \boldsymbol{\Phi})_{\theta=2 \pi N_{o}} . \tag{2}
\end{equation*}
$$

This map will be evaluated by numerical integration of Hamilton's equations with an explicit fourth-order symplectic integrator. ${ }^{6}$ Following accelerator terminology, we call it the map for $N_{o}$ turns. In some examples, including complex accelerator models, it may be possible to represent $\mathcal{M}(\mathbf{I}, \boldsymbol{\Phi})$ by an explicit formula, at least for small $N_{o}$.

Our argument is based on a canonical transformation to new action-angle variables, $(\mathrm{I}, \mathbf{\Phi}) \mapsto(\mathrm{J}, \boldsymbol{\Psi})$, such that the new action J is nearly invariant under the map $\boldsymbol{M}$. The total change in J during $p N_{o}$ turns can be no greater than $\boldsymbol{p}$ times $\delta \mathbf{J}$, where $\delta \mathbf{J}$ is an upper bound for the change in J during $N_{o}$ turns, valid throughout the relevant region of phase space. If $\delta \mathbf{J}$ can be made sufficiently small, by a good choice of the transformation, then the change in J during $p N_{o}$ turns will stay within acceptable limits even for large $\boldsymbol{p}$. In Nekhoroshev's proof, the canonical transformation is obtained by perturbation theory of finite but high order, and is validated only at very small $\epsilon$. Our aim is to construct the transformation by a nonperturbative method, so as to provide a small $\delta \mathbf{J}$ even in cases of strong nonlinearity.

The canonical transformation is induced by a generating function $S(\mathbf{J}, \mathbf{\Phi}, \theta)=$ $\mathrm{J} . \boldsymbol{\Phi}+\mathrm{G}(\mathrm{J}, \boldsymbol{\Phi}, \theta)$ that is $2 \pi$-periodic in $\boldsymbol{\Phi}$ and $\theta$. The equations relating old and new variables are

$$
\begin{equation*}
\mathbf{I}=\mathbf{J}+G_{\boldsymbol{\Phi}}(\mathbf{J}, \boldsymbol{\Phi}, \theta), \quad \mathbf{\Psi}=\boldsymbol{\Phi}+G_{\mathbf{J}}(\mathbf{J}, \boldsymbol{\Phi}, \theta), \tag{3}
\end{equation*}
$$

where subscripts denote partial derivatives. In accord with our viewpoint based on maps, it is sufficient to deal with the transformation at $\theta=0$ only. If the transformation is ideal, so that J is a constant, then the first equation of Eqs. (3) furnishes an explicit representation of an invariant torus of dimension $\boldsymbol{d}+$ i; i.e., it gives । as a $2 \pi$-periodic function of $\boldsymbol{\Phi}$ and $\theta$, with J acting as a parameter to distinguish different tori. The average of I over $\boldsymbol{\Phi}$ is equal to J. At $\theta=0$ we employ the notation $\mathrm{I}=\mathrm{J}+\mathbf{u}(\mathbf{J}, 0)$, with $\mathbf{u}(\mathbf{J}, \boldsymbol{\Phi})=G_{\boldsymbol{\Phi}}(\mathbf{J}, \boldsymbol{\Phi}, 0)$.

For the nonintegrable systems of interest, exact invariant tori exist, if at all, only for values of J on a closed set of Cantor type. Nevertheless, tori that are approximately invariant exist as smooth functions of $J$ in open regions of phase space, and they can be computed numerically. Our computation is based on the fact that a family of approximate tori, corresponding to values of $J$ in an open region, determines a canonical transformation. An efficient numerical method to determine such a family and the resultant generator $G$ is described in Ref. 7 and is reviewed briefly below.

Given a transformation such that the new action J is nearly constant, we can proceed as follows to set limits on the residual variation of J . In a case with $\boldsymbol{d}=\mathbf{2}$, let $\Omega$ be the interior of a rectangle in the $\mathrm{J}=\left(\mathrm{Jr}, J_{2}\right)$ plane, and let $\Omega_{o}$ be the interior of a smaller, concentric rectangle so that $\Omega_{o} \subset \Omega$. Denote by $\Delta J_{i}$ the width of the annulus between $\Omega_{o}$ and $\Omega$ as crossed in the i-th direction. Suppose that the change in $J_{i}$ during $N_{o}$ turns, for any orbit with initial J in $\Omega$, is less than $\delta J_{i}$. Then any orbit with initial J in the smaller region $\Omega_{o}$ cannot reach the outer boundary of $\Omega$ in fewer than $\mathrm{N}=p N_{o}$ turns, where

$$
\begin{equation*}
p=\min _{i^{:}}\left(\Delta J_{i} / \delta J_{i}\right) \tag{4}
\end{equation*}
$$

This observation is useful if $p$ is sufficiently large. Since the largest tolerable excursion $\Delta J_{i}$ is usually imposed by the problem at hand, a large $p$ is to be achieved by making
$\delta J_{i}$ small through a good choice of the canonical transformation.
To determine the canonical transformation, we expand the function representing the torus in a finite Fourier series. We write

$$
\begin{equation*}
\mathrm{I}=\sum_{\mathrm{m}} \mathbf{u}_{\mathrm{m}} e^{i \mathrm{~m} \cdot \mathbf{\Phi}} \tag{5}
\end{equation*}
$$

and determine the coefficients $\mathbf{u}_{\mathbf{m}}$ so that (5) is satisfied at a finite set of points $(\mathbf{I}(\theta), \boldsymbol{\Phi}(\theta)), \boldsymbol{\theta}=0(\bmod 2 \pi)$, all lyingon a single non-resonant orbit. The coefficient $\mathbf{u}_{\mathbf{o}}$ of the constant term is identified with the action J, which varies with the choice of initial condition of the orbit. We repeat the process for various initial conditions, thereby obtaining $\mathbf{u}_{\mathbf{m}}(\mathrm{J})$ on a mesh of points $\mathrm{J}=\mathbf{J}_{\boldsymbol{i}}, \quad i=1,2, \cdots, \boldsymbol{k}$. We then define $\mathbf{u}_{\mathbf{m}}(\mathbf{J})$ as a smooth function of J by interpolating the values at mesh points with cubic spline functions. Thus, we have defined $\mathbf{u}(\mathbf{J}, \boldsymbol{\Phi})$ as a smooth function of J that can be identified with $G_{\boldsymbol{\Phi}}(\mathbf{J}, \boldsymbol{\Phi}, 0)$. Integration of the Fourier series with respect to $\boldsymbol{\Phi}$ yields the generator of the desired canonical transformation. The constant of integration, corresponding to the $m=0$ term in $G$, can be set equal to 0 . The equation $\mathrm{I}=\mathrm{J}+\mathbf{u}(\mathbf{J}, \boldsymbol{\Phi})$ defines implicitly a function $\mathbf{J}(\mathbf{I}, \boldsymbol{\Phi})$ that will be constant over each of the $k$ sets of orbit points used in the above construction. (Here we suppose that the Jacobian matrix $d \mathbf{I} / d \mathbf{J}$ is non-singular, as is the case in the examples we have treated.) If $\mathbf{J}(\mathbf{I}, \mathbf{\Phi})$ is nearly constant on a continuous family of nearby orbits, then $\delta \mathbf{J}$ will be small and it will be possible to establish long-term bounds.

The determination of the Fourier coefficients $\mathbf{u}_{\mathbf{m}}$ from orbit data cannot be done as a simple discrete Fourier transform, since the values of the angles $\boldsymbol{\Phi}(\theta)$ are scattered unpredictably. Instead, the coefficients can be obtained by iterative solution of a set of linear equations in which the unknowns are the values of $I$ on a regular mesh. This method, introduced in Ref. 7, leads to remarkably accurate tori at relatively low computational cost.

The determination of Fourier coefficients fails, as it ought, when the orbit is resonant. To set up the canonical transformation we choose mesh points $J_{i}$ corresponding to non-resonant orbits, and interpolate with spline functions so as to form bridges over intervening resonances. We emphasize that interpolation is an essential part of the argument, not just a feature forced upon us by our reliance on numerical analysis. Indeed, Nekhoroshev's argument also involves a canonical transformation that is defined as a smooth function of J, in spite of the presence of resonances. Smooth interpolations of exact invariant tori are also possible, as was shown in a KAM context by Pöschel. ${ }^{8}$

To compute a bound it remains to determine $\delta J_{i}$. Because of practical limits on computation time, there is some uncertainty in this determination, but with some care the uncertainty can be made rather small. Note that the only failure of rigor in our argument arises at this point. The canonical transformation itself is mathematically well defined, even though it was obtained numerically.

We denote the increment in J over $N_{o}$ turns as J' -J $=\mathcal{D}\left(\mathbf{J}, \boldsymbol{\Phi}, N_{o}\right)$. To compute $\mathcal{D}$ we simply observe the time evolution of J induced by the map $\boldsymbol{M}$ through the change of variable $(J, \boldsymbol{\Phi}) \rightarrow(I, \boldsymbol{\Phi})$ and its inverse. The change of variable is given explicitly in Eq. (3); its inverse is computed easily by Newton's method with J as the first guess for J'. In the example studied below, the function $\mathcal{D}$ has many oscillations as a function of $\boldsymbol{\Phi}$ but relatively few as a function of J on $\Omega$. It is impractical to follow every oscillation in seeking the upper bound $\delta J_{i}$ of $\mathcal{D}_{i}$, but one can do random sampling with statistical estimates of sampling error to find a fairly convincing value of $\delta J_{i}$. The reader may consult Ref. 9 for details on this relatively delicate problem.

For illustration we treat a basic problem of accelerator physics, the so-called betatron motion, which consists of oscillations transverse to the direction of the beam in a cyclic accelerator or storage ring. ${ }^{10}$ The coordinates $x_{i}(i=1,2)$ are transverse
displacements with respect to a closed reference orbit of circumference $2 \pi R$, and the conjugate (dimensionless) momenta are $p_{i}=d x_{i} / d(R \theta)$. The motion is essentially perturbed harmonic motion, with the perturbation arising from sextupole magnets that are used to compensate the momentum dependence of the focusing system. The field of a sextupole is concentrated in a small region of $\theta$, and gives a term in $V$ proportional to $x_{1}^{3}-3 x_{1} x_{2}^{2}$. Thus, the contribution of a sextupole to the map $\boldsymbol{M}$ resembles a four-dimensional quadratic map. After a canonical transformation, similar to the familiar one for harmonic oscillators but a little more involved, the Hamiltonian takes the form

$$
\begin{gather*}
H(\mathbf{I}, \boldsymbol{\Phi}, \theta)=\boldsymbol{\nu}_{o} \cdot \mathbf{I}+\sum_{j=1}^{n} F_{j}(\theta)\left[\left(\beta_{1 j} I_{1}\right)^{3 / 2} \cos ^{3}\left(\Phi_{1}+\xi_{1 j}\right)\right.  \tag{6}\\
\left.-3\left(\beta_{1 j} I_{1}\right)^{1 / 2} \cos \left(\Phi_{1}+\xi_{1 j}\right) \cdot \beta_{2 j} I_{2} \cos ^{2}\left(\Phi_{2}+\xi_{2 j}\right)\right]
\end{gather*}
$$

The tunes $\nu_{o i}$ (winding numbers) are the unperturbed frequencies normalized to the beam revolution frequency. The function $F_{j}(\theta)$ is nonzero only over the extent of the $j$-th sextupole, where it has a constant value. Between sextupoles there is linear propagation at constant I. The constants $\beta_{i j}$ and $\xi_{i j}$ are determined by linear aspects of the magnets that guide and focus the beam. The action $I_{i}$ is measured in units of length, being the usual action divided by the momentum of the beam.

We derive a bound for two-dimensional betatron motion in a configuration of magnets that corresponds to one cell of the Berkeley Advanced Light Source (ALS); four sextupoles are involved. The parameters of Eq. (6) are given in Ref. 11. This example allows relatively fast computation, but involves nonlinear phenomena similar to those in large hadron colliders. The sextupoles are so strong as to drive highorder resonances such as those excited by high-order multipoles in superconducting magnets. We work in a region $\Omega$ of substantial nonlinearity, about halfway to the short-term dynamic aperture in the ( $x_{1}[\max ], x_{2}[\max ]$ ) plane. (The short-term dy-


Figure 1: Plot of orbit points $\left(I_{1}, \Phi_{1}, \Phi_{2}\right)$ on a torus with $\mathbf{J}$ in the region $\Omega$ defined in Eq. (7). The origin is at $I_{1}=0, \Phi=0$.
namic aperture is defined loosely as a boundary beyond which the motion is unstable after a few thousand turns.) With actions measured in units of $10^{-6}$ meters the region $\Omega$ is given by

$$
\begin{equation*}
2.51<J_{1}<2.82, \quad 1.34<J_{2}<1.64 . \tag{7}
\end{equation*}
$$

Orbit points on a typical invariant surface with J in this region are shown in Figure 1. We plot in three dimensions the points $\left(I_{1}, \Phi_{1}, \Phi_{2}\right)$; according to the first component of Eq. (3), these points should lie on a two-dimensional surface.

By using our canonical formalism to map resonances from the frequency plane into the J plane, we find that $\Omega$ contains the resonances shown in Figure 2. This figure shows the images of all resonant lines $m_{1} \nu_{1}+m_{2} \nu_{2}=\mathrm{n}$ with $\left|m_{i}\right| \leq 20$, where the perturbed tunes (winding numbers) are denoted by $\nu_{i}$; the $m_{i}$ and $n$ are integers. The stars mark the mesh points $\mathbf{J}_{\boldsymbol{i}}$ used to set up the canonical transformation as a smooth function of J. The transformation as represented in (5) involves up to


Figure 2: The image in the $\left(J_{1}, J_{2}\right)$ plane of all resonance lines $m_{1} \nu_{1}+m_{2} \nu_{2}=\boldsymbol{n}$ with $\left|m_{i}\right| \leq 20$, for the region $\Omega$ defined in Eq. (7). Each line is labeled by $\left(m_{1}, m_{2}\right)$. The stars indicate the mesh points $\mathrm{J}_{i}$ used to set up the canonical transformation.

20 Fourier modes in each angle $\Phi_{i}$.
By means of the procedure outlined above, we find the following values for the numbers $\delta J_{i}$ that bound $\mathcal{D}_{i}$, for $N_{o}=10^{4}$ :

$$
\begin{equation*}
\left(\delta J_{1}, \delta J_{2}\right)=(2.8,4.0) \cdot 10^{-6} \tag{8}
\end{equation*}
$$

The various tests used to certify these values are described in Ref. 9. The corresponding values for $N_{o}=10^{k}, k=0,1,2,3,4$, have similar magnitudes. Let us choose $\Delta J_{i}$ of Section 2 so that $q=\Delta J_{i} / \delta J_{i}=10^{4}$, with $N_{o}=10^{4}$. Then the subset $\Omega_{o}$ of $\Omega$ is defined by

$$
\begin{equation*}
2.54<J_{1}<2.79, \quad 1.38<J_{2}<1.60 \tag{9}
\end{equation*}
$$

Any orbit beginning in $\Omega_{o}$ will stay within the slightly larger region $\Omega$ for at least $p N_{o}=10^{8}$ turns. This result is quite satisfactory, since a stability time of $10^{8}$ turns is in a range of practical interest, and far beyond the range accessible by direct tracking. Recall that we have obtained the result by tracking for $10^{4}$ turns from many initial conditions, a technique that implies good control of rounding error.

All resonances in the region $\Omega$ defined above are weakly excited, and have little effect. The variation of J on the resonance lines is hardly stronger than elsewhere in the region. In other regions, at comparable amplitudes, we encounter strong resonances that are associated with larger variations of J. This does not necessarily imply a degradation of the time for stability, since oscillations on a well isolated resonance can be stable for a long time and not associated with fast transport to nearby resonances, even if the amplitude of oscillation is fairly large. The derivation of long-term bounds in this situation is discussed in Ref. 9.

We have demonstrated the feasibility of bounds on Hamiltonian motion for very long time intervals under conditions of strong nonlinearity. The scheme is quite general, and proceeds in the same way for any Hamiltonian system. Points on orbits of the-time evolution map $M$ are the only data required to establish bounds. Although the method was motivated by problems in accelerator theory, it should be of interest as well for stability questions in other fields such as plasma theory and celestial mechanics.

We are indebted to J. Irwin and É. Forest for encouragement and good ideas.

## REFERENCES

1. G. Gallavotti, The Elements of Mechanics (Springer, New York, 1983).
2. A. J. Lichtenberg and M. A. Lieberman, Regular and Stochastic Motion (Springer, New York, 1983).
3. V. I. Arnol'd, Sov. Math.-Doklady 5, 581 (1964).
4. N. N. Nekhoroshev, Russian Math. Surveys 32:6, 1 (1977).
5. G. Benettin, in Nonlinear Evolution and Chaotic Phenomena, eds. G. Gallavotti and P. F. Zweifel (Plenum, New York, 1988).
6. R. D. Ruth, IEEE Trans. Nucl. Sci. 30, 2669 (1983); É. Forest and R. D. Ruth, Physica 43D, 105 (1990).
7. R. L. Warnock, SLAC-PUB-5327, submitted for publication.
8. J. Pöschel, Comm. Pure Appl. Math. 35, 653 (1982).
9. R. L. Warnock and R. D. Ruth, SLAC-PUB-5267, in preparation.
10. R. D. Ruth, in Physics of Particle Accelerators, Vol. 1, AIP Conf. Proc. 153, eds. M. Month and M. Dienes (Amer. Inst. Phys., New York, 1987).
11. W. E. Gabella, R. D. Ruth, and R. L. Warnock, Proc. 1989 IEEE Part. Accel. Conf., Chicago, March 20-23, 1989, p. 1310, eds. F. Bennett and J. Kopta (IEEE, New York, 1989).

[^0]:    (a) Work supported by Department of Energy contract DE-AC03-76SF00515.

