Ann. Rev. Nucl. Part. Sci. 1986. 36 : 287–325

NONLINEAR PROBLEMS IN ACCELERATOR PHYSICS

S. G. Peggs and R. M. Talman[1]

Superconducting Super Collider Central Design Group, Lawrence Berkeley Laboratory, Berkeley, California 94720

CONTENTS

[1] Permanent address: Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853.

0163–8998/86/1201–0287$02.00

1. INTRODUCTION

Certain nonlinear problems of accelerator physics are both important for successful operation of accelerators and interesting as problems in their own right. This review attempts to illustrate both of these aspects. Since many readers are assumed to have little technical knowledge of accelerators, we simplify (even oversimplify) in the interest of suppressing unessential complication. We concentrate instead on features that are legitimately motivated by accelerator performance requirements.

No systematic attempt is made to survey accelerator experience as an observational "laboratory" in which curious behavior predicted by nonlinear mechanics is to be sought, though a case can be made that such study is justified. Unfortunately, as a theme, "nonlinear" is about as unifying as "bizarre" would be. Fortunately though, a few of the most important effects in accelerators are well described by methods known in other areas of physics or mathematics (typically their development has been independent.) To make this review of more general interest, such aspects are stressed in preference to uniquely accelerator physics issues.

In the interest of making the paper self-contained, a discussion of accelerator essentials is given in Section 2. As well as establishing terminology, a basis is established for concepts like transfer maps, concatenation, and Liouville's theorem, though comments on them are deferred to later sections. Section 3 forms the best example of the sort described in the previous paragraph. Equations first solved in an entirely mathematical context are applied to the analysis of an accelerator containing precisely one nonlinear element. The results will be referred to repeatedly in the sequel. A method of exact solution is described that illustrates essential nonlinear features and that will serve as a basis for discussion of more complicated modern methods. The important and interesting problems of resonant extraction and resonances in general are treated largely as a specialization of this problem in Section 4. Again the description of accelerator behavior can be compared with descriptions used in other areas.

The following sections are more technical and less explicit, relating back where possible to the previous example. Included is discussion of nearly linear behavior, which though less bizarre is probably more important for accelerator design than is the potentially catastrophic large amplitude behavior. Some other modern areas and methods are also described. Other authors would have given at least one of these areas far greater prominence. For example, the entire subject can be formulated in Hamiltonian terms, or in terms of the Krylov-Bogoliubov method.

The problem of beam-beam interactions is discussed in Section 6. Of the problems discussed this is the most important, at least as a mechanism

that limits performance of colliding-beam facilities. It has been the object of intense study, experimentally, theoretically, and numerically. If this article has a unifying theme it is to explain, in as simple terms as possible, ideas useful in the analysis of this problem.

Problems of the collective motion of the essentially infinite number of particles circulating in an accelerator are perhaps the most challenging, important, and difficult faced in accelerator physics but they are not discussed in this review.

Regrettably, since the publication of research in accelerators has been spotty, it is necessary to refer to laboratory reports that are not readily available. Without claiming to be systematic, references to review-type articles are also given. Since most of the results have been derived independently by different authors using different methods, the chronology is confusing. There are several useful general references (1–4).

2. ACCELERATOR ESSENTIALS

2.1 *Deflection by Thin Elements*

Consider a particle of mass m and charge e traveling approximately along the s axis of a locally Cartesian x, z, s coordinate frame through a purely magnetic element of length L. The field is assumed to be two dimensional, independent of s. The most general field of this sort can be expanded into polynomials in x and z according to

$$L(B_z+iB_x) = LB_0 \sum_{n=0}^{M} (b_n+ia_n)(x+iz)^n. \qquad 1.$$

Any accelerator can be broken into elements of this sort. To simplify this review, and without much loss in generality, all elements are treated as "thin." This is accomplished by allowing the length L to approach zero while holding the "field integrals" LB_x and LB_z finite. This may entail artificially breaking up into shorter lengths elements that are actually long. The particle deflection in passing through such a field is governed by the equation

$$\mathrm{d}\mathbf{v}/\mathrm{d}s = c^2/E\,\mathrm{d}t/\mathrm{d}s\,\mathrm{d}\mathbf{p}/\mathrm{d}t = (c^2/E)(e/v_s)\mathbf{v}\times\mathbf{B}, \qquad 2.$$

where the quantities $\mathbf{p}$, E, $\mathbf{v}$, and c have their usual meanings.

In passing through this element the particle will suffer an impulsive deflection (a kink) given by

$$\Delta v_x/v = -B_z Le/p \equiv F_x \qquad 3.$$

$$\Delta v_z/v = B_x Le/p \equiv F_z. \qquad 4.$$

Table 1 "Forces" exerted by standard elements, and common notation for the coefficients

	n	R_n	I_n	a'_n	b'_n	F_x	F_z
Horizontal bend	0	1	0	0	$\Delta\theta_x$	$-\Delta\theta_x$	0
Vertical bend	0	1	0	$\Delta\theta_z$	0	0	$\Delta\theta_z$
Erect quadrupole	1	x	z	0	q	$-qx$	qz
Skew quadrupole	1	x	z	q_s	0	q_sz	q_sx
Erect sextupole	2	x^2-z^2	$2xz$	0	$S/2$	$S(z^2-x^2)/2$	Sxz
Skew sextupole	2	x^2-z^2	$2xz$	$S_s/2$	0	S_sxz	$S_s(x^2-z^2)/2$

The longitudinal velocity component v_s will also be altered; it can be worked out from the requirement that v remain constant in a magnetic field.

The most important accelerator components can be identified with the leading terms in the expansion Equation 1. Introducing the notation

$$a'_n, b'_n = (LB_0e/p)a_n, b_n \qquad 5.$$

one has

$$F_x = \sum_{n=0}^{M} (-b'_nR_n + a'_nI_n)$$
$$F_z = \sum_{n=0}^{M} (b'_nI_n + a'_nR_n), \qquad 6.$$

where

$$(x+iz)^n = R_n + iI_n. \qquad 7.$$

Examples of the leading terms in these expansions are given in Table 1. For a circular accelerator it is the horizontal bends $\Delta\theta_x$ that guide the particle. By defining a central closed orbit and letting x and z be deviations from that orbit, the leading deflections are due to quadrupoles, which we will assume are centered on the closed orbit. Since these forces are linear in x and z their effect is usually represented in matrix notation;

$$\begin{pmatrix} x \\ x' \end{pmatrix}_2 = \begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_1 \qquad 8.$$

$$\begin{pmatrix} z \\ z' \end{pmatrix}_2 = \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix} \begin{pmatrix} z \\ z' \end{pmatrix}_1 \qquad 9.$$

which represents propagation from point 1 just before the quadrupole to point 2 just after it. The prime stands for d/ds and the normally extremely

good "paraxial" approximation

$$v \approx v_s \qquad 10.$$

has been made.

2.2 *Linear Lattice Design*

Most of accelerator design is based on the paraxial approximation and the resultant linearized equations. In preparation for the main topic of this review, namely the subsequent nonlinear terms, it is important to understand the conventional linear formalism. This is called "optics" since it is no different from the geometric optics of light, using ordinary lenses.

As well as accounting for deflections in the lenses it is necessary to describe propagation through field-free drift spaces. For length l one has

$$\begin{pmatrix} x \\ x' \end{pmatrix}_2 = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_1 \qquad 11.$$

Figure 1 shows the basic FODO cell on which most accelerators are based. (FODO stands for focus-drift-defocus-drift.) Neglecting the dipoles and sextupoles for now, the linear elements can be "concatenated" by matrix multiplication to obtain the "transfer matrix" M_{21}, which relates the state vector $X_2 = (x, x')_2$ at the end of the cell to the state vector X_1 at the beginning,

$$X_2 = M_{21}X_1. \qquad 12.$$

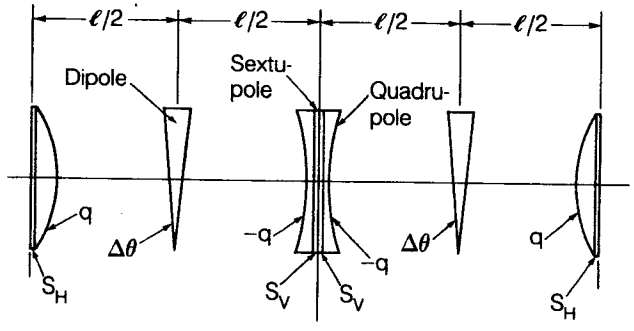

Figure 1 Dimensions, magnetic elements, and strength parameters for a FODO accelerator cell.

It is

$$M_{21} = \begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix}\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2q & 1 \end{pmatrix}\begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -q & 1 \end{pmatrix} \quad 13.$$

$$= \begin{pmatrix} \cos 2\phi & \beta \sin 2\phi \\ -\sin 2\phi/\beta & \cos 2\phi \end{pmatrix}$$

where

$$\phi = \sin^{-1} lq; \qquad \beta = [(1+lq)/(1-lq)]^{1/2}/q. \quad 14.$$

This special case has been worked out both as an example and because it gives an excellent description of most sections of most accelerators. But, in greater generality, in the linear approximation, the transverse coordinate x will execute a "betatron oscillation" satisfying an equation of motion

$$d^2x/ds^2 + K(s)x = 0. \quad 15.$$

The "focusing function" $K(s)$ can represent arbitrary variation of the focusing as a function of the longitudinal coordinate s, including thin lenses, in which case $K(s)$ would be a sum of delta functions.

There is a method of solving Equation 15 that resembles the WKB method, but is exact. It is to assume two solutions of the form

$$x_{\pm}(s) = aw(s)\exp[\pm i\psi(s)], \quad 16.$$

where a is an arbitrary constant factor. It is also customary, and consistent with the usage in Equation 14, to define the so-called beta function according to

$$w(s) = \beta^{1/2}(s). \quad 17.$$

From Equation 16 the so-called pseudo-harmonic solution

$$x(s) = a\beta^{1/2}(s)\cos[\psi(s) - \psi(s_1)] \quad 18.$$

can be constructed. Substituting it into Equation 15 and demanding independence of $\psi(s_1)$ one obtains

$$(\beta\psi')' = 0 \quad 19.$$

$$2\beta\beta'' - \beta'^2 - 4\beta^2\psi'^2 - 4\beta^2 K(s) = 0. \quad 20.$$

Integrating Equation 19, and making a conventional choice for the integration constant, yields

$$\psi' = 1/\beta, \quad 21.$$

which can be integrated to give

$$\psi(s) = \int_{s_1}^{s} ds/\beta(s). \quad 22.$$

Further, ψ' can be eliminated from Equation 20 using Equation 21. For a linear equation like Equation 15, propagation from s_1 to s_2 can be described by a transfer matrix

$$M_{21} = \begin{pmatrix} C_{21} & S_{21} \\ C'_{21} & S'_{21} \end{pmatrix} \quad 23.$$

where $C_{21} \equiv C(s_2, s_1)$ is that "cosine-like" solution having unit value and zero slope at s_1 and S_{21} is the "sine-like" solution having zero value and unit slope at the same point. They are given by

$$C_{21} = (w_2/w_1)\cos\psi_{21} - w_2 w_1' \sin\psi_{21} \quad 24.$$

$$S_{21} = w_1 w_2 \sin\psi_{21}, \quad 25.$$

where $\psi_{21} = \psi_2 - \psi_1$.

Especially important is the situation in which $s_2 - s_1$ is equal to the circumference C of the machine, as $K(s)$ is necessarily periodic

$$K(s+C) = K(s). \quad 26.$$

(A refinement that we do not pursue here is that, if the machine has "superperiodicity," then Equation 26 may also hold for sub-multiples of C.) When subjected to this periodicity condition, Equation 15 is known as the Hill equation. It is then sensible and possible to demand that β have the same periodicity, which with Equation 20 fixes it uniquely as a function of s. The transfer matrix around the whole machine is given by

$$M(s) = \begin{pmatrix} \cos\mu + \alpha(s)\sin\mu & \beta(s)\sin\mu \\ -\gamma(s)\sin\mu & \cos\mu - \alpha\sin\mu \end{pmatrix}, \quad 27.$$

where, comparing with Equations 24 and 25

$$\alpha(s) = -ww' = -\beta'(s)/2 \quad 28a.$$

$$\beta\gamma - \alpha^2 = 1 \quad 28b.$$

$$\mu = \oint ds/\beta(s) \equiv 2\pi Q. \quad 28c.$$

From Equation 28b it can be seen that the determinant of $M(s)$ is equal to 1—a feature discussed below. The "tune" Q defined in Equation 28c,

which according to Equations 18 and 21 is the phase advance divided by 2π, is the number of oscillations in going once around the machine.

2.3 *Sources and Effects of Nonlinearity*

Nonlinear forces are present in accelerators both intentionally and unintentionally. The latter result when magnetic elements that would ideally be pure dipoles or pure quadrupoles have, as they are actually constructed, field non-uniformities requiring further terms in the expansion Equation 1 to give an adequate representation of the field. This source of nonlinearity has become relatively more important as smaller coil diameters have been required for reasons of economy. For the Superconducting Super Collider (5), this source is expected to be dominant, even though the error terms appearing in Equation 1 can be held below one part in a thousand, or even better with extreme care in the manufacturing process. It is because there are thousands of magnets that such small terms can be important. This also makes theoretical analysis of the nonlinear effects especially difficult.

Sextupoles are the predominant source of nonlinearity in most existing accelerators. They are intentionally installed in order to control or correct the machine "chromaticity." A particle having fractional momentum excess δ will be more weakly focused, $q \to q(1-\delta)$, by the quadrupoles making up the accelerator lattice. As a result there will be momentum dependence $Q(\delta)$ of the tune. The quantity $\mathrm{d}Q/\mathrm{d}\delta$ is called the chromaticity. Because of the previously mentioned weakening it naturally has a negative value, which, coming from the FODO sections making up the regular arc or arcs of the accelerator, is approximately $-Q$. The presence of "low-β" intersection regions in storage rings can double or triple this value.

This chromaticity is typically unacceptable for two reasons. Beams of vanishingly small momentum spread are both unachievable and, for reasons of high current stability, undesirable. To avoid resonances (see below) the spread of tunes must be kept small and this requires that $\mathrm{d}Q/\mathrm{d}\delta$ be held near zero. Actually, to suppress the "head-tail effect," a particular current-dependent effect, $\mathrm{d}Q/\mathrm{d}\delta$ must be held somewhat positive.

Fortunately, the momentum offset δ leads to a spatial offset owing to the lattice "dispersion" $\eta(s)$. A higher momentum particle follows an orbit that is systematically at a larger radius. The outward displacement is symbolized by $\eta\delta$. If the accelerator were constructed entirely out of the FODO cells of Figure 1, the dispersion $\eta(\text{peak})$ at the ends of the cell would be given by

$$\eta(\text{peak}) = (\Delta\theta/lq^2)(1+lq/2). \qquad 29.$$

A sextupole of strength q/η, if superimposed on a regular arc quadrupole of strength q, will render that quadrupole achromatic, as can be seen by

extracting the coefficient of $x\delta$ in the expression for the deflection in the quadrupole-sextupole combination, which is

$$\Delta x' \approx q(1-\delta)x + (q/2\eta)(x+\eta\delta)^2. \qquad 30.$$

Effectively, the off-momentum particle, by virtue of its displacement $\eta\delta$, feels a quadrupole field stronger by just that factor needed to compensate for its increased inertia. In practice the sextupole may be somewhat stronger than this to compensate for chromaticity introduced elsewhere in the ring.

One beneficial application of nonlinear elements has been described, but otherwise their effects are mainly bad because they degrade beam quality by diluting the density of particles in phase space. According to Liouville's theorem, equivalent to Equation 28b for linear systems, and true also in the presence of nonlinearities, the microscopic phase space density is invariant. But owing to "filamentation" (4a), caused by nonlinearities, the macroscopic or average phase space density may be reduced. This is especially serious during beam manipulations such as injection or extraction. Even with steady conditions, particle oscillation amplitudes can increase to such large values, normally because of some resonance phenomenon, that some or all of the particles are lost.

3. LATTICE WITH ONE NONLINEAR ELEMENT

3.1 *Transfer Map and Difference Equation*

So far the discussion has been restricted to particles having betatron amplitudes small enough to justify dropping quadratic and higher powers in the multipole expansion, Equation 1. For larger amplitudes at least some of these terms must be retained. When the equation of betatron motion (Equation 15) is appropriately generalized, it becomes nonlinear and, as a result, much harder to solve. Some, but not all, of the difficulties can be illustrated by analyzing a system with one thin nonlinear element in an otherwise linear lattice. This example, which can be solved exactly for modest amplitudes, is addressed next.

It is known (6–8) that any conservative dynamical system with two degrees of freedom can, using Poincaré surfaces-of-section (defined below), be reduced to a transfer map such as Equation 12 as regards small deviation from a repetitive motion, and to a generalized nonlinear transfer map for larger amplitudes. Problems falling into this category include the three-body problem of astronomy (9), satellite orbits (10) and oscillations (11), and magnetically trapped particles (12, 13). Hénon (14) has plausibly distilled all such maps down to a unique simplest nontrivial form, and studied that map both theoretically and numerically. It had previously

been studied (15, 16), but this work was not particularly influential in accelerator physics, perhaps because it was too specialized. Here, following an independent development (17, 18), we specialize appropriately to obtain equations of Hénon's form so that we can use his results directly.

An observer stationed at a fixed point in an accelerator observing one-dimensional (say horizontal) motion can plot the phase space point x_t, x'_t for successive turns t, to make a Poincaré plot. The symbol t has been chosen because the turn number is a kind of quantized time, with the time unit being the revolution period in an accelerator. It will, however, always be an integer. Suppose that there is a nonlinear element situated at the observation point administering a deflection $\Delta x'$ that acts half before and half after that point. Propagation around the rest of the ring is described by Equation 27. Combining, the turn-to-turn propagation is described by

$$\begin{pmatrix} x \\ x'-\Delta x'/2 \end{pmatrix}_{t+1} = \begin{pmatrix} C_0+\alpha S_0 & \beta S_0 \\ -\gamma S_0 & C_0-\alpha S_0 \end{pmatrix} \begin{pmatrix} x \\ x'+\Delta x'/2 \end{pmatrix}_t, \qquad 31.$$

which gives the state vector on the $(t+1)$st passage, and where we have abbreviated

$$C_0 = \cos\mu_0; \quad S_0 = \sin\mu_0; \quad C = \cos\mu \text{ etc.} \qquad 32.$$

In what follows C_0 will emerge as a more convenient variable than μ_0. The subscript 0 has been added to μ so that it will be notationally possible to distinguish between μ_0, the unperturbed value, and μ, which includes any frequency shift for which the perturbation is responsible.

For reasons discussed below $\Delta x'_t$ is assumed to be a function only of x_t. Equation 31 is a nonlinear transfer map or, in other words, two coupled first-order difference equations in the two variables x_t and x'_t. An equivalent second-order difference equation in the one unknown x_t can be obtained. To eliminate x'_t, write the equation analogous to Equation 31 but expressing the time-reversed propagation from t to $t-1$. Adding this to Equation 31 yields

$$x_{t+1}-2C_0x_t+x_{t-1} = \beta S_0 \Delta x'_t. \qquad 33.$$

Further algebraic manipulation yields the formula

$$x'_t = \frac{x_{t+1}-x_{t-1}-2\alpha S_0 x_t}{2\beta S_0} \qquad 34.$$

by which the slope can be found once the position is known.

Using Equation 1 we now write an explicit formula for $\Delta x'_t$:

$$\Delta x'_t = b'_2 x_t^2 + b'_3 x_t^3, \qquad 35.$$

where the first term yields the Hénon map. One further term is retained

to illustrate a respect in which the Hénon map is not typical of all linearities. In order to obtain precisely the Hénon map one takes

$$\alpha = 0; \quad \beta = 1; \quad b'_2 = 1; \quad b'_3 = 0. \qquad 36.$$

Much of this specialization results merely from choosing units such that $b'_2\beta^2 = 1$. One is left with a two-dimensional quadratic map depending on one parameter, μ_0, which in our context is called 2π times the tune. Typical results from Hénon are shown in Figure 2. These are Poincaré maps for

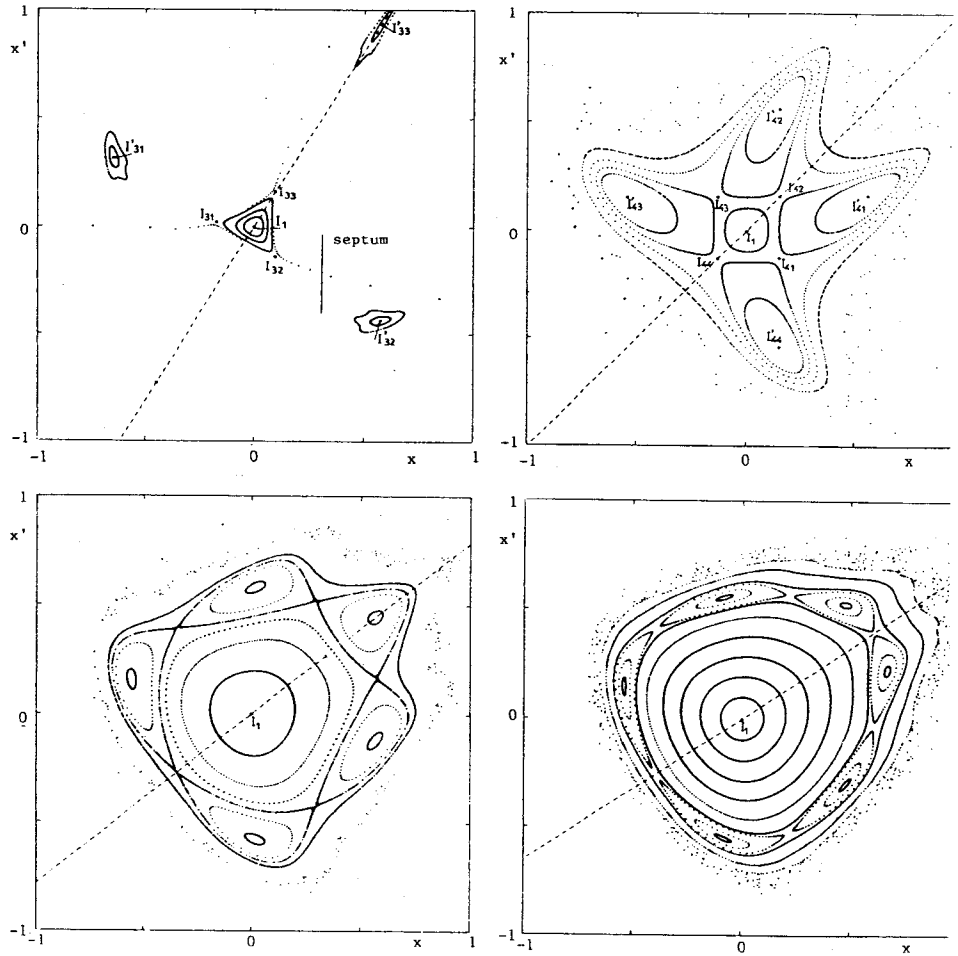

Figure 2 Numerical results from Hénon (14) from iterating his map. (*a*) $Q_0 = 0.324$, close to 1/3. About ten chaotic large amplitude points have been removed from this plot to avoid confusion in the discussion of resonant extraction. (*b*) $Q_0 = 0.2516$, close to 1/4. (*c*) $Q_0 = 0.211$, close to 1/5. (*d*) $Q_0 = 0.185$, close to 1/6.

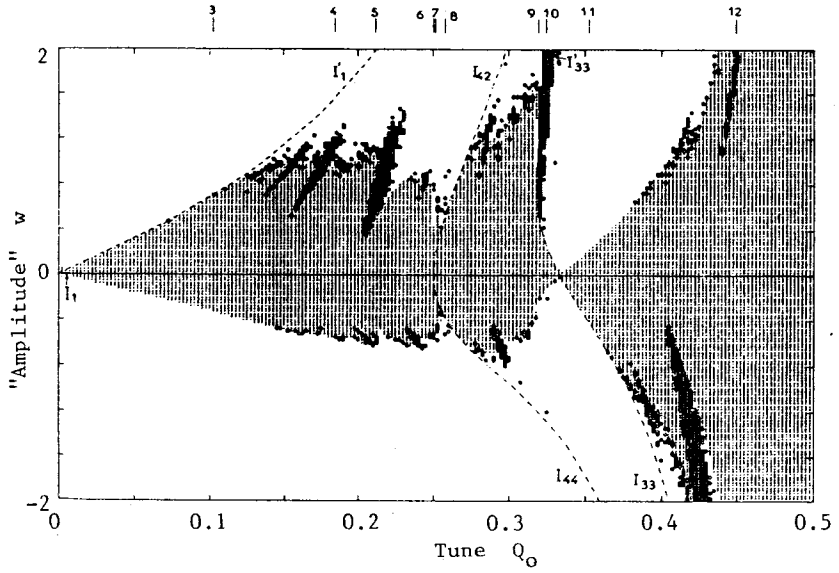

Figure 3 Synopsis of Hénon's numerical maps, showing the dependence on tune of region I (*shaded*), region II (*black*), and region III (*blank*).

particular tune values. There are several more such plots in the original paper. Three qualitatively different types of motion are observed, corresponding to regions of parameter space that can be labelled I, II, and III, nearly (but not completely) in order of ascending amplitude. The qualitative features are that for all time the points either

I. lie on regular closed smooth curves;
II. lie on islands, jumping from island to island on successive turns; or
III. follow chaotic trajectories, jumping around erratically and sometimes diverging eventually to infinity.

These features will be referred to repeatedly. The way the regions change as the tune Q_0 is varied is shown in Figure 3. Regions of type I are indicated by shaded, type II by black, and type III by blank regions. In this plot it is w, the distance from the origin along the axis of symmetry of the map, that is used to characterize the particle amplitude; it is plotted as the vertical coordinate. Acceptable accelerator operation would be anticipated in region I, and possibly in region II.

3.2 *Iteration to an Exact Solution*

Features of this and other maps can be inferred by analysis of the more general map (Equation 33). By judicious choice of the starting label the

general unperturbed motion can be written

$$x_t = a \cos t\mu \quad t = 0, 1, 2, \ldots, \tag{37.}$$

where

$$\mu = \mu_0, \tag{38.}$$

which satisfies Equation 33 with $b'_2 = b'_3 = 0$. To allow for a possible frequency shift due to the perturbation, Equation 33 can be rewritten as

$$x_{t+1} - 2Cx_t + x_{t-1} = \beta S_0 \Delta x'_t + 2(C_0 - C)x_t. \tag{39.}$$

Here a term proportional to x_t, which could have been incorporated on the left-hand side of the equation, has instead been included in the perturbation, i.e. the right-hand side.

To begin an iterative scheme, $\Delta x'_t$ on the right side of Equation 39 can be approximated by substitution of Equation 37 into Equation 35. Since this is a function of $\cos t\mu$ it can be expanded in a Fourier series of terms $\cos rt\mu$, where r is an integer. The right side of Equation 39 becomes

$$\begin{aligned}&\beta S_0 a^2 b'_2/2 + [3\beta S_0 a^3 b'_3/4 + 2(C_0 - C)a] \cos t\mu \\ &\quad + \beta S_0 a^2 b'_2/2 \cos 2t\mu + \beta S_0 b'_3 a^3/4 \cos 3t\mu.\end{aligned} \tag{40.}$$

Similar expansions will be possible in subsequent iterations, with μ not equal to μ_0 in general. The equation to be solved at each stage is

$$x_{t+1} - 2Cx_t + x_{t-1} = \sum_{r=0}^{N} c_r \cos rt\mu, \tag{41.}$$

where c_r, $r = 0, 1, \ldots, N$ are known coefficients. Equation 40 terminates after a few terms, but for some perturbations it will be necessary to truncate the series after a "sufficiently large" number N of terms. Numerically, the coefficients c_r can be determined by straightforward application of standard fast Fourier transform (FFT) routines. (In a more general formulation terms $\sin rt\mu$ would also appear.)

It is natural to seek a solution of Equation 41 as a similar series

$$x_t = \sum_{r=0}^{N} a_r \cos rt\mu \quad t = 0, 1, 2 \ldots. \tag{42.}$$

The coefficient a_r can be regarded as the "response" to the "drive" c_r. It is given directly by

$$a_r = \frac{c_r/2}{\cos r\mu - \cos \mu} \quad r = 0, 1, \ldots, N \tag{43.}$$

as can be checked using standard trigonometric formulae. Of these equations, the one with $r = 1$ clearly requires special treatment as the denomi-

nator vanishes identically. This difficulty is related to the problem of "secular" terms in ordinary oscillators, and the remedy, formulated in that case by Linstedt (18a), is to adjust the fundamental frequency to make c_1 vanish. From Equation 40 this yields a formula from which the perturbed tune can be extracted

$$C = C_0 - (3/8)\beta S_0 a^2 b'_3. \qquad 44.$$

Having obtained the coefficients a_r, Equation 42 can be regarded as an improvement upon Equation 37. It can be used to obtain an improved Fourier expansion of the perturbation (using FFTs as already stated).

An iterative loop has been described which, barring unforeseen complications, will "converge" to the exact solution. The scheme works for any perturbation, and it is straightforward to incorporate vertical and longitudinal oscillations. In region I the scheme proves numerically to be very accurate. Since this is the region actually populated by most particles in an accelerator, the method is very practical. However, much of modern analysis amounts to studying failure of this convergence.

Returning to our special perturbation Equation 35 and referring to Equation 44, we can see that the quadratic term yields no frequency shift in the lowest approximation. This is the untypical feature of the Hénon map mentioned previously. In practice, and speaking loosely, this makes the quadratic term especially damaging since "detuning" of the frequency with increasing amplitude tends to have a "stabilizing" influence, due to loss of synchronism. It can also be said that the absence of this term occurs because the decrease in frequency for positive x, where the "restoring force" is weaker, is compensated by the increase for negative x, where it is stronger. Proceeding for one more iteration an improved formula for the tune is obtained,

$$C = C_0 - (3/8)\beta S_0 a^2 b'_3 + \frac{(\beta S_0 b'_2)^2}{8}\left(\frac{2}{1-C} + \frac{1}{\cos 2\mu - C}\right). \qquad 45.$$

4. RESONANCE

4.1 *Conditions for Resonance*

Of the above-mentioned "unforeseen complications" the most important relates to the vanishing of the denominator in Equation 43. For our example, corresponding denominators appear in Equation 45. In Figure 3 it can be seen that the stable region vanishes for $Q = 0$ and $Q = 1/3$, which relate respectively to the vanishing of the two denominator factors of Equation 45. These are the so-called integer and 1/3-integer resonances. Since the former resonance is already present in the linear theory, it is the latter that is characteristic of nonlinearity, and in particular, sextupoles. This resonance is dominant in Figure 2*a*. We return to it below.

To aid in contemplation of the possible vanishing of the denominator expression in Equation 43, it is rewritten as

$$\cos r\mu - \cos \mu = -2 \sin[(r+1)\mu/2] \sin[(r-1)\mu/2], \qquad 46.$$

which vanishes if

$$p\pi = (r \pm 1)\mu/2 \equiv n\mu/2, \qquad 47.$$

where p is any integer and the integer $r \pm 1$ has been renamed n; in other words if

$$Q = \mu/(2\pi) = p/n. \qquad 48.$$

Since p and n can take on any integer values there can be a resonance arbitrarily close to any tune; this makes the convergence of our series at least questionable. According to the theory of Kolmogorov, Arnold & Moser (KAM) (15, 16, 19), convergence is assured for some, but not all, sufficiently small amplitudes. In other words it is guaranteed that region I is not vanishingly small. In this region most points lie on regular "KAM curves." For small amplitudes the truncated series in Equation 42 gives an accurate description of the motion—with Equation 34 it can be regarded as the equation of the KAM curve. This is typical behavior in region I, but even in this region there are initial conditions that lead to chaotic motion not describable by Equation 42. Such solutions can be inferred to be "unimportant" or "improbable" in region I because, with the granularity of Hénon's study, none were detected.

For sufficiently large perturbation strength, Equation 42 is not valid and the motion is chaotic or divergent. This is Hénon's region III. It is unacceptable for accelerator operation because the particles go to sufficiently large amplitude that they are lost. Of some interest is the problem of finding the largest regular KAM curve, as that defines the "dynamic aperture" inside which accelerator operation is at least potentially possible. The outermost smooth curves in Figure 2 are approximately such curves.

In region II the tune locks onto a resonance value that causes the island structure. (By tune here we mean average tune; after a sufficient number of turns the state vector has, on average, made p revolutions in phase space every n steps.) In this region the qualitative motion of two particles of the same initial amplitude but different initial phases can be different, as is discussed below in this section and in Section 6.

4.2 *Resonant Extraction*

To understand resonant extraction (20–24) one need only look at Figure 2*a*. Since the tune is close to 1/3, the final denominator factor in Equation 45 is small, as already discussed. By adjusting the tune, the accelerator

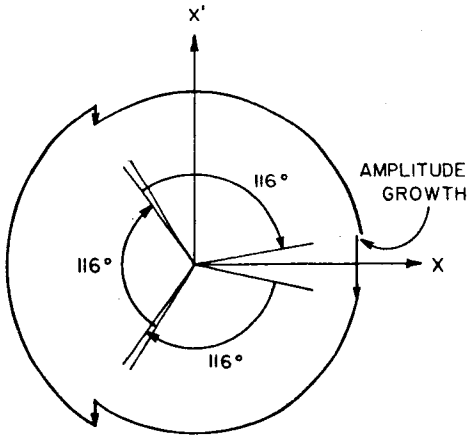

Figure 4 Phase space plot showing the tune being pulled onto the 1/3-integer resonance.

operator can make the stable region I as small as desired. Particles outside this region proceed to large amplitude. Sketched on the figure is a septum magnet needed to complete a controlled extraction of the beam. Any particle jumping over the septum finds itself in a new environment, i.e. an extraction beam line. Note that the particles do not proceed isotropically to large amplitude in phase space—that would violate Liouville's theorem, and would not give an acceptable external beam. They proceed close to a one-dimensional curve called a separatrix. In Figure 2*a* there are three such separatrices, and each particle jumps from one to the next on successive turns. Figure 4 indicates the progress of the state vector for three successive turns. Between impulses the state vector travels on a circle. This will clearly be valid if $K(s)$ in Equation 15 is constant, and can in any event be achieved by judicious transformation (2). The kicks pull the tune "exactly" onto 1/3, and their nonlinear nature causes the length of the state vector, invariant in the linear approximation, to vary.

4.3 *Expansion About Fixed Points*

Having seen that operation close to a resonance may be desirable and/or unavoidable, it is worthwhile setting up a prescription, a variant of "secular perturbation theory," capable of treating that case. We analyze behavior near the 1/3 integer resonance, as in Figure 2*a*. The points labelled I_{31}, I_{32}, and I_{33} are fixed points of the cube of the map Equation 31, for which the tune $Q = \mu/(2\pi)$ has been shifted by the perturbation to be exactly 1/3. For $Q_0 = 1/3$ these points would coalesce into a single point, labelled I_1, which is also the origin. We introduce a small frequency deviation $\Delta\mu_3$

defined by

$$3\Delta\mu_3 = 3\mu_0 - 2\pi. \tag{49.}$$

The linear part of the cubed map is given by

$$\begin{pmatrix} \cos 3\mu_0 & \sin 3\mu_0 \\ -\sin 3\mu_0 & \cos 3\mu_0 \end{pmatrix} \approx \begin{pmatrix} 1 & 3\Delta\mu_3 \\ -3\Delta\mu_3 & 1 \end{pmatrix}. \tag{50.}$$

Dropping all terms beyond quadratic in x and x' and beyond linear in $\Delta\mu_3$ and b_2' the cubed map becomes

$$\begin{aligned} x_3 &\approx x + 3\Delta\mu_3 x' + b_2' \sum_{k=0}^{2} -S_k(C_k x + S_k x')^2 \\ x_3' &\approx -3\Delta\mu_3 x + x' + b_2' \sum_{k=0}^{2} C_k(C_k x + S_k x')^2, \end{aligned} \tag{51.}$$

where

$$S_k = \sin k2\pi/3 = -\sin(3-k)2\pi/3 \tag{52.}$$

$$C_k = \cos k2\pi/3 = \cos(3-k)2\pi/3. \tag{53.}$$

Each term in the summation represents linear propagation for k turns (to get the amplitude needed for substitution into Equation 35 to calculate the kth deflection) and then propagation of the resulting sine-like trajectory for the remaining $3-k$ turns. In this approximation it is easy to incorporate any number of sextupoles, label A, of strength $b_2'(A)$, at arbitrary locations in the accelerator, specified by the betatron phase $\psi(A)$. To do this (17, 18) the phases in Equations 52 and 53 would be shifted by $\psi(A)$ and Equation 51 would also be summed over A.

4.4 *Introduction of the Resonant Invariant*

After three turns a starting state vector x, x' will be shifted by the small quantities δx, $\delta x'$ given by

$$\delta x = x_3 - x; \quad \delta x' = x_3' - x'. \tag{54.}$$

Equation 51 can be expressed in "discrete Hamiltonian" form

$$\delta x = 3\,\partial H/\partial x' \tag{55.}$$

$$\delta x' = -3\,\partial H/\partial x, \tag{56.}$$

where the function $H(x, x')$, also called the resonant invariant, is given by

$$3H(x, x') = (3\Delta\mu_3/2)(x^2 + x'^2) - (b_2'/3)\sum(C_k x + S_k x')^3. \tag{57.}$$

In this formalism the difference operator δ takes the place of d/dt, since time is replaced by the discrete time index t. Factors of three have been inserted artificially into Equations 55–57, corresponding to letting the unit of time be the revolution time. Apart from certain theoretical virtues of a Hamiltonian treatment (see below), the main benefit is that $H(x, x')$ is a constant of the motion. This follows from

$$\delta H \approx (\partial H/\partial x)\delta x + (\partial H/\partial x')\delta x' = 0. \quad 58.$$

In the regular mechanics of a particle having coordinate q and momentum p, the equation of energy conservation, $H(q,p) = E$, gives a relation between q and p, independent of time. In the same way,

$$H(x, x') = H_0 = \text{constant} \quad 59.$$

gives a relation between x and x' independent of t, which in our example, performing the sum in Equation 57, becomes

$$(\Delta\mu_3/2)(x^2 + x'^2) - (b'_2/12)(x^3 - 3xx'^2) = H_0. \quad 60.$$

This resonant invariant can be regarded as the equation satisfied by the various curves in Figure 2*a*, both stable (drawn continuous) and unstable (dotted). The three fixed points I_{31}, I_{32}, and I_{33}, are found by solving the equations

$$\partial H/\partial x = \partial H/\partial x' = 0. \quad 61.$$

An especially instructive procedure is to factorize Equation 60 according to

$$-12H_0/b'_2 + (1/2)(4\Delta\mu_3/b'_2)^3$$
$$= (x + \sqrt{3}x' - 4\Delta\mu_3/b'_2)(x - \sqrt{3}x' - 4\Delta\mu_3/b'_2)(x + 2\Delta\mu_3/b'_2). \quad 62.$$

Each factor on the right-hand side, when set to zero, gives the equation of a straight line. Their intersections are at the fixed points. Points inside this triangle are stable; points outside are unstable. This is obvious from Hénon's numerical analysis and it can be demonstrated using the equations in this section. For example, by expanding about, say, I_{31}, it can be seen to be an unstable fixed point since Equation 60 becomes hyperbolic there. Note that I_1 is stable since Equation 60 is elliptic (actually circular) there.

The extensions of the sides of the triangle are the separatrices along which the particles leave. The importance of having left out higher order terms in Equation 51 can be judged by the curvature exhibited by the actual exiting particles in Figure 2*a*. Also the stable fixed points I'_{31}, I'_{32}, and I'_{33}, present numerically, are missing in this analytic approximation.

4.5 *Islands and the "Standard Map"*

The iterated map of Figure 2*c* exhibits a well-defined chain of five islands. The reason for the number 5 is presumably that the tune Q_0 is close to 1/5. Since such island chains play an important role in accelerator theories (25–28), it is worth studying them for our simple map. Consider, for example, the 4-island structure of Figure 2*b* for which Q_0 is close to 1/4. Suppose that, instead of the Hénon map ($b'_2 \neq 0$, $b'_3 = 0$, in Equation 35), one were studying the case $b'_2 = 0$, $b'_3 \neq 0$. In Equations 40 and 41 the dominant coefficient in lowest order would have $r = 3$, leading to resonance near $Q_0 = 1/(3+1) = 1/4$. (In accelerator jargon, octupoles cause 1/4-integer resonances.) But Hénon's numerical results leave no doubt that a single sextupole can also cause 4 islands (also 5, 6, 7, or, one supposes, any number of islands) depending on the value of Q_0. In the iterative scheme of Section 3.2 such resonances would appear in higher than first order, with strength proportional to quadratic or higher powers of b'_2. If both b'_2 and b'_3 were nonvanishing it would be unclear, without quantitative analysis, whether the islands were due to b'_3 in lowest order, or b'_2 in higher order, or both. For general nonlinearities the ambiguity is worse, since any term in Equation 6 can be responsible for any island chain.

A related comment can be made about a truncated series such as Equation 51. As stated previously, even the presence of many sextupoles A can be handled by extending the summation over them, provided that only terms linear in $b'_2(A)$ are to be retained. The resulting resonant invariant, like Equation 60, will be a cubic function of x and x', subject to factorization as in Equation 62. Such a form is clearly incapable of describing more than three islands. In short, truncated formulae cannot be expected to reproduce the island structure.

The island-chain feature of Figure 2*c*, with stable and unstable fixed points alternating at equal intervals on an approximate circle, is typical of many maps. Were the circle to be stretched out in a straight line, the phase space structure would resemble that of the simple pendulum. (A 6-island example is shown in Figure 7*a*.) Exploiting this analogy, such circular island chains can be described by the so-called standard mapping (19). Here we sketch heuristically how such a mapping can be obtained using the methods of Sections 4.3 and 4.4, and later, in Section 6, apply it to the beam-beam problem.

Instead of coordinates x, x' the state vector can be specified by action-angle variables J and ϕ; they are related by

$$x = \sqrt{2J}\cos(\phi + \phi_0) \quad 63.$$

$$x' = -\sqrt{2J}\sin(\phi + \phi_0). \quad 64.$$

The variables $\sqrt{2J}, \phi$ will be polar coordinates in the phase space of the 5-times iterated map, since we will concentrate on the case of Figure 2*c*. For the Hamiltonian representation (Equations 55 and 56) of the difference equations (Equation 54) the usual rules for manipulating Hamiltonians are valid. We must then, in Equation 57, substitute for x and x' using Equations 63 and 64. At the same time we must retain terms quadratic in b_2' if we hope to have 5 islands, and the summation must run from 0 to 5.

The purpose in studying the 5-times iterated map is to "freeze" the motion approximately. An alternative and more common approach is to describe the motion using a rotation frame of reference in phase space. To accomplish this we introduce a "slow phase" ψ, defined by

$$\psi = \phi - 2\pi t\, 1/5 (\equiv \phi - 2\pi t\, p/n). \tag{65.}$$

We analyze the one-turn map from here on, corresponding to t being an integer and increasing by 1. To the extent that Figure 2*c* has 5-fold circular symmetry it will look as before, but now a particle stays always on the same island.

Analogously to Equation 49, we introduce the frequency deviation from resonance

$$\Delta\mu = \mu_0 - 2\pi\, 1/5 (\equiv \mu_0 - 2\pi\, p/n), \tag{66.}$$

or, expressed as a tune,

$$\Delta Q = Q_0 - 1/5 (\equiv Q_0 - p/n). \tag{67.}$$

To give the pattern observed in Figure 2*c* the resonant invariant must take the approximate form

$$H(\psi, J)/(2\pi) = \Delta Q J + \xi V_n(J) \cos n\psi + \xi U(J). \tag{68.}$$

Here, for later use, we have generalized the notation to a resonance with n islands, but for now n is equal to 5. The parameter ξ, to be defined later, can for now be regarded as a small expansion parameter. The first term can be derived as in Equation 60. The functions $V_n(J)$ and $U(J)$ are called in Section 6 the resonance strength function and the amplitude detuning function respectively. We could also have expected to have terms with periodic angular dependence other than 5ψ. In fact the island chain of Figure 2*c* is not quite circular, which would require the presence of such terms, but we ignore that complication. In idealizations of more complicated systems, a Hamiltonian of this form sometimes is declared to be the result of averaging over "fast" variables, whose frequencies are not commensurate with the resonance frequency. The locations of the fixed points of the mapping are determined by solving the equations

$$\partial H/\partial J|_{J_n} = \partial H/\partial\psi|_{J_n} = 0, \tag{69.}$$

where the action coordinate of these fixed points is J_n. Instead of J we now use ΔJ, the deviation from J_n,

$$\Delta J = J - J_n, \tag{70.}$$

as the action variable. Using Equation 69 and dropping constant terms, the Hamiltonian becomes

$$H(\psi, \Delta J)/(2\pi) = \xi V_n(J_n) \cos n\psi + \xi U''(J_n) \Delta J^2/2 \tag{71.}$$

Hamilton's equations are then given by

$$\delta\Delta J = -\partial H/\partial\psi = 2\pi n \xi V_n(J_n) \sin n\psi \tag{72.}$$

$$\delta\psi = \partial H/\partial\Delta J = 2\pi\xi U''(J_n)\Delta J. \tag{73.}$$

Except for constant factors, this is known as the "standard mapping." The motion in phase space is the same as that of a pendulum, even for large amplitudes. "Newton's second law" is obtained by "differentiating" Equation 73 and substituting from Equation 72

$$\delta^2\psi = 2\pi\xi U''(J_n)\delta\Delta J = (2\pi)^2 n \xi^2 V_n(J_n) U''(J_n) \sin n\psi. \tag{74.}$$

The "frequency" of small oscillations is $\xi |U''(J_n) V_n(J_n)|^{1/2}$. The half-width of the island-chain, ΔJ_n, can be obtained from the separatrix trajectory; this is analogous to the pendulum motion, which just stops short of going past the position of unstable equilibrium. At that point, where the "energy" is entirely "potential," $H(\psi, \Delta J)$ can be evaluated; it yields $\xi V_n(J_n)$ as the "energy." The maximum "momentum" occurs at the position of stable equilibrium, $\psi = 0$, where the "potential energy" has been reversed and yields

$$\Delta J_n(\text{half-width}) = 2|V_n(J_n)/U''(J_n)|^{1/2} \tag{75.}$$

as the island half-width. This can also be used to define a kind of "range of applicability" of the Hamiltonian, Equation 68, running from $J_n - \Delta J_n$ to $J_n + \Delta J_n$.

5. SPECIALIZED TOPICS

5.1 *Nearly Linear Description: Distortion Functions*

We have seen, in Section 2.2, that judicious transformation of the variables can force the formula describing betatron motion to be identical to that for simple harmonic motion. The transformations required were changing the independent variable (Equation 22) and factoring out the s-dependent modulating factor $\sqrt{\beta(s)}$. What is left in Equation 18 are the arbitrary, but constant, phase $\psi(s_1)$ and amplitude a. Note that the meaning of ψ in this section is not the same as in the last section. The idea of distortion

functions is to account approximately for extra perturbing effects by allowing a and $\psi(s_1)$ to be distorted by the addition of small s-dependent terms $\delta a(s)$ and $\delta\psi(s)$. To first order in the strength of the nonlinear perturbation, then, the motion is described by

$$x(s) = [a+\delta a(s)]\sqrt{\beta(s)}\cos[\psi(s)-\psi(s_1)+\delta\psi(s)]. \qquad 76.$$

Much in the way the differential equation (Equation 20) satisfied by $\beta(s)$ was found, differential equations can be found for $\delta a(s)$ and $\delta\psi(s)$ that permit $x(s)$, as given by Equation 76, to satisfy Equation 15 with extra perturbing terms present. To be useful, $\delta a(s)$ and $\delta\psi(s)$ are assumed, like $\beta(s)$, to be periodic, as in Equation 26, and they must be expressible in terms of functions independent of a. Functions having all these properties are said to be lattice functions.

Before specifying these distortion functions more explicitly, we make a short historical digression. Transformations like these are by no means original in accelerator physics. In celestial mechanics the method is known as "variation of the elements" where, for planetary orbits, the orbit elements are constants of the unperturbed motion such as major axis a, orientation of the ellipse axis ψ, and four others. (The symbols have been intentionally chosen to be ambiguous so that the formulae will apply to our actual problem.) A perturbing potential $R[x(a,\psi),z,\ldots]$ causes motion that can be described by allowing $a,\psi,\ldots$ to vary with time according to the differential equations

$$\{a,a\}\dot{a}+\{a,\psi\}\dot{\psi}+\ldots = \partial R/\partial a$$
$$\{\psi,a\}\dot{a}+\{\psi,\psi\}\dot{\psi}+\ldots = \partial R/\partial\psi \qquad 77.$$
etc.

The coefficients are Lagrange brackets, which can be shown to be constants of the unperturbed motion. As a result, the equations can be written by matrix inversion in the more convenient form

$$\dot{a} = [a,a]\langle\partial R/\partial a\rangle+[a,\psi]\langle\partial R/\partial\psi\rangle\ldots$$
$$\dot{\psi} = [\psi,a]\langle\partial R/\partial a\rangle+[\psi,\psi]\langle\partial R/\partial\psi\rangle\ldots \qquad 78.$$
etc.

These coefficients are the Poisson brackets, which are an important ingredient of the Lie algebraic methods discussed briefly below. In writing Equation 78, an extra step was performed for the sake of brevity. The right-hand sides have been averaged over one complete period of the unperturbed motion. Such averaging (characteristic of the Krylov-Bogoliubov method discussed below, and other methods) along with the invariance of the Poisson brackets, causes the right-hand sides of

Equation 78 to be independent of $x,z,\ldots$. This permits $a,\psi,\ldots$ to be found by direct integration. These equations are called the Lagrange planetary equations (they are 200 years old.)

In the accelerator physics context, distortion functions have been derived independently by various authors. They were used in the design of the storage ring CESR (17), derived using difference equations, and later (29), when the term distortion function was coined, derived by ad hoc methods. They have been derived using Hamiltonian methods (30) and Lie algebra (31), and the equivalence of different representations has been demonstrated (32).

The idea of distortion functions can be illustrated using the Hénon map and formulae of Section 3. To lowest order in b_2' the motion is given by

$$x_t = a_0+a\cos t\mu+a_2\cos 2t\mu, \quad t=0,1,2,\ldots, \qquad 79.$$

where the procedure for finding a_0 and a_2 has already been given. They are proportional to a^2. The slope is given by substitution into Equation 34. The quantity $(x_t^2+x_t'^2)^{1/2}$ is an invariant of the unperturbed motion as well as being the distance from the origin in the phase space plots of Figure 2. Because of the perturbation, it is not invariant and the phase space trajectory is not circular. The perturbed invariant is a, the coefficient of the term varying as $\cos t\mu$. This can be obtained by Fourier projection, or if one insists on generating a circular phase space plot, by subtracting the known perturbing terms from x_t and x_t' before plotting them. This will yield a circle for modest amplitudes, although the description will begin to deteriorate as higher harmonics become important.

The presence of many sextupoles can be accounted for as described after Equation 53. For the most general distribution of sextupoles the coefficients a_0 and a_2 can be written in terms of four quantities

$$(A_1,B_1) = \sum_A b_2'(A)[\beta(A)]^{3/2}[\sin\psi(A),\cos\psi(A)] \qquad 80.$$

$$(A_3,B_3) = \sum_A b_2'(A)[\beta(A)]^{3/2}[\sin 3\psi(A),\cos 3\psi(A)]. \qquad 81.$$

These are the desired distortion functions. They form pairs that behave like "phasors." In regions with no sextupoles, the phasor (A_1,B_1) rotates, with the phase advancing at the same rate as betatron oscillations, while (A_3,B_3) rotates at three times that rate. For coupled motion other distortion functions such as

$$\sum_A b_2'(A)\sqrt{\beta_x(A)}\beta_z(A)\begin{matrix}\sin\\ \cos\end{matrix}[2\psi_z(A)\pm\psi_x(A)] \qquad 82.$$

are needed for a similar description.

With enough correction sextupoles, it is possible to adjust all these

distortion functions to zero. When that is done, then, according to Equation 79 and to lowest order only, the phase space trajectory becomes a perfect circle. As explained at the end of Section 3.2, it is necessary to go to the next order (quadratic in b_2') to obtain the tune shift caused by sextupoles. By straightforward generalization of Equation 45 the tune shift can be written in terms of the distortion functions.

5.2 *Hamiltonian Methods, Stochasticity, and Long-Term Stability*

It is typical for a high energy particle to circulate for many hours in a storage ring, performing tens of betatron oscillations every turn, with each turn taking some microseconds. That makes some 10^{11} oscillations. It is a formidable challenge to make any statement that remains valid after that many cycles. Nor are such questions of purely academic interest, as the difference between a ten-minute and a one-hour lifetime can spell the difference between success and failure. The only formalism with sufficient mathematical discipline to have any reasonable expectation of succeeding is the Hamiltonian (or, as one says, the canonical) approach.

A common theme in this research has been to simplify an otherwise intractable problem and then to re-express it in Hamiltonian form, for which strong statements about stochasticity and long-term stability can be made. This amounts to redefining (some might say emasculating) rather than "solving" the original problem, as far as mathematical rigor is concerned, but it constitutes reasonable "physics." An example of this was given in Section 4.4 and others are mentioned below. Reference (33) contains much material on long-term stability.

Hamiltonian methods have also been useful in analyzing more mundane situations in which description of the behavior over a more modest number of cycles, perhaps a hundred or a thousand, is required (34–36).

Hamiltonian perturbation techniques proceed by finding canonical transformations that make the Hamiltonian approximately time independent. This follows a tradition of Hamilton-Jacobi, Poincaré, and Birkhoff. The most essential difference between these methods and those emphasized in this paper is the difference between series expansion and iteration. In series methods, functions are expanded in Fourier series based on the unperturbed period; in iteration, such series are based on a perturbed period.

5.3 *Particle Tracking and Symplectification*

When one attempts to "improve" a transfer map like Equation 31 to account for the effect of finite magnet thickness, it is not automatic for symplecticity to be preserved.

(Symplecticity is a concrete mathematical attribute of a map that is present if and only if the theory is canonical. A necessary condition is that the determinant of the linearized part of the map be 1. For example, this is satisfied by Equation 31 but not by Equation 51.)

There is a standard "TRANSPORT" formalism (37) for representing a section of an accelerator by a quadratic map like Equation 51. The whole lattice can be described by concatenating such maps. Though the resulting map may be quite accurate for a few turns, the long-term behavior will be ruined because the map is not symplectic. One popular, if suspect, procedure is to allow blemishes to develop while making approximations, but then to "paint over them" by resymplectifying. This is easy for the linear part of the map (37a), but a quadratic map will inevitably be nonsymplectic. Apart from going to higher order, which only delays and compounds the problem, the only solution is to find an implicitly defined map that agrees with the quadratic map to quadratic order (38, 39). To track a particle it is then necessary to invert numerically the implicit function relating the state vector before and after each turn.

Instead of using approximate formulae to perform tracking through "exact" (i.e. thick) elements, it is possible to perform exact tracking through approximate (i.e. thin) elements. The approximation can be improved by breaking thick elements into several thin elements. Long-term precision in such a program is only compromised by round-off error in the computer. Being exact, such a calculation is necessarily globally symplectic, though it is not locally symplectic (unless extended to six dimensions). A program called TEAPOT uses this method (40).

5.4 *Lie Algebraic Methods*

Several of the features mentioned are most elegantly handled by Lie algebraic methods. These include (*a*) concatenation of nonlinear maps, (*b*) systematic treatment of higher order perturbation theory, (*c*) preservation of symplecticity after truncation, and (*d*) distortion functions. These methods are applied in celestial mechanics (41), and they have been developed for accelerators by Dragt (42) and his coworkers (38, 39) in the computer program MARYLIE. Here we give only the general idea.

Hamilton's equations can be written in Poisson bracket form

$$\mathrm{d}\mathbf{a}/\mathrm{d}t = -[H, \mathbf{a}] \equiv -:H:\mathbf{a} \qquad 83.$$

where, to simplify the formulae, the Hamiltonian will be assumed to have no explicit time dependence. The state vector $\mathbf{a}$ is made up of coordinates and momenta that are invariants of the unperturbed motion. The close analogy to Equation 78 should be noted. The symbol :H: is just a notation expressing the Poisson bracket relation as an operator, as defined by Equation 83.

The evolution of the state vector from its initial value $\mathbf{a}(t_0)$ to its value

at time t is given by

$$\mathbf{a}(t) = \left(1 - \int_{t_0<t} :\mathrm{H}:\mathrm{d}t' + \int_{t_0<t'<t} :\mathrm{H}:\mathrm{d}t' \int_{t_0<t''<t'} :\mathrm{H}:\mathrm{d}t'' - \ldots\right)\mathbf{a}(t_0) \quad 84.$$

as can be checked by substitution into Equation 83. This time evolution can also be represented by a transfer map

$$\mathbf{a}(t) = M(t, t_0)\mathbf{a}(t_0). \quad 85.$$

Taken together, and with proper attention to noncommuting operators, these yield

$$M(t, t_0) = \exp\left(-\int_{t_0}^{t} :H:\mathrm{d}t'\right). \quad 86.$$

This is purely formal since the operator $:H:$ appearing in the various integrands depends implicitly on t through its dependence on the coordinates and momenta. However, these formulae lend themselves to solution using a series of powers of the time.

The map in Equation 86 is a Lie transformation. That is to say, it has the form

$$M = \exp(:f:), \quad 87.$$

where $f(\mathbf{a})$ is any function. The key theorem says that any such map is symplectic. To exploit this, a sophisticated calculus for approximating transfer maps in the form of Equation 87 has been developed (42). The four features listed above follow naturally. In particular, symplecticity is preserved in succeeding orders of approximation.

When the Lie method is applied to a simple system like a gravity pendulum (8), the results are disappointing. The description of small amplitudes is excellent but, for a given precision, the slow increase in range of applicability in proceeding from order to order does not seem to justify the effort. An extreme position on high-order perturbation methods was stated by Taf (43): "Beyond first-order results I know of no useful result from perturbation theory in celestial mechanics . . . ." These contraindications probably do not apply to accelerator applications in the nearly linear regime. They certainly do not apply to the common circumstance, e.g. Equation 45, where the frequency is unshifted in lowest order. Nevertheless, it seems that the main virtue of the Lie algebraic method is its systematic handling of the distortion functions.

5.5 *Superconvergence*

There is a procedure (44, 45) by which the slow convergence, lamented in the previous section, can be speeded up. This is of considerable theoretical importance and might be said to be the main ingredient of the so-called KAM theory. So far, however, such methods appear to have had little practical impact, perhaps because of their complexity.

5.6 *Krylov-Bogoliubov Methods*

Some physicists find the canonical methods too formal and poorly motivated. The averaging methods of Krylov and Bogoliubov (46) form a powerful and, some would say, a more physical alternative.

6. THE BEAM-BEAM PROBLEM

6.1 *Introduction*

In the simplest type of colliding storage ring, each of two counter-rotating beams circulating in a purely magnetic guide field contains a single bunch, with a population of order 10^{10} particles, which remains virtually intact, one hopes, for several hours. The bunch in one beam contains electrons or protons, while the other beam contains positrons or antiprotons. In the absence of electric fields the particle and antiparticle bunches follow the same trajectories in opposite directions, so that they pass through each other at two diametrically opposed points, where experimental high energy physics detectors are located. On comparatively rare occasions hard particle-antiparticle collisions are observed in the detectors, but for the most part the tenuous beams pass right through each other. In all the "weak-strong" models of the beam-beam interaction described below, a single test particle in one bunch collides softly, only once per turn, with the macroscopic electromagnetic fields caused by the other bunch. We first consider betatron motion restricted to a single transverse dimension, the horizontal. In Section 6.5, simultaneous vertical motion is also included.

Consider a round incident bunch, Gaussian in all coordinates, with an rms transverse size, $\sigma = \sigma_x = \sigma_z$, much smaller than its length. Its charge density distribution is given by

$$\rho(r) = (ne/2\pi\sigma^2)\exp(-r^2/2\sigma^2), \quad 88.$$

where $r^2 = x^2 + z^2$, and $n(s)$ is the number line-density along the bunch. The total beam-beam impulse imparted during one passage through the incident bunch to a test particle of charge $-e$ displaced by x is

$$\Delta x' = -(4\pi\xi/\beta^*)(2\sigma^2/x)[1 - \exp(-x^2/2\sigma^2)], \quad 89.$$

where β^*, the beta function at the crossing point, is taken to be much larger than the bunch length. In the highly relativistic limit that is assumed, the magnetic and electrostatic beam-beam impulses are equal in magnitude and add constructively. The strength of the total impulse has been parameterized by ξ, the "beam-beam tune shift parameter," which (generalizing

to elliptical beams) is given by

$$\xi_{x,z} = \frac{Nr_c}{(2\pi\gamma)} \frac{\beta^*_{x,z}}{\sigma_{x,z}(\sigma_x+\sigma_z)}. \qquad 90.$$

where N is the total bunch population, r_c is the classical radius of the particle, and γ is the usual relativistic factor.

For small amplitude particles, Equation 89 reduces to

$$\Delta x' = -(4\pi\xi/\beta^*)x; \qquad x \ll \sigma. \qquad 91.$$

This is a linear perturbation, acting just like a thin focusing quadrupole. Substituting Equation 91 into 33 and rearranging terms, it can be seen that the effect of such a quadrupole is to shift the tune from Q_0 to Q; their relation being

$$\cos(2\pi Q) = \cos(2\pi Q_0) - 2\pi\xi \sin(2\pi Q_0). \qquad 92.$$

This confirms that the tune shift for small amplitude particles is in fact ξ, and that they will be stable. Large amplitude particles, with $x \gg \sigma$ for most of their interactions, experience impulses that drop off like $1/x$, so that they also tend to be stable, with tune shifts approaching zero. This is in contrast with the nonlinearities discussed above, from magnets, where the tune shift increases with amplitude, and amplitudes above some dynamic aperture are lost. These differences occur because the magnetic fields are caused by current sources external to the vacuum chamber, while the beam-beam interaction sources are localized within the incident bunch. It is therefore not surprising that external nonlinear correction elements are of little use in stabilizing against the beam-beam effect.

Betatron oscillations can be described using variables J, ϕ as in Equations 63 and 64 or in terms of a normalized amplitude α:

$$\begin{aligned} x &= (2J)^{1/2}\sigma\cos\phi \equiv \alpha\sigma\cos\phi \\ x' &= -(2J)^{1/2}(\sigma/\beta^*)\sin\phi \equiv -\alpha\sigma/\beta^*\sin\phi. \end{aligned} \qquad 93.$$

It can be shown (27), that the perturbed tune of a nonresonant particle is given by

$$Q(J) = Q_0 + \xi U'(J) \qquad 94.$$

where

$$U'(J) = (2/J)[1-\exp(-J/2)I_0(J/2)]. \qquad 95.$$

Here I_0 is a modified Bessel function. While Equation 95 is written in terms of J, it is natural to plot U' in terms of α, the normalized amplitude; Figure 5*a* is such a plot.

Electron colliders are observed to have a beam-beam tune shift limit ξ_{max} roughly in the range 0.02 to 0.05 per collision (47), depending somewhat on the betatron motion damping time (47a), and depending strongly on the operating point (Q_x, Q_z) in the tune plane (48). The transverse distribution of an unperturbed electron bunch is determined by a dynamic balance between quantum emission of photons in the horizontal plane and radiative damping in both transverse planes, with a characteristic damping time of order 10^3 accelerator turns. This leads to a naturally flat beam, with a relatively small vertical size; $\sigma_z \ll \sigma_x$. However, as the beam currents are increased equally, both beams increase their vertical sizes substantially, which limits the value of ξ, or worse, causes particle losses that reduce the stored beam lifetime of one or both beams. An important practical electron beam-beam problem is to understand quantitatively the resonance features of the tune plane, in order to find the optimum operating point. The most successful solutions to this problem rely heavily on the use of numerical simulations (49–51).

At the time of this writing, the only (bunched) proton beam-beam data available are from the SPS at CERN (27), where, with six collisions of strength $\xi \approx 0.004$, resonances of order ten (i.e. n is 10 in Equation 48) and less must be avoided in the tune plane for acceptable storage lifetimes. One reason proton colliders are more sensitive than electron colliders to higher order resonances is that the larger proton mass essentially stops the proton from radiating. This makes the damping time almost infinite, with the corollary that proton beams tend to be round and the Hamiltonian description of round beam-beam interactions pursued below appropriate. The proton collider operating point is forced close to the tune diagonal,

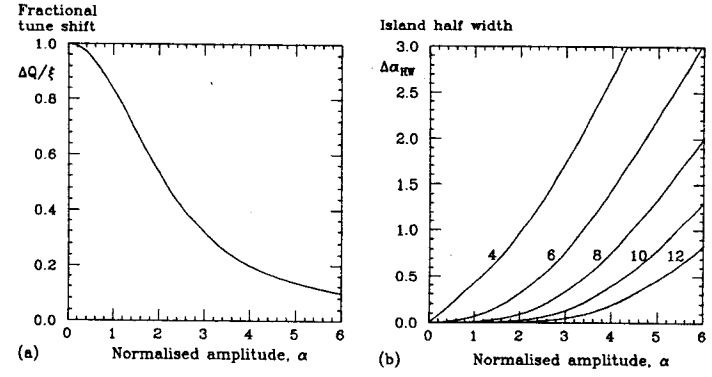

Figure 5 Response of a test particle to a single round beam-beam collision per turn. (*a*) Tune shift, $\Delta Q/\xi = dU/dJ$, as a function of amplitude. (*b*) Resonant island half-widths for various resonance orders as a function of amplitude, as predicted by Equations 75, 95, 97.

$Q_x = Q_z$ (though the integer parts of the tunes need not be equal), where high-order resonances are most sparse. The unperturbed tune is chosen so that the actual tune, shifted and spread according to Equation 94, avoids resonances for values of n less than some critical value n_{critical} for all amplitudes. The impossibility of finding such a resonance-free space to accommodate a total tune spread larger than about $6\xi \approx 0.024$ between tenth and lower resonances makes the SPS beam-beam limited. Thus the most pressing proton beam-beam problem is to find the value of n_{critical} as a function of ξ and the other parameters.

6.2 *The Resonant Hamiltonian and Chaos*

Various workers (25, 25a, 26), following the work of Chirikov (19), showed how to apply the results of Section 4.5 to the beam-beam problem. The Hamiltonian that describes approximately the motion of the test particle can be shown (27) to be

$$H/(2\pi) = (Q_0 - p/n)J + \xi U(J) + \xi V_n(J)\cos n\psi \qquad 96.$$

as already written in Equation 68. The detuning function is now known explicitly through Equation 95. The resonance strength function $V_n(J)$ is given by integrating the equation

$$V_n'(J) = -(-1)^{n/2}(4/J)\exp(-J/2)I_{n/2}(J/2). \qquad 97.$$

Odd resonances are absent by symmetry, though they could be caused by symmetry-breaking effects such as non-head-on bunch collisions.

As discussed in Section 4.5, this Hamiltonian describes motion around a chain of resonance islands, with the island half-width given by Equation 75 and plotted in Figure 5*b*. Note that it is independent of the strength of the beam-beam interaction since both perturbing terms in Equation 96 are proportional to ξ.

According to the Chirikov overlap criterion (19), chaotic motion is expected if two neighboring chains of resonance islands overlap. Technically, chaotic regions of phase space are characterized by neighboring trajectories diverging exponentially, whereas they diverge only linearly in regular regions. As illustrated in Section 3, chaotic and regular trajectories can usually be distinguished visually in a phase space plot, by seeing whether or not they lie on a continuous curve. A more quantitative distinction occurs in the Fourier spectrum (see e.g. Equation 42) of the trajectory; it shows a finite number of peaks for regular motion and a broad continuous spectrum for chaotic motion.

Suppose that the next important island chain has tune Q_+. The tune separation $Q_+ - Q_n$ will depend on the values of p and n for both chains. Using Equation 94 this separation can also be expressed as a separation

in action

$$\Delta J_n(\text{separation}) = (Q_+ - Q_n)/[\xi U''(J_n)] \qquad 98.$$

assuming that the detuning function $U(J)$ is dominant. Combining this with Equation 75, the Chirikov criterion can be expressed as a stability condition on the tune shift parameter ξ; it should be less than

$$\xi_{\max} = (Q_+ - Q_n)[16V_n(J_n)U''(J_n)]^{-1/2}. \qquad 99.$$

Numerical simulations support the validity of the Chirikov criterion in this form, as a predictor of the onset of chaos, but the numerical value of $\xi_{\max}$ is typically greater than 0.1, which makes it more than ten times larger than the limit actually observed in accelerators (25a, 27a). Furthermore, what has been written does not yet really constitute a theory since the factor $Q_+ - Q_n$ has not been calculated. This model will nevertheless form the basis for a more refined theory that removes both defects. In the spirit of resonance overlap, one should look for sources of more closely spaced resonances, and some sort of modulation mechanism is thought to be the most promising candidate (25, 51a). For protons the variation of tune with momentum seems to be important, as described next.

For electrons other sources of modulation are important, though not through resonance overlap. The effect of longitudinal crossing point oscillations is described in Section 6.4. It is especially important to introduce a second transverse coordinate for electron beams because of their ribbon-like shape. The dominant effect is parametric modulation of vertical motion by the horizontal motion (Section 6.5).

6.3 *Tune Modulation*

One source of tune modulation is ripple in the current supplied to some of the guide field magnets. A more fundamental source is the tune variation accompanying energy oscillations that occurs when the chromaticity is not exactly zero, in which case the modulation frequency is the "synchrotron" frequency Q_s, about 0.005 in the SPS. The modulation depth q might typically be 0.001 or larger.

We suppose then, that owing to an external modulating source, the unperturbed betatron tune is given by

$$Q = Q_0 + q\sin 2\pi Q_s t \qquad 100.$$

which causes the Hamiltonian in Equation 96 to be replaced by

$$\bar{H}/2\pi = (Q_0 + q\sin 2\pi Q_s t - p/n)\bar{J} + \xi U(\bar{J}) + \xi V_n(\bar{J})\cos n\bar{\psi}, \qquad 101.$$

where the variables have been given bars in preparation for a canonical

transformation using the generating function

$$W(\psi, \bar{J}, t) = [\psi + (q/Q_s)\cos 2\pi Q_s kt - 2\pi Q_s t/n]\bar{J}. \quad 102.$$

Here k is an integer that will be specified later. The new variables are given by

$$J = \partial W/\partial\psi = \bar{J} \quad 103.$$

$$\bar{\psi} = \partial W/\partial\bar{J} = \psi + (q/Q_s)\cos 2\pi Q_s t - 2\pi Q_s kt/n, \quad 104.$$

and the new Hamiltonian is given by

$$\begin{aligned} H/(2\pi) &= \bar{H}/(2\pi) + (\partial W/\partial t)/(2\pi) \\ &= (Q_0 - p/n - Q_s k/n)J + \xi U(J) \\ &\quad + \xi V_n(J)\cos[n\psi - 2\pi Q_s kt + (nq/Q_s)\cos 2\pi Q_s t]. \end{aligned} \quad 105.$$

The final term can be expanded as a series of synchrobetatron side bands with spacing Q_s/n, as is done for frequency-modulated radio signals; the coefficients are Bessel functions $J_k(qn/Q_s)$. Each of these terms can be treated as an independent resonance provided that the perturbation strength is weak enough that they do not overlap. We analyze them separately, choosing the appropriate value of k for each. From Equation 104 it can be seen that this amounts to viewing each side band from an appropriately rotating frame. We use the resulting Hamiltonian H_k only to estimate the width of the kth side band, and as a result the signs of the coefficients do not matter. For qualitative discussion then, the Bessel function $J_k(K)$ can be approximated by zero when $K < k$ and by $(\pi K)^{-1/2}$ otherwise. We obtain

$$H_k/(2\pi) = (Q_0 - p/n - kQ_s/n)J + \xi U(J) + \xi[Q_s/(\pi qn)]^{1/2} V_n(J)\cos n\psi \quad 106.$$

for

$$k < k_{\max} \equiv qn/Q_s. \quad 107.$$

Note that, as k has been chosen, the Hamiltonian is independent of time and is hence conserved.

This shows that only about $2k_{\max}$ lines in the range $\pm q$ have significant strength, corresponding to the observation that a particle will feel the resonance strongly only if its tune is modulated across the central resonance frequency. With the numbers given above, at least five side bands are expected around a tenth-order resonance.

The widths of these lines can be obtained by applying Equation 75 again. Then the Chirikov criterion can be employed to give a threshold value $\xi_{\max}$ above which chaos is expected (27):

$$\xi_{\max} = (1/4)(\pi q)^{1/4}(Q_s/n)^{3/4}[V_n(J_n)U''(J_n)]^{-1/2}. \quad 108.$$

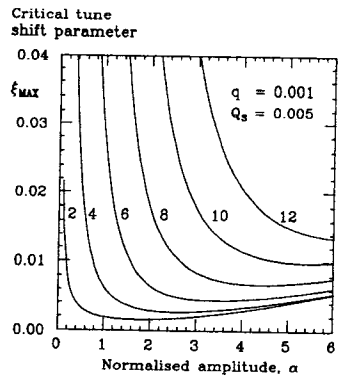

Figure 6 The critical tune shift parameter, $\xi_{\max}$ in Equation 108 for various resonance orders. The resonance overlap criterion predicts that chaos ensues for a single collision per turn above this strength.

Values of $\xi_{\max}$ are plotted in Figure 6 as a function of amplitude α for various orders n. The simulation results shown in Figure 7 confirm a sample quantitative prediction derived from Figure 6 that the onset of chaos near a sixth-order beam-beam resonance should occur for a tune shift parameter $\xi_{\max}$ of about 0.0045. Simulation results (28) tend to support the accuracy of Equation 108, even when more than one collision per turn and other realistic features are properly incorporated.

For the purpose of making various comments, let us now consider the dependence on Q_s, holding all other variables constant. The leading trend is that chaos sets in sooner as Q_s is reduced, because the side-band spacing varies linearly with Q_s, while according to Equations 106 and 75 the side-band width varies as the fourth root. The fact that the critical tune shift varies as $Q_s^{3/4}$ makes this mechanism less important for electrons than for protons. For electrons $Q_s \approx 0.05$, an order of magnitude higher than for protons, a consequence of the high power RF accelerating system needed to replenish energy lost by radiation. There appears to be unphysical limiting behavior as Q_s approaches zero; chaos is always present with no modulation. That this is not contradictory just points up the limitation of a model that predicts the onset of chaos without calculating the resulting diffusion rate in phase space.

Diffusion results when a particle finds itself on the "wrong" side of a "moving separatrix." Rates for this to occur can be estimated (25), but as yet there is not a dynamic theory. As Q_s is reduced, there arise two impediments to massive outward displacement in phase space: the time rate of resonance crossing falls linearly, and there are more "separatrices" to be crossed. For sufficiently low values of Q_s, considerations of adiabaticity are more useful. If a particle goes around a fundamental resonance island

many times in the time it takes for the island to move its own width due to the tune modulation, the particle can be expected to remain trapped and to move with the island. This is the phenomenon of "resonance trapping" (48). Expressing the above condition quantitatively yields a critical frequency Q_{sc} below which resonance trapping occurs;

$$Q_{sc} = n\xi^2 V_n(J_n)U''(J_n)/q. \qquad 109.$$

This is plotted in Figure 8. For example, with the parameters $\xi = 0.006$ and $q = 0.001$ used in Figure 7, the modulation frequency $Q_s = 0.005$ is almost an order of magnitude above the critical value for resonance trapping on sixth-order resonances.

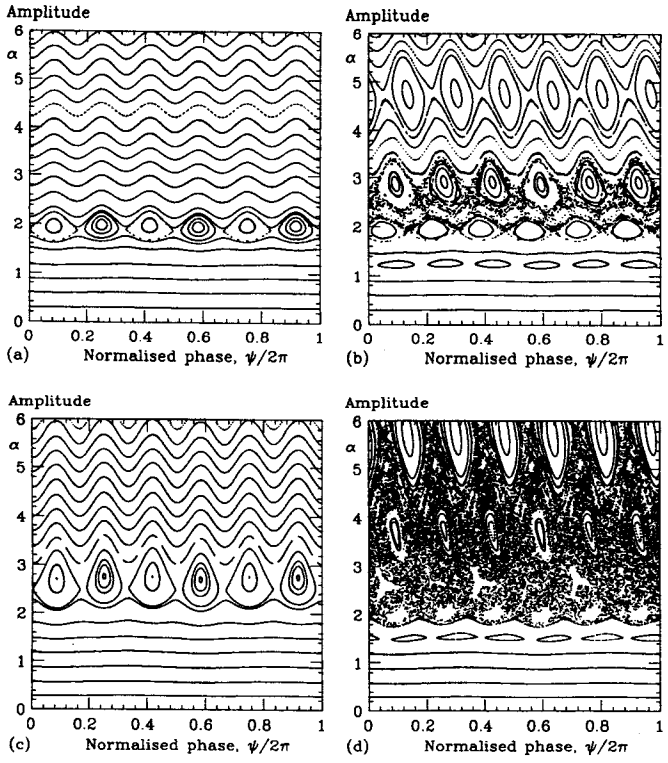

Figure 7 Simulated trajectories tracked for 2000 synchrotron periods, with $Q_S = 0.005$, and an unshifted tune of 0.331, near a sixth-order beam-beam resonance. The two left figures have no tune modulation, while the two right figures have a modulation amplitude $q = 0.001$. The two top figures have a tune shift parameter of $\xi = 0.0042$, slightly under $\xi_{max} \approx 0.0045$, while the two bottom figures have a value $\xi = 0.006$, slightly above ξ_{max}. Side bands $k = +1$, 0, -1, and -2, visible in (*b*) at increasing amplitudes, overlap and are submerged in a chaotic sea in (*d*).

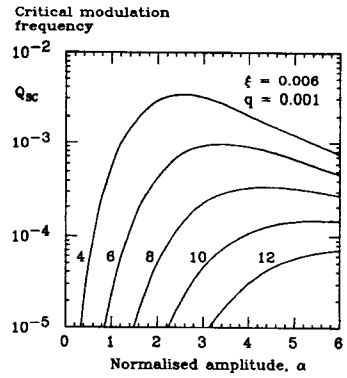

Figure 8 The critical tune modulation frequency, Q_{SC} in Equation 109 below which resonance trapping is expected. The parameters used correspond to those in Figures 6 and 7.

It can be observed in Figure 5*a* that $dQ/d\alpha$ becomes small for large α, a fact that, both for this and the chaotic mechanism, encourage large excursions of the action coordinate of a particle as in Figure 7. One of these excursions may prove to be fatal for the particle.

6.4 *Longitudinal Collision Point Oscillations*

Energy oscillations are also accompanied by longitudinal oscillations of the test particle, $s = a_s \cos(2\pi Q_s t)$, relative to the bunch center; this causes the collision point to oscillate longitudinally with amplitude $a_s/2$. This can equivalently be regarded as an oscillation of the tune of the accelerator; the formula is obtained by differentiating Equation 22:

$$Q = Q_0 + (Q_s a_s/2\beta)\cos 2\pi Q_s t. \qquad 110.$$

Furthermore, the denominator factor β varies with t, since its longitudinal dependence in the vicinity of a minimum of the beta function, such as is normal at a collision point where $\beta = \beta^*$, is

$$\beta = \beta^*[1 + (s/\beta^*)^2]. \qquad 111.$$

This follows from Equation 20. For a low-beta intersection point, typical values of the parameters satisfy

$$a_s/2 \approx \sigma_s \approx \beta^*. \qquad 112.$$

In that case the depth of modulation from this source is $q \approx Q_s$, and because of Equation 111 modulation at all odd harmonics of Q_s is present (52). This is likely to be the dominant source of the tune modulation effects analyzed in the previous section.

For electron accelerators these longitudinal oscillations lead to another modulation mechanism (49, 51a, 52). Accompanying the beta-function dependence of Equation 111 is an effective dependence on s of the vertical tune shift parameter,

$$\xi_z(s) = \xi_z(0)\,[1+(s/\beta_z^*)^2]^{1/2}. \qquad 113.$$

For large s, the tune shift parameter becomes markedly larger. That the affected motion is primarily vertical is due to the previously described asymmetry of the transverse beam dimensions; optimal design has $\beta_z^* \ll \beta_x^*$. Since there is no such effect for round beams, the dominant resonances in electron and proton accelerators are different.

Certain colliders have nonzero crossing angles, which leads to other resonances (51) similar to those mentioned here.

6.5 *Motion in Two Transverse Dimensions*

Especially when the colliding beams are not round, the description must be extended to another transverse dimension. In that case rigorous theoretical results are even more rare. It is known that regular KAM surfaces continue to exist, but topologically they are incapable of isolating regions of good behavior, such as region I described above, and regions of bad behavior, such as region III. All the chaotic trajectories (of a fixed total Hamiltonian "energy") are connected together in an "Arnold web," along which "Arnold diffusion" takes place (19). In principle this permits a particle to embark on grand excursions from small amplitudes to large amplitudes, threatening the long-term stability in an unfortunate way. Fortunately, however, the diffusion rates obtained by calculation and by simulation still appear to be negligibly small for realistic beam-beam parameters (8, 53, 54). For round beams, two-dimensional simulations do not in general demonstrate qualitatively different, or more sensitive, behavior than one-dimensional simulations (28). The possibility exists, however, that dramatic effects remain to be uncovered by simulation, if more fully developed theoretical models suggest the right behavior for which to look.

The vertical blow-up of flat electron beams, which occurs on the comparatively very rapid time scale of the radiation damping time, is well described by the "parametric oscillator" model of Peggs & Talman (50). In this model the horizontal motion is taken as inexorable, acting, in effect, as an external modulation of the parameters of the vertical motion. For horizontal ribbon-shaped beams the vertical deflection is given, instead of Equation 89, by

$$\Delta z' = -(4\pi\xi_z\sigma_z/\beta_z^*)\sqrt{(\pi/2)}\,\mathrm{erf}\,[z/(\sigma_z\sqrt{2})]\exp[-x^2/(2\sigma_x^2)], \qquad 114.$$

where erf is the error function. Note the factorization into the product of two factors, each of which depends only on one coordinate. For modest amplitudes the motion is analogous to that of a pumped pendulum, with a small displacement z, and a small perturbation of its length at a frequency $2Q_x$. Such a parameterically driven linear system is well known to be self-exciting on a family of resonances $Q_z = pQ_x$, corresponding to lines in the tune plane (Q_x, Q_z). Many other such families appear when the restriction to small amplitudes is removed.

The general motion can be accurately solved, in a two-dimensional version of the method described in Section 3.2, by iteratively refining the coefficients of an expansion of the vertical displacement, in a double Fourier sum in the tunes Q_x and Q_z. (Again, to avoid secular terms, an improved tune must be employed in each stage of iteration.) For small amplitudes, an accurate description is given by a few terms, much as in Equation 79. For larger amplitudes, resonance islands and chaos set in, preventing convergence of the Fourier series. Such solutions agree well with flat beam simulations (50) in predicting the critical tune shift parameter and the location of dangerous resonances, when systematic scans of the tune plane are performed. Only relatively minor changes are observed when horizontal beam-beam interactions are also included, so that "energy" can flow in both directions between horizontal and vertical oscillations. This is not surprising, considering the relatively enormous source of "energy" in the horizontal motion that is available for pumping the vertical motion.

In conclusion, it is worth commenting on the role that simulations have frequently played in understanding nonlinear effects in colliders. The "ultimate" general simulation program, capable of reproducing all effects simultaneously, has not been written, but neither would it be worth writing, even ignoring the slow performance that would inevitably result. Rather, simulations are most useful in studying individual mathematical models in detail, with none of the approximations inevitably necessary on paper; or they are useful as true simulations of realistic colliders, selecting only dominant features reasonably expected to be present. While simulations are invaluable when properly used, they lose their value when removed from the context of theoretical models and/or real experience.

Literature Cited

1. Bruck, H. *Accelerateurs Circulaires de Particules*. Paris: Presses Univ. France. (In French) (1966)
2. Courant, E., Snyder, H. *Ann. Phys.* 3: 1 (1958)
3. Kolomensky, A., Lebedev, A. *Theory of Cyclic Particle Accelerators*. Amsterdam: North Holland (1966)

4. Blewett, M., ed. In *Proc. Int. Sch. Part. Accel., CERN 77-13.* Geneva: CERN (1977)
4a. Lichtenberg, A. *Phase Space Dynamics of Particles.* New York: Wiley (1979)
5. Fisk, H., et al. *SSC-7,* SSC Central Design Group, Berkeley, Calif. (1985)
6. Birkhoff, G. *Dynamical Systems.* New York: AMS Press (1927)
7. Poincaré, H. *Les Methods Nouvelles de la Mechanique Celeste.* Paris: Gautier-Villars. (In French) (1892)
8. Lichtenberg, A., Lieberman, M. *Regular and Stochastic Motion.* New York: Springer-Verlag (1983)
9. Hénon, M. *Ann. Astrophys.,* No. 28, pp. 499–511, 992–1007 (1965)
10. Kyner, W. T. *Commun. Pure Appl. Math.* 17: 227 (1964)
11. Zlatoustov, V. A., et al. *Kosm. Issledov.,* No. 2, pp. 657–66 (In Russian) (1964)
12. Dragt, A. J. *Rev. Geophys.* 3: 255–98 (1965)
13. Northrop, T. G. *The Adiabatic Motion of Charged Particles.* New York: Interscience (1963)
14. Hénon, M. *Q. Appl. Math.,* No. 3, p. 291 (1969)
15. Moser, J. *Nachr. Akad. Wiss. Gottingen, Math. Phys. K1,* p. 1 (1962)
16. Siegel, C., Moser, J. *Grund. Math. Wiss. Bd. 187.* Berlin: Springer-Verlag (1971)
17. Talman, R. *Nonlinear perturbation of a cyclic accelerator lattice; exact and approximate solutions,* Cornell LNS Rep., Ithaca, New York (1976)
18. Talman, R. In *Proc. Workshop on Accelerator Orbit and Particle Tracking,* Upton, NY: BNL (1982)
18a. Linstedt, M. *Astron. Nach.* 103: 211 (1883)
19. Chirikov, B. *Phys. Rep.* 52: 265 (1979)
20. Symon, K. *FNL Notes FN-130, 134, 140 & 144.* Batavia, Ill: Fermilab (1968)
21. Kobayashi, Y. *Nuclear Instrum. Methods* 83: 77 (1970)
22. Month, M. *BNL Intern. Rep. 11614.* Upton, NY: BNL (1972)
23. Edwards, D. *FNL Intern. Rep. 7-75.* Batavia, Ill: Fermilab (1975)
24. Talman, R. In *Physics of High Energy Particle Accelerators,* ed. M. Month, p. 691. New York: AIP (1982)
25. Tennyson, J. *AIP Conf. Proc. No. 57,* p. 158. New York: AIP (1979)
25a. Izrailev, F., Misnev, S., Tumaikin, G. *Preprint 77-43,* Inst. Nucl. Phys., Novosibirsk (1977)
26. Courant, E. *BNL-28186,* Upton, NY: BNL (1980); *ISABELLE Tech. Note No. 163,* Upton, NY: BNL (1980)
27. Evans, L. *CERN SPS/81-2.* Geneva: CERN (1981); Evans, L., Gareyte, J. *IEE Trans. Nucl. Sci.* NS-30: 4 (1982); Evans, L. *CERN SPS/83-38.* Geneva: CERN (1983)
27a. Chirikov, B., Keil, E., Sessler, A. *Statist. Phys.* 3: 307 (1971)
28. Peggs, S. *Part. Accel.* 17: 11–50 (1985)
29. Collins, T. *Fermilab Tech. Note 84/114.* Batavia, Ill: Fermilab (1984)
30. Ohnuma, S. *Proc. Steamboat Springs Conf. on Particle Interactions.* Batavia, Ill: Fermilab (1984)
31. Dragt, A. *Phys. Pub. 85-004.* College Park: Univ. Maryland (1984)
32. Ng, K. Y. *Fermilab Intern. Rep. TM-1281.* Batavia, Ill: Fermilab (1984)
33. Horton, C., Reich, L., Szebehely, V. *Long-Time Prediction in Dynamics.* New York: Wiley (1983)
34. Schoch, A. *CERN 57-21.* Geneva: CERN (1958)
35. Guignard, G. *CERN 78-11.* Geneva: CERN (1978)
36. Courant, E., Ruth, R., Weng, W. In *Physics of High Energy Particle Accelerators,* ed. M. Month, p. 294. New York: AIP (1983)
37. Brown, K., et al. *SLAC-91.* Palo Alto, Calif: Stanford Linear Accel. Ctr. (1977)
37a. Furman, M. *SSC-TM-4001.* SSC Central Design Group, Berkeley, Calif. (1984)
38. Douglas, D., Dragt, A. In *Proc. 12th Int. Conf. on High Energy Accelerators.* Batavia, Ill: Fermilab (1983)
39. Forest, E. PhD thesis. Univ. Maryland (1984)
40. Schachinger, L., Talman, R. *SSC-52.* SSC Central Design Group, Berkeley, Calif. (1985)
41. Giacaglia, G. *Perturbation Methods in Nonlinear Systems, Appl. Math. Sci. No. 8.* New York: Springer-Verlag (1972)
42. Dragt, A. In *Physics of High Energy Particle Accelerators,* ed. M. Month, p. 147. New York: AIP (1981)
43. Taf, L. G. *Celestial Mechanics.* New York: Wiley (1985)
44. Kolmogorov, A. *Dokl. Akad. Nauk. SSSR* 98: 527 (In Russian) (1954)
45. Bogoliubov, N., Mitropolsky, Yu., Samoilenko, A. *The Method of Rapid Convergence in Nonlinear Mechanics.* Kiev: Naukova dunka (In Russian) (1969)
46. Bogoliubov, N., Mitropolsky, Yu. *Asymptotic Methods in the Theory of Nonlinear Oscillations.* New York: Gordon & Breach (1961)
47. Seeman, J. *SLAC-PUB-3182.* Palo Alto, Calif: Stanford Linear Accel. (1983)
47a. Keil, E., Talman, R. *Part. Accel.* 14: 1–2, 109–18 (1983)
48. Month, M. *IEEE Trans. Nucl. Sci.* NS-22: 1376–80 (1975)

49. Myers, S. *LEP Note 362.* Geneva: CERN (1982)
50. Peggs, S., Talman, R. *Phys. Rev. D* 24: 2379 (1981); Peggs, S. PhD thesis. Cornell Univ. (1981)
51. Piwinski, A. *IEEE Trans. Nucl. Sci.* NS-24: 3 (1977)
51a. Izrailev, F., Vasserman, I. *Preprint 81-60.* Inst. Nucl. Phys., Novosibirsk (1981)
52. Peggs, S. *IEEE Trans. Nucl. Sci.* NS-30: 4 (1983)
53. Vivaldi, F. *Rev. Mod. Phys.* 56: 4, 737 (1984)
54. Chirikov, B., Ford, J., Izrailev, F. *Preprint 81-70.* Inst. Nucl. Phys., Novosibirsk (1981)