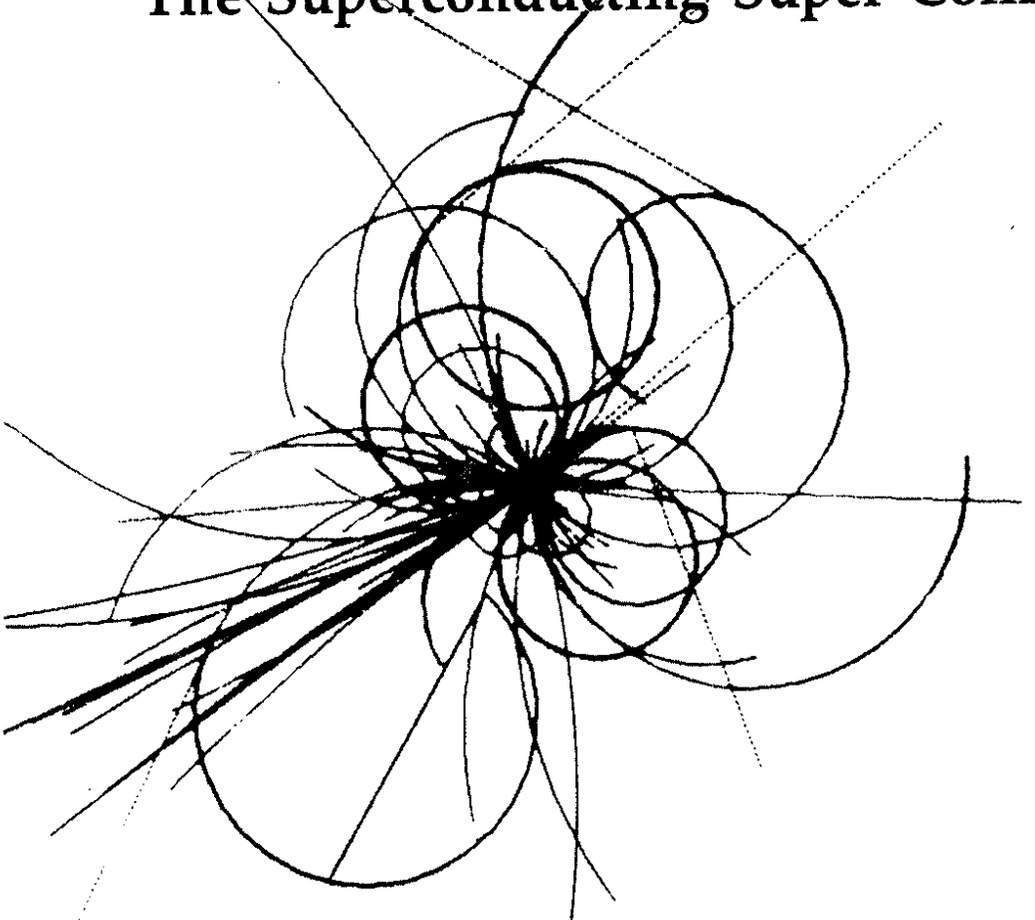
SSC-144

The Superconducting Super Collider



Iteration and Accelerator Dynamics

S. Peggs
SSC Central Design Group

October 1987

SSC-144

ITERATION AND ACCELERATOR DYNAMICS*

S. Peggs
SSC Central Design Group†
c/o Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720

October 1987

*Lecture given at Joint US/CERN Accelerator School, Topical Course on Frontiers of Particle Beams, South Padre Island, October 1986

†Operated by Universities Research Association for the U. S. Department of Energy

Iteration and Accelerator Dynamics

S. Peggs
SSC Central Design Group,
1 Cyclotron Road, Berkeley, CA 94720

"The Mersenne prime problem was a highly artificial, if ingenious, application of the growing Manchester computer. Only from the autumn of 1949 could it be applied to 'regular' problems. Besides those of Alan Turing himself, ... , it was used for optical calculations, tracking rays through systems of lenses, and for some mathematical work in connection with guided missiles."[1]

"If the only tool you have is a hammer, all your problems look like nails."[2]

Introduction

Computers have been used to model physical systems from their earliest days. The behavior of these models is subtly but profoundly affected by the sequential and discrete nature of all digital computers, in which one step in an algorithm follows another, just as machine cycle follows machine cycle. Analytic tools make problems look like differential equations, while numerical tools make them look like difference equations. Which representation is more appropriate depends on the nature of the system itself, and on the questions which need to be answered - some physical systems are inherently continuous, and some, like accelerators, are inherently discrete.

For example, the motion of a pendulum swinging under the continuous influence of gravity is exactly described by a differential equation. The numerical representation breaks the motion into two steps, which are iterated. In the first step the pendulum 'drifts' with a constant angular velocity for a time Δt, and in the second step the action of gravity is approximated by an impulsive 'kick' which changes the angular velocity. This numerical model more naturally describes the longitudinal oscillations of a test particle, relative to the center of its own bunch, which encounters a single short radio frequency cavity once per turn of an accelerator.

It is often faster and easier to answer questions about a physical system by writing a short program containing an approximate numerical model, even if the

system is inherently continuous. Take the case of a charged particle injected at rest into the periphery of a radially symmetric betatron. The purely vertical magnetic field increases uniformly from a non-zero value at injection time, inducing concentric horizontal electric fields which accelerate the particle. What does the particle do? Trying to solve this "simple" problem using both approaches is an illuminating exercise.

With luck, it is possible to find the analytic solution to differential equations - although even the pendulum problem becomes non-trivial when finite angle oscillations are considered. It is much rarer to solve non-trivial difference equations analytically. This is the main reason that traditional mathematical methods are almost all implicitly continuous, and that iteration only began to be used extensively with the advent of the computer. In some cases traditional methods of analysis break down more or less completely, and numerical techniques are not just advantageous, but essential. For example, Hamiltonian analyses of particle dynamics are usually insufficient to determine the stable aperture of an accelerator accurately enough. Instead, particles must be numerically tracked around a machine.

Some iterative algorithms have been around for a long time. The best known example is probably the Newton search method for solving the equation $f(x) = 0$

$$\begin{aligned} &x = \text{guess} \\ &\text{until converged} \quad \{ \\ &\qquad x = x - \frac{f(x)}{f'(x)} \\ &\} \end{aligned} \tag{1}$$

where $f'(x)$ is the differential of the function. The following algorithm, based on Newtons method, has been used for centuries to find the square root of y in the absence of calculators and tables of logarithms

$$\begin{aligned} &x = 1 \\ &\text{until converged} \quad \{ \\ &\qquad x = (x + \frac{y}{x}) / 2 \\ &\} \end{aligned} \tag{2}$$

This algorithm converges remarkably rapidly - the square root of 10 is correct to 10 decimal places after only 6 iterations - leading to its continued use in the mathematical library functions installed on computers. Good algorithms die hard.

A common problem with Newton searches is that the initial guess must be close enough to the right answer for the solution to converge. Pan and Reif recently

(1985) rediscovered the result of Ben-Israel (1966) that the initial guess in the algorithm

$$
\begin{aligned}
&X = \tilde{Y} \;/\; [\; \max_i \Sigma_j |A_{ij}| \;.\; \max_j \Sigma_i |A_{ij}| \;] \\
&\text{until converged } \{ \\
&\qquad X = X + X(I - XY) \\
&\}
\end{aligned}
\tag{3}
$$

for inverting the matrix Y guarantees convergence[3-6]. Here $\tilde{Y}$ is the transpose of Y, and I is the identity matrix. The most difficult requirement of the algorithm, which again is based on a Newton search, is that the user be able to multiply matrices. This algorithm is not the fastest way to invert matrices, but it is surely the easiest to understand and construct. In solving a problem numerically there is often a choice, as here, between a more or less direct transcription of an analytic method, and a less traditional iterative method. An iterative approach frequently results in a program which is easier to understand, is shorter, and is more flexible, even if it performs somewhat slower.

Four examples of iteration in accelerator dynamics are studied in detail below. The first three show how iterations of the simplest maps reproduce most of the significant nonlinear behavior in real accelerators. Each of these examples can be easily reproduced by the reader, at the minimal cost of writing only 20 or 40 lines of code. The fourth example outlines a general way to iteratively solve nonlinear difference equations, analytically or numerically.

The standard map

How significant are the differences between the dynamics of discrete and continuous systems - between analogous representations of the pendulum, for example? The continuous motion of a pendulum of unit length is described by

$$\theta'' = -g \sin(\theta) \tag{4}$$

where θ is the angle from the vertical, g is the acceleration due to gravity, and the prime denotes differentiation with respect to time. Note that the unit of time may be adjusted to make $g = 1$, a condition which is assumed from here on. The map which is the discrete analog of the pendulum system has received so much attention, and occurs in so many contexts, that it has come to be called the 'standard' map[7].

$$
\begin{aligned}
&\text{until bored} \qquad \{ \\
&\qquad \theta \;=\; \theta \;+\; \Delta t\ \theta' \\
&\qquad \theta' \;=\; \theta' \;-\; \Delta t \sin(\theta) \\
&\}
\end{aligned}
\tag{5}
$$

Here Δt is a finite time step.

It has already been remarked that equation (5), with a suitable coordinate transformation, describes the relative longitudinal motion of a particle in a storage ring with one short radio frequency cavity. In the limit of small amplitude oscillations the motion is written as

$$\theta \;=\; \theta_0 \cos(2\pi Q_0\, n) \tag{6}$$

where n is the iteration or turn number, and Q_0 is the linear 'tune'. The longitudinal 'synchrotron' tune in real accelerators varies from about 0.005 in proton storage rings to about 0.1 in electron rings. Comparing equations (5) and (6) in the low amplitude limit gives

$$\cos(2\pi Q_0) \;=\; 1 \;-\; \frac{\Delta t^{\,2}}{2} \tag{7}$$

showing that even small amplitude motion is unstable if $\Delta t > 2$.

Equation (4) is formally solved by associating canonical coordinates q and p with θ and θ', respectively, and finding a Hamiltonian

$$H \;=\; \frac{1}{2}p^2 \;-\; \cos(q) \tag{8}$$

so that

$$
\begin{aligned}
\frac{dq}{dt} &= \frac{\partial H}{\partial p} \\
\frac{dp}{dt} &= -\frac{\partial H}{\partial q}
\end{aligned}
\tag{9}
$$

Since H is an explicitly conserved quantity, trajectories follow contours of the Hamiltonian function (8), as drawn in figure 1a. The rate of progress along a contour depends only on the local slope of the function. It takes an infinite amount of time to move once around the separatrix, since the slope is zero at the 'unstable fixed points' where the pendulum is inverted and motionless.

The fundamental difference between the systems described by expressions (4) and (5) is that the restoring force is time independent in the first 'autonomous' case, and time dependent in the second. This is illustrated by rewriting the standard map as a differential equation

$$\theta'' \quad = \quad -\delta(T - n\,\Delta t)\; g\,\Delta t\,\sin(\theta) \qquad (10)$$

where the delta function is non zero every Δt units of time. It is rigorously true for an autonomous system that two neighboring trajectories which start infinitesimally close together diverge from each other linearly in time. Non-autonomous systems show both linear and exponential divergence.

A trajectory responding to the standard map can behave quite differently from its continuous cousin, depending on its initial coordinates and on Δt (or Q_0), the only control parameter. A trivial difference is that only discrete values of time exist, so it is only possible to plot one phase space point per iteration of the map, on a 'Poincare surface of section'. If enough of these dots are drawn, they often appear to form a continuous line, a 'KAM surface', looking quite like the contours of the continuous Hamiltonian[7-10]. This is shown for the case with $Q_0 = 0.06$ in figure 1b. If two neighboring trajectories are both 'regular' in this sense, then they diverge linearly with time, just as in the continuous case.

In contrast, some trajectories are 'chaotic', with dots which appear to be scattered randomly within regions of phase space that are bounded by KAM surfaces. To be more formal, neighboring regular trajectories diverge linearly with time, while neighboring chaotic trajectories diverge exponentially. Chaotic regions are found everywhere in phase space, associated with resonant trajectories whose perturbed tunes are rational fractions. (A general definition of tune, valid for nonlinear motion, is "the average ratio of phase space turns to map turns".) Just as irrational numbers fit into a number line which is dense in rational fractions, so also regular trajectories commingle with chaotic phase space regions. Fortunately, the size of the chaotic regions shrinks rapidly with the rational fraction denominator - the 'order' of the associated resonance.

Chaos is most apparent in figures 1c and 1d near the (now nonexistent) separatrix, where the tune approaches zero, a first order resonance. Figure 1 demonstrates a general feature of nonlinear systems - that the situation is more stable if the nonlinear force is more continuous. For example, the radio frequency system in a collider should be broken up into several cavities when the synchrotron tune is large. Electron colliders like LEP need high power radio frequency systems. While this requirement drives the synchrotron tune to relatively vulnerable high values, it also

causes the many meters of radio frequency cavities to be well distributed. Longitudinal chaos is seldom a problem in practice. As a second example, the maximum 'space charge' tune shift, due to the continuous interactions of a particle with the electromagnetic fields of its own bunch, can be as high as about 0.25 before beam losses and emittance blow up become significant. However, the 'beam-beam' tune shift, due to occasional localized interactions with a counter rotating bunch, may only be as high as about 0.005 .

So, to answer the original question, there can be dramatic differences between analogous discrete and continuous physical systems, of which the numerical modeler must beware. On the one hand, large time steps must be avoided while modeling a gravity pendulum. On the other hand, continuous models of longitudinal motion in an accelerator are inherently flawed.

The Henon map

Although the experimental, theoretical, and numerical study of chaos started to explode about two decades ago, this was not because the phenomenon was a new discovery - Poincare mentioned the topic in the late 19th century[11]. The explosion occurred because of, and hand in hand with, the availability of relatively powerful computers. Henon, an astrophysicist, was one of the first to investigate nonlinear maps numerically[12]. He found that the map which now bears his name

$$
\begin{aligned}
&\begin{pmatrix} x \\ x' \end{pmatrix} = \text{initial values} \\
&\text{until bored } \{ \\
&\qquad \begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} \cos(2\pi Q_0) & \sin(2\pi Q_0) \\ -\sin(2\pi Q_0) & \cos(2\pi Q_0) \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix} \\
&\qquad x' = x' + x^2 \\
&\}
\end{aligned}
\tag{11}
$$

"exhibits all the typical properties of more complicated mappings and dynamical systems". This map is directly relevant to accelerator physics, as it describes the horizontal betatron motion of a particle circulating a storage ring, of linear tune Q_0, which contains a single sextupole of unit strength. In this case the prime denotes differentiation with respect to the azimuthal coordinate, so that x' is the angle that a trajectory makes relative to the nominal closed orbit.

Typical results from Henon are shown in figure 2, where the motion of several trajectories with different initial conditions have been followed for four values of Q_0, the control parameter. Four kinds of trajectories can be loosely distinguished, corresponding to behavior which is observed - and avoided or exploited - in accelerators.

A Regular non-resonant trajectories. At low amplitudes, near the origin of each figure, the motion is regular, and the trajectories form roughly circular continuous lines. Given enough time, a particle in this region will come arbitrarily close to any given phase angle. Under normal stable accelerator operation the beam fills a region around the origin with an area proportional to its emittance.

B Regular resonant trajectories. The noncircular distortions increase as the amplitude gets larger, until the motion breaks up into a chain of resonance islands. A trajectory launched in the middle of one of the five islands in figure 2c, for example, skips successively from island to island. Although this motion is still regular, some phase angles are inaccessible - the motion is phase locked, or resonant. The islands in figures 2c and 2d are all quite similar to one another. If one of them is replotted in polar rather than cartesian coordinates, and is then isolated from its companions, the resemblance to the standard map structure in figure 1 is striking.

C Rapidly divergent regular trajectories. Three arms of widely spaced dots are visible in figure 2a, corresponding to a trajectory whose amplitude increases rapidly from turn to turn. One way to extract particles from an accelerator is to gradually move the linear tune closer to 1/3, so that the stable area in figure 1a shrinks, squeezing particles out onto such divergent trajectories. Once the particles have crossed a septum, they are steered into an external beam line. Strictly speaking, this motion is no different from type A behavior, since the trajectories are regular. After reaching very large amplitudes they eventually return close to their initial conditions.

D Chaotic trajectories. Some of the trajectories at the largest amplitudes in figures 2b, 2c and 2d appear to be randomly placed dots, following no obvious pattern. These trajectories usually diverge very rapidly, but may be confined to a bounded region of phase space. (Some chaotic points have been removed from figure 2a for the sake of clarity.)

Phase space plots can also be produced from real accelerators, by kicking the beam with a pulsed magnet to induce oscillations, and then recording the turn-by-turn displacement of the beam in two neighboring beam position monitors. These techniques are still under development, but it is already clear that the perturbed tune and the noncircular distortion, or 'smear', can be accurately measured. The major

difficulty in making these measurements is due to the finite size of the beam, which is usually not negligible compared to the size of the features being investigated. Nonlinearities cause the tune to vary with amplitude, so that there is a tune spread across the distribution of the beam, causing larger amplitude particles in the kicked beam to rotate faster (say) in phase space than low amplitude particles. The beam filaments into a hollow shell, and the center of charge signal picked up by the position monitors decoheres. However, if some of the beam is kicked into a resonant island, a fraction of the signal is phase locked, and does not decohere. The frequency of this persistent signal identifies the resonance, while the amplitude of the signal offers a direct measurement of the resonance strength.

Suppose that the linear tune is close to 1/3, as in figure 2a, so that

$$\delta q \;=\; Q_0 - \frac{1}{3} \tag{12}$$

is a small parameter, and so that the net movement of the trajectory after three iterations is relatively small. Then it can be shown, with a little algebra, that the three turn motion is approximated by a new map[13]

$$\Delta x \;=\; \frac{\partial H_3}{\partial x'} \qquad\qquad \Delta x' \;=\; -\frac{\partial H_3}{\partial x} \tag{13}$$

through the introduction of a 'discrete Hamiltonian' H_3

$$H_3 \;=\; 3\;2\pi\;\delta q\;\frac{1}{2}(x^2 + x'^2) \;-\; \frac{\varepsilon}{3}\sum_{k=1}^{3}(c_k\,x \;+\; s_k\,x')^3$$

$$c_k \equiv \cos(k\;2\pi/3) \qquad\qquad s_k \equiv \sin(k\;2\pi/3) \tag{14}$$

Here ε is the strength of the sextupole, which has so far been set equal to one. The value of H_3 is explicitly preserved under its <u>continuous</u> application as a Hamiltonian, so H_3 is often called the 'resonant invariant'. It is not fully invariant under iterations of either the Henon map, or the approximate map in expression (13) .

When the contours of H_3 are plotted for the case shown in figure 1a, the small amplitude behavior is accurately reproduced - the triangular structure has about the right size, the arms point in the right directions, and the rate of escape of divergent particles is well estimated. The large amplitude structure, however, is incorrect - the

outlying islands are totally absent, and divergent trajectories do not return, but continue off to infinity. If the tune is shifted to the conditions shown in figure 2d, then it is the six turn motion which is small, and an H_6 discrete Hamiltonian must be constructed. In principle it is possible to do this analytically, but in practice it is very tedious and unenlightening, since the lowest order nonlinear term is cubic in ε, the sextupole strength. And, to quote Taf, "Beyond first order results I know of no useful result from perturbation theory in celestial mechanics... "[14]. Analytic Hamiltonian theory is only useful in the case of the Henon map when the amplitudes are small or moderate, and the tune is close to 1/3 .

A discrete Hamiltonian can also be constructed empirically, even when the one turn map contains many nonlinearities, by fitting data produced by a tracking program[15]. This 'empirical Hamiltonian' may then be used to speed up tracking significantly, making long term tracking studies possible for contemporary accelerators which contain thousands of nonlinear elements. It can also be used as the basis for further theoretical investigations.

The round beam-beam map, and tune modulation

The task of accurate dynamic aperture determination has probably been the most intensely studied nonlinear problem in accelerator physics. The only plausible contender for this title is the beam-beam interaction, an effect which limits the useful luminosity performance of almost all electron and proton storage rings. When a test particle in (say) an antiproton bunch passes through the macroscopic electromagnetic field of a counter-rotating proton bunch, it receives a transverse beam-beam impulse which modifies its angle. If the opposing beam has a round gaussian transverse profile, and if the vertical displacement of the test particle is zero, then the horizontal angular kick is

$$\Delta x' \;=\; -4\pi\xi\;\frac{2}{x}\;\left[1 - e^{\left(-\frac{x2}{2}\right)}\right] \qquad (15)$$

where the unit of distance in these normalized coordinates is the beam size, and where ξ is the beam-beam tune shift parameter. In an accelerator with an unperturbed linear tune of Q_0, and a single such interaction per turn, the perturbed tune becomes $Q_0 + \xi$ in the zero amplitude limit, but remains Q_0 in the large amplitude limit.

The lifetime of the beam decreases dramatically if the tune shift parameter becomes too large - if there is an attempt to store too much charge in the opposing bunches. For example, ξ must be held to less than about 0.004 in the CERN Super Proton Synchrotron (SPS) storage ring. The simplest numerical model of the beam-

beam interaction replaces the x^2 term in the Henon map (11), with the $\Delta x'$ kick in (15). This model fails, however, because ξ has to be increased to 0.1 or so before severe trajectory distortions and chaos are seen[16,17].

The model is significantly improved by including tune modulation in the map[18-21]. For example,

$$
\begin{aligned}
&n = 0 \\
&\text{until bored} \qquad \{ \\
&\qquad n = n + 1 \\
&\qquad Q = Q_0 + q \sin(2\pi Q_m n) \\
&\qquad \begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} \cos(2\pi Q) & \sin(2\pi Q) \\ -\sin(2\pi Q) & \cos(2\pi Q) \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix} \\
&\qquad x' = x' - 4\pi\xi \frac{2}{x} \left[1 - e^{\left(-\frac{x2}{2}\right)}\right] \\
&\qquad \text{if } Q_m n = \text{integer, plot } (x,x') \text{ in polar coordinates} \\
&\}
\end{aligned}
\tag{16}
$$

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 with energy that occurs when the net chromaticity is not exactly zero. In this case the modulation tune Q_m is the synchrotron tune Q_s, about 0.005 in the SPS, and the modulation amplitude q is typically 0.001 or larger.

Figure 3 follows several trajectories through the map (16) for 2000 modulation periods of exactly 194 turns, plotting the phase space location once per period in polar coordinates of transverse amplitude and phase. The modulation period is chosen to be an integer so that the plot is not blurred in the phase coordinate. The tune shift parameter ξ is 0.0042 in the two top figures, 3a and 3b, while it is 0.006 in the two bottom figures. There is no tune modulation, $q = 0$, in the two left hand figures, while $q = 0.001$ in the two right hand figures. In all cases the unperturbed linear tune is $Q_0 = 0.331$.

Several important properties of the beam-beam interaction are visible. There is no sign of a beam-beam dynamic aperture in the absence of tune modulation. There are no rapidly divergent trajectories, as there are under the Henon map in figure 2, because the beam-beam angular kick decreases with increasing displacement, while the sextupole kick increases with displacement. The electromagnetic field source is localized inside the vacuum chamber in the beam-beam case, but is outside in the sextupole case.

The tune of a resonance must be within the range Q_0 to $Q_0 + \xi$ in order for its island chain to appear. A sixth order resonance, 2/6, is the lowest order resonance to satisfy this condition in figure 3 - the 1/3 resonance is excluded because the nonlinear kick is an odd function of x, exciting only even order resonances. The central amplitude a_r of the island chain is determined by the condition that

$$Q(a_r) = Q_0 + \xi\, U(a_r) = \frac{2}{6} \qquad (17)$$

where U is a known universal function with $U(0) = 1$, which asymptotically approaches zero as the amplitude goes to infinity. The location of the island chain changes between figures 3a and 3c because the tune shift parameters are different.

The amplitude width of an unmodulated resonance island chain is determined by a competition between the "strength" of the resonance and the local slope of tune with amplitude. Both of these quantities are proportional to the tune shift parameter, so the island width (at constant central amplitude) is independent of ξ, except at very large values. If the linear tune in 3c is lowered to about 0.330, the central amplitude is restored from about 2.75 to about 1.95, and figures 3a and 3c become essentially indistinguishable. This is why the simple beam-beam model without tune modulation fails.

Synchrobetatron sidebands are visible in figures 3b and 3d as extra chains of islands, separated in tune from the fundamental by $k\, Q_m/6$, where k is an integer. They appear when tune modulation is introduced, and have significant strength over a tune range of about $\pm q$ around the fundamental. Their central amplitudes are found by solving equation (17) with the resonance tune 2/6 replaced by $2/6 + k\, Q_m/6$.

The sidebands extend over an amplitude range which is skewed towards larger amplitudes, because the slope of tune with amplitude decreases with increasing amplitude. While the tune separation of neighboring sidebands is constant, their amplitude separation varies inversely with the slope of the tune, and hence with the tune shift parameter. The 40% increase of ξ in going from 3b to 3d is sufficient to bring the sideband islands close enough together to overlap and partially destroy each other, surrounding themselves with a sea of chaos. Trajectories in figure 3d with an initial amplitude of 2 reach an amplitude of 6, and beyond, in a few hundred synchrotron oscillations. In a real storage ring this corresponds to 2 sigma particles being lost in a few seconds, rather than being stored for several hours.

It is somewhat fortuitous that this model reproduces approximately the correct critical tune shift parameter for the SPS. A more realistic model would include the more stringent need to survive several collisions per turn, and the less stringent

need to avoid only resonances as strong as about tenth order - as well as including vertical motion. Unlike the Henon sextupole problem, the beam-beam problem is well described in an analytically tractable Hamiltonian (resonant invariant) form. This is because the beam-beam kick (15) is not a low order monomial in x, but is a polynomial of infinite order, so that any given resonance is driven to first order in ξ, the strength parameter. Numerical modeling, analytic theory, and practical experience agree quite well in predicting and measuring the critical tune shift parameter in proton storage rings[20,21].

The beam-beam interaction behaves somewhat differently for electron storage rings, where synchrotron radiation in the plane of the dipole bends is important. Quantum emission excites the transverse motion almost entirely in the horizontal plane (assuming the ring is flat), while indirectly damping the motion in both planes, resulting in an equilibrium beam shape which is quite flat. As already noted, the synchrotron tune in electron colliders is about an order of magnitude larger than in proton colliders. The end result of modifying the beam-beam equations of motion to include transverse damping and diffusion, flat beams, a larger synchrotron frequency, and other related effects, is to raise the critical beam-beam tune shift parameter to between 0.02 and 0.05, depending on the machine under consideration[22-25].

Iterative solutions to nonlinear maps

When the phase space motion of a trajectory under the influence of a nonlinear map is regular and nonresonant, the motion is conveniently described by a series expansion in harmonics of the fundamental tune. A general method exists for determining the coefficients of this series, without having to resort to tracking[13,24,26]. Consider again one dimensional motion in an accelerator with a single nonlinearity,

$$\Delta x' = f(x) = \sum_{p=0}^{\infty} b_{2p+1} x^{2p+1} \qquad (18)$$

where it is assumed for the sake of definiteness and simplicity that the function $f(x)$ is a polynomial in odd powers of x, as with the round beam-beam kick described in (15).

The equation of motion is conveniently written as a single second order difference equation

$$x_{n+1} - 2\,C_0\,x_n + x_{n-1} = f(x_n)$$

$$C_0 \equiv \cos(2\pi\,Q_0) \tag{19}$$

Here Q_0 is the tune of an unperturbed trajectory, whose displacement on turn n is taken to be

$$x_n = a_1\,\cos(2\pi\,Q_0\,n) \tag{20}$$

by a judicious choice of initial phase. Equation (19) is modified, for reasons which will become apparent below, by adding a term linear in x_n to both sides, so that

$$x_{n+1} - 2\,C\,x_n + x_{n-1} = f(x_n) + 2(C_0 - C)\,x_n$$

$$C \equiv \cos(2\pi\,Q) \tag{21}$$

The first iteration loop begins by substituting (20) into (18), and expanding $f(x_n)$ as a series in harmonics of the tune, ready for substitution into the right hand side of equation (21).

The general iteration loop begins by substituting the generalization of (20),

$$x_n = \sum_{m=0}^{M} a_m\,\cos(2\pi\,m\,Q\,n) \tag{22}$$

which is necessarily truncated after a finite number of terms M, into (18), to give

$$f(x_n) = \sum_{p=0}^{P} b_{2p+1}\,\left[\sum_{m=0}^{M} a_m\,\cos(2\pi\,m\,Q\,n)\right]^{2p+1} \tag{23}$$

This expression may be reduced by using the identities

$$\cos^{2p+1}(A) \equiv \frac{1}{2^{2p}} \sum_{k=0}^{p} \frac{(2p+1)!}{k!\,(2p+1-k)!}\cos((2p+1-2k)A)$$

$$\cos(A)\,\cos(B) \equiv \frac{1}{2}\,[\,\cos(A+B) + \cos(A-B)\,] \tag{24}$$

to generate a purely cosine series

$$f(x_n) \quad = \quad \sum_{m=0}^{M} c_m \cos(2\pi\, m\, Q\, n) \qquad (25)$$

which is also truncated after M terms. The net result so far is to generate a new set of "drive" coefficients c_m from an old set of "response" coefficients a_m .

The iterative loop is closed by equating the left and right sides of (21), to generate a new set of response coefficients a_m from the new drive coefficients

$$a_m(\text{new}) \quad = \quad \frac{\frac{c_m}{2} + 2(C_0 - C)\, a_m(\text{old})}{\cos(2\pi\, m\, Q) - \cos(2\pi\, Q)} \qquad (26)$$

There is a strong response at the frequency mQ if the corresponding denominator on the right of (26) is small, that is, if the resonance condition

$$Q \quad = \quad \frac{i}{(m \pm 1)} \qquad (27)$$

is approached. The additional term on both sides of equation (21) is needed because the denominator always vanishes for $m = 1$, and the only way to ensure that a_1 remains finite - constant at its unperturbed value - is by adjusting $C = \cos(2\pi\, Q)$ so that

$$\cos(2\pi\, Q) \quad = \quad \cos(2\pi\, Q_0) + \frac{c_1}{4\, a_1(\text{old})} \qquad (28)$$

This manoeuvre amounts to adjusting the perturbed tune to remove secular terms.

In simple situations it is possible to get interesting results from this method using only a pencil and paper - for example, the octupolar tune shift due to a single sextupole is easily obtained. As the mappings get more complicated, like the round beam-beam kick, for example, and as greater accuracy is called for, more powerful tools must be used. Analytic results for the coefficients of the series, which in general will contain both sine and cosine terms, are reliably obtained by using an algebraic manipulation package. Even so, the analytic expressions get to be quite opaque when several iterations are necessary.

The most difficult procedure in the solution is the analytic expansion of $f(x_n)$, from the double sum in (23) to the series in (25) . This can be avoided by using equation (23) directly to generate a time sequence of $f(x_n)$ data, for $n = 1, 2....M$, which is then simply Fourier analyzed. The disadvantage is that the coefficients are

then known only numerically, and not analytically. A double Fourier analysis, necessary for the two dimensional flat beam-beam problem, has been successfully carried out in this way, and shows good agreement with tracking and with electron storage ring behavior[24].

The method fails to find a convergent series for chaotic trajectories, which are aperiodic and have broad noisy spectra. It also fails for rapidly divergent trajectories. More disturbing, however, is its failure on regular resonant trajectories, at least in the simple form presented here. The method is nonetheless useful in many practical situations.

Conclusion

Computers play an essential role in the study and practice of accelerator physics. Contemporary accelerators are so complex, and respond over such a wide range of time scales, that they would be impossible to operate without control computers. Lattice design and tracking programs must be used for small synchrotron light rings and for huge proton colliders alike, in order to arrange the linear optics, and to calculate how badly the inevitable optical nonlinearities affect trajectory dynamics. Modeling software is used to set hardware and software specifications for an accelerator before construction, and to empirically describe it, and tune it, after construction. Finite element codes are used to solve Maxwells equations in electromagnets, permanent magnets, radio frequency cavities, and vacuum chamber bellows.

The level of intelligence of accelerator software varies enormously. For example, control system software is used, in order of increasing intelligence, to turn things on and off, to accumulate and analyze large amounts of data, and to invoke an empirical model to retune the accelerator. Accelerator physicists and applied mathematicians disagree amongst themselves about how much intelligence numerical models of physical systems contain, and about how much intellectual respect they deserve. Common skeptical complaints include that

- it is hard to construct code guaranteed to be 'bug free'
- computers are just powerful calculators, producing mere numbers as output
- numerical models offer no new insight into physical processes

Despite the coming of age of computer science as an intellectual discipline, numerical techniques are sometimes felt to not be *bona fide* mathematical tools, on a par with analytic manipulations.

The first of these complaints is true, and too weakly stated - it is practically impossible to guarantee any code as bug free. However, it is also true that human conceptual and typographical errors enter analytical expressions written on paper. Crises of confidence, whether analytical or numerical, can be alleviated by running test cases with known results through a model, or by reproducing the same result in two independent ways. Even after the most careful checking, problems often appear when an old model is used for a new purpose - bugs are never eliminated completely, they just get harder to find.

It is not true that computers only manipulate numbers. Symbolic algebra packages like MACSYMA and REDUCE, for example, have been used to find otherwise undetected errors in analytic expressions in hand compiled tables of integrals. Relational databases and languages like LISP and PROLOG understand and manipulate the relations and hierarchies natural to almost all logical and data structures. It is true, though, that most every day numerical techniques lose the formal properties of the objects they manipulate. For example, the strength of a beam-beam resonance involves the integral of a reduced Bessel function. In order to understand how the strength varies with resonance order and trajectory amplitude, it is useful to know the recursion relations between reduced Bessel functions, and to know their asymptotic behavior. All the computer knows (unless a symbolic algebra package is being used) is how to evaluate the functions and their definite integrals.

It is unequivocally clear that computer graphics have provided a powerful new insight into physical processes, as demonstrated by the central role that figures have played above. To continue the resonance strength example further, a single graph showing the resonance strengths for several different resonance orders, as a function of particle amplitude, provides a detailed description of the situation. Whether or not such a graphical solution answers practical accelerator questions depends on the details of those questions.

Different tools are useful in different situations. Arguments about the propriety of numerical techniques are rather fatuous when these techniques solve physical problems which would otherwise have to be ignored. It is more productive to speculate on what new techniques the future will bring.

Acknowledgment

I am glad to acknowledge the extent to which H. van Heijaar has influenced this work.

References

1. Hodges, A. (1983). *Alan Turing, the Enigma,* Simon and Schuster, New York.
2. Brazier, J. (1987). *Personal Communication.*
3. Ben-Israel, A. (1966). *Math. Comput.* 20 : 439.
4. Pan, V. and Reif, J. (1985). *Proc. of 17th Ann. ACM Symp. on Theory of Computing.* Providence, R. I.
5. Phipps, T. (1986). *BYTE,* April : 181.
6. Stewart, G.W. (1987). *Science,* April 24 : 461.
7. Chirikov, B. (1979). *Phys. Rep.* 52 : 265.
8. Moser, J. (1962). *Nachr. Akad. Wiss. Gottingen, Math. Phys.* K1 : 1.
9. Siegel, C. and Moser, J. (1971). *Grund. Math. Wiss. Bd. 187,* Springer-Verlag, Berlin.
10. Lichtenberg, A.J. and Lieberman, M.A. (1983). *Regular and Stochastic Motion,* Springer-Verlag, New York.
11. Poincare, H. (1892). *Les Methods Nouvelle de la Mechanique Celeste.* Gautier-Vilars, Paris.
12. Henon, M. (1969). *Q. Appl. Math.* 3 : 291.
13. Peggs, S. and Talman, R. (1986). *Ann. Rev. Nucl. Part. Sci.* 36 : 287-325.
14. Taf, L.G. (1985). *Celestial Mechanics,* Wiley, New York.
15. Peggs, S. and Talman, R. (1986). *Proc. of the Summer study on the Physics of the SSC,* DPF, Snowmass or
    SSC-84, SSC Central Design Group, Berkeley.
16. Izrailev, F., Misnev, S., Tumaikin, G. (1977). *Preprint 77-43,* Inst. Nucl. Phys., Novosibirsk.
17. Chirikov, B., Keil, E., Sessler, A. (1971). *Statist. Phys.* 3 : 307.
18. Tennyson, J. (1979). *AIP Conf. Proc.* 57 : 158, AIP, New York.
19. Courant, E. (1980). *BNL-28186* and *ISABELLE Technical Note 163,* BNL, Upton.
20. Evans, L. (1981). *CERN SPS/81-2,* CERN, Geneva.
    Evans, L. and Gareyte, J. (1982). *IEEE Trans. Nucl. Sci.* NS-30 : 4.
    Evans, L. (1983). *CERN SPS/83-38,* CERN, Geneva.
21. Peggs, S. (1985). *Part. Accel.* 17 : 11-50.
22. Seeman, J. (1983). *SLAC-PUB-3182,* SLAC, Palo Alto.
23. Keil, E., and Talman, R. (1983). *Part. Accel.* 14 : 1-2, 109-18.
24. Peggs, S., and Talman, R. (1981). *Phys. Rev. D* 24 : 2379.
    Peggs, S. (1981). *PhD Thesis,* Cornell University, Ithaca.
25. Myers, S. (1982). *LEP Note 362,* CERN, Geneva.

26. Talman, R. (1976). *Nonlinear perturbation of a cyclic accelerator lattice; exact and approximate solutions,* Cornell LNS Report, Ithaca.

Talman, R. (1982). *Proc. Workshop on Accelerator Orbit and Particle Tracking,* BNL, Upton.

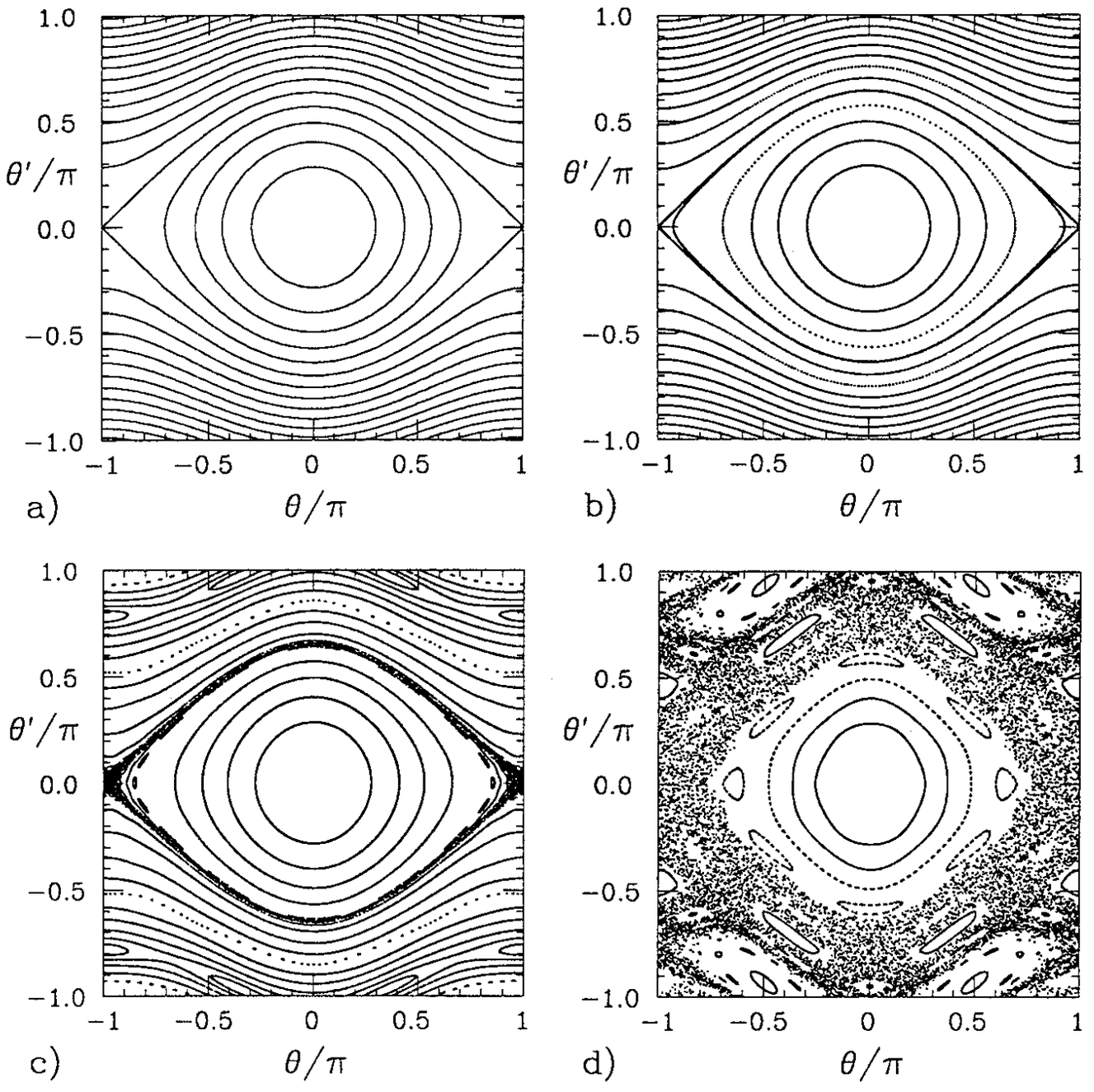


Figure 1. Standard map trajectories, with different time steps and linear tunes, Δt and Q_0. The separatrix only strictly exists in the continuous case, a).

a) Contours of the Hamiltonian, $H = 1/2\ \theta'^2 - \cos(\theta)$ [that is, $Q_0, \Delta t \rightarrow 0$].

b) $Q_0 = 0.06$. Essentially indistinguishable from a), with no sign of chaos, even close to the 'separatrix'.

c) $Q_0 = 0.12$. A narrow chaotic region appears near the 'separatrix', and some island structure appears.

d) $Q_0 = 0.18$. A large fraction of phase space is chaotic, surrounding complex island structures.

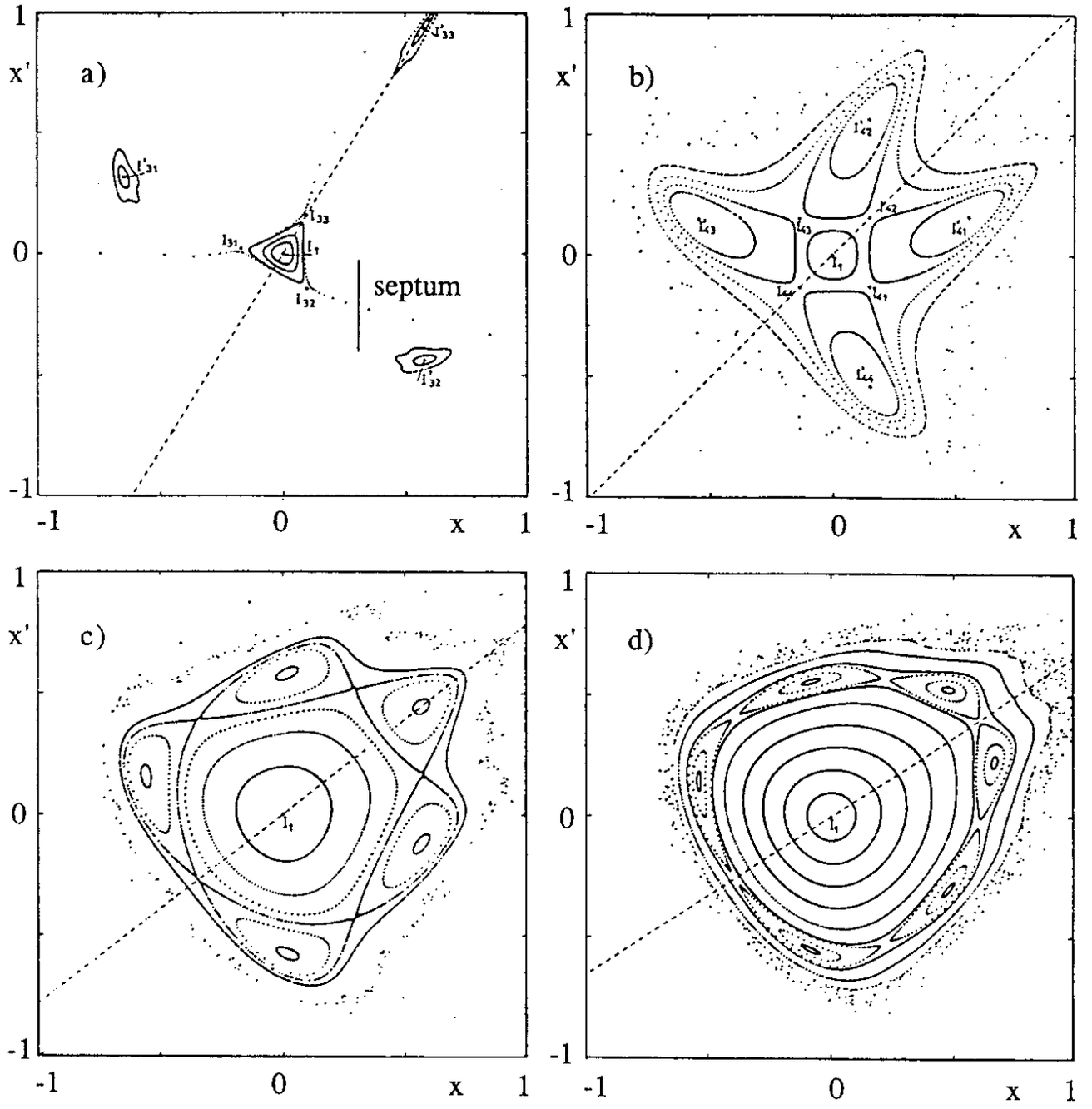


Figure 2. Trajectories obtained by Henon[12] from iterating his map with different linear tunes Q_0. When $Q_0 \approx 1/M$, M resonance islands appear.

a) $Q_0 = 0.324$, close to 1/3. The stable triangle and the divergent arms are well described by simple Hamiltonian theory, but the outlying islands are not.

b) $Q_0 = 0.2516$, close to 1/4. Four big islands, with Q_0 very close to resonance.

c) $Q_0 = 0.211$, close to 1/5. Five islands, surrounded by a stable KAM surface.

d) $Q_0 = 0.185$, close to 1/6. The islands are almost rotationally symmetric - they resemble each other and the standard map structure.

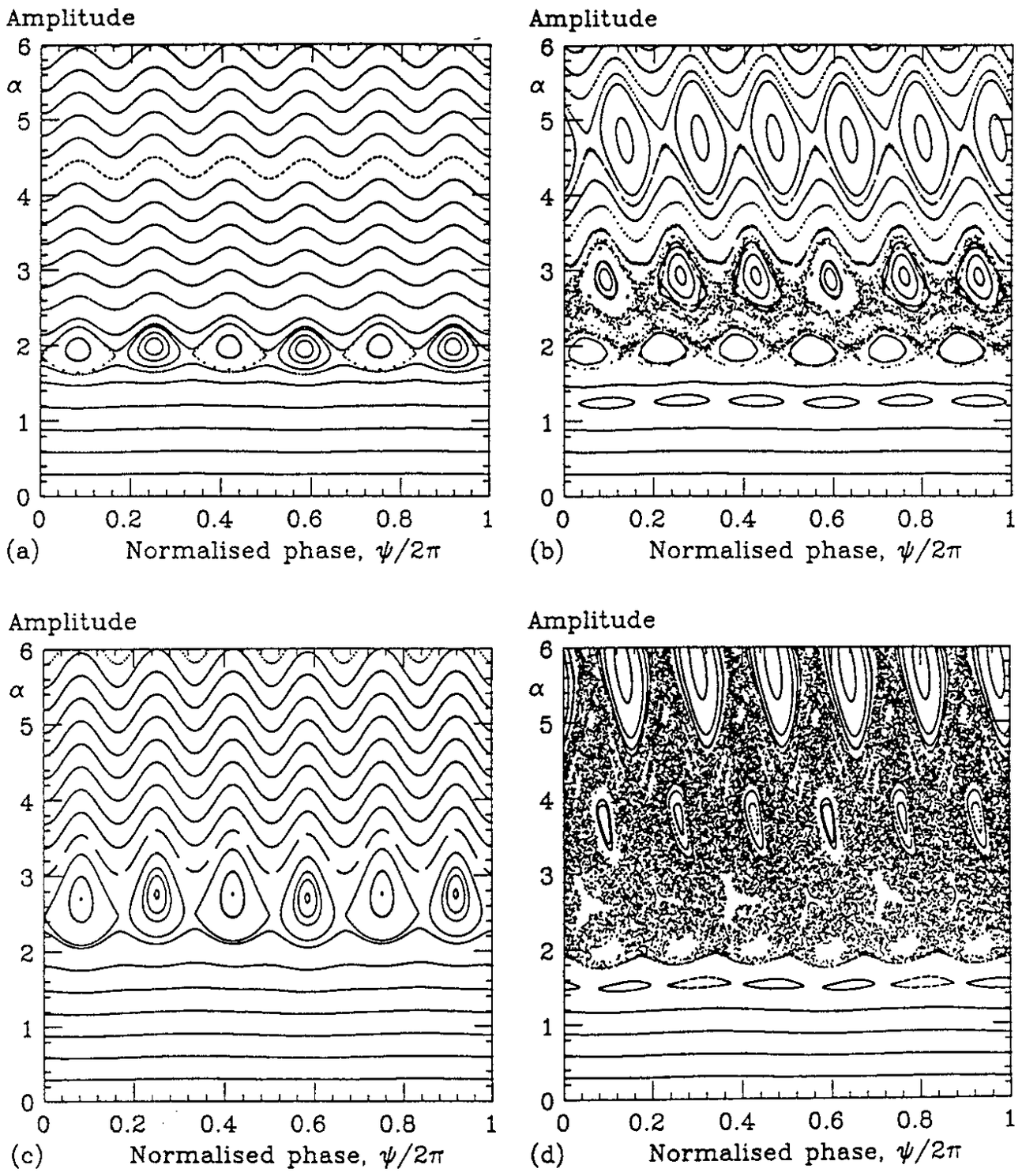


Figure 3. Phase space structure due to a map with one round beam-beam kick of strength ξ, and tune modulation of amplitude q . One phase space point is plotted for each modulation period of 194 turns, corresponding to typical proton collider synchrotron tunes of $Q_S \approx 0.005$. The two left figures have no tune modulation, while the two right figures have $q = 0.001$. The two top figures have a weaker value of $\xi = 0.0042$, while the two bottom figures have a stronger value of $\xi = 0.006$. Sidebands are visible when the tune modulation is on. They overlap and are partially submerged in a chaotic sea in d).