On the analytic treatment of non-integrable difference equations

1983 J. Phys. A: Math. Gen. 16 9

(http://iopscience.iop.org/0305-4470/16/1/011)

J. Phys. A: Math. Gen. **16** (1983) 9–24. Printed in Great Britain

On the analytic treatment of non-integrable difference equations

D S Broomhead and G Rowlands

Department of Physics, University of Warwick, Coventry, CV4 7AL, England

Received 24 May 1982

Abstract. Guided by general features of the solution of non-integrable difference equations that emerge from KAM theory, a local perturbation theory is developed. This captures, both qualitatively and quantitatively, all the essential features of the solution. The method is applied to the 'standard' map of Chirikov.

1. Introduction

There has been a growing realisation, particularly over the last decade, that a whole range of simple equations can sustain very complicated dynamical behaviour, without the *ad hoc* introduction of stochastic terms. This has led to a considerable interest in the physical implications of nonlinearity as it occurs in a wide variety of fields of study. For example, the three coupled first-order ordinary differential equations of Lorenz (1963) have solutions corresponding to motion on a so-called strange attractor. Whilst first-order difference equations, for example the logistic equation, have solutions showing chaotic behaviour. The 'standard' map of Chirikov, namely,

$$P_{n+1} = P_n - K \sin \theta_n, \qquad \theta_{n+1} = \theta_n + P_{n+1} \tag{1.1}$$

can, for suitable values of K ($K \geqslant 1$), give rise to solutions which are pseudo-random. (This latter example is a model of a conservative or area-preserving system.)

All these equations have the common feature of being nonlinear and deterministic while for some range of parameters their solutions are pseudo-random or chaotic.

It is important to realise that apart from their intrinsic mathematical interest these equations arise naturally in non-trivial applications. The Lorenz equations are model equations describing turbulent motion of liquids and gases and the presence of pseudo-randomness in the solutions has important implications for weather forecasting (Lorenz 1963). The logistic equation

$$y_{n+1} = \lambda y_n (1 - y_n)$$

describes the population density, y_n, of a self-limiting species which has non-overlapping generations. Though the biological significance of this equation is limited there are a number of important ecological problems which may be posed in the form of sets of coupled difference equations (see for example May 1976). The standard map itself arises in the study of the motion of charged particles in non-uniform magnetic fields. Then the pseudo-random behaviour leads to enhanced diffusion which is important in the design of fusion reactors (Cohen and Rowlands 1981). In fact

0305-4470/83/010009+16$02.25

Chirikov refers to the map as the 'standard' map because it arises whenever one attempts to study non-integrable Hamiltonian systems (Chirikov 1979).

The common feature of these models is that a change in the value of the parameters, λ, K, and the Reynolds number in the case of the Lorenz model, leads to a qualitative change in the nature of the solution. A relatively simple obviously deterministic solution changes to one which is extremely complicated. Such solutions, although arising from deterministic equations, have the appearance of being random and are termed pseudo-random or chaotic.

A unification in the study of such equations can be brought about by considering them as examples of dynamical systems. The object of the present paper is to study those dynamical systems which are area preserving. Such systems have a strong connection with classical Hamiltonian systems. For example in the standard map, (θ, P) may be thought of as the canonically conjugate coordinate-momentum coordinates. In the study of conservative systems the 'integrable' ones play a special role. The solutions, or trajectories, of such systems may in principle be expressed in closed analytic form, are deterministic and decidedly non-random. However, the vast majority of systems are non-integrable. They have been studied over a long period of time and in particular various perturbative schemes have been introduced attempting to relate them to adjacent integrable systems. Unfortunately the question of the convergence of these methods is far from simple being generally plagued with the problem of resonant denominators. Further, the solution of non-integrable systems is expected to be extremely complicated, witness the phase plane portrait first given by Arnold and reproduced by many authors. (See for example the review articles Jorna (1978) and Whiteman (1977). These references give excellent introductions to the basic problem discussed in this paper.)

However, though the solutions are complicated, certain qualitative features exist and these have been studied in detail. Arnold has shown that for integrable systems the solution is confined to lower dimensional regions, imbedded within the phase space, having the topology of tori. To each torus one can associate a coordinate system consisting of a set of angles (one angle per distinct irreducible closed loop) and the motion on the torus is then specified by a set of winding numbers (frequencies) corresponding to these angular coordinates. Theorems arising from the work of Kolmogorov, Arnold and Moser (KAM) tell one that for sufficiently small non-integrable perturbations the qualitative features of the solution remain unchanged, except in the neighbourhood of those tori where the winding numbers contain rational relationships. On these troublesome (resonant) tori the Poincaré–Birkhoff theorem tells one that, under the perturbation, a finite set of periodic orbits remain which, when viewed using a Poincaré section, form a string of alternating elliptic and hyperbolic periodic points. Associated with each hyperbolic point are trajectories of immense complexity which wind into tighter and tighter wiggles while still remaining in a finite region of the phase space. Repeated application of these theorems shows that each elliptic point is in reality just a microcosm of the whole.

In summary the complexity of the solution arises from two distinct mechanisms. Firstly the perturbation produces a breakup of regions surrounding elliptic fixed points into finer and finer arrays of elliptic and hyperbolic fixed points, and secondly it produces in the region of these hyperbolic points, trajectories of immense complexity.

A glance at diagrams such as that given by Arnold gives one the impression that any analysis would be prohibitively complicated. However, by a judicious choice of perturbation methods, guided by the qualitative features one expects from KAM theory,

one can obtain analytic results which are relatively simple. It is the purpose of the present paper to illustrate this procedure with reference to the standard map. The methods are, as will be appreciated, much more general.

A straightforward extension of the many-time method of perturbation theory as applied to differential equations is used here to study difference equations. (For an introduction to many-time methods see for example Nayfeh (1973).) The first point that emerges from such a study is that the method can adequately describe one of the main qualitative features described above. Namely, a resonant trajectory associated with an integrable system can, when perturbed, give rise to a string of elliptic and hyperbolic periodic points each surrounded by its own characteristic trajectories. Surrounding this array are trajectories which have been distorted in a continuous manner by the perturbation. A basic and extremely important aspect of the method is that not only are those trajectories associated with resonances treated correctly, but so also are those which are near resonance. The method removes the difficulties of both resonant and near-resonant denominators at one and the same time. Indeed we are able, by concentrating our attention on particular resonant tori, to 'filter out', through the ordering of the perturbation expansion, the effects of close higher-order resonances. By relinquishing the need to obtain a globally valid expansion we are able to handle, step by step, each resonance as it becomes significant to the order of the calculation. This is in contrast with the traditional canonical perturbation methods and is a reason for the success of the present method.

The second important characteristic of the method is that the equations describing the 'slower' time-scale variations are themselves nonlinear second-order difference equations. Thus analysis of these equations using the same basic method generates a whole hierarchy of difference equations. This evolution of a hierarchy of equations serves to describe the microcosm of the structure of the solution. It may be stopped at a desired level by replacing the difference equation by an appropriate differential one. In this way this aspect of the complexity of the Arnold diagram can be understood quantitatively.

To understand the complexity associated with each hyperbolic point, the unperturbed system has been taken to be one which is integrable, has a single heteroclinic orbit, and whose solution is known analytically. The perturbation theory shows that the perpendicular displacement of the orbit away from the unperturbed heteroclinic orbit is proportional to a bounded oscillatory term divided by what is effectively the velocity of motion along the unperturbed orbit. For orbits which approach the saddle points associated with the heteroclinic orbit, the velocity becomes small and hence the large perpendicular excursions are to be expected. In particular along the heteroclinic orbit the velocity is proportional to $e^{-\alpha n}$ where α is a constant and n the order of iteration and hence the amplitude of the perpendicular oscillations is proportional to $e^{\alpha n}$. Of course for sufficiently large n the perturbation theory breaks down but it is found that the results obtained are a good approximation for a significant number of oscillations. Using these results it is possible to associate with each unperturbed heteroclinic orbit leaving a saddle point an exponentially growing width in which the true orbit undergoes oscillatory motion with increasing frequency.

Of course, one is very aware of the pitfalls that can arise when applying perturbation theory to the type of problem discussed in this paper. Although we have not attempted to obtain analytic conditions for the range of applicability of the method, which in all probability gives an asymptotic series, we feel that by ensuring that it gives the qualitative behaviour demanded by the KAM and Poincaré–Birkhoff theorems, the

errors will be small. The results have been compared with numerical simulations from which it is concluded that the method yields both good qualitative and quantitative results.

For example, second-order theory as applied to the standard map gives the condition for chaotic orbits to extend across phase space to be $K < 1.09$ which compares well with the numerically obtained value of unity (Lichtenberg 1979).

In the next section the perturbation method is applied to (1.1) and it shows how a string of elliptic and hyperbolic points arise associated with each of the lowest-order resonances. A detailed study of the first few of these resonances gives a phase portrait which shows excellent agreement with the main features obtained numerically. At this stage of the analysis the difference equation obtained describing the *slow* variation of θ_n and P_n is approximated by replacing it by a differential equation. This procedure is valid except near the hyperbolic points which, as mentioned above, are known to be associated with chaotic behaviour.

A more rigorous study of the behaviour near a hyperbolic point is made in § 3. This analysis is based on the fact that the difference equation

$$\theta_{n+1} + \theta_{n-1} = \alpha\theta_n/(1-\theta_n^2),$$

has an exact solution

$$\theta_n = A \tanh(nx + a)$$

where a is an arbitrary constant, $A = \tanh x$ and $\tanh^2 x = 1 - \alpha/2$. This solution corresponds to an isolated heteroclinic orbit which shows smooth variation with n and the phase parameter a. Our study of the behaviour of orbits about a string of hyperbolic points is based on a perturbation method which takes the above solution as its lowest-order contribution. In this way an estimate of the perpendicular displacement, d_n, of the solution from the above unperturbed orbit has been obtained. This is used to obtain the width of a region about the heteroclinic orbit in which the actual orbit is assumed to be chaotic.

2. The standard map

The many-time perturbation theory as applied to differential equations (Nayfeh 1973) is extended in this section to treat difference equations. In particular the standard map (1.1) is treated in detail. This map may be written as

$$\theta_{n+1} + \theta_{n-1} - 2\theta_n = -K \sin \theta_n. \tag{2.1}$$

A formal expression in powers of K leads to lowest order

$$\theta_n^{(0)} = a + nP, \tag{2.2}$$

where, to this order, P is the constant momentum P_n, and a is a constant. To next order we find

$$L\theta_n^{(1)} \equiv \theta_{n+1}^{(1)} + \theta_{n-1}^{(1)} - 2\theta_n^{(1)} = -K \sin(a + nP), \tag{2.3}$$

which has the solution

$$\theta_n^{(1)} = K \sin(a + nP)/4 \sin^2(P/2). \tag{2.4}$$

This expansion obviously breaks down when $P = 2\pi m$, m an integer, and this corresponds to the primary resonance.

Motivated by the success of many-time perturbation theories we consider a to be a 'slowly' varying function of n and use this variation to remove singular behaviour; we formally write $\theta_n = a_n + nP$, $\theta_{n+1} = a_{n+1} + (n+1)P = a_n + (n+1)P + (a_{n+1} - a_n)$ and treat $a_{n+1} - a_n$ as of order $\sqrt{K}$. This gives rise to an extra term, $-(a_{n+1} + a_{n-1} - 2a_n)$, on the right-hand side of (2.3) which we treat as of order K. For $P = 2\pi m$ this equation may be solved in the spirit of the many-time formulation by ignoring the n variation of the a_n's. In this way one obtains

$$\theta_n^{(1)} = -\tfrac{1}{2}[K \sin a_n + (a_{n+1} + a_{n-1} - 2a_n)]n^2,$$

and to avoid non-physical behaviour of the momentum $P_n^{(1)}$ we must satisfy the consistency condition

$$a_{n+1} + a_{n-1} - 2a_n = -K \sin a_n. \tag{2.5}$$

This is of course the same form as the original equation for θ_n but now it refers to the slow n variation of the a_n's. We relegate a discussion of this equation until we have studied the higher-order resonances.

Before so doing it is extremely important to stress that the above treatment not only applies for $P \equiv 2m\pi$ but to all values of P which are to $O(\sqrt{K})$ in the vicinity of $2m\pi$, since the difference of order $\sqrt{K}$ may be absorbed into the slow variation of a. Thus the perturbation method treats all unperturbed orbits in the vicinity of the resonance and not just the isolated resonant one.

If $P \neq 2m\pi + O(\sqrt{K})$, then (2.4) is applicable with a constant and one may proceed to next order to find

$$\theta_n^{(2)} = \tfrac{1}{32}K^2 \sin[2(a + nP)]/\sin^2(P/2) \sin^2(P).$$

Now the method breaks down if $P = 2\pi(l/2)$, the second-order resonance condition. For $P = 2\pi(l/2) + O(K)$ we proceed as above and allow a to vary on a K scale. The new consistency condition is

$$a_{n+1} + a_{n-1} - 2a = -\tfrac{1}{8}K^2 \sin(2a_n),$$

which, except for a normalisation of the a's, is of the same form as (2.5) with, however, an effective value of K, K_e say, given by $K_e = K^2/4$. Note that with this consistency condition satisfied we have $\theta_n^{(2)} \equiv 0$ so that for $l = 1$

$$\theta_n = (a_n + \pi n) + \tfrac{1}{4}K \sin(a_n + \pi n) + O(K^3), \tag{2.6}$$

from which the corresponding value of $P_n (= \theta_n - \theta_{n-1})$ is readily calculated.

The next-order resonance is treated in the same way, but now $P = 2\pi(l/3)$, and for $l = 1$,

$$\theta_n = (a_n + 2\pi n/3) + \tfrac{1}{3}K \sin(a_n + 2\pi n/3) + \tfrac{1}{18}K^2 \sin[2(a_n + 2\pi n/3)] + O(K^4). \tag{2.7}$$

Finally, if we write $a_n = \bar{a}_n/3$, $\bar{a}_n$ satisfies (2.5) with an effective K given by $K_e = K^3/8$.

The higher resonances may be treated in the same way, and in particular for the Nth resonance where $P = 2\pi l/m$ (m an integer having no common factor with l) one obtains an expansion in powers of K but with a term of order K^N missing. The consistency condition (2.5) is satisfied by $\bar{a}_n = ma_n$ with an effective value of K given by $K_e = (K/\sigma)^m$, where σ is a function of m. For $m = 1, 2, 3, 4, 5$, $\sigma = 1, 2, 2, 1.76$, and 1.66.

Using the above results one may eliminate n and express $P_n(\equiv\theta_n-\theta_{n-1})$ as a function of θ_n and a_n. For the three lowest resonances we obtain

$$P_n = a_n - a_{n-1}, \tag{2.8}$$

$$P_n = \pi + \tfrac{1}{2}K\sin\theta_n - \tfrac{1}{16}K^2\sin 2\theta_n + (a_n - a_{n-1})(1-\tfrac{1}{4}K\sin\theta_n) + \mathrm{O}(K^3), \tag{2.9}$$

and for $P = 2\pi l/m$, $m = 3$, $l = 1, 2$,

$$\begin{aligned} P_n = {} & P + [K/2\sin(P/2)]\cos(\theta_n - P/2) \\ & + [K^2/16\sin(P)\sin^2(P/2)][2\cos^2(P/2) - \cos 2\theta_n] \\ & + (a_n - a_{n-1})[1 + K/4\sin^2(P/2)\cos(\theta_n - P)] + \mathrm{O}(K^3). \end{aligned} \tag{2.10}$$

For small enough values of K, one expects that a reasonable approximation to (2.5) may be obtained by replacing it by an equivalent differential equation. The simplest approximation is to replace (2.5) by the differential equation

$$\mathrm{d}^2a/\mathrm{d}n^2 = -K\sin a, \tag{2.11}$$

whose solution in the form of a phase-plane portrait is shown in figure 1. For a particular choice of the integration constant, which distinguishes the various orbits in the phase plane, one obtains a heteroclinic orbit. From the discussion in the introduction of the present paper, one expects the replacement leading to (2.11) to break down in the vicinity of this orbit. The nature of the solution in the vicinity of this orbit is discussed in detail in the next section where it is shown that, though it behaves very irregularly, the complexity is restricted to a narrow band about the separatrix. This suggests that an approximation to a_n, where a_n is replaced by a smooth function of n, will give good results except in the neighbourhood of the separatrix. This type of approximation will be better the smaller the value of K_e, and hence better for the higher resonances.

A smoothing approximation to an equation of the form (2.5) is discussed in appendix 1 where analytic expressions for $(a_n - a_{n-1})$ as a function of θ_n, K_e and a constant of integration E are obtained. The constant E distinguishes the orbits in the phase plane with $E = K_e(1 + 5K_e/48)$ corresponding to the separatrix. Using this

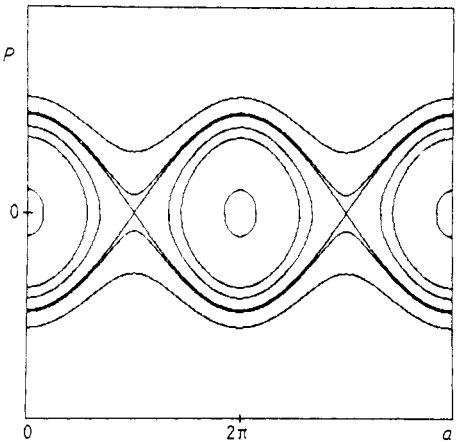

Figure 1. The phase portrait of the complete pendulum (equation (2.11)).

result, together with the appropriate definition of $\bar{a}_n$ and K_e, we may obtain expressions for P_n as functions of θ_n. We now drop the suffix n in which case $P(\theta)$ gives an analytic expression for the orbits in the (θ, P) phase-plane.

The final results are:

for the first-order resonance ($m = 1$)

$$P(\theta) = \tfrac{1}{2}K \sin\theta \pm (1 - \tfrac{1}{6}K\cos\theta)[2E + 2K(1 + \tfrac{1}{6}E)\cos(\theta + \tfrac{1}{8}K^2\cos 2\theta)]^{1/2}, \tag{2.12}$$

for the second-order resonance ($m = 2$)

$$P(\theta) = \pi + \tfrac{1}{2}K\sin\theta \pm (\tfrac{1}{2} - \tfrac{1}{8}K\cos\theta)[2E + \tfrac{1}{2}K^2\cos(2\theta - \tfrac{1}{2}K\sin\theta)]^{1/2} + O(K^3), \tag{2.13}$$

and for $P = 2\pi l/m$, $m \geq 3$, l not a factor of m,

$$\begin{aligned} P(\theta) = P + K\cos(\theta - P/2)/2\sin(P/2) + K^2[2\cos^2(P/2) - \cos 2\theta]/16\sin P\sin^2(P/2) \\ \pm (1/m)[1 + K\cos(\theta - P)/4\sin^2(P/2)] \\ \times [2E + 2(K/2)^m \cos\{m\theta - mK\sin\theta/4\sin^2(P/2)\}]^{1/2} + O(K^3). \end{aligned} \tag{2.14}$$

For $m > 4$ the last term does not contribute, so that to $O(K^3)$ the orbits in the phase plane have no heteroclinic points.

These results are shown in figure 2 as a phase plot of P against θ for K values of 0.5 and 0.97, and should be compared with results obtained by direct numerical evaluation of (2.1) and given in figure 3. Except for the random behaviour apparent in the numerical results in a narrow band about the separatrices, the agreement is excellent. Even after applying the smoothing approximation to the a_n's, the present perturbation method describes the breakup of resonant tori into strings of elliptic and hyperbolic points, one of the two major qualitative features demanded by the general theory of dynamical systems.

The results obtained above may be used to calculate a critical value of K, K_c say, above which there is overlap of the various separatices. Chirikov takes the condition $K > K_c$ as being necessary for the existence of stochastic orbits no longer confined by KAM surfaces but connected across phase space. It is apparent from the analytic

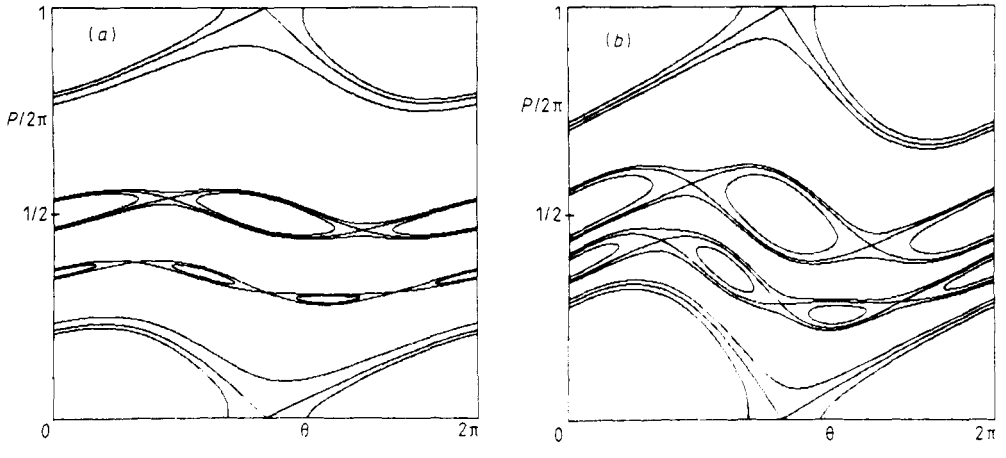

Figure 2. Smoothed analytic approximations to represent trajectories close to the three lowest-order resonances of the Chirikov map: (*a*) $K = 0.5$; (*b*) $K = 0.97$.

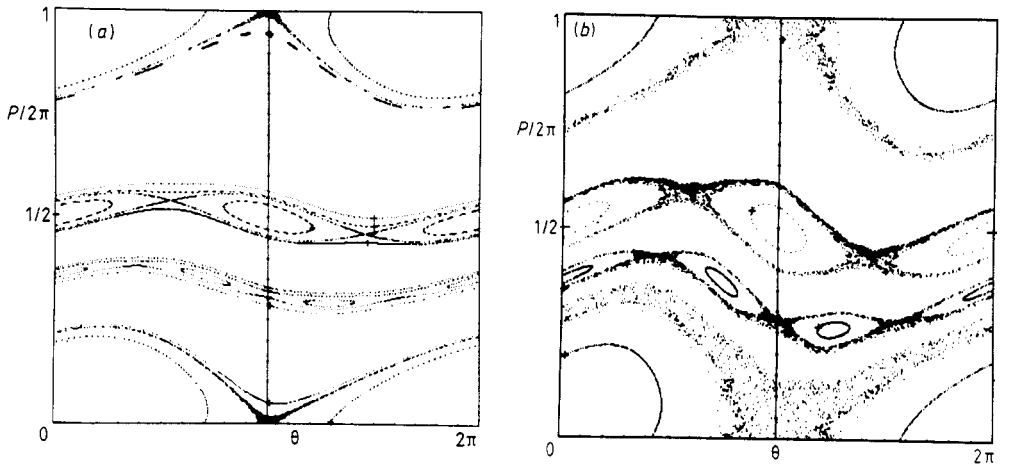

Figure 3. Numerical iteration of the Chirikov map close to the same three resonances as shown in figure 2: (*a*) $K = 0.5$; (*b*) $K = 0.97$.

results shown in figure 2 that the orbits corresponding to the various resonances come closest together near $\theta = 0$. Thus we adopt as a criterion to calculate K_c that the sum of the widths of the separatrices equals 2π. Taking the first four resonances only and neglecting terms of $O(K^{5/2})$ we obtain a value $K_c = 1.09$ (this is of course an upper bound), which compares most favourably with values obtained by other authors (see Lichtenberg 1979) and the numerically obtained value of 1.0.

It will be also noticed from the phase-plane plots of the analytic results that the orbits associated with the various resonances tend to avoid one another. This phenomenon has been discussed by Escande (1979) using the super convergence method of Kolmogorov to take into account the interaction of resonances. However, this effect arises naturally in the present theory even though we treat each resonance separately.

3. Heteroclinic orbits

It was shown in the last section that many-time perturbation theory applied to the standard map reproduces quantitatively many of the features of the numerical results. The major discrepancy is in the vicinity of the separatrix and arises because of the smoothing approximation used to treat the difference equation (2.5). In this section we look more closely at the region around the separatrix by comparing equation (2.5) with an equation which also has a separatrix but for which an analytic solution is known. This equation

$$\phi_{n+1} + \phi_{n-1} = \alpha\phi_n/(1-\phi_n^2), \tag{3.1}$$

where α is a constant, has the exact solution

$$\phi_n = k\ \mathrm{sn}(x, k)\ \mathrm{sn}(nx + b, k). \tag{3.2}$$

Here $\mathrm{sn}(x, k)$ is a Jacobian elliptic function in the notation of Whittaker and Watson (1952), and x satisfies $\mathrm{cn}(x, k)\,\mathrm{dn}(x, k) = \alpha/2$. The constants k and b are the two integration constants associated with (3.1). For $k = 1$ the solution reduces to a

heteroclinic orbit given by

$$\phi_n = \tanh x \tanh(nx + b), \tag{3.3}$$

with $\operatorname{sech}^2 x = \alpha/2$.

McMillan (1970) showed that the solutions of (3.1) formed a set of invariant curves corresponding in the present notation to different values of k. However, the fact that an analytic solution can be given, as by (3.2) above, is apparently new.

We write the basic equation of interest in the form

$$\bar{a}_{n+1} + \bar{a}_{n-1} - \alpha \bar{a}_n/(1 - \bar{a}_n^2 s^2) = g(\bar{a}_n), \tag{3.4}$$

with

$$g(\bar{a}_n) = -K_e \sin \bar{a}_n - \alpha \bar{a}_n/(1 - \bar{a}_n^2 s^2) + 2\bar{a}_n, \tag{3.5}$$

and treat $g(\bar{a}_n)$ as a small correction. Here s is a constant to be determined later. Writing $\bar{a}_n = \phi_n/s + \delta a_n$, with ϕ_n given by (3.3), then first-order ordinary perturbation theory gives

$$\delta a_{n+1} + \delta a_{n-1} - \alpha[(1 + \phi_n^2)/(1 - \phi_n^2)^2]\delta a_n = g(\phi_n/s). \tag{3.6}$$

A solution to the homogeneous equation is $\partial\phi_n/\partial b$, as is readily seen by differentiating (3.1) with respect to b. If one multiplies (3.6) by $\partial\phi_n/\partial b$ and sums from $n = -\infty$ to m one finds

$$\delta \vec{a}_{m+1} \frac{\partial \phi_m}{\partial b} - \delta \vec{a}_m \frac{\partial \phi_{m-1}}{\partial b} = \sum_{-\infty}^{m} \frac{\partial \phi_n}{\partial b} g(\phi_n/s), \tag{3.7}$$

where we have taken δa_m to remain finite, possibly going to zero as $m \to -\infty$.

From the general theory of heteroclinic orbits it is known that the effect of introducing a perturbation is that the orbit bifurcates such that the unstable orbit emerging from one hyperbolic point can no longer be identified with the stable orbit converging to the other hyperbolic point (Holmes 1979). The procedure leading to (3.7) corresponds to summing along the unstable orbit with the hyperbolic point being approached as $m \to -\infty$. For this reason we may take $\delta \vec{a}_m$ at least finite in this limit. Further, we expect $\delta \vec{a}_m$ to behave more and more wildly as m increases, that is as the orbit approaches the other hyperbolic point. The arrow over the δa_m designates the direction of motion along the unperturbed orbit. If we consider the variation of δa_n along an orbit which converges to this other hyperbolic point, then by analogy with (3.7) we have

$$-\delta \overleftarrow{a}_{m+1} \frac{\partial \phi_m}{\partial b} + \delta \overleftarrow{a}_m \frac{\partial \phi_{m+1}}{\partial b} = \sum_{m+1}^{\infty} \frac{\partial \phi_n}{\partial b} g(\phi_n/s),$$

with $\delta \overleftarrow{a}_m$ remaining finite as $m \to +\infty$. Combining the above two equations gives

$$(\delta \vec{a}_{m+1} - \delta \overleftarrow{a}_{m+1}) \frac{\partial \phi_m}{\partial b} - (\delta \vec{a}_m - \delta \overleftarrow{a}_m) \frac{\partial \phi_{m+1}}{\partial b} = h(b) \equiv \sum_{-\infty}^{+\infty} g(\phi_n/s) \frac{\partial \phi_n}{\partial b}. \tag{3.8}$$

If we define a perturbed momentum by $\delta P_n = \delta a_n - \delta a_{n-1}$ then this equation may be rearranged into the form

$$(\delta \vec{P}_m - \delta \overleftarrow{P}_m) \frac{\partial \phi_m}{\partial b} - (\delta \vec{a}_m - \delta \overleftarrow{a}_m) \frac{\partial P_m}{\partial b} = h(b), \tag{3.9}$$

where $P_m = \phi_m - \phi_{m-1}$ is the unperturbed momentum.

The left-hand side of equation (3.9) is simply related to the perpendicular distance between points in the phase plane and the unperturbed orbit. If we define d_m as the difference between the perpendicular distance of the orbit connected to one of the hyperbolic points and that connected to the other, both relative to the unperturbed orbit, then

$$d_m = h(b)/[(\partial\phi_m/\partial b)^2 + (\partial P_m/\partial b)^2]^{1/2}. \tag{3.10}$$

This is the main result of this section.

In a study of a class of nonlinear autonomous differential equations subject to non-autonomous perturbations, Mel'nikov has introduced a quantity Δ which measures the perpendicular distance between heteroclinic orbits. The quantity d_m is the analogous quantity to Δ but now for autonomous difference equations. The relationship between Mel'nikov's work and perturbation theory as applied to differential equations has recently been discussed by the authors (Broomhead and Rowlands 1982). The analysis described in the present section is the extension of this work to difference equations.

The evaluation of $h(b)$ for the particular perturbation given by (3.5) is discussed in appendix 2 where it is shown that to reasonable accuracy

$$h(b) = \lambda \sin(2\pi b/x)$$

where λ is a known function of x and s but is independent of b. The value of b depends on the initial starting point in the iteration of the map. For the unperturbed map, (3.1), a particular choice will lead to a set of uniquely specified points in the phase plane. An expression for d_m can however be given which is independent of the initial point by using (3.3) to express b in terms of ϕ_m. In this manner we obtain

$$d_m = (\lambda/D)\sin\{(2\pi/x)[\tanh^{-1}(\phi_m/\tanh x)]\}, \tag{3.11}$$

where

$$D^2 = (1-\phi_m^2/T^2)^2[2(1-\phi_m)^4 - 2\,\mathrm{sech}^2 x(1-\phi_m)^2 + \mathrm{sech}^4 x]/(1-\phi_m)^4$$

and $T = \tanh x$.

Before one can apply the above results directly to the Chirikov map it is necessary to discuss the relationship between the parameters s, α and K_e. We insist that the fixed points of the map given by (3.1) agree with those of (2.5). This gives

$$s\pi = \tanh x.$$

The other criterion we impose is that the eigenvalues of the linear theory about the centre are the same. This gives $K_e = 2-\alpha \equiv 2\tanh^2 x$ so that

$$s = K_e/2\pi^2.$$

With this identification one can evaluate d_m as given by (3.10), as a function of $\bar{a}_m (= \phi_m/s)$. The form of the separatrix is now obtained by adding to the unperturbed form as given by (3.3) the perpendicular displacement as given by d_m. The final result is shown in figure 4, for the case of the primary resonance ($\bar{a} \equiv a = \theta$, $K_e = K$) for $K = 0.97$. This is to be compared with the form of the separatrix obtained by the direct numerical mapping of a line segment using the map as defined by (2.5), and as shown in figure 5. The agreement is surprisingly good. The general features of an oscillatory form for the separatrix whose amplitude and wave number increase as one approaches the hyperbolic point is reproduced by the analytic result.

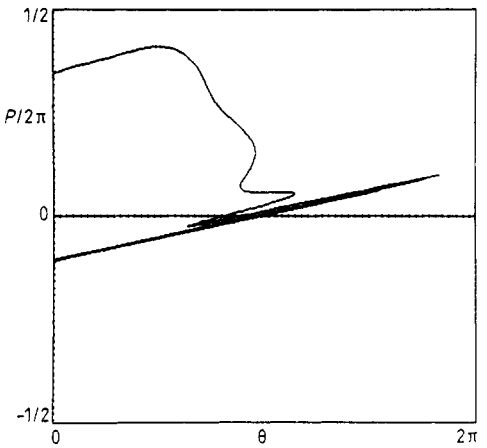

Figure 4. Analytic approximation to the distortion of an unstable manifold associated with the primary resonance of the Chirikov map with $K = 0.97$.

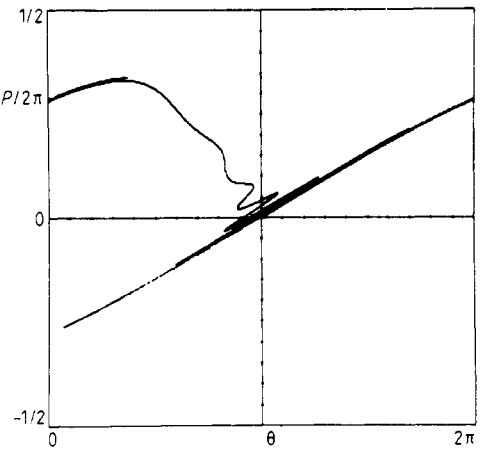

Figure 5. The result of 20 iterations of a linear segment of the same unstable manifold initially close to the hyperbolic field point of the Chirikov map ($K = 0.97$).

The above form for d_m was of course calculated using linear theory and this will break down as the orbit gets too close to the hyperbolic point. Further, a detailed solution in the form obtained above is not usually required when the shape of the orbit gets complicated. It is in most cases sufficient to obtain a bound for the region where the orbits go chaotic, that is where the convolutions in the separatrix become large.

We propose to obtain an estimate of this boundary by using orbits of the continuous approximation to the Chirikov map which are close to the orbit corresponding to the separatrix. The details are given in appendix 3 where it is shown for the continuous map that $d(\bar{a})$, the perpendicular distance between points and orbits in the neighbourhood of the separatrix, is given by

$$d(\bar{a}) = \frac{2(1-k)\sqrt{K_e}}{\{[1-\sin^2(\bar{a}/2)][1+K\sin^2(\bar{a}/2)]\}^{1/2}}.$$

Here k labels the orbits, and $k \equiv 1$ corresponds to the separatrix.

We now define k, and hence the particular adjacent orbit which bounds the chaotic region, by equating $d(\bar{a})$ to the envelope of d_m, namely λ/D, in the vicinity of the centre. That is we put $\bar{a} \equiv 0$ in the above and $\phi_m \equiv 0$ in the expression for D. This is done since it is in this vicinity that the expressions for $d(\bar{a})$ and d_m are valid having been obtained by linear theory. This finally gives

$$k = 1 \pm \lambda/\{2\sqrt{K_e}[1+(K_e/2)^2]^{1/2}\},$$

the $\pm$ signs corresponding to orbits on either side of the separatrix. For $K_e \ll 1$, $x^2 = K_e$ and we see from the form of λ given in appendix 2 that $k - 1 \simeq \pm\exp(-2\pi^2/\sqrt{K_e})$.

The main idea behind the above procedure is that, away from the separatrix, the perturbation will not destroy the smooth orbits. We use the existence of such orbits to obtain boundaries for the chaotic region.

This result has been used to determine orbits on either side of the separatrix which serve as boundaries for the chaotic regions. The results are shown in figure 6, where these boundaries are given and superimposed on the figure are points obtained by direct numerical evaluation of the Chirikov map. This shows that the above estimates for the chaotic regions are surprisingly good.

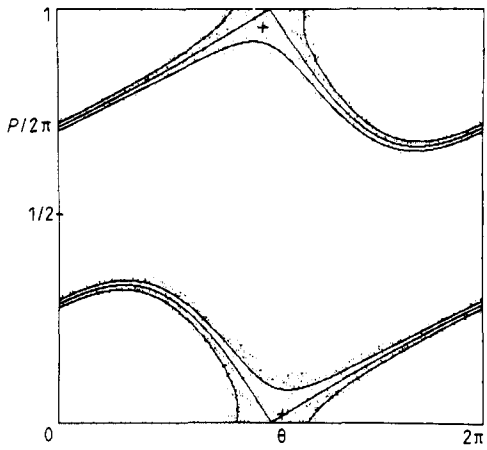

Figure 6. Comparison of the analytical bands for the stochastic layer of the primary resonance with the result of numerical iteration of the Chirikov map ($K = 0.97$).

4. Conclusions

A method of perturbation theory has been developed and applied to the 'standard' map which gives good quantitative agreement with all the main features of the numerical solution. The success of the method depends on two properties. Firstly, the perturbation method is local in the sense that different expressions apply to different regions of phase space. This allows one to deal with the problem of resonant denominators in a successive manner. In fact, the various resonances are treated independently and it is assumed there is no interaction. Secondly, it incorporates the essential qualitative features one knows about such solutions from KAM theory, the Poincaré–Birkhoff theorem and the work of Mel'nikov.

The results are obtained in the form of smooth orbits in the (θ, P) phase-plane, and for the lowest resonances are given by (2.12), (2.13) and (2.14). On these is superimposed a band which represents the region where chaotic motion is expected. The structure of such bands is described in detail in § 3. Comparison between the analytic expressions and numerical simulations can be made by referring to the figures.

Finally we consider the relationship between the present work and that of Escande and Doveil (1981). These authors considered the motion of a particle governed by the non-autonomous Hamiltonian

$$H = v^2/2 - M\cos x - P\cos[k(x-t)].$$

They solve this problem by a renormalisation procedure which at each stage reproduces

a Hamiltonian of the above form but with renormalised coefficients. The stages of this procedure follow closely the sequences of denominators in a continued fraction approximation to an irrational winding number. It is then claimed that if the renormalised coefficients approach zero as the continued fraction approaches the irrational winding number, then there exists a KAM surface associated with this number. The onset of chaos, allowing particle motion throughout the phase plane, is seen as the condition that all such KAM surfaces are destroyed.

The renormalisation procedure of Escande and Doveil is an approximate one, since at each stage only the primary resonances closest to the given irrational surface are retained, and the Hamiltonian is expanded so as to include only linear and quadratic terms in the deviation of the action from that specifying this surface. However, they are able to carry out the renormalising procedure to all orders and hence obtain a criterion for the destruction of KAM surfaces.

In the present work we have studied the lowest-order resonances only and obtained a criterion for chaos throughout the phase plane following Chirikov by imposing a resonance overlap condition. It is important to note that this criterion as applied in the present paper requires inclusion of the stochastic width about each separatrix since adjacent resonances tend to avoid overlap. This latter point is discussed at the end of § 2. The criterion we have given leads to an estimate in good agreement with numerical computation. However, in the light of Escande and Doveil's work we can reinterpret our method. The method of solution is not restricted to the low-order resonances. In principle we can apply the perturbation method to any order resonance and find a solution in the form of a closed curve superimposed on which is a set of solutions which satisfy a suitably renormalised standard map (unlike the procedure of Escande and Doveil our renormalisation produces exactly the standard map at each step and one does not have to neglect higher harmonics). Thus in principle one could apply this renormalisation to a sequence of resonances which eventually approach an irrational limit. Then one could apply the simple criterion of non-overlap of adjacent resonances in the sequence to define the existence of a KAM surface. Unfortunately, although the renormalised standard map can be seen to be reproduced at each stage, the exact rescaling properties of the parameter are difficult to determine so that the above procedure has not, as yet, been carried out.

Acknowledgments

One of us (D S Broomhead) would like to acknowledge the support of the Science and Engineering Research Council during the prosecution of this work. The figures for this paper were prepared using an HP 9845 C micro-computer belonging to the Synergetics Group in the University of Warwick.

Appendix 1.

The purpose of this appendix is to discuss a 'smooth' approximation to a difference equation of the form (2.5). An equation of this form, but with a_n replaced by $\bar{a}_n$ and K by K_e, was generated as a consistency condition by the formal perturbation theory discussed in § 2. For the mth order resonance $\bar{a}_n = ma_n$ and $K_e = (K/\sigma)^m$, where σ is a function of m.

If we assume $\bar{a}_n$ to be a smooth function of n and Taylor series $\bar{a}_{n+1}$ about $\bar{a}_n$ then we have

$$d^2\bar{a}/dn^2 + \tfrac{1}{12}d^4\bar{a}/dn^4 + \ldots = -K_e \sin \bar{a}. \tag{A1.1}$$

Introduction of the momentum $\bar{P} = d\bar{a}/dn$ allows us to write this as

$$(d/d\bar{a})\bar{P}^2 + \tfrac{1}{12}(d/d\bar{a})\bar{P}^2[(d^2P^2/d\bar{a}^2) - (d\bar{P}^2/d\bar{a})] + \ldots = -2K_e \sin \bar{a},$$

which may be solved by perturbation theory treating $\bar{P}^2 = O(K_e)$. In this way we obtain

$$\bar{P}^2 = 2E + 2K_e \cos \bar{a}(1 + \tfrac{1}{6}E) + \tfrac{1}{8}K_e^2 \cos(2\bar{a}) + O(K_e^3) \equiv 2(E - V(\bar{a})). \tag{A1.2}$$

Here E is the integration constant which specifies the particular orbit.

Using this result, the appropriate Taylor series, and noting that the sum of the even order derivation may be summed using (A1.1), we find

$$\bar{a}_n - \bar{a}_{n-1} = \bar{P}[1 - (K_e/6)\cos \bar{a}] + (K_e/2)\sin \bar{a} + O(K_e^{5/2}). \tag{A1.3}$$

Finally, we may express $\bar{a}_n - \bar{a}_{n-1}$ in terms of θ_n rather than $\bar{a}_n$. For the primary resonance $\bar{a}_n = a_n$, $\theta_n = a_n + 2n\pi$ so that $V(\bar{a}) = V(\theta_n)$, and (A1.3) is now an expression for $a_n - a_{n-1}$ as a function of θ_n. For the second-order resonance, $\bar{a}_n = 2a_n$, $K_e = K^2/4$ and a_n and θ_n are related by (2.6). Simple algebraic manipulation yields

$$2(a_n - a_{n-1}) = \pm[2E + (K^2/2)\cos(2\theta_n - (K/2)\sin \theta_n)]^{1/2} + (K^2/8)\sin(2\theta_n) + O(K^3).$$

Similar treatment of the third-order resonance yields

$$3(a_n - a_{n-1}) = \pm[2E + (K^3/8)\cos(3\theta_n - K \sin \theta_n)]^{1/2} + (K^3/16)\sin 3\theta_n + O(K^4).$$

It is essential to maintain the symmetry of $V(\bar{a})$ with respect to $\bar{a}$ when making the transformation from $\bar{a}$ to θ_n. For this reason terms of the form $\cos(3\theta_n - K \sin \theta_n)$ that occur in the above expressions must not be expanded in powers of K.

Appendix 2.

The basic problem is to evaluate $h(b)$ as defined by (3.8) for ϕ_n given by (3.3) and $g(\phi_n)$ by (3.5). To illustrate the method used we consider in detail one particular term, namely

$$h_1(b) = \sum_{-\infty}^{+\infty} \frac{\partial \phi_n}{\partial b}\phi_n = \tanh^2 x \sum_{-\infty}^{+\infty} \frac{\tanh(nx + b)}{\cosh^2(nx + b)}.$$

Use of the Poisson summation formula gives

$$h_1(b) = \tanh^2 x \sum_{-\infty}^{+\infty} F(2n\pi/x),$$

where

$$F(l) = e^{-ilb} \int_{-\infty}^{+\infty} \frac{\tanh y\, e^{ily}\, dy}{\cosh^2 y}.$$

The integral may be evaluated using the Cauchy residue method. There are an infinite number of double poles at $y = i(2m + 1)\pi/2$ where $m = 0, 1, 2, \ldots$, with residues

$\exp(-l(2m+1)\pi/2)\,\mathrm{cosech}^2 x$. Summing over the residues is easily accomplished giving

$$h_1(b) = \mathrm{i}\pi \tanh^2 x \sum_{-\infty}^{+\infty} (2n\pi/x)^2 \frac{\exp(-\mathrm{i}2n\pi b/x)}{\sinh(n\pi^2/x)}.$$

Because of the denominator, the terms in this series converge exponentially fast and thus to a good approximation we need only consider the lowest-order term. This gives

$$h_1(b) = [2\pi \tanh^2 x/\sinh(\pi^2/x)]\sin(2\pi b/x).$$

The evaluation of $h(b)$ may be accomplished using the above method if one expands $\sin\phi_n$ as a power series in ϕ_n. The term $\phi_n/(1-\phi_n^2)$ in $g(\phi_n)$ does not contribute to $h_1(b)$. Then to the same approximation as above one finds

$$h_1(b) = \lambda \sin(2\pi b/x)$$

where

$$\lambda = (4\pi/s)[\tanh^2(x)/\sinh(\pi^2/x)]\left(1-(K_e/x)\sum_{m=0}^{+\infty} \frac{A_m(2\pi/x)\tanh^{2m}(x)(1/s)^{2m}}{(2m+1)!(2m+2)!}\right)$$

and

$$A_m(l) = \lim_{x\to 0}[(\partial^{2m+2}/\partial x^{2m+2})(\mathrm{e}^{\mathrm{i}lx}\cosh^{2m+1}(x))].$$

Appendix 3.

To lowest order, the continuous approximation to the Chirikov map takes the form

$$\mathrm{d}^2\bar{a}/\mathrm{d}n^2 = -K_e \sin\bar{a}.$$

This is readily solved to give

$$\sin(\bar{a}/2) = \mathrm{sn}(\eta n + b, k)$$

where $\eta = \sqrt{K_e}/k$ and k is a parameter which labels the various orbits. In particular $k=1$ corresponds to the separatrix. In the vicinity of the separatrix we may write $\bar{a} = a + \delta a$ with $\sin(a/2) = \tanh U_0$ and

$$\delta a = (1-k)\,\mathrm{sech}\,U_0(\sinh U_0 \cosh U_0 + U_0 - 2b)$$

where $U_0 = n\sqrt{K_e} + b$. By analogy with equations (3.9) and (3.10) we define a perpendicular distance $\bar{d}(a)$ in the following manner

$$\bar{d}(a) = (\delta P\partial a/\partial b - \delta a\,\partial P/\partial b)/[(\partial a/\partial b)^2 + (\partial P/\partial b)^2]^{1/2}.$$

Using the above we find

$$\bar{d}(a) = 2\sqrt{K_e}(1-k)/\{[1-\sin^2(a/2)][1+K_e\sin^2(a/2)]\}^{1/2}.$$

References

Broomhead D S and Rowlands G 1982 *Phys. Lett.* **89A** 63
Chirikov B V 1979 *Phys. Rep.* **52** 265

Cohen R H and Rowlands G 1981 *Phys. Fluids* **24** 2295
Escande D F 1979 *Intrinsic Stochasticity in Plasmas* ed G Loval and D Gresillan (distributor: les Edition de Physique, Constaboeuf, Orsay)
Escande D F and Doveil F 1981 *J. Stat. Phys.* **26** 257
Holmes P 1979 *Phil. Trans. R. Soc.* **292** 419, 1963
Jorna S 1978 *AIP Conf. Proc. Series* **46** 173
Lichtenberg A J 1979 *Intrinsic Stochasticity in Plasmas* ed G Loval and D Gresillan (distributor: les Edition de Physique, Constaboeuf, Orsay)
Lorenz E N 1963 *J. Atmos. Sci.* **20** 448
McMillan P W 1970 *Topics in Modern Physics* (London: Adam Hilger) p 219
May R M 1976 *Theoretical Ecology* (Oxford: Blackwell) p 49
Nayfeh A 1973 *Perturbation Methods* (New York: Wiley) ch 23 p 228
Whiteman K J 1977 *Rep. Prog. Phys.* **40** 1055
Whittaker E T and Watson G N 1952 *Modern Analysis* p 491