Exact oscillation-centre transformations

1978 J. Phys. A: Math. Gen. 11 9

(http://iopscience.iop.org/0305-4470/11/1/006)

J. Phys. A: Math. Gen., Vol. 11, No. 1, 1978. Printed in Great Britain. © 1978

Exact oscillation-centre transformations

R L Dewar†

Department of Theoretical Physics, Research School of Physical Sciences, The Australian National University, Canberra, ACT, 2600, Australia

Received 2 May 1977, in final form 5 September 1977

Abstract. The topological constraints imposed on a canonical transformation by the requirement of continuous connection to the identity are considered. In the case of a particle interacting with a periodic potential, it is shown that a transformation to normal form (new Hamiltonian independent of position) is in general impossible using a transformation satisfying this requirement; the best that can be done being a transformation to a Hamiltonian which is almost everywhere in normal form, but which retains potential barriers of infinitesimal width to reflect trapped particles. A new Hamiltonian–Jacobi theory based on Lie operator techniques is presented and its relation to the usual theory established. It is suggested that the new method enables the motion of a particle in a random potential to be transformed into an almost Markovian random walk.

1. Introduction

We have recently (Dewar 1976) set up the mathematical apparatus for systematic perturbative generation of regular canonical transformations, and have sketched out some physical motivations and consequences of making such a transformation in the case of particle motion in a random potential. The basic idea in this application was that there are two fundamentally distinct parts of the particle motion: a 'coherent oscillatory' motion, and a 'purely stochastic' part, and that the purely stochastic part could be isolated by removing the coherent part through a near-identity canonical transformation to new generalised coordinates and momenta, called oscillation-centre variables.

A prescription was suggested for achieving this goal, based on the observation that the non-Markovian collision operator arising in renormalised perturbation theory acts as a *filter*, selecting out resonant terms and allowing them to appear as a residual interaction in the new Hamiltonian. Implicit in this prescription was the assumption that 'resonant' can be equated with 'stochastic'. This is not unreasonable since it is known (Zaslavskii and Chirikov 1971) that stochastic behaviour is associated with overlapping non-linear resonances; but the lack of precise definition of terms must be regarded as a definite defect of our 1976 paper. In the present paper we adopt a more cautious approach and study a model system where some of the questions associated with oscillation-centre theory can be answered without resort to perturbation expansions and with reasonable precision.

Of course, one can, in principle, solve a particle dynamics problem in *any* system of canonical coordinates. What we really want, however, is the 'optimum' system: one

† Present address: Plasma Physics Laboratory, Princeton University, Princeton, NJ 08540, USA.

in which no further simplification of the motion is possible. This leads us to a projective requirement: *the prescription for setting up an oscillation-centre transformation must be such that it leads to no further change when re-applied to the transformed system.* The filtering prescription of our previous paper does not meet this criterion. In this paper we present an alternative prescription, and verify in the case of a periodic stationary potential that the projective requirement is satisfied. Another reasonable requirement is that the new interaction Hamiltonian be *minimal* in some sense. Our new prescription also satisfies this, in that the L^2-norm of the interaction Hamiltonian is not simply minimal, but vanishes altogether for the optimum transformation.

A singly periodic potential is obviously not random, nor can the particle motion be truly stochastic. Nevertheless it does exhibit the property of allowing *irreversibility.* This is because two distinct parts of phase space (the positive and negative velocity parts) become mixed, in a narrow strip of phase space near zero velocity (the 'trapping region'). By irreversibility we mean that if the potential is adiabatically switched on and adiabatically switched off again, then an initially stationary distribution function will not, in general, return to its original value (Dewar 1972). This irreversible behaviour also explains the damping of the amplitude oscillations of a large-amplitude plasma wave (O'Neil 1965, Mazitov 1965). Irreversible behaviour occurs only in the trapping region—outside this region an initially stationary distribution function *will* return to its original value after adiabatic switching on and off, and the motion in this region is termed reversible.

We can now replace the notion of 'purely stochastic' with that of 'purely irreversible', and attempt a reasonable definition consistent with the projective requirement. Outside the trapping region there must certainly be no irreversible part of the motion, and we also wish to remove all the oscillatory information. Therefore we require that in oscillation-centre variables an untrapped particle simply free streams. That is, its momentum is a constant of the motion. In the trapping region we again remove all the oscillatory structure, but to give rise to irreversibility we cannot have the momentum as a strict constant of the motion. Instead we require the momentum to be a constant of the motion *almost everywhere,* except for regions of measure zero (which a particle intersects periodically) where the momentum changes instantaneously. Thus the particle is reflected to keep it in an orbit topologically equivalent to the true orbit, which is sufficient to retain irreversibility. This condition is fulfilled by our new prescription with a continuous Lie generating function. The difference between the true motion and the oscillation-centre motion defines (somewhat tautologically) the coherent part of the motion. We find that secular terms can be avoided in all parts of phase space, and that the projective requirement can be fulfilled.

In § 2 we review the Lie operator method and introduce the new prescription for constructing the generating function. In § 3 we introduce a non-perturbative method for constructing the Lie generating function, and in § 4 and § 5 we use it to construct generating functions for the cases of triangular and sinusoidal waves, respectively. Appendix 1 gives the details of the derivation of the result which allows us to relate conventional and Lie generating functions, and appendix 2 looks more closely than § 3 at the limiting process used to find the optimum transformation.

Although no very advanced mathematics has been used, an attempt has been made to introduce modern mathematical terminology where appropriate. We feel that this improves the precision of the presentation, and might hopefully encourage mathematicians to study the rather subtle question of the general existence of solutions to equation (5), the Hamilton–Jacobi equation for the Lie generating function.

2. Lie operators

We are concerned with Hamiltonian systems in which the Hamiltonian $H(x, p, t; \epsilon)$ is of the form $H_0+\epsilon H_1$ where H_0 is independent of x, and ϵ is a parameter ranging from 0 (unperturbed system) to 1 (fully interacting system). For simplicity, we restrict our attention to one-dimensional systems. We denote by (x^*, p^*) the canonical coordinates corresponding to the motion governed by H, and seek an area-preserving, one-to-one mapping of phase space onto itself. The area-preserving property is guaranteed for canonical transformations, but the second condition is quite restrictive and defines the subclass of *regular* canonical transformations (Sudarshan and Mukunda 1974). The physical point (x^*, p^*) is mapped on the nearby oscillation-centre position (x, p), whose motion is governed by the Hamiltonian $K(x, p, t; \epsilon)$. Actually, we work with the inverse map, $\c{C}(x, p)=(x^*, p^*)$, where $\c{C}$ (which depends parametrically on t and ϵ) is called the *clothing transformation* (the idea being that the oscillation-centre coordinates represent the position of the 'bare particle', whose motion is stripped of irrelevant structure). As well as for denoting the exact coordinates, we shall use x^* and p^* to denote the functions $x^*(x, p, t; \epsilon)$ and $p^*(x, p, t; \epsilon)$ defining the transformation.

We wish the oscillation-centre motion to be *topologically* the same as the true motion, so we require the clothing transformation and its inverse function, the oscillation-centre transformation, to be continuous (in fact, differentiable) functions of the canonical coordinates. (We shall sometimes be discussing discontinuous transformations, but only those which are formed as limiting cases of continuous mappings.) The conditions of being one-to-one and onto, and being bicontinuous are summarised by saying that $\c{C}$ is a *homeomorphism* (Roman 1975, p 221). As we require the stronger requirement of (piecewise continuous) differentiability, $\c{C}$ is in fact a C^1-*diffeomorphism*. Physically, this just means that we can only deform phase space like a rubber sheet, without folding or tearing it. Finally, we require that the transformation lie in the identity component of the full group of canonical transformations (Sudarshan and Mukunda 1974); that is, that it reduce continuously to the identity as $\epsilon \to 0$. These conditions are imposed because we hope eventually to develop a perturbation theory in which ϵ is treated as a small parameter.

The transformation $\c{C}$ can be effected (Dewar 1976) by the use of an invertible linear operator C_W, acting on functions of (x, p). We continue C_W from the identity operator at $\epsilon=0$ by the equation

$$\partial C_W/\partial\epsilon = L_W C_W \tag{1}$$

where L_W is the *Lie derivative* $W_p\partial/\partial x - W_x\partial/\partial p$, the function $W(x, p, t; \epsilon)$ being the *Lie generating function*. The subscripts x and p denote partial differentiation with respect to the first and second arguments, respectively. For both x and p, the transformation is of the form $\xi^* = C_W\xi$, where ξ denotes either x or p. From equation (1), it follows that x^* and p^* obey Liouville-like equations of the form

$$\xi^*_\epsilon = \{\xi^*, W\} \tag{2}$$

where subscript ϵ denotes the partial derivative with respect to the fourth argument and $\{\xi^*, W\}$ is the Poisson bracket between ξ^* and W. Since (2) can be solved by the method of characteristics, it is equivalent to a pair of ordinary differential equations. A sufficient condition that $\underset{\sim}{C}$ and its inverse function be differentiable everywhere is that W be twice differentiable (a less restrictive condition is that W_x and W_p satisfy a Lipschitz condition; cf, Ince 1956, § 3.22).

The oscillation-centre Hamiltonian K is given in terms of H by (Dewar 1976)

$$K = C_W\left(H + \int_0^\epsilon C_W^{-1} W_t \,\mathrm{d}\epsilon\right) \tag{3}$$

where subscript t denotes partial differentiation with respect to the third argument. A different but equivalent form can be obtained by operating on the left with C_W^{-1}, differentiation with respect to ϵ, and left multiplication with C_W:

$$W_t + L_K W = C_W H_\epsilon - K_\epsilon. \tag{4}$$

If we replace $K(x, p, t; \epsilon)$ in the above equation with a function $\bar{K}(p, t; \epsilon)$, independent of x, and denote the solution by $\hat{W}$ we obtain what we shall refer to as the *Hamilton–Jacobi equation for the Lie generating function*:

$$\hat{W}_t - L_{\bar{K}}\hat{W} = C_{\hat{W}} H_\epsilon - \bar{K}_\epsilon. \tag{5}$$

We can construct a solution of equation (5) as a power series in ϵ (cf Deprit 1969), $\bar{K}$ being chosen so as to make $\hat{W}$ non-secular. Provided the series converges, $\hat{W}$ is differentiable and can be identified with W. (More precisely, we need the series formed by the derivatives of the terms of the ϵ-series to converge uniformly.) Comparison of equations (4) and (5) then shows that $K = \bar{K}$, and hence p is a constant of the motion. That is, when the ϵ-series converges the oscillation-centre motion is just free streaming. This is the reversible case mentioned in § 1.

It is just when elementary perturbation theory breaks down that the interesting case of irreversibility occurs, which is why we have been led to seek analytically soluble models. In the vicinity of resonances it is clear that something must break down, since K cannot be strictly independent of x. This previously led us to assume (Dewar 1976) that no $\bar{K}$ can be formed to make $\hat{W}$ non-secular. In the present paper we show this assumption to be false: $\bar{K}$ *can* be found, but $\hat{W}$ is everywhere finite and non-secular (at least in the cases studied). The 'catch' is that $\hat{W}$ has a cornice or cusp singularity on the separatrix, where its derivative with respect to p is infinite, and therefore it is inadmissible as a Lie generating function. We can, however, find a family of smooth functions W_δ approximating $\hat{W}$ arbitrarily closely as a parameter $\delta \to 0$, and henceforth W will be taken to be one of these approximating functions.

As will be seen in appendix 2, the corresponding transformation $\underset{\sim}{C}_\delta$ becomes discontinuous in the limit $\delta \to 0$. Referring to figure 7 we see that, as we follow along the line $p = \text{constant}$, the point (x^*, p^*) traces out the contour $H = \text{constant}$, except that if the contour is closed this behaviour breaks down over a narrow interval in x, of width $O(\delta^2)$. In this interval the point (x^*, p^*) leaves the constant-H contour and jumps to the next trapping region. In the time-independent case equation (3) simply states that $K(x, p) = H(x^*, p^*)$, so we see that K is independent of x except for the narrow region of width $O(\delta^2)$ referred to above. We denote the departure from $\bar{K}$ in this region by $\tilde{K}$. Thus

$$K = \bar{K} + \tilde{K}, \tag{6}$$

where the support of $\tilde{K}$ (i.e. the region where it is non-zero) has zero measure in phase space in the limit as $\delta \to 0$. (This is assuming the system to be confined within a large but finite box, or that we are using a measure $d\mu = f\,dx\,dp$ such that phase space has a finite total measure.) Since $\tilde{K}$ is finite, the integral of $\tilde{K}^2$ over phase space tends to zero as $\delta \to 0$. Hence, using the L^2-norm (Roman 1975, p 399), $\|\tilde{K}\| \to 0$ as $\delta \to 0$. Similarly, the product of $\tilde{K}$ with a smooth test function integrates to zero as $\delta \to 0$. In generalised function theory such a function is said to vanish (Roman 1975, p. xxix). Borrowing from potential theory we term this a *removable singularity*. However its effect on particle dynamics does *not* vanish as $\delta \to 0$ because, in classical mechanics, a potential barrier can be effective in producing scattering, no matter how thin it is.

On the other hand, the corresponding singular contributions in equation (5) *are* negligible because they are integrated over in finding the solution. It is for this reason that we can satisfy the projective requirement. To see this, suppose K is obtained from equation (3) with W a close approximating function to the solution of equation (5). Now set $\epsilon = 1$, and regard $K = \bar{K} + \lambda\tilde{K}$ as the *old* Hamiltonian, that is, $\bar{K}$ replaces H_0, λ replaces ϵ, and $\tilde{K}$ replaces H_1. We now seek *new* oscillation-centre coordinates by solving equation (5) for a new $\hat{W}$ (denoted by $\hat{W}'$) and a new Hamiltonian $\bar{K}'$. Suppose that at $t = 0$ we take $\hat{W}' \equiv 0$ and $\bar{K}' \equiv \bar{K}$. Then the right-hand side of equation (5) is just $\tilde{K}$, which is a removable singularity in the limit $\delta \to 0$. That is, the initial conditions will be propagated on to later times as we integrate equation (5) along the characteristics of the left-hand side. This shows that attempting to reduce the motion still further by a second oscillation-centre transformation will have no effect.

3. Non-perturbative solution

The conventional way (Goldstein 1950) of performing a canonical transformation is by means of a generating function $F(x, p, t; \epsilon)$ such that the mapping is found by solving the equations

$$x = F_p(x^*, p, t; \epsilon), \qquad p^* = F_x(x^*, p, t; \epsilon) \tag{7}$$

for x^* and p^* as a function of x and p. The generating function obeys the general Hamilton–Jacobi equation

$$F_t(x^*, p\ \ t; \epsilon) + H(x^*, p^*, t; \epsilon) = K(x, p, t; \epsilon) \tag{8}$$

which defines F if K, and appropriate initial conditions, are specified. Because our x^*, p^*, H correspond to the usual q, p, H while our x, p, K correspond to the usual Q, P, K, we recognise F as a generating function of the second type (Goldstein 1950, p 241).

In appendix 1 we show that equation (7) leads to the equations

$$\begin{aligned} \partial p^*/\partial\epsilon &= -\{p^*, F_\epsilon(x^*, p, t; \epsilon)\} \\ \partial x^*/\partial\epsilon &= -\{x^*, F_\epsilon(x^*, p, t; \epsilon)\} \end{aligned} \tag{9}$$

where $\{f, g\}$ denotes the Poisson bracket between f and g. Provided that $F = xp$ (the generator of the identity transformation) when $\epsilon = 0$, that a real F exists for all values of its arguments, and that $F_{xp} \neq 0$ everywhere, we can, by comparing equations (1), (2) and (9), make the identification

$$W(x, p, t; \epsilon) = -F_\epsilon(x^*, p, t; \epsilon) \tag{10}$$

with x^* understood to be the function of x, p, t and ϵ obtained from equation (7). It can also be verified that differentiation of equation (8) with respect to ϵ gives rise to equation (4). As there is a large literature on the Hamilton–Jacobi equation, we have in equation (10) a potentially powerful means of finding non-perturbative solutions.

The one-point principal function $S(x^*, p, t; \epsilon)$, obtained as the solution to the Hamilton–Jacobi equation (8) with $K \equiv 0$, fails as a candidate for F. The main problem is that S has a secular component, so it cannot generate a near-identity transformation at all times. For instance, when $\epsilon = 0$ we have $S = xp - H_0(p)t$. One might hope that this problem could be solved by taking $K \equiv \bar{K}(p, t; \epsilon)$, with the function $\bar{K}$ chosen to remove the secularity. That is, that we could define F by the modified Hamilton–Jacobi equation

$$F_t(x^*, p, t; \epsilon) + H(x^*, p^*, t; \epsilon) = \bar{K}(p, t; \epsilon). \tag{11}$$

When H is independent of t we can choose $\bar{K}(p; \epsilon)$ such that time-independent solutions of equation (11) exist. With $F_t = 0$ equation (11) becomes the equation for the one-point characteristic function (Goldstein 1950, p 280), which is used for obtaining the action-angle variable transformation. Unfortunately this version of Hamilton–Jacobi theory also fails, because the action-angle transformation is not in general a homeomorphism from phase space onto itself. We shall however be able to find a simple modification of the conventional theory which provides a suitable generating function.

To see the problem more clearly we consider the case of a non-relativistic particle of unit mass, $H_0 = \frac{1}{2}p^2$, and suppose that H_1 is a periodic function of x, but is independent of p and t. In this case, phase space has the natural topology of a Euclidean metric space. It is consistent to assume that $F_t = 0$, so that equations (7) and (11) yield

$$F_x(x^*, p; \epsilon) = \pm 2^{1/2}(\bar{K}(p; \epsilon) - \epsilon H_1(x^*))^{1/2}. \tag{12}$$

The problem arises with trapped particles, for which $\bar{K}$ is less than the maximum of H_1, so that the argument of the square root in equation (12) passes through zero. Now it can be shown (see appendix 1) that the real function F_x must exist for all x^* and p if F is to generate a diffeomorphism, a condition which is not fulfilled by Hamilton's characteristic function as defined by equation (12).

The solution to this problem is suggested by equation (6). Instead of requiring K to be independent of x *everywere*, we allow it to have an x-dependent component $\tilde{K}$ such that

$$\tilde{K}(x, p; \epsilon) = (\Delta(p; \epsilon, \delta) + \epsilon H_1(x^*) - \bar{K}(p; \epsilon))_+ \tag{13}$$

where Δ is a positive function tending to zero as $\delta \to 0$ and the '+' operation is defined for an arbitrary function f by

$$(f(x))_+ = f(x)\theta f(x) \tag{14}$$

with $\theta(x)$ the Heaviside step function. The motivation for the choice (13) will become clear when we calculate F_x. For an untrapped particle $\tilde{K}$ vanishes identically as $\delta \to 0$, while even for trapped particles we shall show that its support is of zero measure in phase space in the limit as $\delta \to 0$. Nevertheless it plays the crucial role of reflecting a trapped particle.

Substituting $K = \bar{K} + \tilde{K}$ in equation (8) and assuming $F_t = 0$ we find that

$$F_x(x^*, p; \epsilon) = 2^{1/2} s_p[(\bar{K}(p; \epsilon) - \epsilon H_1(x^*) - \Delta)_+ + \Delta]^{1/2} \tag{15}$$

where s_p is +1 or −1 accordingly as $p \geq 0$ or $p < 0$. The crucial difference between equations (12) and (15) is that F_{xp} now exists for all x^*. The reason for introducing Δ is to ensure that F is single valued and that F_{xp} is strictly positive (see appendix 1). To make F reduce to the identity generator x^*p as $\epsilon \to 0$ uniformly for all x^*, we require that F differ from x^*p only by a periodic function of x^*. If the periodicity length of the system is λ, then this requirement implies

$$p = 2^{1/2} s_p \lambda^{-1} \int_0^{\lambda} [(\bar{K} - \epsilon H_1 - \Delta)_+ + \Delta]^{1/2} \, dx^* \tag{16}$$

which determines $\bar{K}$ implicitly. Note that, since the transformation is time independent, $\bar{K}$ is just the total energy of the system. Thus, in the limit $\delta \to 0$, λp is the familiar adiabatic invariant $\oint p \, dx$ of an untrapped particle, and is one half of the adiabatic invariant for a trapped particle.

Because F now satisfies all the appropriate conditions we can use equation (10) to solve equation (4), if K is defined by equations (13) and (16). We denote the limit of F as $\delta \to 0$ by the symbol $\hat{F}$. As we show in appendix 2, the support of $\tilde{K}$ is of measure zero in the limit $\delta = 0$, whence we can use equation (10) to associate $\hat{F}$ with $\hat{W}$, the solution of equation (5). Note that initial conditions have been assumed to be such that $\hat{F}_t$ and $\hat{W}_t$ are zero, which corresponds physically to adiabatic switching on of the interaction. Although this is not the most general solution, it is the most interesting because it leads to p being an adiabatic invariant.

4. Triangular waveform

The simplest analytically soluble waveform is the triangular waveform depicted in figure 1. Taking the wavelength $\lambda = 2\pi$, we can express this waveform analytically by

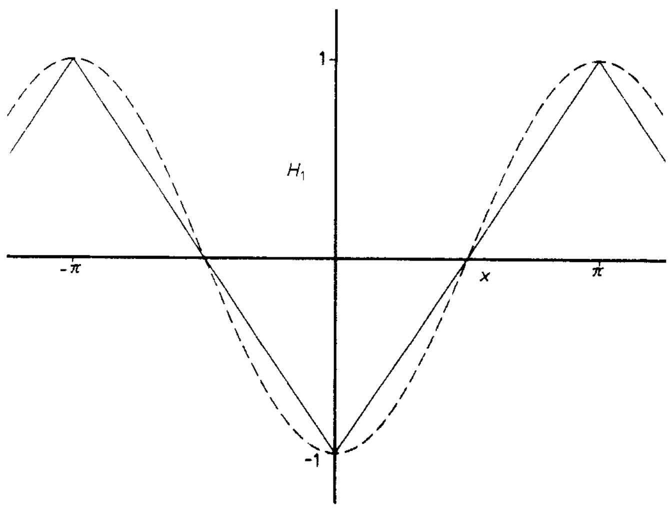

Figure 1. Two periodic, stationary interaction potentials plotted as functions of x. The triangular waveform of § 4 is shown by the full line, and the sinusoidal waveform analysed in § 5 is shown by the broken curve.

the equation

$$H_1(x)=2|[x]|/\pi-1, \tag{17}$$

where $[x]\equiv x-2n\pi$, with $n(x)$ an integer chosen such that $-\pi<[x]\leq\pi$.

Integrating equation (15) with $\delta=0$ we find

$$\hat{F}(x,p;\epsilon)=(x-[x])p+\frac{2^{1/2}\pi s_p\sigma_x}{3\epsilon}\left[(\bar{K}+\epsilon)^{3/2}-\left(\bar{K}+\epsilon-\frac{2\epsilon|[x]|}{\pi}\right)_+^{3/2}\right], \tag{18}$$

where σ_x is $+1$ or -1 according as $[x]\geq 0$ or $[x]<0$, and $\bar{K}(p;\epsilon)$ is determined from equation (16), which gives

$$|p|=2^{1/2}(3\epsilon)^{-1}[(\bar{K}+\epsilon)^{3/2}-(\bar{K}-\epsilon)_+^{3/2}]. \tag{19}$$

The function $\bar{K}$ is graphed in figure 2, together with H_0, and $\bar{K}$ for a sinusoidal wave.

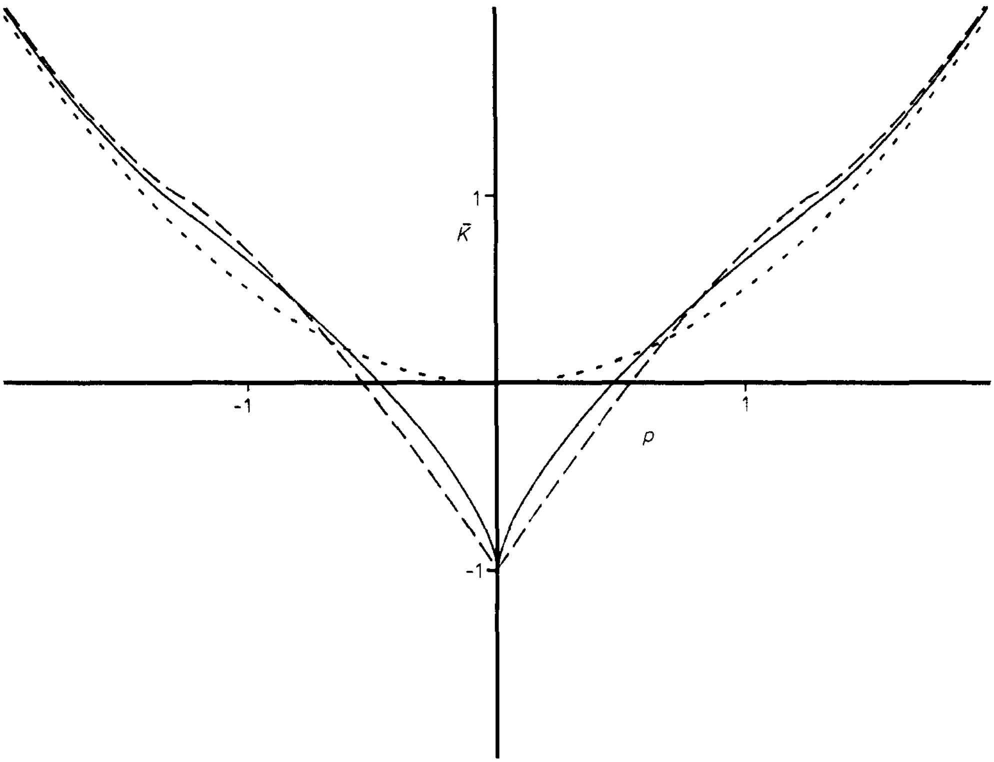

Figure 2. The spatially averaged part $\bar{K}$ of the oscillation-centre Hamiltonian as a function of p for the triangular waveform (full curve) and the sinusoidal waveform (broken curve). The fully interacting case $\epsilon=1$ is assumed. Note the non-analytic behaviour at $p=0$ and, more weakly, at $p=\pm p_s$ for both waveforms. The short dashes represent $H_0\equiv p^2/2$.

From equations (7) and (18) the transformation $\mathcal{C}_0$, the limit of $\mathcal{C}_\delta$ as $\delta\to 0$ is given by

$$\begin{aligned}[x^*]&=2^{-1/2}|p|'[x]\{2(\bar{K}+\epsilon)^{1/2}-2^{1/2}\epsilon|p|'|[x]|/\pi\}\\ p^*&=2^{1/2}s_p\{(\bar{K}+\epsilon)^{1/2}-2^{1/2}\epsilon|p|'|[x]|/\pi\}\end{aligned} \tag{20}$$

where $|p|'$ denotes $\partial|p|/\partial\bar{K}$. In figure 3(*b*) we show the image of the rectangular grid shown in figure 3(*a*) under the mapping (20), the images of the lines p = constant being just the phase-space orbits under H. In figure 3(*a*) the locations of the potential barriers are indicated by vertical bold lines at $x = \pm\pi$. In figure 4 we plot x^* as a function of x for three values of p, one less than $p_s = 4\epsilon^{1/2}/3$ (the value on the separatrix), one equal to p_s, and one greater than p_s.

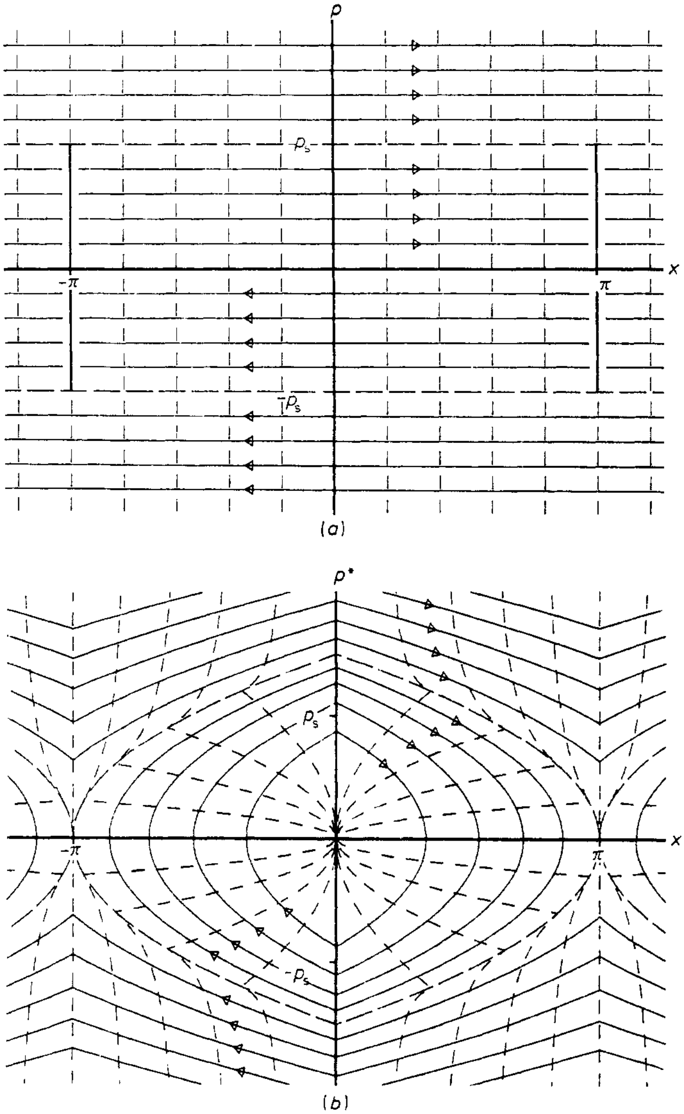

Figure 3. (*a*) A rectangular grid defined on phase space. The arrows indicate the oscillation-centre trajectories, while the 'cuts' at $x = \pm\pi$ show the positions of the thin potential barriers which keep the trapped-particle trajectories topologically circular. The separatrix is shown by the line with long dashes. (*b*) The image of the rectangular grid of (*a*) under the transformation C_0 defined by equation (20). The arrows indicate the exact trajectories.

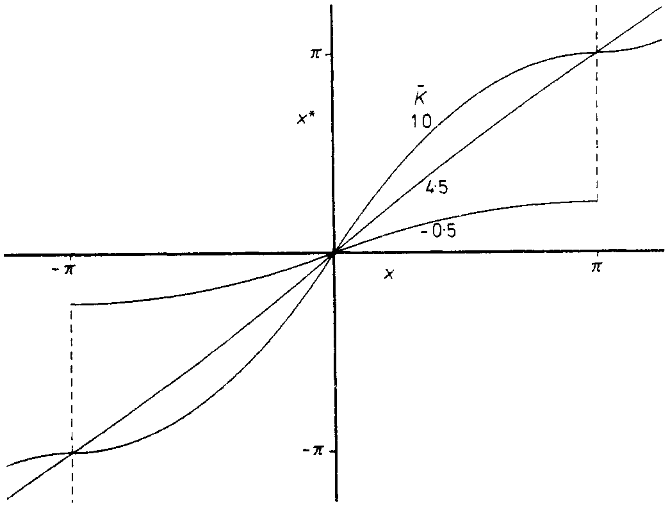

Figure 4. The exact position x^* as a function of the oscillation-centre position x for values of p such that $\bar{K} = -\epsilon/2$ (deeply trapped particle), $\bar{K} = \epsilon$ (particle on separatrix), and $\bar{K} = 4{\cdot}5\epsilon$ (untrapped particle). ($\epsilon = 1$ assumed.)

It is seen that, for $|p| < p_s$, the transformation is discontinuous at $x = (2n+1)\pi$, $n = 0, \pm 1, \pm 2 \ldots$, x^* jumping from $(2\pi n + \pi(\bar{K}+\epsilon)/2\epsilon)$ to $(2\pi(n+1) - \pi(\bar{K}+\epsilon)/2\epsilon)$. But this is just the range of x^* for which $\bar{K} \leq H_1$. Comparison with equation (13) shows that this is the region over which $\tilde{K}$ is non-zero. That is to say, the support of $\tilde{K}$ includes the set of points $\{(x, p) | x = (2n+1)\pi, |p| < p_s\}$, which make up the lines dividing one well from another. These infinitely thin potential barriers serve to reflect a trapped particle so that its orbit is rectangular, as indicated on figure 3(a). It is shown in appendix 2 how this may be regarded as the limiting case of a continuous motion. It is also shown there that the support of $\tilde{K}$ includes the line $p = 0$.

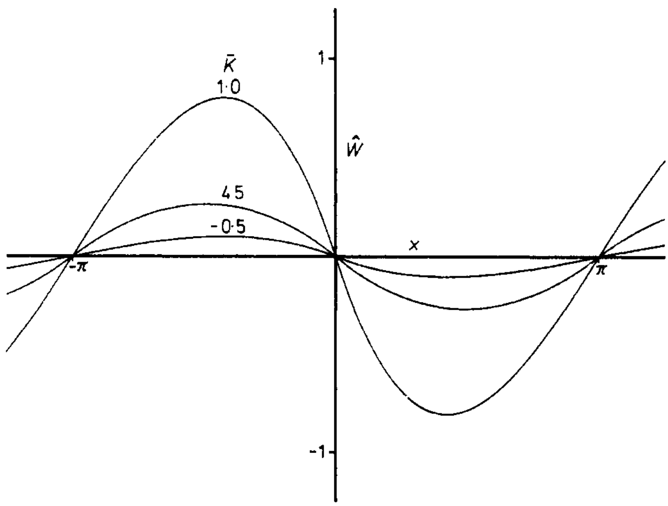

Figure 5. The Lie generating function $\hat{W}$ as a function of x for $\bar{K} = -\frac{1}{2}$, 1, and 4·5 ($\epsilon = 1$ assumed).

It remains to construct $\hat{W}$ using equations (10), (18)–(20). We find

$$\hat{W} = -2^{1/2}3^{-1}\pi s_p\sigma_x\{3\epsilon^{-1}(2^{-1/2}|p| - (\bar{K}-\epsilon)_+^{1/2})|[x]|/\pi$$
$$-3(\bar{K}+\epsilon)^{1/2}(|p|')^2[x]^2/\pi^2 + 2^{1/2}\epsilon(|p|')^3|[x]|^3/\pi^3\}. \quad (21)$$

Note that the factor ϵ^{-1} is multiplied by a term which goes to zero as $\epsilon \to 0$, since $\bar{K} \to H_0$ as $\epsilon \to 0$ (p being held fixed). Thus $\hat{W}$ is regular at $\epsilon = 0$.

We graph $\hat{W}$ as a function of x in figure 5, and as a function of p in figure 6. Note that $\hat{W}$ is a continuous function. However it is singular at $p = \pm p_s$ where the derivative with respect to p becomes infinite, and here $\hat{W}$ violates the differentiability condition (§ 2). This explains how $\hat{W}$ can generate the discontinuities in $\mathcal{C}_0$. Note also that $\hat{W}$ is smooth (in fact analytic) on the line $p = 0$, so that the only points where $\hat{W}$ is not 'sufficiently smooth' are along the lines $p = \pm p_s$. This is important because it shows that the singular points of $\hat{W}$ lie outside the support of $\tilde{K}$ in the limit $\delta \to 0$. Thus the singularities of $L_{\tilde{K}}W$ are 'removable', so that equation (4) does indeed go over into (5) as $\delta \to 0$.

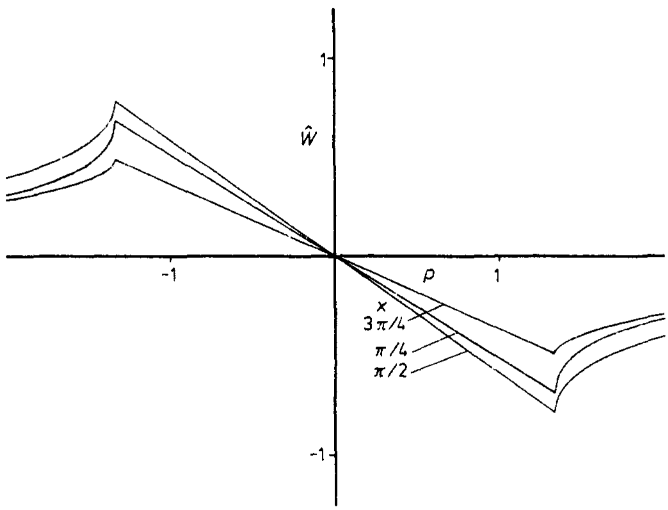

Figure 6. The Lie generating function $\hat{W}$ as a function of p for $x = \pi/4$, $\pi/2$, and $3\pi/4$. Note the singular behaviour at $p = \pm p_s$.

5. Sinusoidal waveform

While the triangular waveform is simple, it is rather unphysical because of the discontinuous behaviour of its x derivative. The simplest infinitely differentiable periodic potential is the sinusoidal wave

$$H_1(x) = -\cos x, \quad (22)$$

which is also depicted in figure 1. From equation (15) the conventional generating function in the limit $\delta = 0$ is found to be

$$\hat{F}(x, p; \epsilon) = 2^{3/2}s_p(\bar{K}+\epsilon)^{1/2}\,\mathrm{Re}\,E(\tfrac{1}{2}x, \kappa), \quad (23)$$

where $E(\phi, \kappa)$ is the elliptic integral of the second kind (Gradshteyn and Ryzhik

1965), with κ determined from the relation $\kappa^2 = 2\epsilon/(\bar{K}+\epsilon)$. For trapped particles we have $\kappa > 1$, but equation (23) remains valid provided we take the real part as indicated.

The relation between the average oscillation-centre Hamiltonian and p is

$$|p| = 2^{3/2}\pi^{-1}(\bar{K}+\epsilon)^{1/2}\,\mathrm{Re}\,\underset{\sim}{E}(\kappa), \tag{24}$$

$\underset{\sim}{E}$ being the complete elliptic integral of the second kind. The value of $|p|$ on the separatrix is $p_s = 4\epsilon^{1/2}/\pi$, which is close to the value $4\epsilon^{1/2}/3$ for the triangular wave. In fact figure 2 shows that the behaviour of $\bar{K}(p;\epsilon)$ is very similar for the two waveforms, except near the points $p = 0$ and $p = \pm p_s$.

The transformation generated by $\hat{F}$ is, for untrapped particles ($|p| > p_s$)

$$\begin{aligned} \frac{x^*}{2} &= \mathrm{am}\, u = \frac{x}{2} + 2\sum_{n=1}^{\infty}\frac{q^n}{1+q^{2n}}\frac{\sin nx}{n}, \\ \frac{|p^*|}{2\epsilon^{1/2}} &= \frac{\mathrm{dn}\, u}{\kappa} = \frac{\pi}{2\kappa\underset{\sim}{K}} + \frac{2\pi}{\kappa\underset{\sim}{K}}\sum_{n=1}^{\infty}\frac{q^n}{1+q^{2n}}\cos nx, \end{aligned} \tag{25}$$

where $u \equiv \underset{\sim}{K}x/\pi$, $q \equiv \exp(-\pi\underset{\sim}{K}'/\underset{\sim}{K})$ (Gradshteyn and Ryzhik 1965). For the trapped particles ($|p| < p_s$) we find $\sin\frac{1}{2}x^* = \kappa_1\,\mathrm{sn}\, u_1$, so

$$\begin{aligned} \frac{x^*}{2} &= \sin^{-1}\left(\frac{2\pi}{\underset{\sim}{K}_1}\sum_{n=1}^{\infty}\frac{q_1^{n-\frac{1}{2}}}{1+q_1^{2n-1}}\sin[\tfrac{1}{2}(2n-1)x]\right), \\ \frac{|p^*|}{2\epsilon^{1/2}} &= \kappa_1\,\mathrm{cn}\, u_1 = \frac{2\pi}{\underset{\sim}{K}_1}\sum_{n=1}^{\infty}\frac{q_1^{n-\frac{1}{2}}}{1+q_1^{2n-1}}\cos[\tfrac{1}{2}(2n-1)x], \end{aligned} \tag{26}$$

where subscript 1 means that the same definition as for untrapped particles is to be used, except that κ is to be replaced by $\kappa_1 \equiv \kappa^{-1}$. As for the triangular waveform, the transformation is discontinuous in the trapping region.

Using equation (10) we find, for $|p| > p_s$

$$\hat{W} = -2s_p\kappa^{-1}\epsilon^{-1/2}Z(u,\kappa) = -\frac{2s_p}{\kappa\epsilon^{1/2}}\frac{2\pi}{\underset{\sim}{K}}\sum_{n=1}^{\infty}\frac{q^n}{1-q^{2n}}\sin nx, \tag{27}$$

where $Z(u,\kappa)$ is Jacobi's zeta function (Abramowitz and Stegun 1964, p 576). For $|p| < p_s$ we find

$$\hat{W} = -2s_p\epsilon^{-1/2}Z(u_1,\kappa_1) = -\frac{2s_p}{\epsilon^{1/2}}\frac{2\pi}{\underset{\sim}{K}_1}\sum_{n=1}^{\infty}\frac{q_1^n}{1-q_1^{2n}}\sin nx \tag{28}$$

with $u_1 \equiv \underset{\sim}{K}_1 x/\pi$ and so on, as in equation (26).

In order to understand better the behaviour near the separatrix, we have derived the approximate results for particles with $|p|$ slightly greater than p_s, and $|[x]| < \pi$:

$$\begin{aligned} [x^*] &\approx 2\,\mathrm{gd}\, u \\ p^* &\approx 2\epsilon^{1/2}s_p\,\mathrm{sech}\, u \\ \hat{W} &\approx -2s_p\epsilon^{-1/2}\left(\tanh u - \frac{[x]}{\pi}\right) \\ u &\approx \pi^{-1}[x]\ln(4/\kappa') \end{aligned} \tag{29}$$

valid for small $\kappa' \approx [4\zeta/\ln(4/\zeta)]^{1/2}$, where $\zeta \equiv (|p|-p_s)/p_s$. Here $\mathrm{gd}\, u$ is the Gudermannian function $2\tan^{-1}(\exp u)-\frac{1}{2}\pi$ (Abramowitz and Stegun 1964). As $|p|\to p_s$, $\kappa'\to 0$, and $[x^*]$ approaches the step function $\pi|x|/x$, while $p^*\to 0$. That is, nearly all points on the lines $|p|=p_s$ are projected by the map $\mathcal{C}_0$ onto the stagnation points $x^*=(2n-1)\pi$, $p^*=0$. This is due simply to the fact that, for any smooth waveform, a particle on the separatrix takes an infinite time to reach a crest of the wave. Since the oscillation-centre velocity $\bar{K}_p$ is a constant of the motion (between reflections), the oscillation-centre interval Δx is proportional to the time Δt spent by a particle in traversing this part of its orbit. This weighting implies that an infinitesimal interval Δx^* in the neighbourhood of a stagnation point corresponds to a finite interval Δx in oscillation-centre picture. A corollary of this is that $\bar{K}_p\to 0$ as $p\to p_s$.

Note that $\hat{W}$ approaches a sawtooth profile as $|p|\to p_s$. The complicated nature of the behaviour on the separatrix in the sinusoidal case is another reason for studying the simpler triangular waveform first. As for the triangular waveform $\hat{W}$ is continuously differentiable on the line $p=0$, although, unlike the previous case, it is not analytic there.

6. Discussion

Although the method we have used in this paper to obtain Lie generating functions is limited to the rare cases where the Hamilton-Jacobi equation for Hamilton's characteristic function can be solved analytically, it is important because it has allowed us to establish the *existence* of solutions to the 'Hamilton–Jacobi equation for the Lie generating function', equation (5), outside the radius of convergence of the elementary ϵ-series. Moreover, we have shown that the existence of cusp-type singularities in the generating function leads to a residual interaction in the form of thin potential barriers which preserve the topological character of the 'invariant tori' to which the trapped-particle orbits are confined. More interesting would be a 'non-integrable' system in which the invariant tori are disrupted, leading to 'stochastic' regions of phase space (Zaslavskii and Chirikov 1971). If equation (5) can also be solved for such a system, this raises a very exciting prospect. It means that we can find a new Hamiltonian for which scattering occurs only off potential barriers of infinitesimal width, the motion elsewhere being in straight lines. Since the scattering events are instantaneous in this picture, it is reasonable to suppose that the process is describable by a Boltzmann-like equation. Speculating further, we conjecture that the highly non-linear nature of the transformation will make the positions of the potential barriers extremely unpredictable in space and time, so that the oscillation-centre motion will be essentially Markovian.

The elementary perturbation theory we have referred to is the expansion in terms of the bare propagator $(\partial_t+L_{H_0})^{-1}$. There may be more sophisticated perturbation theories which converge everywhere, and even if this is not the case there are techniques for handling divergent series. The existence of an analytic solution will be very useful for checking the validity of such techniques. Even elementary perturbation theory is still useful, of course, for calculating phase-space averages, such as densities and currents, since the convergence is much improved by the integration over phase space.

Another promising approach towards establishing the nature of the general solution of equation (5) is that of computer solution, since the Lie transform method lends itself to numerical implementation.

Appendix 1

In this appendix we derive equation (9) from equation (7). For generality we consider a Hamiltonian with N degrees of freedom, and use the Einstein summation convention. The general form of equation (7) is

$$q_i = \left(\frac{\partial F}{\partial p_i}\right)_{q^*,p} \tag{A.1}$$

$$p_i^* = \left(\frac{\partial F}{\partial q_i^*}\right)_{q^*,p} \tag{A.2}$$

where F denotes $F(q^*, p, t; \epsilon)$, and the 'thermodynamic' notation for partial differentiation is used, in which the variables to be held constant during the differentiation are indicated as subscripts to the parentheses surrounding the derivatives. Thus, the subscripts q^*, p indicate that the variables q_j^* and p_j $(j = 1, 2, \ldots, N)$ are to be held fixed (with the exception of the variables with respect to which the derivative is being taken). As it is always clear whether or not ϵ and t are fixed, we have no need to indicate this explicitly.

Differentiation of equation (A.1) with respect to ϵ, with q and p being fixed, yields, on using (A.2) to eliminate $\partial F/\partial q_j^*$,

$$\left(\frac{\partial p_j^*}{\partial p_k}\right)_{q^*,p}\left(\frac{\partial q_j^*}{\partial \epsilon}\right)_{q,p} = -\left(\frac{\partial^2 F}{\partial \epsilon\, \partial p_k}\right)_{q^*,p}.$$

Contracting with $(\partial p_k/\partial p_i^*)_{q^*,p^*}$ we find

$$\left(\frac{\partial q_i^*}{\partial \epsilon}\right)_{q,p} = -\left(\frac{\partial}{\partial p_i^*}\left(\frac{\partial F}{\partial \epsilon}\right)_{q^*,p}\right)_{q^*,p^*}.$$

We recognise the right-hand side as the Poisson bracket $-\{q_i^*, F_\epsilon\}_{q^*,p^*}$ taken with respect to q^* and p^*. But as Poisson brackets are invariant under canonical transformation (Goldstein 1950, p 254) we have found the first of the sought-for equations,

$$\left(\frac{\partial q_i^*}{\partial \epsilon}\right)_{q,p} = -\{q_i^*, F_\epsilon(q^*, p, t; \epsilon)\}_{q,p}. \tag{A.3}$$

To proceed further, we need the lemma

$$\left(\frac{\partial q_i^*}{\partial \epsilon}\right)_{q,p} = -\left(\frac{\partial q_i^*}{\partial q_k}\right)_{q,p}\left(\frac{\partial^2 F}{\partial \epsilon\, \partial p_k}\right)_{q^*,p} \tag{A.4}$$

which is proved by differentiating (A.1) with respect to ϵ, using (A.1) to eliminate $\partial F/\partial p_i$, and contracting with $(\partial q_i^*/\partial q_k)_{q,p}$.

We now differentiate (A.2) with respect to ϵ and use the lemma (A.4) to find

$$\left(\frac{\partial p_i^*}{\partial \epsilon}\right)_{q,p} = \left(\frac{\partial F_\epsilon}{\partial q_i^*}\right)_{q^*,p} - \left(\frac{\partial^2 F}{\partial q_i^*\, \partial p_j^*}\right)_{q^*,p}\left(\frac{\partial q_j^*}{\partial q_k}\right)_{q,p}\left(\frac{\partial F_\epsilon}{\partial p_k}\right)_{q^*,p}.$$

In the last term of the above we recognise the factor $(\partial p_i^*/\partial q_k)_{q,p} = -(\partial p_k/\partial q_i^*)_{q^*,p^*}$, by the theorem on invariance of Poisson brackets. Thus the right-hand side of the above is $(\partial F_\epsilon/\partial q_i^*)_{q^*,p^*}$, whence

$$\left(\frac{\partial p_i^*}{\partial \epsilon}\right)_{q,p} = -\{p_i^*, F_\epsilon(q^*, p, t; \epsilon)\}_{q,p} \tag{A.5}$$

which is the second of the equations to be proved.

If equations (A.3) and (A.5) hold for all q and p, with $F(q^*, p, t; 0) \equiv q_i^* p_i$ and with F_ϵ being a sufficiently smooth function of q and p that the solutions q^* and p^* depend differentiably on the initial values q and p (Ince 1956), then we can make the identification (10) uniquely, to within a constant function of q and p.

It can be shown (cf Sudarshan and Mukunda 1974, p 58) that the differentiability condition is satisfied if and only if

$$\det \|F_{q_i p_j}(q^*, p, t; \epsilon)\| \neq 0$$

for all values of the arguments. (For the one-dimensional case of this paper, this becomes $F_{xp} \neq 0$). If the determinant is never infinite, the condition of continuous connection with the identity implies that it is strictly positive ($F_{xp} > 0$).

As $F(q^*, p, t; \epsilon)$ depends only implicitly on q, the significance of the condition that it exist for *all* q is not immediately apparent. An intuitive understanding of the situation in the one-dimensional Cartesian case can be obtained by considering the mesh formed by the lines $x = \text{constant}$, $p = \text{constant}$ (cf figure 7). The continuity condition and the area-preserving nature of the transformation imply that the lines $p = \text{constant}$ never touch or break (and similarly for the lines $x = \text{constant}$), and that they cover the entire phase space. Thus x^* and p must be a suitable coordinate system over the whole of phase space and $F(x^*, p, t; \epsilon)$ must exist for all x^* as well as all x. Note that F can become multivalued if the lines are sufficiently twisted (signalled by F_{xp} becoming infinite), but that it must have an odd number of branches. Since the square root in equation (12) leads to an even number of branches, we must modify (12) in such a way as to avoid the branch point. Thus we have been led to introduce

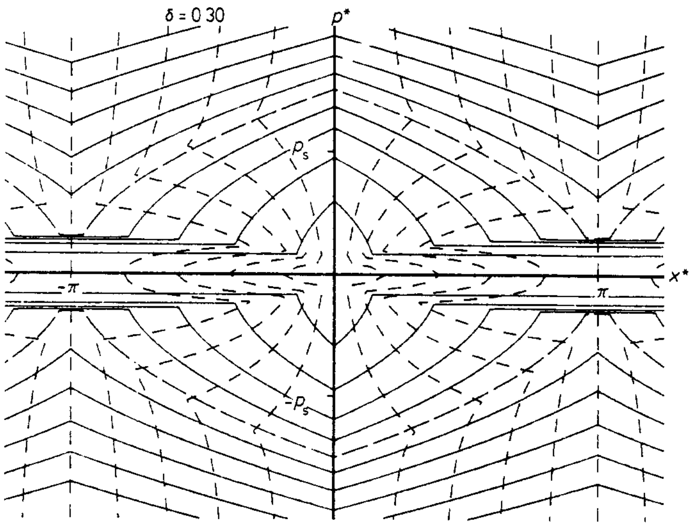

Figure 7. The image of the rectangular grid of figure 3(a) under the map $\mathcal{C}_\delta$. The case $\delta = 0{\cdot}3$ is shown.

the positive function Δ into equation (15) to keep F_x single valued. As will be seen in appendix 2, Δ can be chosen so as to keep F_{xp} strictly positive. This is especially significant at the separatrix $|p| = p_s$, where (12) leads to F_{xp} being zero for all x^* (in the case of smooth H_1).

Appendix 2

In this appendix we investigate the properties of the generating function F defined by equations (15) and (16) for small but finite δ. For definiteness we consider the triangular waveform of § 4. Without loss of generality we can restrict initial attention to the half period $0 \leqslant x^* \leqslant \pi$. It is first of all necessary to determine the value of x^* at which $\tilde{K}$ becomes non-zero. Denoting this value by $X^*(p; \epsilon, \delta)$ we find from equation (13)

$$X^* = \min[\pi(\bar{K} + \epsilon - \Delta)/2\epsilon, \pi]. \tag{A.6}$$

The integral (16) can now be performed by noting that λ can be replaced by π, owing to the symmetry about $x^* = \pi$, and then by breaking the interval $[0, \pi]$ into the subintervals $[0, X^*]$ and $[X^*, \pi]$, in which the integrands are respectively $(\bar{K} - \epsilon H_1)^{1/2}$ and $\Delta^{1/2}$. We find

$$|p| = \begin{cases} 2^{1/2}(3\epsilon)^{-1}(\bar{K}+\epsilon)^{3/2} + (2\Delta)^{1/2}[1-(\bar{K}+\epsilon)/2\epsilon] + 2^{1/2}(6\epsilon)^{-1}\Delta^{3/2} \\ 2^{1/2}(3\epsilon)^{-1}[(\bar{K}+\epsilon)^{3/2} - (\bar{K}-\epsilon)^{3/2}] \end{cases} \tag{A.7}$$

for $|p| < p_c$ and $|p| \geqslant p_c$, respectively. Here p_c, the momentum at which $X^* = \pi$, is implicitly defined by

$$p_c \equiv 2^{1/2}(3\epsilon)^{-1}[(2\epsilon + \Delta_c)^{3/2} - \Delta_c^{3/2}] \tag{A.8}$$

with $\Delta_c \equiv \Delta(p_c; \epsilon, \delta)$. In terms of X^* and p_c the support $\mathscr{S}_\delta$ of $\tilde{K}$ is defined as the set

$$\mathscr{S}_\delta = \{(x, p) | \, |[x^*]| > X^*(p; \epsilon, \delta), |p| < p_c\}.$$

From equation (15) we see that Δ cancels outside the support of $\tilde{K}$ (that is for $(x, p) \in \mathscr{S}_\delta^c$, where $\mathscr{S}_\delta^c$ is the complement of $\mathscr{S}_\delta$). In this region the only dependence of F on Δ is implicitly, through $\bar{K}$. Thus equation (20) remains valid in $\mathscr{S}_\delta^c$, provided (19) is replaced by (A.7). Within $\mathscr{S}_\delta$, on the other hand, we have $F_x = \Pi(p; \epsilon, \delta)$, where

$$\Pi \equiv (2\Delta)^{1/2} s_p. \tag{A.9}$$

By matching to the solution in $\mathscr{S}_\delta^c$ we have, if x is near $(2n-1)\pi$, $(n = 0, \pm 1, \pm 2, \ldots)$,

$$F = (2n-1)\pi p + \Pi[x^* - \pi] \tag{A.10}$$

for $(x, p) \in \mathscr{S}_\delta$. We shall denote the transformation so determined by the symbol $\mathcal{C}_\delta$.

So far we have refrained from specifying Δ beyond saying that it is positive and that $\Delta \to 0$ as the parameter $\delta \to 0$. We have also assumed for simplicity that it is independent of x^*. It is more convenient to specify Π, which is somewhat arbitrary with the exception of its behaviour near $p = 0$. The reason that $p = 0$ is special is that s_p is discontinuous there, and Π must be specially chosen to remove the effect of this. With Π independent of x^* the only way we can ensure continuity is by demanding that $\mathcal{C}_\delta$ be locally equal to the identity transformation in the neighbourhood of $p = 0$. This is ensured by including the x axis in $\mathscr{S}_\delta$ and by demanding $\Pi_p = 1$ at $p = 0$. In order to ensure that F_{xp} is strictly positive within $\mathscr{S}_\delta$, we must also demand that Π_p be positive for all p. A simple choice satisfying these conditions is

$$\Pi = \epsilon^{1/2}\,\delta \tan^{-1}(p/\epsilon^{1/2}\delta). \tag{A.11}$$

With this choice we can give a more explicit description of $\mathscr{S}_\delta$ as the set

$$\mathscr{S}_\delta = \left\{(x,p)\middle|(\pi - |[x]|) < \frac{\epsilon\delta^2(\pi - X^*)}{p^2+\epsilon\delta^2},\ |p| < p_c\right\}. \tag{A.12}$$

It is seen that as $\delta \to 0$, $\mathscr{S}_\delta$ collapses onto the union of the lines $p = 0$ and $\{(x, p)|x = (2n+1)\pi, |p| < p_s\}$, $n = 0, \pm 1, \pm 2, \ldots$.

In figure 7 we show the image of a rectangular grid under the map $Ç_\delta$. As $\delta \to 0$ the situation approaches that depicted in figure 3(*b*). The trapping regions are linked by 'tubes', into which the lines p = constant are constricted, and which become thinner and thinner as $\delta \to 0$. The area preserving property implies that the lines x = constant are squeezed out the ends of these tubes, thus giving rise to the discontinuity seen in figure 4. In figure 8 we show the image of a rectangular grid under the map $Ç_\delta^{-1} \circ Ç_0$. The images of the lines p = constant are just the phase-space orbits under the Hamiltonian $K(x, p; \epsilon, \delta)$, the region where they deviate from straight lines being the support $\mathscr{S}_\delta$ of $\tilde{K}(x, p; \epsilon, \delta)$.

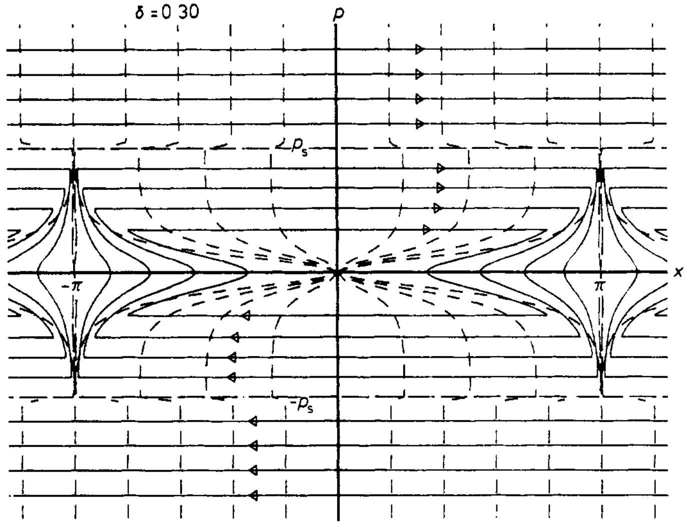

Figure 8. The image of the rectangular grid of figure 3(a) under the map $Ç_\delta^{-1} \circ Ç_0$, or of the curvilinear grid of figure 3(b) under the map $Ç_\delta^{-1}$. The case $\delta = 0{\cdot}3$ is shown. The arrows indicate the oscillation-centre trajectories under the Hamiltonian $\bar{K} + \tilde{K}$ determined by equations (13) and (16).

Acknowledgments

It is a pleasure to acknowledge useful and stimulating conversations with Professor K J Le Couteur, Professor H A Buchdahl, Dr T K Donaldson, Dr L F Peterson, Professor A N Kaufman and Mr J Cary.

References

Abramowitz M and Stegun I A 1964 *Handbook of Mathematical Functions* (Washington: NBS)
Deprit A 1969 *Celestial Mech.* **1** 12–30
Dewar R L 1972 *Phys. Fluids* **15** 712–4
—— 1976 *J. Phys. A: Math. Gen.* **9** 2043–57
Goldstein H 1950 *Classical Mechanics* (Reading, Mass.: Addison-Wesley)
Gradshteyn I S and Ryzhik I M 1965 *Table of Integrals, Series and Products* (New York: Academic)
Ince E L 1956 *Ordinary Differential Equations* (New York: Dover) chap. 3
Mazitov R K 1965 *Zh. Prikl. Mekh. Tekh. Fiz.* **1** 27–31 (1965 *J. Appl. Mech. Tech. Phys.* **1** 22–5)
O'Neil T 1965 *Phys. Fluids* **8** 2255–62
Roman P 1975 *Some Modern Mathematics for Physicists and Other Outsiders* (New York: Pergamon)
Sudarshan E C G and Mukunda N 1974 *Classical Dynamics: A Modern Perspective* (New York: Wiley)
Zaslavskii G M and Chirikov B V 1971 *Usp. Fiz. Nauk* **105** 3–39 (1972 *Sov. Phys.–Usp.* **14** 549–672)