Improved accuracy of the Birkhoff-Gustavson normal form and its convergence properties

1992 J. Phys. A: Math. Gen. 25 5311

(http://iopscience.iop.org/0305-4470/25/20/013)

J. Phys. A: Math. Gen. 25 (1992) 5311–5327. Printed in the UK

Improved accuracy of the Birkhoff–Gustavson normal form and its convergence properties

M Kalužat and M Robnik

Centre for Applied Mathematics and Theoretical Physics, University of Maribor, Krekova 2, 62000 Maribor, Slovenia

Received 4 December 1991, in final form 1 April 1992

Abstract. We study the generalized Hénon–Heiles system by systematic exploration and analysis of the Birkhoff–Gustavson normal form and the associated formal integral of motion up to and including the 14th order, which is one order higher than published before. At low energies the formal integral of motion is still an excellent approximation to the exact integral. The convergence properties of the formal integral have been analysed in regular and irregular regions. Strictly convergent behaviour is found in some regions of chaotic motion. No obvious example of divergent behaviour is found. Regions of strict convergence correspond either to regular motion, or to weakly unstable chaotic motion with short-time clustering, characterized by a small value of the finite-time analogue of the Lyapunov exponent.

1. Introduction

A typical classical Hamiltonian system exhibits three possible regimes of motion: regular, chaotic and mixed. In the mixed regime, there are regions of phase space where the motion is regular and regions where it is chaotic. Regular trajectories exhibit weak dependence on initial conditions and are predictable while trajectories in chaotic regions show sensitive dependence on initial conditions and on roundoff errors and are thus unpredictable in the long run. One way to investigate the degree of chaoticity or unpredictability of trajectories is to study Lyapunov exponents. A Lyapunov exponent is defined as an asymptotic property of orbits, namely as an average exponent of the divergence of two exact neighbouring trajectories, averaged over *infinite time*. As defined, it is difficult to calculate reliably, just because of the sensitive dependence on initial conditions. It is not directly related to the short-time behaviour of trajectories, which is the only behaviour predictable without enormously increased numerical efforts.

The Birkhoff–Gustavson normal form (Birkhoff 1927, Gustavson 1966) is a possible method of perturbative investigation of the motion of Hamiltonian systems in the neighbourhood of an equilibrium point. Using a series of canonical transformations, a series of approximate integrals of motion, which are polynomial functions of coordinates and momenta, are constructed order by order in coordinates and momenta. The series of successive approximations is asymptotic around the equilibrium

† Present address: Institut für Theoretische Physik, Universität Heidelberg, Philosophenweg 19, W-6900 Heidelberg, Federal Republic of Germany.

0305-4470/92/205311+17$07.50

point and is in general expected to diverge. Nevertheless, the first few terms should approximate the exact integrals of motion in the regular regime. Furthermore, we cannot exclude the possibility that the Birkhoff–Gustavson normal form is useful in studying the short-time dynamics even in the chaotic regime. To investigate this in relation to the convergence properties is the primary motivation of the present work.

A related work with a similar goal has been published by Shirts and Reinhardt (1982), in which they have analysed the convergence and divergence properties of the formal integrals as calculated by the method of the Birkhoff–Gustavson normal form. They investigate the Padé approximants to the formal integral and examine their singularity structure. Their results are interesting but still largely qualitative and the method is probably more complex than necessary to draw the conclusions. In the present paper we re-examine similar questions by using an approach that is as elementary as possible and obtain somewhat stronger conclusions.

The goals of the present paper can be summarized as follows.

(i) To provide a general algorithm encoded in a computer program (in REDUCE) for an *algebraic computation* of the Birkhoff–Gustavson normal form. To our knowledge the only published and available computational scheme is that by Gustavson (1966) and by Giorgilli (1979), both being written in FORTRAN, and therefore subject to numerical errors as a result of the accumulation of roundoff errors. We will show by comparing our results with Gustavson (1966) that sometimes the resulting errors can be very large, so that the general algebraic computation is undoubtedly not only more modern and elegant, but also much more accurate (since all operations are exact and the numerical evaluation—if necessary—takes place only at the final step) and is thus superior. Our scheme (cf Kaluža 1991) for calculating the Birkhoff–Gustavson normal form is as general as possible, i.e. the number of freedoms and the highest order desired are arbitrary.

(ii) To use this program in order to calculate the normal form and the formal integral of motion for the (generalized) Hénon-Heiles (1964) system to the highest possible order (on the given computer). Indeed, we succeeded in going one order beyond the results of Shirts and Reinhardt (1982), i.e. we calculated all orders up to and including the 14th.

(iii) To analyse the convergence properties of the classical normal form and of the formal integral of motion and to explain in this manner the striking and until now not properly understood difference between the semiclassical quantization methods of Swimm and Delos (1979), Noid and Marcus (1977) and Jaffe and Reinhardt (1982) on the one side, and of Robnik (1984) on the other side. These authors largely agree among each other, although Swimm and Delos (1979) definitely have errors in their calculations as explained in detail by Robnik (1984). The method of algebraic quantization by Robnik (1984) yields results which are much more accurate and almost exact. This is surprising as Swimm and Delos quantized the normal form including terms up to at least eighth order, while Robnik's method was applied to just the fourth-order normal form, and yet it yields much better results. The difference in the quality of the results can be explained either (A) by the fact that Robnik's method is considerably more accurate, or (B) by the fact that at orders higher than the fourth order the classical normal form already diverges (thereby also spoiling the semiclassical results), or else (C) by the fact that there is some error in the classical and/or semiclassical calculations of other workers. We carefully checked and tested our algorithm and the computer program and compared the results with Gustavson (1966). Apart from the (occasionally considerable) roundoff errors (see section 3) a

very good agreement was found, and we are absolutely confident that our program is flawless. By examining the convergence properties we exclude the option (B) and conclude that (A) is correct but cannot exclude (C) as well. We emphasize the importance of the various checks of accuracy of such encoded algorithms, and do not exclude the possibility that such tests were not always satisfactory in the papers of other workers mentioned above.

(iv) Finally—as already explained—to use our computational scheme and the computer program (in REDUCE) as a tool to generate the normal form and the formal integral of motion for the Hénon-Heiles Hamiltonian up to and including the 14th order, which is one order higher than previously published. By using elementary methods we analyse the convergence properties of the formal integral and relate them to the properties of motion (regularity and chaoticity). Our main conclusions by using elementary methods are similar to but somewhat stronger than those of Shirts and Reinhardt (1982).

2. The general procedure

2.1. The Birkhoff–Gustavson normal form

The general presentation of the Birkhoff–Gustavson normal form will be introduced along the lines of Gustavson (1966) and Robnik (1984).

Consider a *polynomial* Hamiltonian of a system with N degrees of freedom

$$H(x,y)=\sum_{j=2}^{\infty}H^{(j)}(x,y) \tag{1}$$

where $x=(x_1,\ldots,x_N)$ are the coordinates and $y=(y_1,\ldots,y_N)$ are the momenta. The harmonic part of (1) is assumed to be

$$H_0(x,y)=H^{(2)}(x,y)=\sum_{k=1}^{N}\frac{\omega_k}{2}(x_k^2+y_k^2) \tag{2}$$

where ω_k are the frequencies of oscillation. The terms $H^{(j)}$ are homogeneous polynomials of order j, that is

$$H^{(j)}(x,y)=\sum_{|k|+|l|=j}h_{kl}x^k y^l \tag{3}$$

where h_{kl} are real coefficients. The multi-index notation $k=(k_1,\ldots,k_N)$, $|\,k\,|=k_1+\cdots+k_N$ and $x^k=x_1^{k_1}\ldots x_N^{k_N}$ is used.

The Hamiltonian (1) is in normal form to order s if

$$D_{(x,y)}H^{(j)}(x,y)=0 \qquad j=2,\ldots,s \tag{4}$$

where D is

$$D_{(x,y)}=\sum_{k=1}^{N}\omega_k\left(x_k\frac{\partial}{\partial y_k}-y_k\frac{\partial}{\partial x_k}\right). \tag{5}$$

There is an *r-fold resonance* at the equilibrium point $(x, y) = (0,0)$, if there are r independent commensurability conditions between the frequencies, i.e.

$$\sum_{k=1}^{N} a_{ik}\omega_k = 0 \qquad i = 1, \ldots, r \tag{6}$$

where a_{ik} is an integer matrix with rank r. The equilibrium point is *non-resonant* if there are no such commensurability conditions, or $r = 0$.

That the Hamiltonian (1) can be brought to normal form by canonical transformations for both resonant and non-resonant cases was shown by Gustavson (1966). The procedure will be presented in the following section. If the Hamiltonian is in normal form to all orders, Gustavson showed that:

(a) For an r-fold resonance there are $N - r$ independent formal integrals of motion

$$I_l = \sum_{k=1}^{N} \mu_{lk}\tau_k \qquad l = 1, \ldots, N - r \tag{7}$$

where $\tau_k = \frac{1}{2}(x_k^2 + y_k^2)$ for $k = 1, \ldots, N$ and μ_{lk} are components of the $N - r$ kernel vectors of the commensurability matrix a_{ik}

$$\sum_{k=1}^{N} a_{ik}\mu_{lk} = 0 \qquad i = 1, \ldots, r. \tag{8}$$

In this case the Hamiltonian itself is an integral independent of all I_l.

(b) In the non-resonant case there are N independent integrals $\tau_1, \ldots, \tau_N$ and the Hamiltonian is a function of them, i.e. $H = H(\tau_1, \ldots, \tau_N)$.

2.2. *Calculating the normal form*

Let $H(x, y)$ be in normal form to order $s - 1$. Then there exists a canonical transformation $(x, y) \to (q, p)$, $H(x, y) \to \tilde{H}(q, p)$, where q and p are new coordinates and momenta, respectively, and $\tilde{H}$ is in normal form up to order s. The generating function of the canonical transformation is

$$G^{(s)}(x, p) = xp + W^{(s)}(x, p) \tag{9}$$

where $W^{(s)}$ is a homogeneous polynomial of order s and the canonical transformations are

$$q = x + \partial W^{(s)}(x, p)/\partial p \qquad y = p + \partial W^{(s)}(x, p)/\partial x\,. \tag{10}$$

By considering the expansion in powers of x and p of the following equation (Goldstein 1956):

$$H(x, p + \partial W^{(s)}(x, p)/\partial x) = \tilde{H}(x + \partial W^{(s)}(x, p)/\partial p, p) \tag{11}$$

one sees that the terms which are already in normal form remain unchanged

$$H^{(j)}(\xi, \eta) = \tilde{H}^{(j)}(\xi, \eta) \qquad j = 2, \ldots, s - 1\,. \tag{12}$$

Here (ξ, η) stand for arguments of the functions. The order s of (11) gives

$$D_{(x,p)}W^{(s)}(x,p) = H^{(s)}(x,p) - \tilde{H}^{(s)}(x,p). \tag{13}$$

$D_{(x,p)}$ is a linear operator on the space of all polynomials in variables x and p. The linear space of all such polynomials can be decomposed in a direct sum of kernel and range subspaces of the operator $D_{(x,p)}$. In order for (13) to have solutions for $W^{(s)}$, its right-hand side must be in the range of the operator $D_{(x,p)}$. This requires $\tilde{H}^{(s)}(x,p)$ to be a kernel component of $H^{(s)}(x,p)$. Therefore

$$D_{(x,p)}\tilde{H}^{(s)}(x,p) = 0 \tag{14}$$

and $\tilde{H}(x,p)$ is in normal form up to order s. The solution $W^{(s)}$ of (13) is determined up to an arbitrary polynomial from the kernel of $D_{(x,p)}$. Following Gustavson, one makes $W^{(s)}$ unique by requiring its kernel component to vanish. Note that this choice of $W^{(s)}$ is not unique. The freedom to add any term from the kernel of $D_{(x,p)}$ at will may eventually be used to speed up the convergence properties of the integral—see discussions later on.

Equation (13) is most easily solved for $W^{(s)}$ by introducing the complex coordinates

$$z_k = \frac{1}{\sqrt{2}}(x_k + \mathrm{i}p_k) \qquad z_k^* = \frac{1}{\sqrt{2}}(x_k - \mathrm{i}p_k). \tag{15}$$

The operator D becomes diagonal in the basis of monomials

$$Dz^m z^{*n} = \omega(m-n)z^m z^{*n} \tag{16}$$

where the eigenvalue is $\omega(m-n) = \sum_{k=1}^{N} \omega_k(m_k - n_k)$.

Using $\tilde{H}^{(s)}$ and $W^{(s)}$ thus obtained, one can express all higher-order terms $\tilde{H}^{(i)}$, $i = s+1, s+2, \ldots$ as

$$\tilde{H}^{(i)}(x,p) = H^{(i)}(x,p) + \sum_{\substack{1 \leqslant |j| \leqslant l < i \\ l \geqslant 2, s \geqslant 3}} \frac{1}{j!}\left[\frac{\partial^{|j|}H^{(l)}}{\partial p^j}\left(\frac{\partial W^{(s)}}{\partial x}\right)^j - \frac{\partial^{|j|}\tilde{H}^{(l)}}{\partial x^j}\left(\frac{\partial W^{(s)}}{\partial p}\right)^j\right] \tag{17}$$

where $l = i - |j|(s-2)$ and the multi-index notation is $j! = j_1! \ldots j_N!$ and

$$\frac{\partial^{|j|}H^{(l)}}{\partial p^j} = \frac{\partial^{|j|}H^{(l)}(x,p)}{\partial p_1^{j_1} \ldots \partial p_N^{j_N}} \tag{18}$$

$$\left(\frac{\partial W^{(s)}}{\partial x}\right)^j = \left(\frac{\partial W^{(s)}}{\partial x_1}\right)^{j_1} \cdots \left(\frac{\partial W^{(s)}}{\partial x_N}\right)^{j_N}. \tag{19}$$

Note that the nonlinear appearance of $W^{(s)}$ in (17) makes it difficult to see what the consequences would be of the different choices of $W^{(s)}$ mentioned above.

By assumption, $H^{(2)}$ is already in normal form. By applying the successive canonical transformations as described, one can generate the normal form to an arbitrary order. If one fixes the maximum order desired to be M, then at each canonical transformation the transformed Hamiltonian $\tilde{H}$ has to be calculated to order M.

2.3. Expressing the final coordinates in terms of the original coordinates

To calculate the approximate integrals of motion (7) up to order M, each canonical coordinate transformation with $s = 3, \ldots, M$, has to be accurate up to order $M-1$. By inverting the $(s-2)$th canonical transformation (10), the new coordinates and momenta (q, p) are expressed as polynomial functions of the old coordinates and momenta (x, y). The inverses are done correctly up to order $M-1$ using an iterative procedure.

In the first step, one concentrates on the canonical transformation of momenta

$$p = y - \partial W^{(s)}(x, p)/\partial x \,. \tag{20}$$

The zero-order approximation to p as a function of y and x is

$$p^0 = y \,. \tag{21}$$

The $(i+1)$th approximation p^{i+1} is

$$p^{i+1} = y - \partial W^{(s)}(x, p)/\partial x \,|_{p=p^i} \qquad i = 0, 1, \ldots \,. \tag{22}$$

It turns out that the procedure converges since the lowest power of x and y of the difference $p^{i+1} - p^i$ is at least of order $(i+1)(s-2)+1$, which is always a strictly increasing function of i. To obtain the new momenta p correct up to order $M-1$ one assigns

$$p = p^{\lfloor (M-2)/(s-2) \rfloor} \tag{23}$$

where $\lfloor a \rfloor$ is the largest integer not exceeding a.

In a second step, one inserts the expressions obtained for p into the formulae for the canonical transformation of coordinates (10):

$$q = x + \partial W^{(s)}(x, p)/\partial p \tag{24}$$

and truncates it at the order $M-1$.

The two-step procedure is repeated for each canonical transformation with generating functions $G^{(3)}$ to $G^{(M)}$.

The formal integrals of motion (7) are quadratic functions of final coordinates and momenta. The expressions for final coordinates and momenta are inserted into the expressions for the integrals of motion (7). These are truncated at the order M. The truncated integrals of motion $I^{(M)}$ should approximate the formal integrals defined in (7).

3. Results for the Hénon–Heiles system

The Hénon–Heiles model (Hénon and Heiles 1964) is a classic example of a non-integrable Hamiltonian system with two degrees of freedom. Its Hamiltonian is

$$H = \tfrac{1}{2}(x_1^2 + x_2^2 + y_1^2 + y_2^2) + \lambda(x_1^2 x_2 + \eta x_2^3) \tag{25}$$

where x_1, x_2 are the coordinates and y_1, y_2 are their conjugate momenta and $\lambda = 1$, $\eta = -\frac{1}{3}$. We shall call the Hamiltonian (25) the generalized Hénon–Heiles Hamiltonian. The unperturbed frequencies are $\omega_1 = \omega_2 = 1$, so there is a 1-fold resonance. The commensurability matrix has the matrix elements $a_{11} = 1$, $a_{12} = -1$. According to Gustavson (7) the formal integral of motion is

$$I = I_1 = \tau_1 + \tau_2 = \tfrac{1}{2}(\tilde{x}_1^2 + \tilde{y}_1^2 + \tilde{x}_2^2 + \tilde{y}_2^2) \tag{26}$$

where $\tilde{x}_i$ and $\tilde{y}_i$ are the transformed final coordinates and momenta, respectively. As is customary in the literature (Gustavson 1966, Shirts and Reinhardt 1982), we define the integral of motion as

$$K = I - H\,. \tag{27}$$

We have used the symbolic algebra program (Kaluža 1991), written in the REDUCE programming language (Hearn 1987) to calculate the series of the normal form Hamiltonian and the approximate integrals of motion for the Hénon–Heiles system. The program works for any Hamiltonian system with polynomial potential with non-vanishing harmonic part and for any number of degrees of freedom. The calculation for the two-dimensional Hénon–Heiles system provides a good test of the codes. We have checked all numerical coefficients given in Gustavson (1966) and found a perfect agreement. There is a single sign error in the coefficient $I(185)$ for the integral of motion in his table 4, and a number of differences in coefficients of the normal form Hamiltonian and the approximate integral in higher orders, which probably come from the round-off errors in his numerical procedure. As one example, his coefficient $I(427) = 70.817255$ should correctly read 70.817274.

The integral $K^{(4)}$ has the following expression in terms of original variables on the surface of section, defined by $x_1 = 0$ and $y_1 > 0$:

$$K^{(4)}\,|_{\mathrm{sos}} = \frac{\lambda^2}{48}[45\eta^2(x_2^2 + y_2^2)^2 + 6\eta(7x_2^2 + 5y_2^2)y_1^2 + (-4x_2^2 + 20y_2^2 + 5y_1^2)y_1^2]\,. \tag{28}$$

We have been able to obtain the approximate integrals up to and including the 14th order in powers of coordinates and momenta. This is one order beyond the previously published results for the Hénon–Heiles system (Shirts and Reinhardt 1982).

We are interested in the convergence properties of the successive approximate integrals of motion obtained from the normal form. The approximate integral of motion $K^{(M)}$ of given order M is an analytic function of coordinates and momenta and is thus isolating. In the regular regime, where the exact isolating integrals of motion exist, the integrals $K^{(M)}$ can be expected to approximate to the exact integrals of motion, at least in the asymptotic sense. In the chaotic regime, however, there are no exact isolating integrals. The approximate integrals $K^{(M)}$, obtained from the normal-form procedure have a more difficult task to approximate the non-isolating exact integrals. What are the consequences of this fact for the convergence of the integrals $K^{(M)}$ in the chaotic regime?

In figure 1 the convergence of the approximate integrals of motion $K^{(M)}$ of the Hénon–Heiles system on the Poincaré surface of section is presented at the energy $E = \frac{1}{12}$. This energy is in the regular regime of the Hénon–Heiles system.

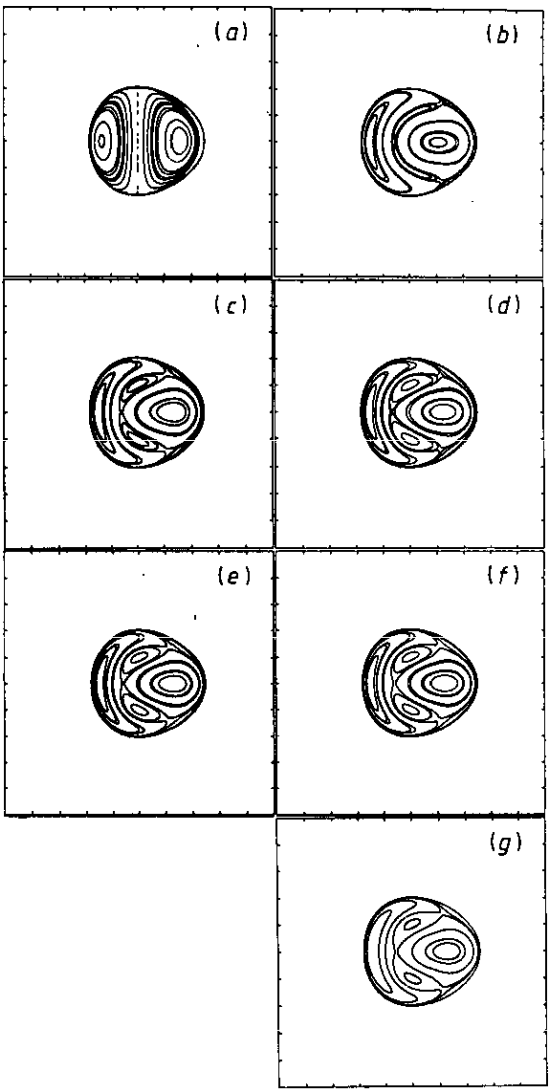

Figure 1. Convergence of the approximate integrals of motion of the Hénon–Heiles system at the energy $E = \frac{1}{12}$. In plot (*a*) the curves of constant value of the approximate integral of motion $K^{(4)} = I^{(4)} - H$ on the surface of section are drawn. Plots (*b*), (*c*), (*d*), (*e*) and (*f*) correspond to integrals $K^{(M)}$ for $M = 6$, 8, 10, 12 and 14, respectively. In each plot (*a*)–(*f*) the values of the approximate integrals of motion on the curves were chosen such that there is a curve going through each of the starting points listed below. Plot (*g*) is a numerically calculated Poincaré surface of section. The trajectories needed in (*g*) are obtained using the fourth-order Runge–Kutta integration method with the time step $h = 0.001$. The starting points of trajectories were chosen on the surface of section with $(x_2, y_2) = (0,0)$, $(-0.11,0)$, $(-0.13,0)$, $(-0.2,0)$, $(-0.25,0)$, $(0.1,0)$, $(0.15,0)$, $(-0.05,0.2)$, $(-0.05,-0.2)$, $(0.25,0.3)$. The parameters of the Hénon–Heiles Hamiltonian are $\lambda = 1$, $\eta = -\frac{1}{3}$. The surface of section is defined by the conditions $x_1 = 0$ and $y_1 > 0$. In all plots, the x-axis represents the coordinate x_2 and the y-axis represents the momentum y_2. The range of both quantities is $[-1.0, 1.0]$. Note the clear convergence of the contour plots to order $M = 14$. Within the graphical resolution, plots (*e*) and (*f*) do not differ. The curves in plot (*g*) consist of individual points of intersection of the trajectories with the surface of section. These are identical to the corresponding curves of constant value of $K^{(14)}$ in plot (*f*).

Figures 1(*a*)–(*f*) represent the curves of constant value of the successive approximate integrals of motion K on the surface of section. Up to the last order calculated, a perfect convergence of the curves can be seen. Furthermore, a perfect agreement can be seen with the exact curves in figure 1(*g*). This figure is the Poincaré surface of section obtained from the numerically exact trajectories. The starting points of these trajectories lie exactly on the level lines of the approximate integrals of figures 1(*a*)–(*f*).

In figure 2, the Hénon–Heiles system is studied at a higher energy, $E = \frac{1}{8}$, which is close to the critical energy. The critical energy is defined as the energy at which the surface of section is evenly divided into regular and irregular regions (Hénon and Heiles 1964). The convergence of the approximate integrals of motion (figures 2(*d*)–(*f*)) is now worse than the convergence at $E = \frac{1}{12}$. An overall impression is that figures 2(*d*) and 2(*e*) are most similar to each other and that the overall best approximate should be $K^{(12)}$. Figure 2(*g*) represents the numerically calculated Poincaré surface of section. One observes a number of small islands of stability around the new periodic orbits appearing at the outer boundary of the larger islands of regular behaviour. These do not seem to be present in figures with the approximate integrals. A novel feature in figures 2(*a*)–(*f*) is slow convergence inside the largest regular island on the right-hand side. This comes as a surprise. The possibility is not excluded, that higher orders would also exhibit convergence in this region. On the other hand the smaller stability regions up/down seem to be in a convergent manner approximated by the integrals of the largest orders.

Figures 3(*a*)–(*f*) represent the behaviour of the approximants at the energy $E = \frac{1}{6}$. This is exactly the ionization energy and the orbits are not bound to the potential well at any slightly higher energy. The successive approximants show again different convergence behaviour to those at lower energies. In general, the convergence can be said to be worse than at both lower energies. In part, this is due to more and more complex contours corresponding to the approximate integrals of higher orders—see in particular figure 3(*f*). There are still two larger regions on the surface of section where the successive approximants show quite convergent behaviour which is surprising since these regions lie in the chaotic regime. Figure 3(*g*) is the corresponding Poincaré surface of section. Almost the entire surface of section seems to be chaotic and the plot is not very informative in the chaotic region. As we shall see in more detail later, the convergence properties of the approximate integrals reveal much more information in the chaotic region than the surface of section plot 3(*g*). Two small stable islands surround a bifurcated stable periodic orbit. The oscillating behaviour of the formal integrals tries to capture this feature.

In figure 4 the convergence properties of the approximate integrals for a generalized Hénon–Heiles system (25) for $\lambda = -1$ and $\eta = 2$ are presented. The system (25) is integrable for any λ at $\eta = 2$. Naively one might expect that in this integrable case the formal integrals of motion will be convergent, or even exact. The exact integral of motion for $\lambda = -1$ and $\eta = 2$ reads (Bountis *et al* 1982):

$$I_e = x_1^4 + 4x_1^2x_2^2 + 4y_1(y_1x_2 - y_2x_1) - 4x_1^2x_2 + 3(y_1^2 + x_1^2). \qquad (29)$$

Note that it is a polynomial of coordinates and momenta. Figures 4(*a*), (*b*), (*c*) and (*d*) present the approximate integrals of orders 4, 6, 8 and 10, respectively. The contours presented in these figures converge. The figures corresponding to orders 12 and 14 are not presented because they are equivalent to figure 4(*d*) within the

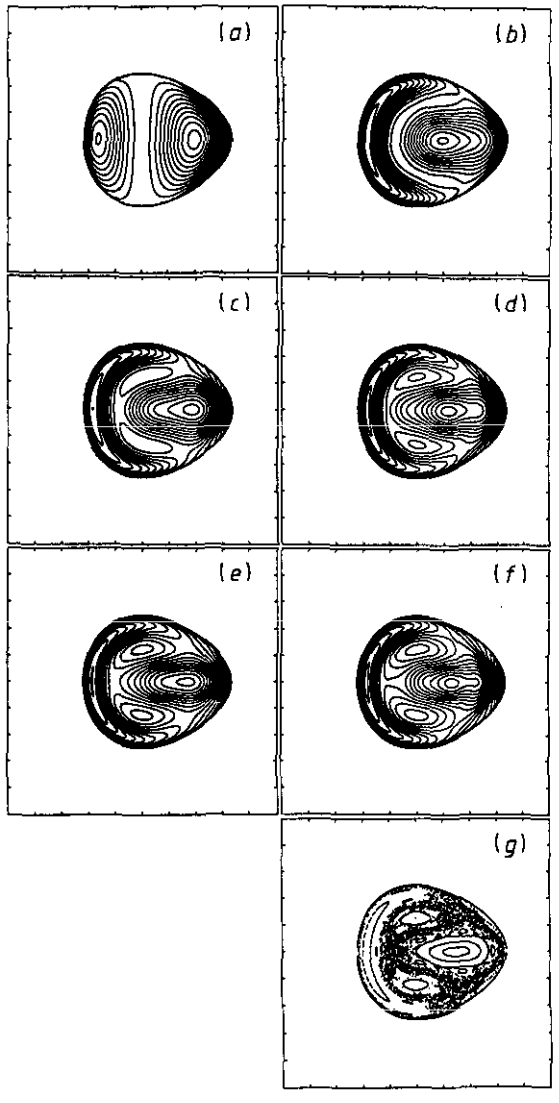

Figure 2. Convergence of the approximate integrals of motion of the the Hénon–Heiles system at the energy $E = \frac{1}{8} \approx$ critical energy. In plot (*a*) the curves of constant value of the approximate integral of motion $K^{(4)} = I^{(4)} - H$ on the surface of section are drawn. The contours with the values of $K^{(4)}$ from -0.020 to 0.020 in steps of 0.001 are drawn. Plots (*b*), (*c*), (*d*), (*e*) and (*f*) correspond to integrals $K^{(M)}$ for $M = 6$, 8, 10, 12 and 14, respectively. Plot (*g*) is a numerically calculated Poincaré surface of section. The trajectories needed in (*g*) are obtained using the fourth-order Runge–Kutta integration method with the time step $h = 0.001$. The starting points of the trajectories in plot (*g*) were chosen such that the trajectories cover both regular and irregular regimes on the surface of section. The parameters of the Hénon–Heiles Hamiltonian are $\lambda = 1$, $\eta = -\frac{1}{3}$. The surface of section is defined by the conditions $x_1 = 0$ and $y_1 > 0$. In all plots the x-axis represents the coordinate x_2 and the y-axis represents the momentum y_2. The range of both quantities is $[-1.0, 1.0]$. Note that the convergence of the contour plots is less pronounced than in figure 1. Nevertheless, the shapes of the regular regions in the exact plot are well reproduced with higher-order approximate integrals.

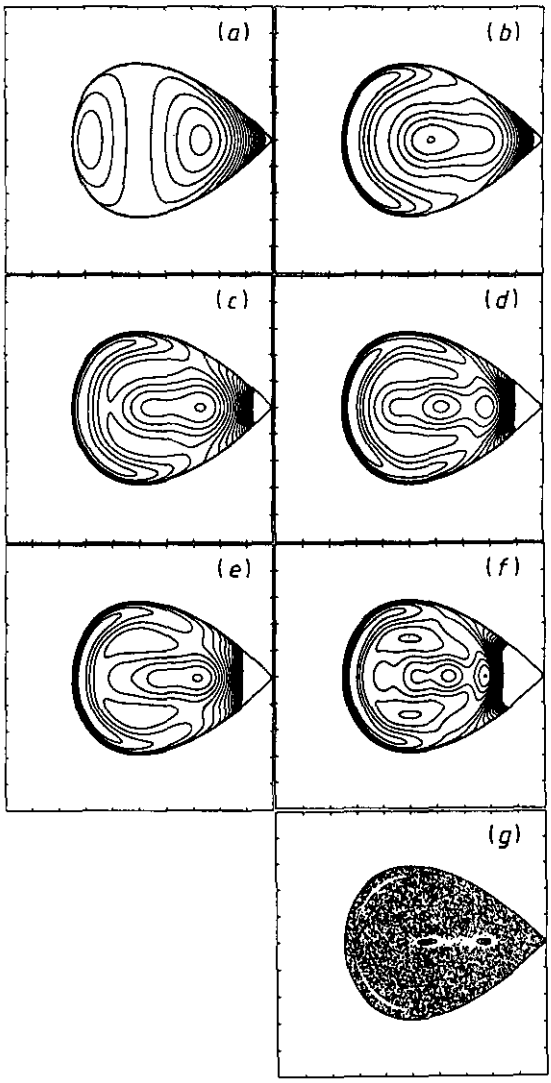

Figure 3. Convergence of the approximate integrals of motion of the the Hénon–Heiles system at the energy $E = \frac{1}{6}$ (ionization energy). In plot (*a*) the curves of constant value of the approximate integral of motion $K^{(4)} = I^{(4)} - H$ on the surface of section are drawn. The contours with the values of $K^{(4)}$ from -0.080 to 0.080 in steps of 0.005 are drawn. Plots (*b*), (*c*), (*d*), (*e*) and (*f*) correspond to integrals $K^{(M)}$ for $M = 6, 8, 10$, 12 and 14, respectively. Plot (*g*) is a numerically calculated Poincaré surface of section. The trajectories needed in (*g*) are obtained using fourth-order Runge-Kutta integration method with the time step $h = 0.001$. The starting points of the trajectories in plot (*g*) were chosen on the surface of section with $(x_2, y_2) = (0,0)$, $(-0.11, 0)$, $(-0.13, 0)$, $(-0.2, 0)$, $(-0.25, 0)$, $(0.1, 0)$, $(0.15, 0)$, $(-0.05, 0.2)$, $(-0.05, -0.2)$, $(0.25, 0.3)$, $(0, 0.22)$, $(0, 0.42)$, $(0.05, 0)$, $(-0.4, 0)$, $(-0.38, 0)$, $(-0.36, 0)$. The parameters of the Hénon–Heiles Hamiltonian are $\lambda = 1$, $\eta = -\frac{1}{3}$. The surface of section is defined by the conditions $x_1 = 0$ and $y_1 > 0$. In all plots the x-axis represents the coordinate x_2 and the y-axis represents the momentum y_2. The range of both quantities is $[-1.0, 1.0]$. The convergence of the curves of constant values of approximate integrals depends on the position on the surface of section. It seems that the curves converge in some parts of both regular and chaotic regions.

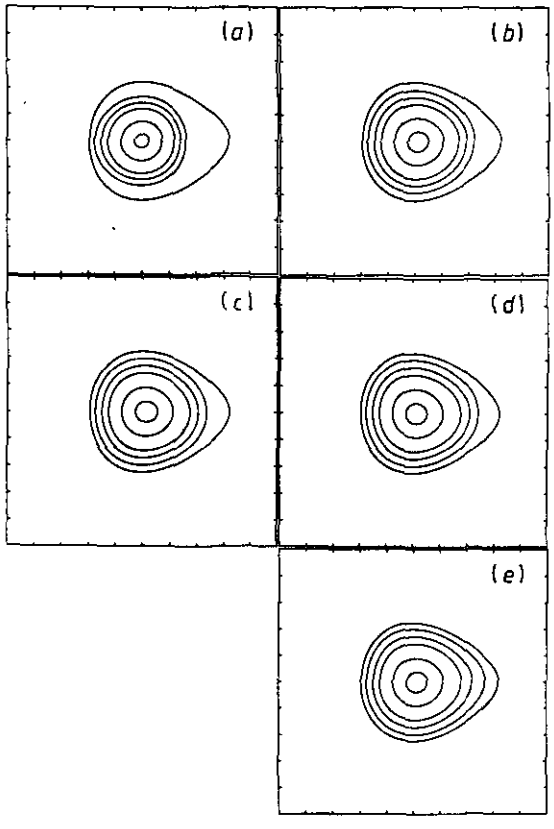

Figure 4. Convergence of the approximate integrals of motion of the generalized Hénon–Heiles system (25) with $\lambda = -1$ and $\eta = 2$ at the energy $E = 0.004$. This system is integrable. In plot (*a*) the curves of constant value of the approximate integral of motion $K^{(4)} = I^{(4)} - H$ on the surface of section are drawn. The values of the integral are chosen so that the curves of constant value of the integral go through the points $(x_2, y_2) = (-0.01, 0)$, $(-0.03, 0)$, $(-0.05, 0)$, $(-0.6, 0)$ and $(-0.07, 0)$ on the surface of section. Plots (*b*), (*c*) and (*d*) correspond to integrals $K^{(M)}$ for $M = 6$, 8 and 10, respectively. Plot (*e*) represents the contour plots of the exact integral of motion of (29). The surface of section is defined by the conditions $x_1 = 0$ and $y_1 > 0$. In all plots, the x-axis represents the coordinate x_2 and the y-axis represents the momentum y_2. The range of both quantities is $[-0.2, 0.2]$. The convergence of the contours is apparent up to order 10. However, as seen when comparing plot (*d*) with plot (*e*), only inner contours of plot (*d*) have converged to the exact contours of plot (*e*).

graphical resolution. In figure 4(*e*) the contours of the exact integral of motion of (29) are drawn. By comparing this figure to figure 4(*d*), one observes that only the innermost contours have converged to the exact curves, while the outermost ones have not. This is a slight disappointment. Perhaps we have reached the point of optimal approximation already at the order $M = 10$–12 and further orders diverge. We cannot give firm evidence for this since, as mentioned above, we have seen the stagnation at orders 10 to 14.

The expression for the exact integral of motion I_e of (29) on the surface of section $x_1 = 0$, $y_1 > 0$ is extremely simple. The successive approximants are more

complicated. What one does here is to approximate the simple polynomial by a series of higher and higher order polynomials, which perhaps is the main reason for the slow convergence. Of course, the formal integral may be generally written as a series of two invariants—the Hamiltonian and the additional integral—and it is this function which obviously converges slowly or perhaps even diverges. In general, the integrability does not necessarily imply convergence of the normal form and the formal integral of motion, as has been shown e.g. by Ali *et al* (1986). In other words, the divergence of the Birkhoff–Gustavson normal form series is not sufficient to maintain the non-integrability of the underlying Hamiltonian system.

There are two possible avenues of escape in this situation. One can try to use methods of resummation which presumably would give better convergence in the integrable case presented. In addition, one might use the non-uniqueness of the procedure of choosing the generating function for the canonical transformations W, as noted before. We have left these possibilities for consideration in further work.

For the reasons related to the semiclassical quantization of the Birkhoff–Gustavson normal form (Robnik 1984, Swimm and Delos 1979, Jaffe and Reinhardt 1982, Noid and Marcus 1977), we have done a calculation for the generalized Hénon–Heiles Hamiltonian (25) with parameters $\lambda = 1/\sqrt{80}$, $\eta = -\frac{1}{3}$. We have found an excellent convergence up to and including the 10th order, but do not show the results. Because of the homogeneity of the generalized Hénon–Heiles Hamiltonian of (25), there is a scaling property of the system: by rescaling the momenta, coordinates, and λ

$$x_i \to \mu x_i \qquad y_i \to \mu y_i \qquad \lambda \to \frac{1}{\mu}\lambda \tag{30}$$

the dynamics of the system remains the same at the scaled energy

$$E \to \mu^2 E\,. \tag{31}$$

Thus, the dynamics of the system are equivalent to those of the Hénon–Heiles Hamiltonian with parameters $\lambda = 1$, $\eta = -\frac{1}{3}$ and energy $E = \frac{1}{80}$. This energy is far below the lowest energy presented in figure 1, with the same λ and η and the excellent convergence of the approximate integrals of motion is not surprising, so we do not show the results, especially since the agreement is perfect. Even the lowest, fourth-order integral $K^{(4)}$ is of better quality than fourth-order integrals at higher energies, presented in figures 1(*a*), 2(*a*) and 3(*a*). This probably accounts for the quality of the quantum results for the energy levels, obtained by Robnik (1984), using the fourth-order normal-form Hamiltonian. however, it is not clear why the quantum results of other workers (Swimm and Delos 1979, Jaffe and Reinhardt 1982, Noid and Marcus 1977), who quantize the eighth-order normal form, are worse than those of Robnik (1984). At this point we can conclude that their lack of success as compared to Robnik (1984) cannot be due to the onset of the divergent behaviour already at the classical level. As mentioned above, the formal integrals of motion show perfectly convergent behaviour up to and including the 10th order at the energy of the quantum ground state. As stated in the introduction, we conclude that either Robnik's method of quantization is considerably more accurate than the methods of other workers mentioned above, and/or their calculations are not without errors. As explained in detail by Robnik (1984) the work by Swimm and Delos (1979) is definitely in error in this respect. It should be mentioned that Robnik's method has been successfully applied

to the hydrogen atom in a strong magnetic field (Robnik and Schrüfer 1985), which is a classically non-integrable and chaotic Hamiltonian system as has been shown for the first time by Robnik (1981), and further explained by Robnik (1982). For related results see also Kuwata *et al* (1990), and the review by Hasegawa *et al* (1989).

The Birkhoff–Gustavson normal form is almost always a divergent series, and at best is an asymptotic series to the exact integrals. Shirts and Reinhardt (1982) attempted to improve the convergence properties by using Padé approximants to the polynomial formal integrals of motion obtained by the Birkhoff–Gustavson normal-form procedure. They related the local property of clustering of poles of Padé approximants to chaoticity of the system.

We have tried to find a simpler and transparent criterion for the local convergence and divergence of the approximate integrals, in order to show the expected correlation between the convergence of the approximate integrals and the regularity of motion. Having calculated the values of all even-order approximate integrals of motion up to and including the order 14, probably the simplest but quite strong criterion for the local convergence on the surface of section $x_1 = 0$ is given by

$$C(x_2, y_2) = \begin{cases} 1 & \text{if } \mid K^{(j-2)} - K^{(j)} \mid > \mid K^{(j)} - K^{(j+2)} \mid \text{ for } j = 6,8,10,12 \\ 0 & \text{otherwise .} \end{cases} \tag{32}$$

At low energies, such as $E = \frac{1}{24}$, by using this criterion, we found convergence almost everywhere on the surface of section, which is regular everywhere. In general we have found that the approximate integrals converged in some regions of chaotic regime and did not converge in some regions of regular motion.

In figure 5 the convergence plots of this type are compared to the results obtained with the numerical calculation of trajectories. In figure 5(*a*) the Poincaré surface of section plot is displayed for the Hénon–Heiles Hamiltonian at the energy $E = \frac{1}{8}$. In figure 5(*c*), the convergence function $C(x_2, y_2)$ is presented: the dark points represent $C(x_2, y_2) = 1$. One sees a very good agreement between figures 5(*a*) and 5(*c*): all the largest dark regions in (*c*) correspond to the parts of the regular regions on the surface of section. However, the correspondence is not one to one. There are large regular regions in (*a*) where the convergence function C is 0. This is not unexpected, as the convergence function of (32) contains a strong condition of monotonic absolute convergence of the contributions of the successive approximate integrals. Figure 5(*b*) represents the Poincaré surface of section at the energy $E = \frac{1}{6}$, the ionization energy. Most of the trajectories are chaotic. In figure 5(*d*) the corresponding convergence function is presented. The most surprising feature is that the convergence criterion is still satisfied in large areas inside the chaotic region. Since the convergence criterion is strong, there should be a reason for this behaviour.

We have explored those parts of the chaotic domain where the convergence function is unity. In particular, we have studied the behaviour of trajectories with starting points in regions denoted by A, B, C, D and E in figure 5(*d*). It seems that the regions of convergence predict the short-period weakly unstable periodic orbits, and their weakly unstable neighbourhoods: the latter are located near the centres of pronounced regions of convergence.

It turns out that, although being in the chaotic regions, the trajectories which start close to the regions A, B, C, D and E show a much larger degree of short-time order than the trajectories with starting points elsewhere in the chaotic regions.

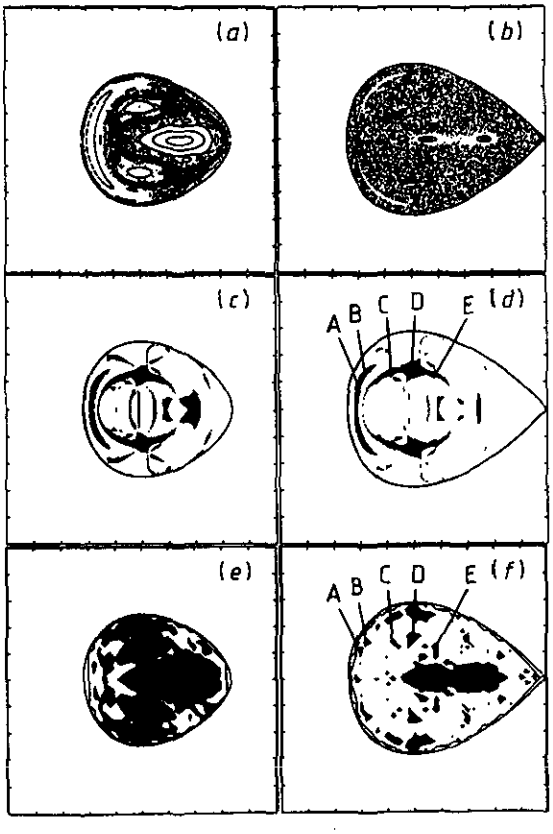

Figure 5. Connection between the finite-time Lyapunov exponent (FTLE), and the convergence behaviour of the consecutive approximate integrals of motion. In plot (*a*) the Poincaré surface of section is presented at the energy $E = \frac{1}{8}$, which is close to the critical energy (Hénon and Heiles 1964). Plot (*c*) is the corresponding convergence plot, calculated on a mesh 250 × 250—see text for the exact definition. Roughly speaking, the dark areas of the plot (*c*) represent regions where the approximate integrals of motion converge. Plot (*e*) is the plot of the FTLE L_6, calculated for six dynamical times—see text for exact definition. The dark areas in plot (*e*) represent the regions on the surface of section where FTLE is less than 0.07. In plot (*b*) the Poincaré surface of section is presented at the ionization energy $E = \frac{1}{6}$. Plot (*d*) is the corresponding convergence plot, calculated on a mesh 250 × 250. The dark areas of plot (*d*) represent regions where the approximate integrals of motion converge. Plot (*f*) is the plot of the FTLE L_6, calculated for six dynamical times, on a mesh 50 × 50. The dark areas in plot (*f*) represent the regions on the surface of section where L_6 is less than 0.1. The parameters of the Hénon–Heiles Hamiltonian are $\lambda = 1$, $\eta = -\frac{1}{3}$. The surface of section is defined by the conditions $x_1 = 0$ and $y_1 > 0$. In all plots, the x-axis represents the coordinate x_2 and the y-axis represents the momentum y_2. The range of both quantities is $[-1.0, 1.0]$. In the case $E = \frac{1}{8}$ the largest dark regions in plot (*c*) lie inside the regular regions on the surface of section in plot (*a*). The dark regions in plot (*c*) also lie inside the regions of surface of section, where the FTLE is smallest (dark regions in plot (*e*)). In the case $E = \frac{1}{6}$, there are just very tiny regions of regular motion left in plot (*b*), so the plot has become useless for studying the regularity properties of the phase space. However, the convergence plot (*d*) still shows the pronounced regions of convergence, even in the chaotic regime. These regions of convergence of the approximate integrals coincide with the dark regions in plot (*f*), where the FTLE is the smallest.

By the short time we mean a few, say 2–10 dynamical times (surface of section return times). To make the previous statement quantitative, we define the finite-time Lyapunov exponent (FTLE) L_n as

$$L'_n = \ln\left(\sqrt{[(\delta x_2(t)]^2 + [\delta y_2(t)]^2)/\delta}\right)/\mid t \mid \tag{33}$$

$$L_n = (L'_n + L'_{-n})/2 \tag{34}$$

where δ is the initial difference between the two neighbouring trajectories, taken in the x_2 direction, and we take time t to be n dynamical times, the time needed for n returns to the surface of section, following the trajectory either forward (L'_n) or backwards (L'_{-n}) in time. We choose $\delta = 10^{-6}$.

In figure 5(*f*) the regions where the FTLE L_6 is less than 0.1 are shown in black. The regions A, B, C, D and E do indeed correspond to the points where FTLE is less than elsewhere in that neighbourhood. In figure 5(*e*) the regions where the FTLE L_6 is less than 0.07 are black. The energy is $E = 1/6$. Again, the regions where the convergence function (32) is unity in figure 5(*c*) correspond to the regions of relatively small FTLE in figure 5(*e*).

One should note that once the expressions for the formal integrals are calculated, it is much faster to calculate the convergence function C than to calculate the FTLEs, which requires precise calculation of trajectories in chaotic regime.

Finally, when talking about the relation between the divergence properties and chaoticity of classical motion, we wish to point out an inconsistency in the qualitative results by Shirts and Reinhardt (1982). They look at the singularities (poles) of the Padé approximants of the formal integral. On the one hand they claim that 'pole regions tend to clump together and become more prominent *precisely* in the regions where chaotic motion is observed'. This is certainly not strictly true: in their figure 7 at $p_2 = 0.10$, $q_2 = 0.20$ and symmetrically there is a pole, or else there is an inaccuracy in the presentation. But on the other hand in this region the motion is regular, as can be seen in their figure 2. Therefore, their conclusions are qualitative too.

4. Summary and perspectives

Let us summarize our main conclusions, the other ones being mentioned in the introduction and in the text.

We have investigated the convergence properties of the Birkhoff–Gustavson normal form in various regimes, regular, chaotic and mixed, of the Hénon–Heiles system. At energies where the motion is regular, the integrals of motion are an excellent approximation to the exact integrals of motion and we have no evidence for divergence. Close to the critical energy, in the mixed regime, the integrals approximate to the exact integrals in the regular regions. In the mostly chaotic regime, the formal integrals may still converge in some regions. It turns out that in the regions where they do converge, the finite-time Lyapunov exponent assumes relatively small values. The convergence of the formal integrals predicts the regions of weak chaos inside chaotic regions.

On the other hand, we were not able to find any clear example of divergence of the approximate integrals of motion, in spite of quite a systematic search. It is thus reasonable to investigate further to what extent the approximate integrals describe

the motion in the chaotic regions for the sufficiently short times. In this regard it would be helpful to obtain even higher orders, for which one could explicitly employ the special symmetries of the Hénon–Heiles system (see Finkler *et al* 1991).

The results suggest the need to study the convergence properties of the approximate integrals by various methods, such as Padé approximants or Borel resummation. According to Bogomolny (1983), the asymptotic expansions of the coefficients of the formal integrals can be used to improve the convergence properties of the approximations. Moreover, Bogomolny (1984) relates the divergence properties of the normal form and of the formal integral of motion to the square root singularities associated with the shortest periodic orbits. Last but not least, although one feels that Gustavson's choice for the generating function W is a privileged choice, it is nevertheless not unique and other choices might provide even more convergent results.

References

Ali M K, Wood W R and Devitt J S 1986 *J. Math. Phys.* **27** 1806

Birkhoff G D 1927 *Dynamical Systems* vol IX (New York: American Mathematical Society Colloquium Publications)

Bogomolny E B 1983 *Sov. J. Nucl. Phys.* **37** 266

—— 1984 *Sov. Phys.–JETP* **59** 917

Bountis T, Segur H and Vivaldi F 1982 *Phys. Rev.* A **25** 1257

Finkler P, Jones C E and Sowell G A 1991 *Phys. Rev.* A **44** 925

Giorgilli A 1979 *Comput. Phys. Commun.* **16** 331

Goldstein H 1956 *Classical Mechanics* (Reading, MA: Addison-Wesley)

Gustavson F G 1966 *Astronom. J.* **71** 670

Hasegawa H, Robnik M and Wunner G 1989 *Suppl. Prog. Theor. Phys.* **98** 198

Hearn A C 1987 *Reduce Manual* Rand Publication CP78

Henon M and Heiles C 1964 *Astron. J.* **69** 73

Jaffe C and Reinhardt W P 1982 *J. Chem. Phys.* **77** 5191

Kaluža M 1991 Report CAMTP/91 *Comput. Phys. Commun.* submitted

Kuwata M, Harada A and Hasegawa H 1990 *J. Phys. A: Math. Gen.* **23** 3227

Noid D W and Marcus A R 1977 *J. Chem. Phys.* **67** 559

Robnik M 1981 *J. Phys. A: Math. Gen.* **14** 3195

—— 1982 *J. Physique* **43** C2 45

—— 1984 *J. Phys. A: Math. Gen.* **17** 109

Robnik M and Schrüfer E 1985 *J. Phys. A: Math. Gen.* **18** L853

Shirts R B and Reinhardt W P 1982 *J. Chem. Phys.* **77** 5204

Swimm R T and Delos J B 1979 *J. Chem. Phys.* **71** 1706