J. Phys. A: Math. Gen. 28 (1995) 1425–1448. Printed in the UK

Normal forms and quasi-integrals for the Hamiltonians of magnetic bottles

U M Engel, B Stegemerten and P Eckelt

Institut für Theoretische Physik I, Universität Münster, Wilhelm-Klemm-Straße 9, 48149 Münster, Germany

Received 23 May 1994, in final form 15 September 1994

Abstract. The well known Birkhoff–Gustavson normal form theory suffers from the restraint that the quadratic part of the Hamiltonian must be of the harmonic-oscillator type. In this paper we describe a generalized normal form concept which can be applied to *any* polynomial Hamiltonian, thus rendering the above restriction to harmonic oscillators unnecessary. As in the classical theory, we can derive an asymptotic expression for a second integral of motion. The truncated formal integral, the *quasi-integral*, exhibits good convergence properties in regions of phase space where the dynamics is regular, whereas in chaotic regions the convergence deteriorates. In order to exemplify these findings we apply the theory to a Hamiltonian describing a particular type of magnetic bottle which cannot be analysed using the Birkhoff–Gustavson normal form. We calculate the quasi-integrals up to and including the 14th order and analyse their convergence properties.

1. Introduction

In this paper we describe a generalization of a very powerful tool for the analysis of Hamiltonian dynamical systems: the theory of normal forms. Originally developed by Birkhoff in 1927 [1], who considered only non-resonant systems near an equilibrium point, the theory was brought into its classical form by Gustavson [2] who showed how to normalize† even in the presence of resonant frequencies. Later important contributions were made by Bryuno [3]. The Birkhoff–Gustavson normal form (BGNF) has received considerable attention because of its utility in finding approximate constants of motion [4–8] and quantizing nonlinear Hamiltonian systems [9–13].

The key idea of the theory is to systematically perform a series of canonical transformations, thus bringing the Hamiltonian into a particularly simple form, its 'normal form'. In this context simplicity means the possibility to read off an expression for a second integral of motion directly from the normal form Hamiltonian. More precisely, for a Hamiltonian H in BGNF the quadratic part of H is proven to be an integral of motion.

Gustavson considered an autonomous Hamiltonian system of n degrees of freedom in the vicinity of a stable equilibrium point, such that in lowest-order approximation

† Throughout this paper, the term 'normalization' refers to the process of transforming a Hamiltonian into its normal form. This is not to be confused with, say, the normalization of a vector.

0305-4470/95/051425+24$19.50

the Hamiltonian can be written as an n-dimensional harmonic oscillator, while the anharmonic parts of the system are given by a power series of order three and higher in the coordinates $q \in \mathbb{R}^n$ and the momenta $p \in \mathbb{R}^n$:

$$H(q,p) = \sum_{l \geqslant 2} H_l(q,p) \tag{1a}$$

$$H_2(q,p) = \sum_{\nu=1}^{n} \frac{\omega_\nu}{2}(q_\nu^2 + p_\nu^2) \tag{1b}$$

with real frequencies $\omega_\nu > 0$ and H_l being a homogoneous polynomial of degree l in q_ν and p_ν. Gustavson's theory necessarily *requires* the quadratic part of H to be of the particular form (1*b*). Only for this special H_2 he can define his normal form, prove normalizability and show how to obtain (an asymptotic expression for) a second invariant. In section 2.1 of this paper we will show in detail how this restriction comes about.

Several attempts have been made to approach the problem from a more general point of view, making it possible to normalize not only Hamiltonians of the Gustavson type [14–18]. However, until recently there has been no general method that could be applied to Hamiltonians with an arbitrary H_2 term. It is this problem that we will address and exemplify in the following sections. Similar (and more general) results have been obtained by Meyer and Hall in [19], though their approach—and especially their proof of what corresponds to our main theorem (cf section 2.3)—is quite different from ours. It is our goal to give a more easily readable account of the theory and to demonstrate in some detail the practical application to a given Hamiltonian and the calculation of the quasi-integral.

In section 2 we formulate the normalization process in terms of Lie operators and Lie transformations and use these techniques to develop a generalized approach, suitable for *any* $H_2 \not\equiv 0$. Section 3 is dedicated to the application of the generalized normal form to a model system that cannot be analysed by means of the Birkhoff–Gustavson theory. As the model system we have chosen a particular type of magnetic bottle that can be used as an ion trap in laboratory experiments. Our main result from the normal form calculations is the derivation of an expression for a formal integral of motion I up to and including the 14th order. The convergence properties of this quasi-integral are analysed in section 3.2; I exhibits a surprisingly rich structure and can be used to reproduce reasonably well the corresponding Poincaré plot.

2. Normal forms

In section 2.1 we give an outline of the BGNF theory using a Lie transformation technique [20, 21]. Sections 2.2 and 2.3 then show how to obtain a generalized normal form and its associated formal integral (quasi-integral). Here we draw mainly from the work of Dragt and Finn [16, 17] and from [22].

2.1. Lie transformations and the Birkhoff–Gustavson normal form

Consider an autonomous Hamiltonian system with n degrees of freedom and a fixed point in the origin. We can always write the Hamiltonian H as a formal power series in the phase-space coordinates q_ν, p_ν, $\nu = 1, 2, \ldots, n$. With $\mathbf{z} =$

$(q_1, \ldots, q_n, p_1, \ldots, p_n) \in \mathbb{R}^{2n}$ we have

$$H(\mathbf{z}) = \sum_{l \geq 2} H_l(\mathbf{z}) \tag{2a}$$

where the H_l are homogeneous polynomials of degree l:

$$H_l(\mathbf{z}) = \sum_{|\mathbf{m}| = l} h_{\mathbf{m}} \mathbf{z}^{\mathbf{m}} \in \mathscr{L}_l \tag{2b}$$

Here $\mathscr{L}_l$ is the $\binom{2n+l-1}{2n-1}$-dimensional vector space of homogeneous polynomials of degree l in $2n$ variables, and we employ the multiindex notation

$$\mathbf{m} \in \mathbb{N}_0^{2n} \qquad |\mathbf{m}| = \sum_{j=1}^{2n} m_j$$

$$z^{\mathbf{m}} = \prod_{j=1}^{2n} z_j^{m_j} \qquad h_{\mathbf{m}} = h_{m_1, \ldots, m_{2n}} \in \mathbb{R}.$$

Note that the dimension of $\mathscr{L}_l$ grows rapidly with l (e.g. for $n=2$ we have $\dim(\mathscr{L}_{14}) = 680$), such that any manipulation of H_l for larger values of l will have to be done by computer algebra rather than 'by hand'. We denote the space of all formal power series beginning with degree 2 by $\mathscr{L} = \bigoplus_{l=2}^{\infty} \mathscr{L}_l$.

The *Lie operator* ad_F adjoint to a power series $F \in \mathscr{L}$ is the Poisson bracket of F with some $G \in \mathscr{L}$:

$$\mathrm{ad}_F(G) := \{G, F\} = \sum_{i=1}^{n} \left(\frac{\partial G}{\partial q_i} \frac{\partial F}{\partial p_i} - \frac{\partial G}{\partial p_i} \frac{\partial F}{\partial q_i} \right). \tag{3}$$

ad_F is a linear operator on $\mathscr{L}$ for all F. The Lie operator adjoint to the quadratic part of the Hamiltonian and restricted to the subspace $\mathscr{L}_m$ is of central importance:

$$\mathscr{A}_m := \mathrm{ad}_{H_2}|_{\mathscr{L}_m} \qquad m \geq 2. \tag{4}$$

Note that $\mathscr{A}_m$ maps monomials of degree m to monomials of degree m.

The *Lie transformation* associated with $F \in \mathscr{L}$ is the exponential of ad_F:

$$\exp(\mathrm{ad}_F) = \sum_{i=0}^{\infty} \frac{1}{i!} \mathrm{ad}_F^i. \tag{5}$$

Lie transformations are an adequate tool for Hamiltonian normal form theory because they are canonical [20].

Let H be a Hamiltonian of type (1). H is in *Birkhoff–Gustavson normal form up to order m* if

$$\mathscr{A}_l(H_l) = 0 \qquad \text{for} \qquad l = 2, 3, \ldots, m. \tag{6}$$

H is in *Birkhoff–Gustavson normal form* if (6) holds for all $l \geq 2$. This definition is motivated by the fact that H_2 is an integral of motion if H is in BGNF:

$$\mathscr{A}_l(H_l) = 0 \quad \forall l \Leftrightarrow \{H, H_2\} = 0.$$

For any given Hamiltonian H of type (1) we can proceed to the BGNF of H in the following way. With some $F_m \in \mathscr{L}_m$ determine a new Hamiltonian $G = \Sigma_{l \geq 2} G_l$ by

$$G = \exp(\mathrm{ad}_{F_m})(H). \tag{7}$$

More explicitly we have

$$G_2+G_3+\ldots=H_2+\ldots+H_m+\mathrm{ad}_{F_m}(H_2)+\mathcal{O}(|\mathbf{z}|^{m+1})$$

where $\mathcal{O}(|\mathbf{z}|^{m+1})$ stands for terms of order $m+1$ and higher. Collecting terms of equal order we get, using $\mathrm{ad}^i_{F_m}(H_l)\in\mathcal{L}_{l+i(m-2)}$:

$$G_l=H_l \qquad \text{for} \qquad 2\leqslant l<m \tag{8a}$$

$$G_m=H_m+\mathrm{ad}_{F_m}(H_2). \tag{8b}$$

Assuming that H is already in BGNF up to order $m-1$, equation (8) allows several important conclusions. First of all, according to (8*a*) the contributions of order less than m remain unchanged under the Lie transformation associated to F_m. This shows that the transformed Hamiltonian G is at least in normal form up to order $m-1$, too. Secondly, (8*b*) indicates how to obtain a Hamiltonian G which is in BGNF even up to order m: From (8*b*) we get

$$H_m=G_m+\mathscr{A}_m(F_m). \tag{9}$$

This *homological equation* [23] must be solved for F_m and G_m under the additional condition

$$\mathscr{A}_m(G_m)=0. \tag{10}$$

In other words, G_m must be in the kernel (or null space) of $\mathscr{A}_m$:

$$G_m\in\mathrm{Ker}(\mathscr{A}_m)=\{P\in\mathcal{L}_m|\,\mathscr{A}_m(P)=0\}.$$

Thus we have the following iterative process for the normalization of H. For all $m\geqslant 3$ we first solve the homological equation for the polynomials G_m and F_m and then obtain the remaining terms $G_{l>m}$ of the new Hamiltonian by evaluating (7). The calculation of the G_l is a tedious but straightforward task that can be left to computer algebra. The non-trivial key point is solving the homological equation.

Let us assume that the vector space $\mathcal{L}_m$ can be decomposed into the direct sum of the kernel and range spaces of $\mathscr{A}_m$,

$$\mathcal{L}_m=\mathrm{Ker}(\mathscr{A}_m)\oplus\mathrm{Im}(\mathscr{A}_m) \tag{11}$$

with $\mathrm{Im}(\mathscr{A}_m)=\mathscr{A}_m(\mathcal{L}_m)$. Then $H_m\in\mathcal{L}_m$ can uniquely be split into its kernel and range components

$$H_m=H'_m+H''_m \tag{12a}$$

with

$$H'_m\in\mathrm{Ker}(\mathscr{A}_m) \tag{12b}$$

$$H''_m\in\mathrm{Im}(\mathscr{A}_m). \tag{12c}$$

Hence, G_m is uniquely determined by the kernel component of H_m:

$$G_m=H'_m. \tag{13}$$

Finally F_m can be obtained by inverting

$$\mathscr{A}_m(F_m)=H''_m. \tag{14}$$

Since there may be several pre-images of H''_m under $\mathscr{A}_m$ (F_m is uniquely determined by (14) up to any element of the null space of $\mathscr{A}_m$) the normalization procedure is not

unambiguous. However, we can always achieve unambiguity by additionally requiring F_m to lie in the range space of $\mathscr{A}_m$.

The key pioint of the above procedure is the splitting (11). By means of the canonical transformation $(q, p) \mapsto (\tilde{q}, \tilde{p})$ with

$$\tilde{q} = \frac{1}{\sqrt{2}}(q - ip)$$
$$\tilde{p} = \frac{1}{\sqrt{2}}(p - iq) \tag{15}$$

Gustavson showed that for a Hamiltonian of type (1) equation (11) holds, since in the new coordinates $\tilde{q}, \tilde{p}$ the corresponding transformed operator $\tilde{\mathscr{A}}_m$ is diagonal. Since $\tilde{\mathscr{A}}_m$ yields the splitting (11), $\mathscr{A}_m$ does as well. This proves the applicability of the BGNF theory to Hamiltonians of the Gustavson type (1): Every such Hamiltonian H can, by means of a formal canonical transformation, be transformed into the equivalent Hamiltonian

$$G = [\ldots \circ \exp(\mathrm{ad}_{F_4}) \circ \exp(\mathrm{ad}_{F_3})](H) \tag{16}$$

where $F_m \in \mathscr{L}_m$ and G is in BGNF. The term 'formal' indicates that we do not yet consider the convergence properties of the power series H, F and G.

2.2. The generalized normal form

We now turn to the generic case where $\mathscr{L}_m$ cannot be decomposed into the direct sum of the kernel and range spaces of $\mathscr{A}_m$. As a trivial example for the way in which this problem arises, consider a particle with a single degree of freedom ($n = 1$) which in lowest-order approximation is 'free':

$$H_2(q, p) = \tfrac{1}{2}p^2. \tag{17}$$

$\mathscr{A}_m$ takes on the form $\mathscr{A}_m = p(\partial/\partial q)$, such that we get $\mathrm{Ker}(\mathscr{A}_3) = \mathrm{span}\{p^3\}$ and $\mathrm{Im}(\mathscr{A}_3) = \mathrm{span}\{p^3, p^2q, pq^2\}$, and obviously

$$\mathscr{L}_3 \neq \mathrm{Ker}(\mathscr{A}_3) \oplus \mathrm{Im}(\mathscr{A}_3).$$

This shows that the normal form considered by Gustavson is not suitable for all types of H_2.

We circumvent this problem by using *Fredholm's alternative* for $\mathscr{L}_m$:

$$\mathscr{L}_m = \mathrm{Ker}(\mathscr{A}_m^*) \oplus \mathrm{Im}(\mathscr{A}_m). \tag{18}$$

Here, as usual, the adjoint operator $\mathscr{A}_m^*$ is defined via $(R|\mathscr{A}_m S) = (\mathscr{A}_m^* R|S) \forall R, S \in \mathscr{L}_m$, where $(\cdot|\cdot)$ is any suitable scalar product. Below we will specify a particular scalar product that will simplify the following expressions as much as possible.

In accordance with (18) it is natural to define a new normal form. Let H be a Hamiltonian of type (1*a*) with an *arbitrary* quadratic contribution H_2. We say that H is in *generalized normal form up to order m* if

$$\mathscr{A}_l^*(H_l) = 0 \qquad \text{for} \qquad l = 3, 4, \ldots, m. \tag{19}$$

H is in *generalized normal form* if (19) holds for all $l \geqslant 3$.

Notice that (19) is not required to hold for $l=2$—in contrast to the corresponding definition (6) of the BGNF. The reason being that in general it is impossible to normalize H_2, since transforming H_2 implies changing $\mathscr{A}_m$ as well. For generic H_2 one has to expect $\mathscr{A}_2^*(H_2)\equiv \mathrm{ad}_{H_2}^*(H_2)\neq 0$. In Gustavson's case, however, (6) is always true for $n=2$ because the Poisson bracket of H_2 with itself vanishes.

In order to complete our definition of a normal form we have to specify the explicit form of the scalar product. For $R(\mathbf{z})=\Sigma_{|\mathbf{j}|=m} r_{\mathbf{j}}\mathbf{z}^{\mathbf{j}}\in\mathscr{L}_m$ and $S(\mathbf{z})\in\mathscr{L}_m$ we set [24, 22]

$$(R|S):=\left(\sum_{|\mathbf{j}|=m}\bar{r}_{\mathbf{j}}\frac{\partial^m}{\partial^{j_1}z_1\partial^{j_2}z_2\ldots\partial^{j_{2n}}z_{2n}}\right)S(\mathbf{z}) \tag{20}$$

where the bar denotes complex conjugation. It is easy to see that this product operation $(\cdot|\cdot)$ indeed has the properties of a scalar product. In [17] a somewhat different scalar product was introduced by choosing a special basis of $\mathscr{L}_m$ and defining it to be orthonormal. These two scalar products are identical up to a normalization factor. However, the definition given here paves way for a more general approach and is much easier to use. This becomes apparent when trying to derive an explicit expression for $\mathscr{A}_m^*$. In [17] this was only achieved for a very restricted case, namely the so-called 'mirror machine' or 'magnetic bottle Hamiltonians'. See section 3 for a discussion of this class of systems.

We first write $\mathscr{A}_m$ in yet another form. Linearizing Hamilton's equations we obtain the Hamiltonian matrix $L=J\,\mathrm{Hess}(H_2)$, with the $2n$-dimensional symplectic matrix

$$J=\begin{pmatrix}0 & \mathrm{id}_n\\ -\mathrm{id}_n & 0\end{pmatrix}$$

and the Hessian $\mathrm{Hess}(H_2)$. Thus we have

$$\mathscr{A}_m(\cdot)=D_{\mathbf{z}}(\cdot)\cdot L\mathbf{z} \tag{21}$$

with the abbreviation $a\cdot b=\Sigma_{i=1}^{2n}a_ib_i$ for $a,b\in\mathbb{C}^{2n}$. In order to find $\mathscr{A}_m^*$ we rewrite its definition $(\mathscr{A}_m^*R|S)=(R|\mathscr{A}_mS)$ as

$$(\mathscr{A}_m^*R|S)=\frac{\mathrm{d}}{\mathrm{d}t}(R(\mathrm{e}^{L^*t}\mathbf{z})|S)|_{t=0}$$

where we have used the relation $(R\circ M^*|S)=(R|S\circ M)$ which holds for any linear mapping M on $\mathbb{C}^{2n}$. Evaluating the time derivative yields $(\mathscr{A}_m^*R|S)=(D_{\mathbf{z}}(R)\cdot L^*\mathbf{z}|S)$, and we obtain $\mathscr{A}_m^*$ as

$$(\mathscr{A}_m^*(\cdot)=D_{\mathbf{z}}(\cdot)\cdot L^*\mathbf{z}. \tag{22}$$

This expression is identical with (21) after conjugating and transposing L.

In the form (22) $\mathscr{A}_m^*$ can easily be used for determining the splitting (18) of $\mathscr{L}_m'$. For the example (17) considered in the beginning of this section we obtain $\mathscr{A}_m^*=q(\partial/\partial p)$ and therefore $\mathrm{Ker}(\mathscr{A}_3^*)=\mathrm{span}\{q^3\}$.

The method for transforming a given Hamiltonian into its generalized normal form is exactly the same as the one described in the previous section; one only has to replace (12*b*) by

$$H_m'\in\mathrm{Ker}(\mathscr{A}_m^*). \tag{23}$$

Since the splitting (18) holds for any H_2, we have proven that *any Hamiltonian can be normalized* according to the generalized definition (19).

Note that for a Hamiltonian of Gustavson's type (1) the two definitions of normal form coincide, because in this case $\mathscr{A}_m^* = -\mathscr{A}_m$. So if H is in BGNF (up to order m) it is in generalized normal form (up to order m), too. The utility of the normal form will become evident in the next section.

2.3. *Quasi-integrals of motion*

By construction, for a Hamiltonian in BGNF H_2 is a formal integral of motion (see section 2.1). We now show how to find an analogous formal integral for a Hamiltonian in generalized normal form. Our results are similar to the findings of Meyer and Hall [19], but the proof differs in some details. We have tried to make the exposition as transparent as possible by focusing on just those aspects that are essential for the reasoning.

We write H_2 as

$$H_2(\mathbf{z}) = \tfrac{1}{2}\mathbf{z} \cdot J^{-1} L\mathbf{z} \tag{24}$$

and decompose L by means of the *Jordan–Chevalley decomposition* [25] into its diagonalizable and nilpotent parts D and N:

$$L = D + N. \tag{25}$$

Existence and uniqueness of this decomposition are assured by the Jordan normal form theorem for matrices. Define the *diagonalizable component* $I(z)$ and the *nilpotent component* $K(\mathbf{z})$ of $H_2(\mathbf{z})$ by

$$\begin{aligned} I(\mathbf{z}) &= \tfrac{1}{2}\mathbf{z} \cdot J^{-1} D\mathbf{z} \\ K(\mathbf{z}) &= \tfrac{1}{2}\mathbf{z} \cdot J^{-1} N\mathbf{z} \end{aligned} \tag{26}$$

such that $H_2(\mathbf{z}) = I(\mathbf{z}) + K(\mathbf{z})$. We are now in the position to prove the main theorem:

Theorem: For a Hamiltonian $H(\mathbf{z})$ in generalized normal form the diagonalizable part $I(\mathbf{z})$ of $H_2(\mathbf{z})$ is a formal integral of motion.

For the proof we must show that the Poisson bracket of I with H_m vanishes for all $m \geqslant 2$. We start with $m = 2$ and then proceed to the case $m > 2$.

By virtue of Jacobi's identity for the Poisson bracket we have

$$\mathrm{ad}_{\{H_2, I\}} = \mathrm{ad}_I \mathrm{ad}_{H_2} - \mathrm{ad}_{H_2} \mathrm{ad}_I. \tag{27}$$

This expression is zero if ad_{H_2} and ad_I commute. In order to show that the latter is the case we first remark that the matrices L, D and N are infinitesimally symplectic [26], i.e. they satisfy $M^t J + JM = 0$ (for $M = L, D, N$). Direct computation shows that for an infinitesimally symplectic matrix M the Lie operator adjoint to the quadratic polynomial $P(\mathbf{z}) = \frac{1}{2}\mathbf{z} \cdot J^{-1} M\mathbf{z}$ can be written as $\mathrm{ad}_P(\cdot) = D_z(\cdot) \cdot M\mathbf{z}$. Thus we have for the Lie operators $\mathscr{D}_m$ and $\mathscr{N}_m$ adjoint to I and K:

$$\begin{aligned} \mathscr{D}_m(\cdot) &:= \mathrm{ad}_I(\cdot) = D_z(\cdot) \cdot D\mathbf{z} \\ \mathscr{N}_m(\cdot) &:= \mathrm{ad}_K(\cdot) = D_z(\cdot) \cdot N\mathbf{z} \end{aligned} \tag{28}$$

It is one of the key advantages of this formulation of the theory that we can characterize all the important operators $\mathscr{A}_m$, $\mathscr{D}_m$ and $\mathscr{N}_m$ which operate in a space of the high dimension $\binom{2n+m-1}{2n-1}$ by matrices of the considerably smaller dimension $2n$: L, D and N.

We now show that for two commuting matrices M_1, M_2 the corresponding Lie operators ad_{P_1}, ad_{P_2} (defined as above) commute as well:

$$\begin{aligned}\mathrm{ad}_{P_1}\mathrm{ad}_{P_2}(\cdot) &= D_{\mathbf{z}}(D_{\mathbf{z}}(\cdot)\cdot M_2\mathbf{z})M_1\mathbf{z}\\ &= \sum_{\mu,\nu=1}^{2n}\frac{\partial^2}{\partial z_\mu\,\partial z_\nu}(\cdot)(M_1\mathbf{z})_\nu(M_2\mathbf{z})_\mu + \sum_{\nu=1}^{2n}\frac{\partial}{\partial z_\nu}(\cdot)(M_2M_1\mathbf{z})_\nu\\ &= \mathrm{ad}_{P_2}\mathrm{ad}_{P_1}(\cdot).\end{aligned}\tag{29}$$

Because $LD-DI=0$, this implies that the right-hand side of (27) is zero, and thus $\{H_2, I\}=0$.

For $m>2$ we proceed in the following way: We show that diagonalizability and nilpotence of the matrices D and N carry over to the corresponding Lie operators $\mathscr{D}_m$ and $\mathscr{N}_m$; these properties then imply that the null spaces of $\mathscr{D}_m$ and $\mathscr{D}_m^*$ coincide and that $\mathscr{D}_m^* I_m=0$, which in turn means $\{H_m, I\}=\mathscr{D}_m H_m=0$.

Consider a unitary matrix T that transforms D into the diagonal matrix $\tilde{D}=TDT^{-1}$. Inserting twice the identity $T^{-1}T$ into the expression for $\mathscr{D}_m$ we get

$$\mathscr{D}_m(\cdot)=D_{\mathbf{z}}(\cdot)\cdot TDT^{-1}\,T\mathbf{z}.$$

With $\tilde{\mathbf{z}}=T\mathbf{z}$, and denoting $\mathscr{D}_m$ in the new coordinates $\tilde{\mathbf{z}}$ by $\tilde{\mathscr{D}}_m$, we obtain

$$\tilde{\mathscr{D}}_m(\cdot)=D_{\tilde{\mathbf{z}}}(\cdot)\cdot\tilde{D}\tilde{\mathbf{z}}\tag{30}$$

Application of this transformed operator to any of the basis monomials $\tilde{\mathbf{z}}^J$ of $\mathscr{L}_m$ yields, because $\tilde{D}$ is diagonal, an eigenvalue equation with the eigenfunction $\tilde{\mathbf{z}}^J$ and a certain eigenvalue μ_J—thus diagonalizability of $\mathscr{D}_m$ is shown.

Now consider any $R\in\mathscr{L}_m$. $R(\mathrm{e}^{Nt}\mathbf{z})$ is a polynomial in t of degree less than or equal to $m(n_0-1)$, since by nilpotence there is some $n_0\in\mathbb{N}$ such that $N^{n_0}=0$. This polynomial is related to $\mathscr{N}_m$ in the following way:

$$\frac{\mathrm{d}}{\mathrm{d}t}R(\mathrm{e}^{Nt}\mathbf{z})=[D_{\mathbf{z}}(R)\cdot N\mathbf{z}]_{\mathrm{e}^{Nt}\mathbf{z}}=\mathscr{N}_m(R)|_{\mathrm{e}^{Nt}\mathbf{z}}.$$

Iterating this expression and evaluating for $t=0$ we get

$$\mathscr{N}_m^{mn_0}(R)=\frac{\mathrm{d}^{mn_0}}{\mathrm{d}t^{mn_o}}R(\mathrm{e}^{Nt}\mathbf{z})|_{t=0}=0\tag{31}$$

which implies nilpotence of $\mathscr{N}_m$, because (31) holds for all R.

Identity of the null spaces of $\mathscr{D}_m$ and $\mathscr{D}_m^*$ is a direct implication of diagonalizability: Application of the diagonalized operator $\tilde{\mathscr{D}}_m$ (cf (30)) yields

$$(\tilde{\mathbf{z}}^J|\tilde{\mathscr{D}}_m^*\tilde{\mathbf{z}}^K)=(\tilde{\mathscr{D}}_m\tilde{\mathbf{z}}^J|\tilde{\mathbf{z}}^K)=\bar{\mu}_J(\tilde{\mathbf{z}}^J|\tilde{\mathbf{z}}^K)=\bar{\mu}_J\delta_{JK}.$$

So the eigenspaces corresponding to the eigenvalue 0 of D_m^* and $\mathscr{D}_m$ are identical.

Finally, we determine how $\mathscr{D}_m^*$ acts on polynomials in $\mathscr{L}_m$. Notice that $\mathscr{N}_m^*$ is nilpotent, because its adjoint is. With $\mathscr{A}_m^*=\mathscr{D}_m^*+\mathscr{N}_m^*$ we obtain for any $H_m\in\mathscr{L}_m$:

$$\begin{aligned}0&=(\mathscr{N}_m^*)^{mn_0}H_m\\ &=(\mathscr{A}_m^*-\mathscr{D}_m^*)^{mn_0}H_m\\ &=\sum_{l=0}^{mn_0}\binom{mn_0}{l}(-\mathscr{D}_m^*)^{mn_0-l}(\mathscr{A}_m^*)^l H_m\end{aligned}$$

because $\mathscr{A}_m^*$ and $\mathscr{D}_m^*$ commute (since the corresponding matrices L_0^* and D_0^* commute; cf (29)). $(\mathscr{A}_m^*)^l H_m$ is zero for $m \geqslant 3$ and $l \geqslant 1$ because H is in generalized normal form. From this it follows for $H_m(\tilde{\mathbf{z}}) = \Sigma_{|\mathbf{j}|=m} h_\mathbf{j} \tilde{\mathbf{z}}^\mathbf{j}$ that

$$0 = (\tilde{\mathscr{D}}_m^*)^{mn_0} H_m = \sum_{|\mathbf{j}|=m} h_\mathbf{j} (\bar{\mu}_j)^{mn_0} \tilde{\mathbf{z}}^\mathbf{j}$$

here, for the sake of notational convenience, we have again turned to the coordinates $\tilde{\mathbf{z}}$ as defined above. Linear independence of the basis monomials $\tilde{\mathbf{z}}^\mathbf{j}$ then gives the result $h_\mathbf{j} \bar{\mu}_\mathbf{j} = 0$ and thus

$$\mathscr{D}_m^* H_m = \sum_{|\mathbf{j}|=m} h_\mathbf{j} \bar{\mu}_\mathbf{j} \tilde{\mathbf{z}}^\mathbf{j} = 0$$

and we have proven the theorem.

3. Normalizing a magnetic bottle

The remaining part of this paper is dedicated to the application of the generalized normal form theory to a class of systems that is not accessible via Gustavson's method.

3.1. The model system

We consider a 'magnetic bottle' which is made up of a homogeneous magnetic (dipole) field with a superimposed octupole contribution. In cylindrical coordinates we have:

$$\mathbf{B}(\rho, z) := B_0 \mathbf{e}_z + B_2[-\rho z \mathbf{e}_\rho + (z^2 - \tfrac{1}{2}\rho^2) \mathbf{e}_z]. \tag{32}$$

A configuration of this kind was used, for example, for very accurate measurements of the g-factor of the electron [27]. Figure 1 shows the field lines of $\mathbf{B}(\rho, z)$ and motivates the term 'bottle': At least in the vicinity of the z-axis, the motion of a charged particle in this type of magnetic field consists of a cyclotron oscillation about the field lines, superimposed to a vertical oscillation along these lines. Since the field lines converge towards the z-axis for larger values of $|z|$ the particle will be reflected at some stage (if it does not move *exactly* on the z-axis all the time), and the resulting motion is bound. So **B** indeed functions as a bottle for charged particles. (This analysis can easily be made rigorous.)

We restrict ourselves to the case $p_\varphi = 0$, such that a particle in the bottle does not encircle the z-axis but continually passes through it. After suitable scaling we arrive (for $B_0 B_2 > 0$) at the Hamiltonian

$$H(\rho, z, p_\rho, p_z) = \tfrac{1}{2}(p_\rho^2 + p_z^2) + V(\rho, z) \tag{33a}$$

with the two-dimensional potential

$$V(\rho, z) = \tfrac{1}{2}\rho^2 + \tfrac{1}{2}\rho^2 z^2 - \tfrac{1}{8}\rho^4 + \tfrac{1}{8}\rho^2 z^4 - \tfrac{1}{16}\rho^4 z^2 + \tfrac{1}{128}\rho^6. \tag{33b}$$

H describes a system with two degrees of freedom, one of which (ρ) corresponds, in lowest-order approximation, to a harmonically bound motion while the dynamics along the z-axis is free in the same approximation. Note that this Hamiltonian cannot be dealt with by the BGNF theory.

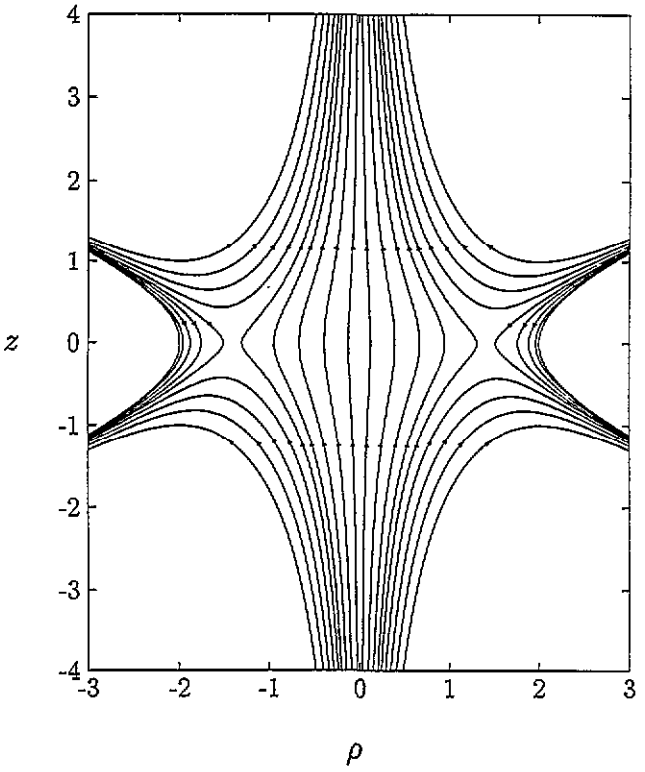


Figure 1. Magnetic field lines of the magnetic bottle as described by (32) with $B_0 = B_2 = 1$. The full three-dimensional picture is obtained by rotation about the z-axis.

In what follows, we do not restrict ρ to positive values but allow for negative values as well. This makes it possible to treat ρ and z simply as cartesian coordinates in two dimensions. An example for the dynamics of our model system, obtained by numerical integration, is shown in figure 2. Similarly, we have numerically calculated Poincaré

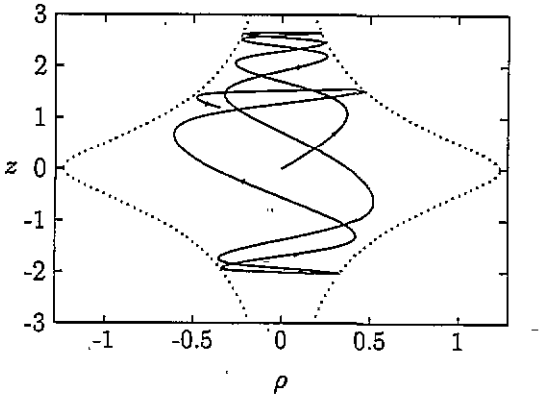


Figure 2. Typical dynamics in the magnetic bottle at the energy $E = 0.5$. The dotted lines show the boundary of the accessible region of the configuration space, defined by the condition $V(\rho, z) \leqslant E$.

plots for several energies E. To this end we have defined Poincaré surfaces of section Σ_E by setting $\rho=0$,

$$\sum_E := \{(0, z, p_\rho, p_z)^T \in \mathbb{R}^4 | p_\rho = \sqrt{2E - p_z^2};\ |p_z| < \sqrt{2E}\}$$

and obtained the corresponding Poincaré plot by recording the points (z, p_z) where the trajectory passes through Σ_E with positive momentum p_ρ (see figure 3). The system exhibits a typical KAM scenario when the control parameter E is increased; at low energies the system is nearly integrable, whereas at higher energies invariant tori break up and the chaotic region of phase space becomes increasingly large.

It is important to realize that for this system a global second integral of motion (the first being the Hamiltonian itself) cannot exist, because the existence of such an integral would render the system integrable. This would be incompatible with the non-integrability demonstrated by the Poincaré plots. Still, the preceding section shows how to construct a formal invariant. The resolution of this ostensible contradiction is that one expects the formal integral to approximate the local exact integrals of motion in the regular regime (where the KAM tori dominate), whereas in the stochastic regions the formal integral is expected to diverge. See [28] for a discussion of these convergence properties.

Rather than discussing just the Hamiltonian (33) we will study the normalization process for the more general class of *magnetic bottle Hamiltonians* which are defined by their quadratic contribution:

$$H_2(\mathbf{z}) = \sum_{\nu=1}^{l} \tfrac{1}{2} z_{n+\nu}^2 + \sum_{\nu=l+1}^{n} \frac{\omega_\nu}{2}(z_\nu^2 + z_{n+\nu}^2). \tag{34}$$

With $l=1$, $n=2$, $z_1=z$, $z_2=\rho$ and $\omega_2=1$ we can write the H_2 of (33) in the form (34).

Transformation of the corresponding Hamiltonian matrix $L=J\mathrm{Hess}(H_2)$ into Jordan normal form yields

$$\tilde{L} = \mathrm{diag}\Bigg(\underbrace{\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \ldots, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}}_{l \text{ times}}, i\omega_{l+1}, \ldots, i\omega_n, -i\omega_{l+1}, \ldots, -i\omega_n\Bigg). \tag{35}$$

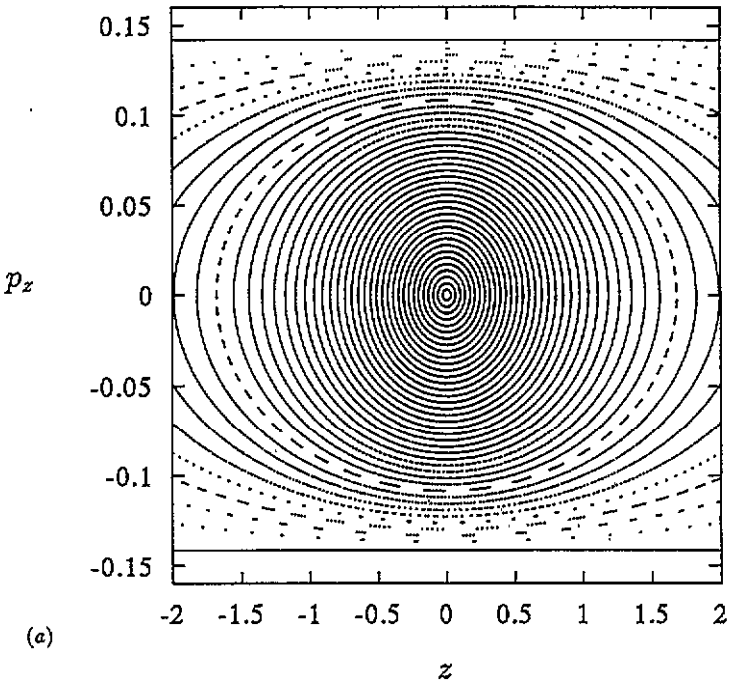
Obviously the frequencies ω_ν mark the diagonalizable component of $\tilde{L}$ (and therefore of L), such that we obtain

$$I_{\mathrm{NF}}(\mathbf{z}) = \sum_{\nu=l+1}^{n} \frac{\omega_\nu}{2}(z_\nu^2 + z_{n+\nu}^2) \tag{36}$$

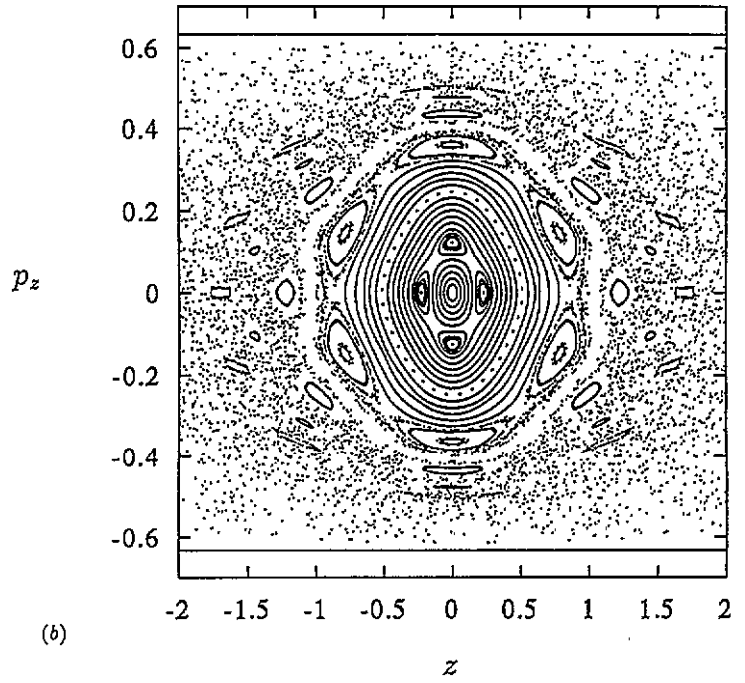
which is a formal integral of motion if $H(\mathbf{z})$ is in generalized normal form. This result was already stated in [17], but here it could be derived with much more ease of computation. For the specific case of (33) we have

$$I_{\mathrm{NF}}(\rho, z, p_\rho, p_z) = \tfrac{1}{2}(\rho^2 + p_\rho^2). \tag{37}$$

The stage is set for application of the normalization process as described in section 2. In the appendix we explain in some detail how the effort needed to normalize magnetic bottle Hamiltonians can be reduced considerably. As the result of these conditions we have obtained the formal integral of the magnetic bottle (33) up to and including



(*a*)



(*b*)

Figure 3. Poincaré plots of the magnetic bottle at the energy E as described in the text. The boundary of the surface of section, defined by $|p_z| = \sqrt{2E}$, shows up as horizontal lines. (*a*) $E=0.01$, (*b*) $E=0.2$, (*c*) $E=0.5$.

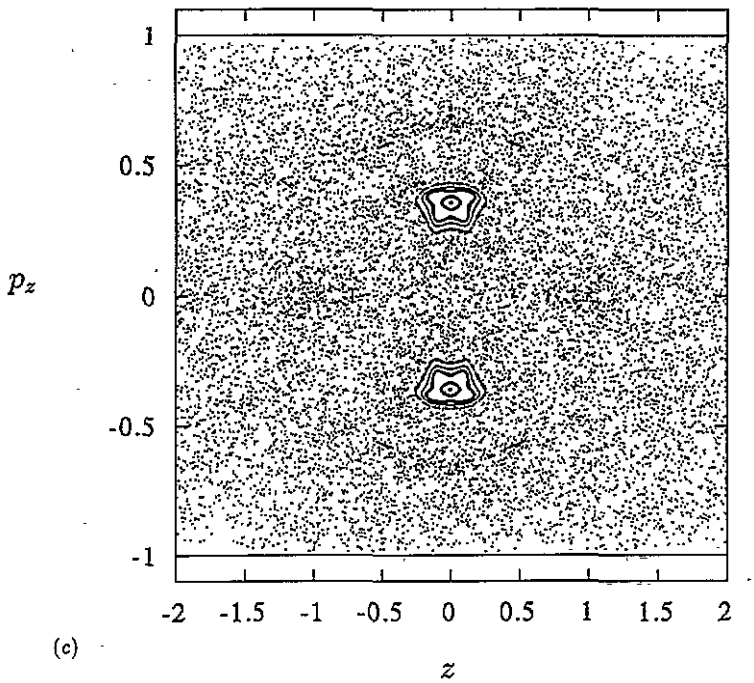


Figure 3. (continued)

the 14th order. The first few terms are

$$\begin{aligned} I(\rho, z, p_\rho, p_z) = {} & 0.5p_\rho^2 + 0.5\rho^2 + 0.046875p_\rho^4 + 0.125p_z^2p_\rho^2 \\ & + 0.09375\rho^2p_\rho^2 - 0.125\rho^2p_z^2 - 0.078125\rho^4 + 0.5z\rho p_z p_\rho \\ & - 0.25z^2p_\rho^2 + 0.25z^2\rho^2 + \mathcal{O}(|\mathbf{z}|^6). \end{aligned} \tag{38}$$

A complete list of all the 415 summands up to order 14 is available on request from the authors. Note that $I(\mathbf{z})$ contains only monomials of even order, because the same is true for the original Hamiltonian (33).

It is important to note the difference between the representations (37) and (38) of the integral of motion. The first formula applies if the Hamiltonian is already in generalized normal form, while the second holds for the non-normalized Hamiltonian (33) in the original coordinates and is obtained from (37) by inverting the normalizing Lie transformations.

To our knowledge, there is only one other example in the literature where normalization for a full Hamiltonian has been carried out up to such a high order [28, 29]. (Discrete mappings, on the other hand, have been normalized up to order 100 and beyond; cf [30].) One has to realize, though, that the Hénon–Heiles Hamiltonian considered in [28, 29] is of the Gustavson type (1), thus rendering $\mathscr{A}_m$ diagonal. As explained in the appendix, for magnetic bottle Hamiltonians $\mathscr{A}_m$ is *not* diagonal, which makes the determination of the splitting (12) and the inversion of (14) a much more difficult task.

3.2. Local and global analysis of the quasi-integral

The characteristic property of an integral of motion is its constancy along trajectories in phase space. As mentioned earlier, one cannot expect this behaviour for a formal integral of a non-integrable system like the magnetic bottle discussed here. But in line with reference [28] (and many others) we still expect convergence of $I(\mathbf{z})$ in regular regions of phase space.

Let us define the quasi-integral $I^{(m)}(\mathbf{z})$ of order m as that approximation to $I(\mathbf{z})$ that contains only monomials of degree m and less:

$$I^{(m)}(\mathbf{z}) = I(\mathbf{z}) + \mathcal{O}(|\mathbf{z}|^{m+1}).$$

In order to check the convergence of $I^{(m)}$ as a function of m, we choose a point $\mathbf{s} \in \Sigma_E$ of the Poincaré surface of section and evaluate the quasi-integral for $\mathbf{z}(t) = \Phi_t(\mathbf{s})$ with $\Phi_t(\mathbf{s})$ being the phase flow of the system (33):

$$I^{(m)}(t; \mathbf{s}) = I^{(m)}(\Phi_t(\mathbf{s})).$$

Figure 4 demonstrates that it depends both on the energy (tantamount to the 'degree of chaoticity' of the system) and on the starting point $\mathbf{s}$ whether convergence of the quasi-integral $I^{(m)}(t; \mathbf{s})$ is observed or not. At the low energy $E = 0.01$ most $\mathbf{s}$ yield fast convergence (figure 4(*a*)). At the energy $E = 0.2$ divergence occurs almost everywhere, but it is typical that for the first few values of m one has 'pseudo-convergence' before the expected divergence takes over (figure 4(*b*)).

The above analysis is *local* in the sense that one has to specify a single point $\mathbf{s}$ for which $I^{(m)}(t; \mathbf{s})$ is evaluated. We now turn to a kind of *global* analysis and show how

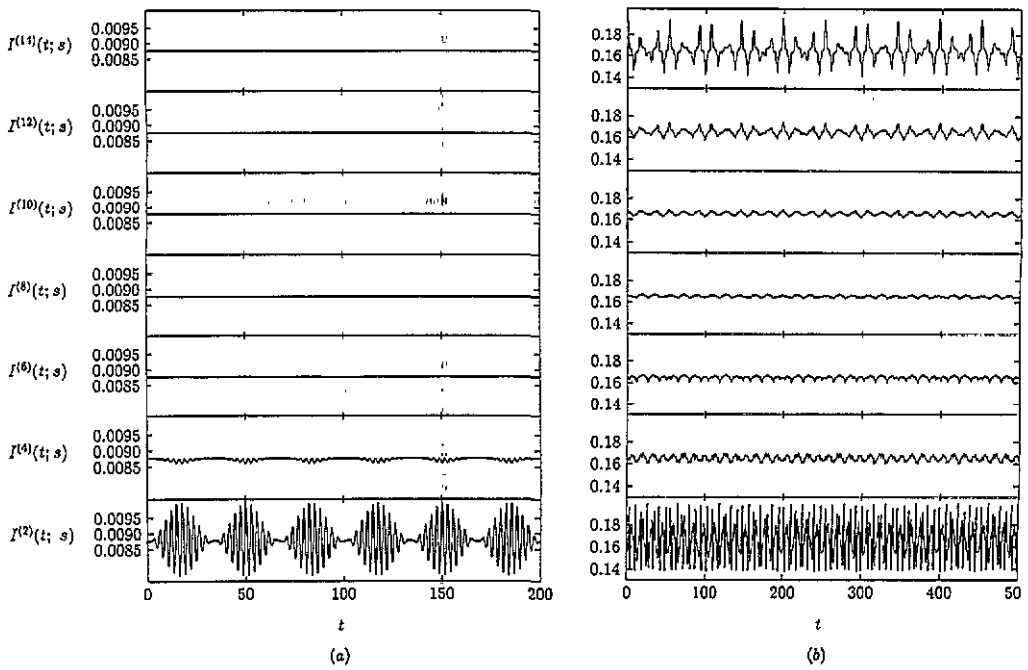


Figure 4. The quasi-integral $I^{(m)}(t; \mathbf{s})$ of the magnetic bottle (33) plotted as a function of time for several different values of m. (*a*) $E = 0.01$, $\mathbf{s} = (0, 0.05)$; (*b*) $E = 0.2$, $\mathbf{s} = (0, 0.3)$.

to obtain a qualitative picture of the convergence properties of $I(\mathbf{z})$ in a larger region of phase space.

In order to have a (though somewhat strong) criterion for convergence we say that $I^{(m)}(t;\mathbf{s})$ is *convergent* at s if

$$|\overline{I^{(m+2)}}(\mathbf{s})-\overline{I^{(m)}}(\mathbf{s})|<|\overline{I^{(m)}}(\mathbf{s})-\overline{I^{(m-2)}}(\mathbf{s})| \qquad \text{for} \qquad m=4,6,8,10,12. \tag{39}$$

Here $\overline{I^{(m)}}(\mathbf{s})=\lim_{T\to\infty}(1/T)\int_0^T I^{(m)}(t;\mathbf{s})$ is the time average of the quasi-integral for a trajectory starting at s. It would be highly desirable to extend the definition to higher orders m of the quasi-integral, but this is limited by the great computational effort needed for this task. Our definition is similar to the one suggested in [28], but we find it appropriate to calculate averages over whole trajectories rather than considering just the behaviour at the point s.

The next step is to define a convergence function $C(\mathbf{s})$ by setting

$$C(\mathbf{s}):=\begin{cases}1\\0\end{cases} \quad \text{if } \overline{I^{(m)}}(\mathbf{s}) \text{ is } \begin{cases}\text{convergent}\\\text{divergent}\end{cases} \quad \text{in the sense of (39).} \tag{40}$$

Though being quite coarse-grained (since it takes into account only the first few approximants of the formal integral) $C(\mathbf{s})$ allows to estimate the convergence properties reasonably well.

In figure 5 we present a convergence plot for the magnetic bottle (33) at the energy $E=0.2$. This picture should be compared with figure 3(*b*). On a 200×200 grid we have marked with black all points where $C(\mathbf{s})=1$ indicates convergence of the quasi-integral.

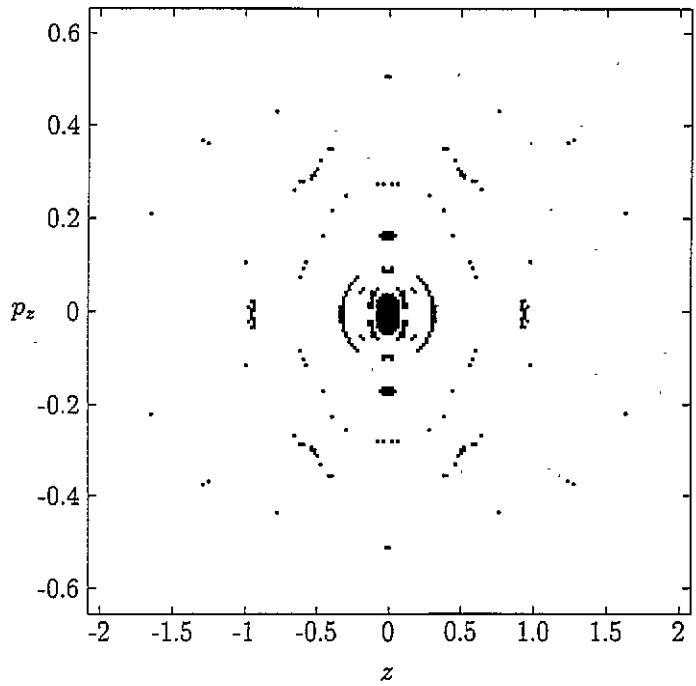


Figure 5. Convergence plot for the magnetic bottle at the energy $E=0.2$. The same Poincaré surface is shown as in figure 3(*b*). On a 200×200 grid the convergence function $C(\mathbf{s})$ has been calculated and the points with $C(\mathbf{s})=1$ have been marked black.

It is interesting to see that although the condition (39) is quite strong there are large convergent regions in Σ_E. Many of the details of the Poincaré plot 3(*b*) show up in the convergence plot as well. As expected, convergence is most frequent in the centre of the picture. Furthermore, the *hyperbolic* periodic points of the two dominant Birkhoff chains (of period four and six, respectively) in figure 3(*b*) are clearly represented in figure 5 by clouds of black marks. This is surprising, because in the neighbourhoods of these hyperbolic points one would expect distinct divergent behaviour, caused by the chaotic dynamics in a heteroclinic tangle scenario.

As an aside we remark that the convergence of the quasi-integral for the magnetic bottle considered here is worse than the convergence for the Hénon–Heiles system discussed in [28]. This seems to be due to the fact that for the latter system the accessible region of phase space is bounded, whereas for the magnetic bottle the dynamical region extends infinitely along the z-axis.

In order to get refined information about the convergence properties of the $I^{(m)}(\mathbf{z})$ as a function of m we now modify the rule for marking points in the convergence plot. As a measure for the deviation of $I^{(m)}(\mathbf{z})$ from its mean value we calculate the standard deviation

$$\sigma^{(m)}(\mathbf{s}) = \sqrt{\lim_{T\to\infty} \frac{1}{T}\int_0^T (\mathrm{I}^{(m)}(t;\mathbf{s}) - \overline{I^{(m)}}(\mathbf{s}))^2 \,\mathrm{d}t}$$

and for normalization (and for comparability of results that belong to different $\mathbf{s}$ and thus to different mean values of the quasi-integral) we divide by $|\overline{I^{(m)}}(\mathbf{s})|$:

$$\eta^{(m)}(\mathbf{s}) := \frac{\sigma^{(m)}(\mathbf{s})}{|\overline{I^{(m)}}(\mathbf{s})|}.$$

This quantity will play a key role in the following. We can take

$$\eta^{(m)}(\mathbf{s}) < \eta^{(m-2)}(\mathbf{s}) \qquad \text{for} \qquad m = 4, 6, 8, \ldots, 14 \tag{41}$$

as new necessary criterion for convergence. This criterion is similar, but not equal to (39). With (41) one can say more about the divergent quasi-integrals: For each of them there is an $m_0(\mathbf{s})$ such that $\eta^{(m_0)}(\mathbf{s}) \leqslant \eta^{(m)}(\mathbf{s})$ for all $m \neq m_0$. A convergent (in the sense of (41)) quasi-integral is characterized by $m_0(\mathbf{s}) = 14$ (the highest order of approximation). If divergence occurs from the beginning, i.e. if the deviation of the quasi-integral from its mean value grows for all available m, then we have $m_0(\mathbf{s}) = 2$. Intermediate values are possible as well. A typical situation (with $m_0 = 8$) is shown in figure 6.

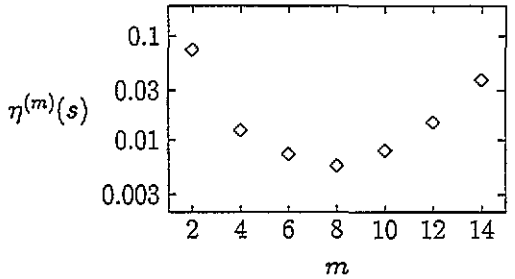
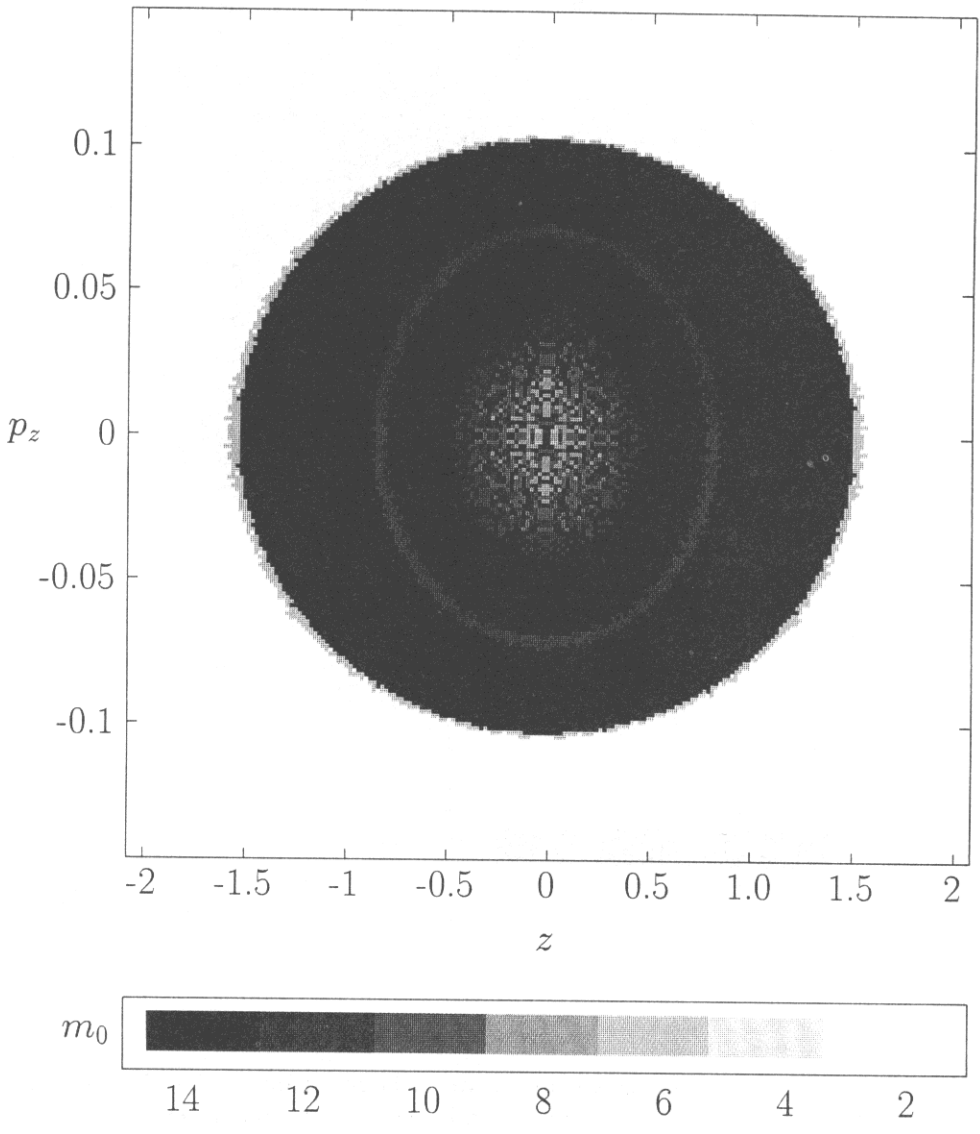


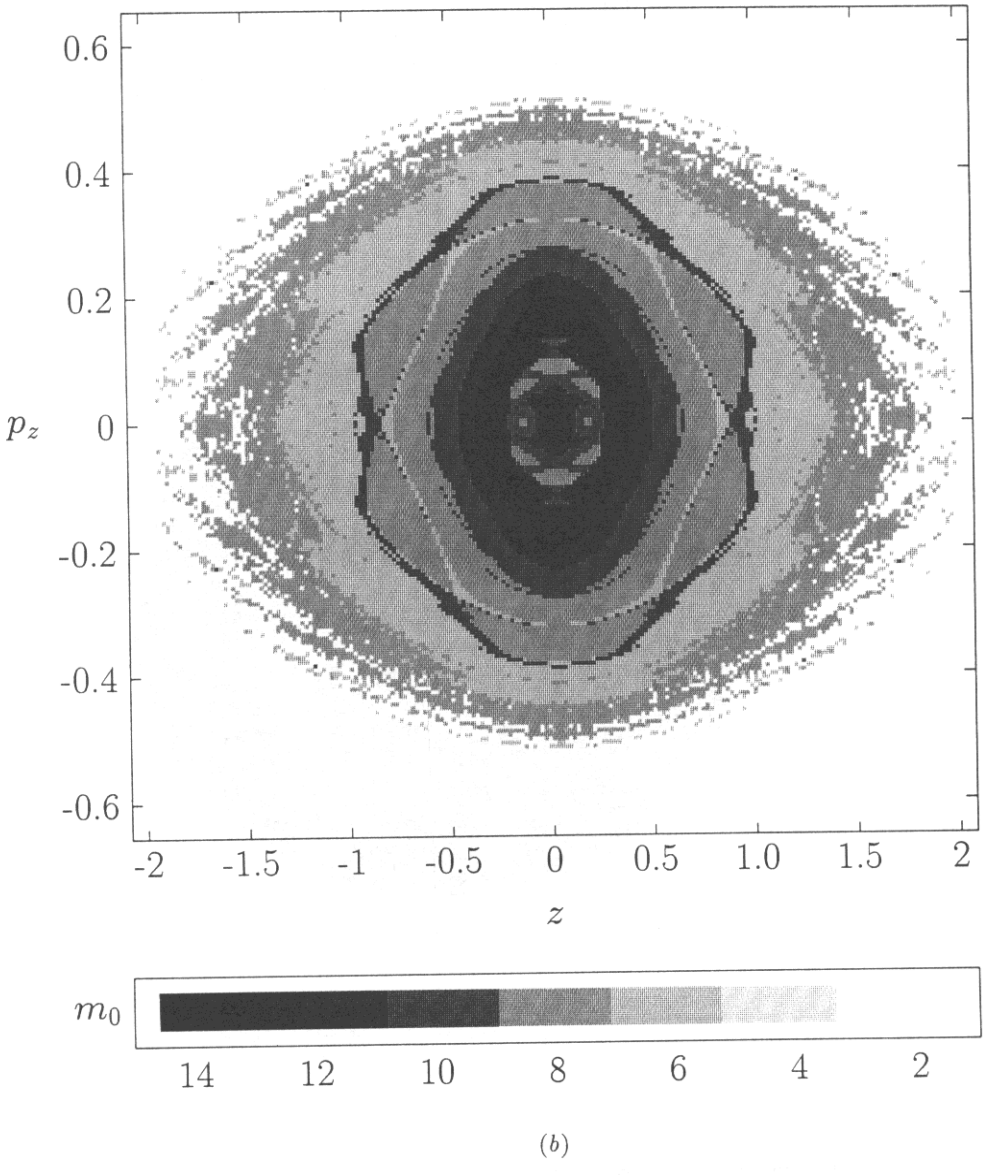
Figure 6. A typical graph of $\eta^{(m)}(\mathbf{s})$. A region of pseudo-convergence and a divergent region are separated by $m_0(\mathbf{s}) = 8$. The parameters for this picture are $E = 0.2$, $\mathbf{s} = (0, 0.3)$.

Again we can mark the points of the surface of section of the magnetic bottle, this time according to the value of $m_0(\mathbf{s})$. The result can be seen in figure 7, which is to be compared with the Poincaré plots of figure 3. The dark regions correspond to larger values of $m_0(\mathbf{s})$, thus indicating—*cum grano salis*—convergence, while points in the light grey or white areas have small $m_0(\mathbf{s})$, which means that $\eta^{(m)}(\mathbf{s})$ increases quite from the beginning. The Poincaré plot 3(*a*) at a very low energy shows regular motion, and



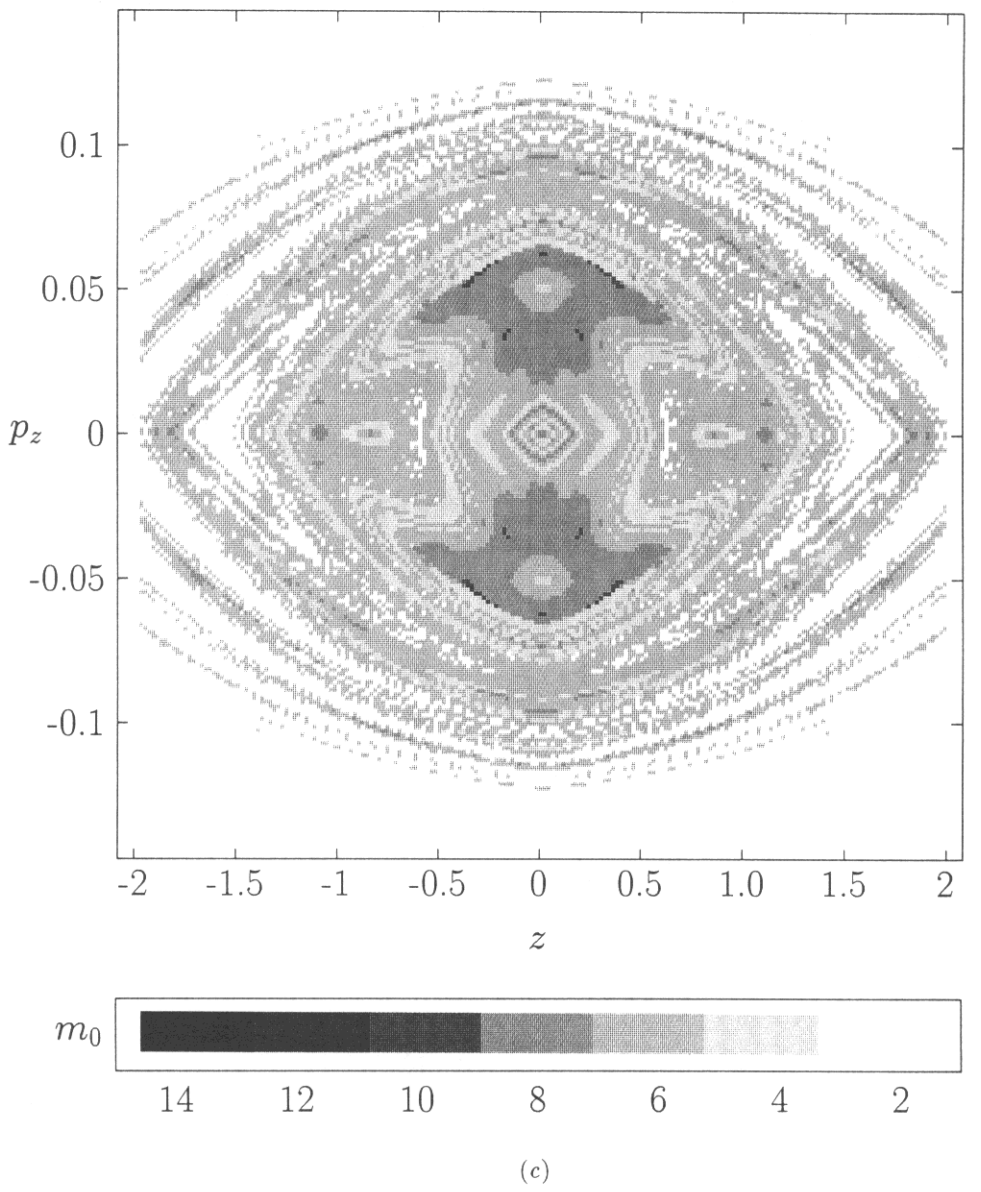
(*a*)

Figure 7. $m_0(\mathbf{s})$-plots for the magnetic bottle at the same energies as in figure 3. Again the Poincaré surface is shown as a 200 × 200 grid of points which are shaded, this time according to their respective values of $m_0(\mathbf{s})$, as shown in the key. (*a*) $E=0.01$, (*b*) $E=0.2$, (*c*) $E=0.5$.



(*b*)

figure 7(*a*) accordingly indicates convergence in large regions of Σ_E. The light grey spots in the centre of the figure must not be mistaken as indicating divergent behaviour. On the contrary, convergence is excellent around the origin, such that very small values of $\eta^{(m)}(\mathbf{s})$ are being compared, and values of $m_0(\mathbf{s})$ larger than two are due to the limited accuracy of the numerical calculation and round-off. Convergence deteriorates with increasing E (and thus increasing chaoticity) as the comparison of figures 3(*b*), (*c*) and 7(*b*), (*c*) shows. In particular, figure 7(*b*) reproduces the content of the Kaluža-Robnik-type figure 5 very well and even adds much more information about the



(*c*)

convergent regions. We conclude that the $m_0(\mathbf{s})$-plots are considerably better suited than the $C(\mathbf{s})$-plots for the analysis of the convergence properties of the quasi-integral.

It is important to keep in mind that we are using quite a special definition of 'convergence' here. Whilst this is useful for the present discussion, comparison with rigorous theoretical results about the divergence mechanism [31, 32] is delicate. Generally, (pseudo-) convergence (in the usual sense) is expected within a disc, which is compatible with figures 5 and 7(*a*). But figure 7(*b*) seems to indicate that convergence has spread into the region between the island chains of period four and six, at variance to the theoretical prediction. The reason being that $m_0(\mathbf{s}) = 2$ or 4 reflects only the

behaviour for the first few orders m and is no direct indicator for true convergence. So the values of $m_0(\mathbf{s})$ can be taken only as a vague phenomenological hint towards true convergence or divergence.

Figure 7(*b*) already reproduces reasonably well the features of the corresponding Poincaré plot, but only in regions not too far away from the origin. In an attempt to enlarge the region that is accessible for the analysis we make the following exponential *ansatz* for the normalized standard deviations as functions of m:

$$\eta^{(m)}(\mathbf{s}) \approx a(\mathbf{s})\, \mathrm{e}^{\alpha(\mathbf{s})m} \tag{42}$$

with $a(\mathbf{s})$ and $\alpha(\mathbf{s})$ to be determined. Here, one is especially interested in the speed of convergence/divergence that is expressed by $\alpha(\mathbf{s})$. For a justification of the approximation (42), we have considered some typical graphs of $\eta^{(m)}(\mathbf{s})$ in figure 8 and determined (with a least-squares method) the corresponding $a(\mathbf{s})$ and $\alpha(\mathbf{s})$. As can be concluded from figures 8(*a*) and (*b*), often the approximation (42) seems to work reasonably well. Figure 8(*c*) shows a case where a transition from convergence to divergence occurs. Taking into account situations like this, we have decided to rely only on the $\eta^{(m)}(\mathbf{s})$

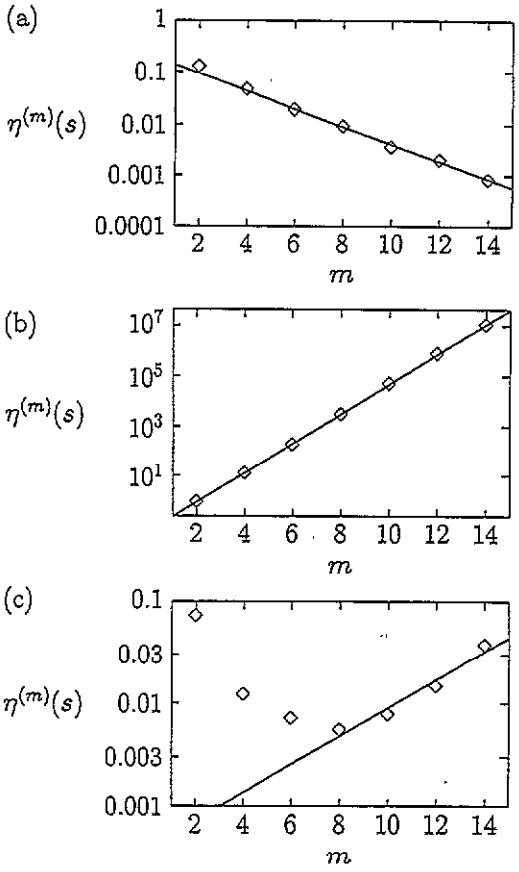


Figure 8. Normalized standard deviations $\eta^{(m)}(\mathbf{s})$ (marked by ◇) and their approximants $a(\mathbf{s})\,\mathrm{e}^{\alpha(\mathbf{s})}$ (solid lines) for the magnetic bottle. The parameters $a(\mathbf{s})$ and $\alpha(\mathbf{s})$ have been computed using the data for $m=8, 10, 12, 14$. (*a*) $E=0.01$, $\mathbf{s}=(0.81, 0.0553)$; (*b*) $E=0.2$; $\mathbf{s}=(1.69, 0.597)$; (*c*) $E=0.2$; $\mathbf{s}=(0, 0.3)$.

with $m=8, 10, 12, 14$ for the calculation of $\alpha(\mathbf{s})$, because we are mainly interested in the convergence properties for larger values of m. With this convention even a behaviour such as that in figure 8(*c*) can be handled reasonably.

It is one of the advantages of the $\alpha(\mathbf{s})$-method that one obtains a continuous spectrum of values of $\alpha(\mathbf{s})$, as opposed to the discrete spectra of $C(\mathbf{s})$ and $m_0(\mathbf{s})$. This becomes apparent in figure 9, where we show again $\Sigma_{0.2}$, now shaded according to $\alpha(\mathbf{s})$. The lighter the grey, the larger $\alpha(\mathbf{s})$ and thus the faster the divergence of the quasi-integral. The central portion of the picture is similar to the one of figure 7(*b*), but now the outer regions show some structure, too. Comparing with the corresponding Poincaré

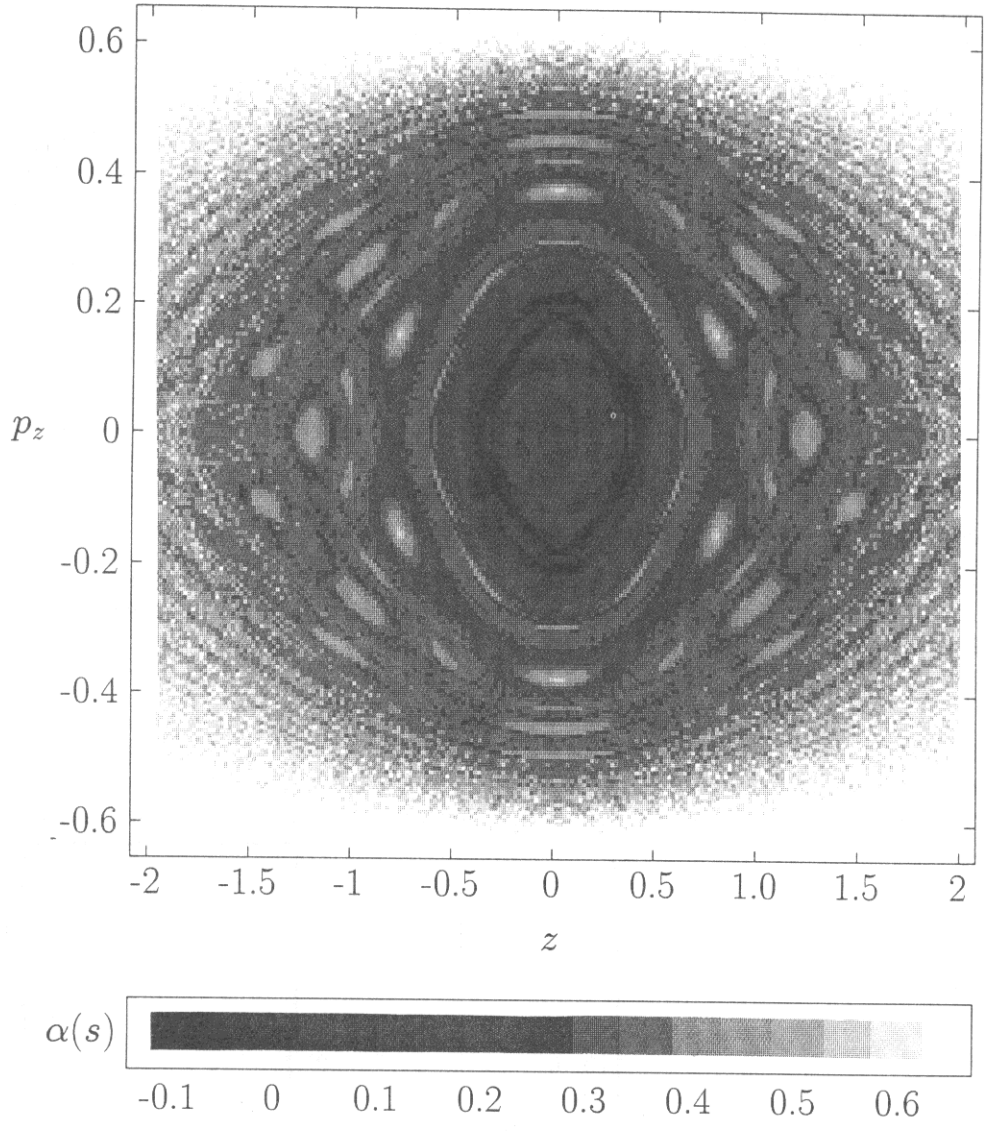


Figure 9. $\alpha(\mathbf{s})$-plot for the magnetic bottle at the energy $E=0.2$. The 200×200 grid points are shaded according to their respective values of $\alpha(\mathbf{s})$. The calculation of these values is based on $\eta^{(m)}(\mathbf{s})$ for $m=8, 10, 12, 14$.

plot (figure 3(*b*)) one sees that the $\alpha(s)$-plot clearly marks the third and fourth largest Birkhoff chains as well. Even more structure can be detected by more careful analysis of the picture. So the ansatz (42) seems to be justified.

4. Concluding remarks

Let us briefly summarize our main results. We have described a generalized version of the powerful tool of normal form theory for Hamiltonian systems. Using this generalized technique, it is now possible to analyze *any* Hamiltonian that is given as a power series in phase-space coordinates. Even if the Hamiltonian is not given in the form of a power series one can always expand H into its Taylor series and normalize the truncated expansion. Thus a large variety of Hamiltonian systems can be analysed in a unified way.

The most important result of a normalization is the derivation of a formal integral of motion that, in general, is different from (and often independent of) the already known integral H. That means that one can obtain substantial new information about the system by normalization. Convergence of this formal integral of motion cannot be taken for granted. Addressing this problem, we have suggested some new methods for analysing the convergence properties of the truncated formal integral. We have found that these quasi-integrals are of physical interest, since their convergence properties reflect many of the characteristics of the corresponding Poincaré plots.

It is certainly necessary to carry the analysis of the convergence of the quasi-integrals further. Many authors [10, 33, 29] have suggested studying the poles of Padé approximations to the quasi-integrals. The location and the number of poles of these approximants then allow us to gauge the properties of $I^{(m)}$.

Appendix. Details of the normalization process for a magnetic bottle

In this appendix we discuss some details of the transformation of a Hamiltonian with a quadratic part (34) into generalized normal form. More specifically, we show how to simplify the Lie operator

$$\mathscr{A}_m(\cdot)=\left[\sum_{\nu=1}^{l} z_{n+\nu}\frac{\partial}{\partial z_\nu}(\cdot)+\sum_{\nu=l+1}^{n}\omega_\nu\left(z_{n+\nu}\frac{\partial}{\partial z_\nu}(\cdot)-z_\nu\frac{\partial}{\partial z_{n+\nu}}(\cdot)\right)\right]_{\mathscr{P}_m}$$

which is adjoint to this particular H_2.

The unitary matrix that by a similarity transformation puts the Hamiltonian matrix

$$L=J\,\mathrm{Hess}(H_2)=\begin{pmatrix}\begin{bmatrix}0 & 0 & \cdots & 0\\ 0 & \ddots & & \\ & 0 & \ddots & \vdots\\ \vdots & \ddots & -\omega_{l+1} & \\ & & & \ddots\ 0\\ 0 & \cdots & & 0\ -\omega_n\end{bmatrix} & 0_n\\ 0_n & \begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & \ddots & & \\ & & 1 & \ddots\ \vdots\\ \vdots & & \ddots\ \omega_{l+1} & \\ & & & \ddots\ 0\\ 0 & \cdots & & 0\ \omega_n\end{bmatrix}\end{pmatrix}$$

into the Jordan normal form $\tilde{L}=MLM^*$ of (35) is

$$M=\begin{pmatrix} \mathbf{e}_1^T \\ \mathbf{e}_{n+1}^T \\ \mathbf{e}_2^T \\ \mathbf{e}_{n+2}^T \\ \vdots \\ \mathbf{e}_l^T \\ \mathbf{e}_{n+l}^T \\ \frac{1}{\sqrt{2}}(\mathbf{e}_{l+1}^T - \mathrm{i}\mathbf{e}_{n+l+1}^t) \\ \vdots \\ \frac{1}{\sqrt{2}}(\mathbf{e}_n^T - \mathrm{i}\mathbf{e}_{2n}^t) \\ \frac{1}{\sqrt{2}}(-\mathrm{i}\mathbf{e}_{l+1}^T + \mathbf{e}_{n+l+1}^T) \\ \vdots \\ \frac{1}{\sqrt{2}}(-\mathrm{i}\mathbf{e}_n^T + \mathbf{e}_{2n}^T) \end{pmatrix}$$

with $\mathbf{e}_i$ being the canonical base vectors of $\mathbb{R}^{2n}$. This formula can easily be derived by performing a certain permutation of the rows and columns first, followed by a transformation similar to (15).

In the new coordinates $\tilde{\mathbf{z}}=M\mathbf{z}$ the Lie operator takes onthe form

$$\tilde{\mathscr{A}}_m=\left[\sum_{\nu=1}^{l}\tilde{z}_{2\nu}\frac{\partial}{\partial\tilde{z}_{2\nu-1}}+\sum_{\nu=l+1}^{n}\mathrm{i}\omega_\nu\left(\tilde{z}_{l+\nu}\frac{\partial}{\partial\tilde{z}_{l+\nu}}-\tilde{z}_{n+\nu}\frac{\partial}{\partial\tilde{z}_{n+\nu}}\right)\right]_{\mathscr{L}_m}.$$

This representation of the Lie operator is advantageous, because here we have collected as many non-zero entries (of the matrix representation) of $\mathscr{A}_m$ on the diagonal as possible. Only the first sum yields an off-diagonal contribution.

We have not yet made any assumptions about the ordering of the monomials $\tilde{\mathbf{z}}^j$ in the basis of $\mathscr{L}_m$. If one chooses the *lexicographical ordering* [2] of the basis monomials, then the matrix representation of $\tilde{\mathscr{A}}_m$ becomes an upper diagonal matrix for all m, and all the manipulations of $\mathscr{A}_m$ that are necessary in the course of the normalization procedure (solving linear equations, inverting $\mathscr{A}_m$, ...) become easier and consume much less computing time.

In the case $l=1$ it is possible to achieve even further simplification by an appropriate ordering of the basis monomials of $\mathscr{L}_m$. One can introduce the so-called *magnetic bottle ordering* of monomials which results in $\tilde{\mathscr{A}}_m$ being bi-diagonal.

References

[1] Birkhoff G D 1966 *Dynamical Systems* 2nd ed (New York: American Mathematical Society)
[2] Gustavson F G 1966 *Astron. J.* **71** 670

[3] Bryuno A D 1971 *Trans. Moscow Math. Soc.* **25** 131
[4] Moser J K 1968 *Mem. Amer. Math. Soc.* **81** 1
[5] Churchill R C and Lee D 1984 Harmonic oscillators at low energies *Classical and Quantum Models and Arithmetic Problems* ed D V Chudnovsky and G V Chudnovsky (New York: Dekker) p 239
[6] Arnold V I, Kozlov V V and Neishtadt A I 1985 *Dynamical Systems III: Mathematical Aspects of Classical and Celestial Mechanics* (Berlin: Springer)
[7] Ozorio de Almeida A M 1988 *Hamiltonian Systems: Chaos and Quantization* (Cambridge: Cambridge University Press)
[8] Hoveijn I and Verhulst F 1990 *Physica D* **44** 397
[9] Swimm R T and Delos J B 1979 *J. Chem. Phys.* **71** 1706
[10] Shirts R B and Reinhardt W P 1982 *J. Chem. Phys.* **77** 5204
[11] Robnik M 1984 *J. Phys. A: Math. Gen.* **17** 109
[12] Eckhardt B 1986 *J. Phys. A: Math. Gen.* **19** 2961
[13] Crehan P 1990 *J. Phys. A: Math. Gen.* **23** 5815
[14] Contopoulos G and Vlahos L 1975 *J. Math. Phys.* **16** 1469
[15] Dragt A J and Finn J M 1976 *J. Geophys. Res.* **81** 2327
[16] Dragt A J and Finn J M 1976 *J. Math. Phys.* **17** 2215
[17] Dragt A J and Finn J M 1979 *J. Math. Phys.* **20** 2649
[18] Cushman R, Deprit A and Mosak R 1983 *J. Math. Phys.* **24** 2102
[19] Meyer K R and Hall G R 1992 *Introduction to Hamiltonian Dynamical Systems and the N-Body Problem* (Berlin: Springer)
[20] Deprit A 1969 *Cel. Mech.* **1** 12
[21] Cary J R 1981 *Phys. Rep.* **19** 129
[22] Elphick C, Tirapegui E, Brachet M E, Coullet P and Iooss G 1987 *Physica* D **29** 95
[23] Arnold V I 1983 *Geometrical Methods in the Theory of Ordinary Differential Equations* (New York: Springer)
[24] Bargmann V 1961 *Comm. Pure Appl. Math.* **14** 187
[25] Humphreys J E 1987 *Introduction to Lie Algebras and Representation Theory* (New York: Springer)
[26] Fomenko A T 1988 *Symplectic Geometry* (New York: Gordon and Breach)
[27] Brown L S and Gabrielse G 1986 *Rev. Mod. Phys.* **58** 233
[28] Kaluža M and Robnik M 1962 *J. Phys. A: Math. Gen.* **25** 5311
[29] Robnik M 1993 *J. Phys. A: Math. Gen.* **26** 7427
[30] Servizi G, Turchetti G, Benettin G and Giorgilli A 1983 *Phys. Lett.* **95A** 11
[31] Giorgilli A and Galgani L 1985 *Cel. Mech.* **37** 95
[32] Bazzani A, Marmi S and Turchetti G 1990 *Cel. Mech.* **47** 333
[33] Ali M K, Wood W R and Devitt J S 1986 *J. Math. Phys.* **27** 1806