Birkhoff-Gustavson normal form in classical and quantum mechanics

1986 J. Phys. A: Math. Gen. 19 2961

(http://iopscience.iop.org/0305-4470/19/15/020)

J. Phys. A: Math. Gen. 19 (1986) 2961-2972. Printed in Great Britain

Birkhoff–Gustavson normal form in classical and quantum mechanics

Bruno Eckhardt
Fachbereich Physik, Universität Bremen, 2800 Bremen 33, West Germany

Received 13 November 1985, in final form 6 February 1986

Abstract. A quantum analogue of the classical transformations to the Birkhoff–Gustavson normal form is derived. With unitary transformations taking over the role of canonical transformations, we find a striking similarity to Lie transformations in the classical case. We suggest that the quantum normal form is identical to the Rayleigh–Schrödinger perturbation series and thus usually only asymptotic to the eigenvalues. The convergence question of both classical and quantum normal forms is discussed in some detail.

1. Introduction

An old problem in quantum mechanics is how to extract semiclassical approximations to eigenvalues of quantum operators from the classical observables. For separable systems, the JWKB method and for more general integrable systems, EBK quantisation provide the solutions. However, as is known today, most Hamiltonian systems are not integrable but instead show a divided phase space. Although it has been possible to extend the EBK procedure and variations thereof to elliptic islands, the general problem of how to treat chaotic regions remains unsolved. For a review see Percival (1977).

One way around this difficulty was pointed out by Swimm and Delos (1979): instead of quantising the original non-integrable Hamiltonian they used integrable approximations to it. From the many possibilities (see, e.g., Rice 1981, Lichtenberg and Liebermann 1983), they chose a truncated Birkhoff–Gustavson normal form around an equilibrium point (Birkhoff 1927, Gustavson 1966). The method has since been extended by Jaffé and Reinhardt (1982), Shirts and Reinhardt (1982) and Robnik (1984), typically yielding good agreement with exact quantum calculations for low lying states, the lowest ones being best approximated. An exception seems to be the work of Williams and Koonin (1982), but see the discussion in § 5.

Truncation of the transformation to normal form is necessary since it must diverge, for otherwise the original system would be integrable (Siegel and Moser 1971). However, it is asymptotic, so the first few terms approximate trajectories of the full system quite accurately for short times (Duistermaat 1984). This hints at one reason for the success of the quantum calculations: if the density of states is written as a sum over periodic orbits (cf Berry and Tabor 1976), then numerical experience (cf Berry 1983) shows that short time periodic orbits dominate at low energies.

Another reason is the close connection to quantum mechanical perturbation theory, which we will develop in this paper. We will show that it is possible to define a quantum normal form with unitary transformations playing the role of the canonical transformations in classical mechanics. This form of perturbation theory is similar to the Fouldy-Wouthuysen transformations in relativistic quantum theory (Bjorken and Drell 1964). The analogy between classical and quantum normal forms becomes evident, if the former are derived in a Lie algebra framework. It can then be shown that the classical normal form, suitably quantised, is the leading order (in a sense made clear below) of the quantum version. We will argue that the quantum normal form to order k is equivalent to kth order Rayleigh-Schrödinger perturbation theory. This brings up the question of convergence, which will be discussed in detail in § 5. We merely point out that we expect the perturbation series to diverge, but that powerful resummation techniques might exist, allowing eigenvalues to be retrieved quite accurately (cf Simon 1982). It is an open question whether the classical normal form can be resummed to something meaningful (for preliminary results see Shirts and Reinhardt (1982)).

After this work was completed we learned of the paper by Ali (1985) where the same connection is found, but no convergence questions are discussed. The outline of this paper is as follows: in § 2 we present the Lie group version of canonical transformations and in § 3 we apply it to derive the Birkhoff-Gustavson normal form for classical Hamiltonians. In § 4 the quantum normal form is introduced. Convergence questions and the relationship to Rayleigh-Schrödinger perturbation theory are discussed in § 5. We conclude with a summary and a few remarks in § 6. Applications to anharmonic oscillators in one and two dimensions are sketched in §§ 3 and 4.

2. Canonical transformations and Lie groups

We begin with a description of canonical transformations in the context of Lie groups and the Lie algebra of vector fields (Dragt and Finn 1976). For the differential geometry involved, see Abraham and Marsden (1981), Arnold (1978) or Thirring (1978).

Let Γ be the $2N$-dimensional phase space with symplectic coordinates $(q, p) = (q_1 \ldots q_N, p_1 \ldots p_N)$ and non-degenerate 2-form $d\Omega = \Sigma_{i=1}^{N} dp_i \wedge dq_i$. The vector fields over Γ form a Lie algebra with the commutator as product structure. The Jacobi identity

$$[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0 \tag{1}$$

holds.

With any smooth function $f(q,p)$ on Γ one can associate a Hamiltonian vector field X_f via

$$X_f = -\sum_{n=1}^{N}\left(\frac{\partial f}{\partial q_n}\frac{\partial}{\partial p_n} - \frac{\partial f}{\partial p_n}\frac{\partial}{\partial q_n}\right). \tag{2}$$

The action of X_f on functions g yields the usual Poisson bracket:

$$(X_f g)(q,p) = -\{f, g\}(q,p). \tag{3}$$

The solution of the system of ordinary differential equations

$$d\varphi/ds = X_f \varphi \tag{4}$$

with $\varphi(s)$ a $2N$-dimensional vector of coordinate functions defines a flow on Γ. Formally, one may write

$$\varphi(s) = \exp(sX_f)\varphi(0) \tag{5a}$$
$$= (1 + sX_f + \tfrac{1}{2}s^2X_f^2 + \ldots)\varphi(0) \tag{5b}$$

where powers of vector fields are defined recursively.

The nomenclature becomes clear when one takes $f = H$, the usual Hamiltonian, for then (4) are Hamilton's equations of motion and the parameter s is ordinary time.

Quite generally, the solutions to (4) generate one-parameter groups of canonical transformations. To first order in the parameter ε this is verified easily. Let (Q, P) be the new coordinates, then

$$\begin{aligned} Q_n &= q_n - \varepsilon \frac{\partial S}{\partial p_n}(q, p) + \mathrm{O}(\varepsilon^2) \\ P_n &= p_n + \varepsilon \frac{\partial S}{\partial q_n}(q, p) + \mathrm{O}(\varepsilon^2) \end{aligned} \tag{6}$$

and

$$\mathrm{d}\Omega = \sum_{n=1}^{N} \mathrm{d}P_n \wedge \mathrm{d}Q_n = \sum_{n=1}^{N} \mathrm{d}p_n \wedge \mathrm{d}q_n + \mathrm{O}(\varepsilon^2) \tag{7}$$

is invariant. Thus, Hamiltonian vector fields generate canonical transformations and, as shown by Dragt and Finn (1976), any canonical transformation continuously connected to the identity ($\varepsilon = 0$) can be obtained this way. The function $S(q, p)$ is called the generator of the transformation. If $\varphi = (q, p)$ and $\Phi = (Q, P)$ denote the old and new coordinates, respectively, then one can write formally

$$\Phi = \exp(\varepsilon X_S)\varphi \qquad \varphi = \exp(-\varepsilon X_S)\Phi \tag{8}$$

where S' has the same functional form as S, but expressed in (Q, P) instead of (q, p).

If S is independent of time, as we have tacitly assumed so far, Hamilton's equations of motion transform according to

$$\exp(-\varepsilon X_S)\, \mathrm{d}\Phi/\mathrm{d}t = X_H \exp(-\varepsilon X_S)\Phi \tag{9}$$

or

$$\mathrm{d}\Phi/\mathrm{d}t = X_{H'}\Phi \tag{10a}$$

with the new vector field

$$X_{H'} = \exp(\varepsilon X_S) X_H \exp(-\varepsilon X_S) \tag{10b}$$

$$X_{H'} = \sum_{n=0}^{\infty} \frac{\varepsilon^n}{n!} \underbrace{[X_S, [X_S, \ldots, [X_S, X_H]]}_{n\ \text{times}} \tag{11}$$

where the last equality follows from series expansion. A final simplification may be achieved by noting that from (3) and (1) one obtains

$$[X_S, X_H](f) = -X_{\{S,H\}}(f) \tag{12}$$

so that the transformed Hamiltonian may be recovered directly:

$$H' = \sum_{n=0}^{\infty} \frac{(-\varepsilon)^n}{n!} \underbrace{\{S, \{S, \ldots \{S, H\}\}}_{n\ \text{times}}. \tag{13}$$

This is our starting point for the discussion of the Birkhoff–Gustavson normal form.

3. Normal forms in classical mechanics

Normal forms for classical Hamiltonians around elliptic fixed points were extensively studied in Birkhoff (1927). Our presentation of resonant cases goes back to Gustavson (1966). Many of the references on semiclassical quantisation of the normal form contain a discussion of the classical case as well.

Assume that the full Hamiltonian H has been expanded around one of its equilibrium points as

$$H = H_0 + \sum_{k=1}^{\infty} \lambda^k H_k \tag{14}$$

where

$$H_0 = \sum_{n=1}^{N} \frac{\omega_n}{2}(p_n^2 + q_n^2) \tag{15}$$

and the H_k are polynomials in $(q_1 \ldots q_N, p_1 \ldots p_N)$ homogeneous of degree $k+2$. Constants in H do not matter and linear terms can be absorbed by a change of coordinates. So this form is quite general.

We say H_k is in normal form, if it is a function of $(p_i^2 + q_i^2)$ alone, i.e. $H_k^{(\mathrm{normal})}(p_1^2 + q_1^2, \ldots, p_N^2 + q_N^2)$. Obviously, the normal forms for H_n with n odd are identically zero, since (p_i, q_i) should enter quadratically.

It was Birkhoff's idea to find a sequence of canonical transformations bringing ever increasing orders of H_k to normal form, i.e. eliminating non-normal terms.

For the first step (assuming H_1 is non-zero), consider equations (13) and (14) to $O(\lambda^2)(\varepsilon = \lambda)$:

$$H' = H_0 + \lambda(H_1 - \{S_1, H_0\}) + O(\lambda^2). \tag{16}$$

Since H_1 is of odd order, we need to find S_1 such that

$$\{S_1, H_0\} - H_1 = 0. \tag{17}$$

This is possible only if (H_1) is in the range of the operator $\mathscr{D} = \{\cdot, H_0\}$ with domain, for example, all smooth functions.

Deferring the question of the existence of S_1 for a moment, let us proceed to the next step. We now start from

$$H' = H_0 + \lambda^2 H_2' + \lambda^3 H_3' + O(\lambda^4). \tag{18}$$

Applying a canonical transformation with $\varepsilon = \lambda^2$ and generator S_2, we obtain

$$H'' = H_0 + \lambda^2(H_2' - \{S_2, H_0\}) + \lambda^3 H_3' + O(\lambda^4). \tag{19}$$

This shows two things. First of all, the existence of S_2 is again linked to the operator $\mathscr{D}$, namely the non-normal parts of H_2' have to be in the range of $\mathscr{D}$. Second, since H_3' is not affected by the transformation, we can add a term λS_3 to S_2, thus obtaining a 'superconvergent' procedure. Without writing out the details, it should be clear that in the next step one can take $\varepsilon = \lambda^4$ and a generator $S_4 + \lambda S_5 + \lambda^2 S_6 + \lambda^3 S_7$, thus taking care of all terms up to order λ^8. Here 'superconvergent' is used in agreement with standard nomenclature (Lichtenberg and Liebermann 1983) to refer to a procedure yielding normal form to order 2^m after m steps. Jaffé and Reinhard (1982) call it the Birkhoff–van Vleck method. It does not imply convergence as a power series (see § 5).

Now we return to a discussion of the operator $\mathscr{D}$. To this end, we change to complex coordinates according to

$$z_n = \frac{1}{\sqrt{2}}(q_n + \mathrm{i}p_n) \qquad z_n^* = \frac{1}{\sqrt{2}}(q_n - \mathrm{i}p_n). \tag{20}$$

The Poisson bracket now becomes

$$\{f, g\} = \mathrm{i} \sum_{n=1}^{N} \left(\frac{\partial f}{\partial z_n^*} \frac{\partial g}{\partial z_n} - \frac{\partial f}{\partial z_n} \frac{\partial g}{\partial z_n^*} \right). \tag{21}$$

In these new coordinates, H_k is normal, if it depends on $z_n^* z_n$ and powers thereof.

The operator $\mathscr{D}$ becomes

$$\mathscr{D} = \mathrm{i} \sum_{n=1}^{N} \omega_n \left(z_n^* \frac{\partial}{\partial z_n^*} - z_n \frac{\partial}{\partial z_n} \right). \tag{22}$$

It is now convenient to introduce multi-indices: $l = (l_1 \ldots l_N)$ denotes a vector of integers, $|l| = \Sigma_{n=1}^{N} l_n$ and $z^l = z_1^{l_1} \ldots z_N^{l_N}$. Similarly for m and z^{*m}. Finally,

$$\omega l = \sum_{n=1}^{N} \omega_n l_n.$$

Then

$$\mathscr{D}(z^{*l} z^m) = \mathrm{i}\omega(l - m) z^{*l} z^m \tag{23}$$

i.e. monominals in z^* and z are eigenfunctions of $\mathscr{D}$.

The kernel of $\mathscr{D}$ contains all monomials with eigenvalue zero, the range those with non-zero eigenvalue. Thus, what is left over after the canonical transformations will will be in the kernel of $\mathscr{D}$. Obviously, normal monomials ($l = m$) are elements of $\ker \mathscr{D}$. In systems with two or more degrees of freedom, it may happen that there is an integer resonance between the frequencies, i.e. there is a set of integers $(j_1 \ldots j_N)$ not all zero such that

$$\sum_{n=1}^{N} \omega_n j_n = 0. \tag{24}$$

Then the kernel is much larger and can be organised as follows. Let there be r ($1 \leq r < N$) relations of the form (24) (rth-order resonance) with integers $(j_{11} \ldots j_{1N}), \ldots, (j_{r1} \ldots j_{rN})$. Then one can define resonant monomials

$$K_s = u_1^{|j_{s1}|} \ldots u_N^{|j_{sN}|} \qquad s = 1, \ldots, r \tag{25}$$

with

$$u_n = \begin{cases} z_n^* & \text{if } j_{sn} > 0 \\ z_n & \text{if } j_{sn} < 0 \\ 1 & \text{if } j_{sn} = 0. \end{cases} \tag{26}$$

Any element of $\ker \mathscr{D}$ can now be written in the form

$$\sum_l f_l(z_1^* z_1, \ldots, z_N^* z_N)(K^l + K^{*l}) \tag{27}$$

where $l = (l_1 \ldots l_r)$ (Robnik 1984, Gustavson 1966).

The resonant case has only been included for completeness and for later reference in the quantum case and will not be discussed further.

So assuming the unperturbed frequencies ω_n are not in resonance, we can completely solve the transformation problem. At each step of the approximation we have to solve an equation of the form

$$\{S_k, H_0\} - H_k = (\text{normal}). \tag{28}$$

If

$$H_k = \sum_{|l|+|m|=k} h_{lm} z^{*l} z^m \tag{29}$$

then, by (23),

$$S_k = -\mathrm{i} \sum_{|l|+|m|=k}{}' \frac{h_{lm}}{\omega(l-m)} z^{*l} z^m \tag{30}$$

provides a solution. Here, the prime denotes omission of all terms with vanishing denominator. The requirement that S should not contain terms in ker $\mathscr{D}$ renders it unique.

Example 1. One-dimensional oscillator
Let

$$H = \tfrac{1}{2}(p^2 + q^2) + \lambda q^4 \tag{31}$$

or, in complex coordinates,

$$H = z^*z + \lambda[\tfrac{3}{2}(z^*z)^2 + \tfrac{1}{4}(z^{*4} + 4z^{*3}z + 4z^*z^3 + z^4)] \tag{32}$$

$$= H_0 + \lambda(H_{2N} + H_{2R}) \tag{33}$$

where $H_{2N} = \frac{3}{2}(z^*z)^2$ is already in normal form. So the aim of the first transformation is to eliminate H_{2R}. We have, to $\mathrm{O}(\lambda^4)$,

$$\begin{aligned} H' = H_0 &+ \lambda[H_{2N} + H_{2R} - \{S, H_0\}] - \lambda^2[\{S, H_{2N}\} + \{S, H_{2R}\} - \tfrac{1}{2}\{S, \{S, H_0\}\}] \\ &+ \lambda^3[\tfrac{1}{2}\{S, \{S, H_{2N}\}\} + \tfrac{1}{2}\{S, \{S, H_{2R}\}\} - \tfrac{1}{6}\{S, \{S, \{S, H_0\}\}\}] + \mathrm{O}(\lambda^4). \end{aligned} \tag{34}$$

If S is determined from

$$\{S, H_0\} - H_{2R} = 0 \tag{35}$$

then we obtain

$$\begin{aligned} H' = H_0 &+ \lambda H_{2D} - \lambda^2[\{S, H_{2N}\} - \tfrac{1}{2}\{S, H_{2R}\}] \\ &+ \lambda^3[\tfrac{1}{2}\{S, \{S, H_{2N}\}\} + \tfrac{1}{3}\{S, \{S, H_{2R}\}\}] + \mathrm{O}(\lambda^4). \end{aligned} \tag{36}$$

An argument similar to that following equation (19) shows that (35) contains the normal form of Hamiltonian (32) up to $\mathrm{O}(\lambda^4)$. With the solution of equation (35),

$$S = -\tfrac{1}{4}\mathrm{i}(\tfrac{1}{4}z^{*4} + 2z^{*3}z - 2z^*z^3 - \tfrac{1}{4}z^4) \tag{37}$$

we finally obtain

$$H = I + \tfrac{3}{2}\lambda I^2 - \tfrac{17}{4}\lambda^2 I^3 + \tfrac{375}{16}\lambda^3 I^4 + \mathrm{O}(\lambda^4) \tag{38}$$

where

$$I = z^*z.$$

To obtain information about the convergence of (38), note that it is an algebraic way of expanding H in terms of the action I, defined by

$$I = \frac{1}{2\pi} \oint (2E - q^2 - 2\lambda q^4)^{1/2}\, \mathrm{d}q. \tag{39}$$

However, trajectories near the origin are bounded only for $E\lambda \geqslant -\frac{1}{16}$, i.e. if $\lambda > 0$ for all positive energies and for $\lambda < 0$ only for those between zero and the tip of the barrier. Now λH is monotonic in λI, so we can expect (38) to converge for

$$|\lambda I| \leqslant \min(I_+, I_-) \tag{40}$$

where

$$I_\pm = \frac{1}{2\pi} \oint (\tfrac{1}{8} - q^2 \mp 2q^4)^{1/2}\, \mathrm{d}q. \tag{41}$$

The point to be made here is that a maximal radius of convergence for (38) is determined by the requirement that for both $+\lambda$ and $-\lambda$, trajectories have to be bounded.

Example 2. Two-dimensional oscillator
Let

$$H = \tfrac{1}{2}\omega_1(q_1^2 + p_1^2) + \tfrac{1}{2}\omega_2(q_2^2 + p_2^2) + \alpha(q_1^4 + q_2^4) + \gamma q_1^2 q_2^2 \tag{42}$$

or, in complex coordinates (z^*, z, u^*, u),

$$H = H_0 + H_{2N} + H_{2R} \tag{43a}$$

with

$$H_0 = \omega_1 z^* z + \omega_2 u^* u \tag{43b}$$

$$H_{2N} = \tfrac{3}{2}\alpha(z^*z)^2 + \tfrac{3}{2}\alpha(u^*u)^2 + \gamma(z^*z)(u^*u) \tag{43c}$$

$$H_{2R} = \tfrac{1}{4}\alpha(z^{*4} + 4z^{*3}z + u^{*4} + 4u^{*3}u)$$
$$+ \tfrac{1}{4}\gamma(z^{*2}u^{*2} + z^{*2}u^2 + 2z^*zu^{*2} + 2z^{*2}u^*u) + \text{complex conjugate}. \tag{43d}$$

Here, too, equations similar to (34)-(36) hold. The generator for the canonical transformation is given by

$$S = -\frac{\mathrm{i}\alpha}{4}\left(\frac{1}{4\omega_1} z^{*4} + \frac{2}{\omega_1} z^{*3} z + \frac{1}{4\omega_2} u^{*4} + \frac{2}{\omega_2} u^{*3} u\right)$$
$$-\frac{\mathrm{i}\gamma}{4}\left(\frac{z^{*2}u^{*2}}{2(\omega_1+\omega_2)} + \frac{z^{*2}u^2}{2(\omega_1-\omega_2)} + \frac{z^*z}{\omega_2} u^{*2} + \frac{u^*u}{\omega_1} z^{*2}\right)$$
$$+ \text{complex conjugate}. \tag{44}$$

One sees that for this system the first generator is non-singular, if $\omega_1 \neq \omega_2$. Higher-order restrictions will only appear in higher orders. The result is

$$H = \omega_1 I + \tfrac{3}{2}\alpha I^2 - \tfrac{17}{4}\alpha^2 I^3 + \omega_2 J + \tfrac{3}{2}\alpha J^2 - \tfrac{17}{4}\alpha^2 J^3 + \gamma IJ$$
$$-\frac{3\alpha\gamma}{\omega_1} I^2 J - \frac{3\alpha\gamma}{\omega_2} IJ^2 - \frac{\gamma^2}{4}\left(\frac{2}{\omega_2} - \frac{\omega_2}{\omega_1^2 - \omega_2^2}\right) I^2 J$$
$$-\frac{\gamma^2}{4}\left(\frac{2}{\omega_1} + \frac{\omega_1}{\omega_1^2 - \omega_2^2}\right) IJ^2 \tag{45}$$

where $I = z^*z$, $J = u^*u$. For a discussion of convergence see § 5.

4. Quantum mechanical normal form

We now develop similar ideas in quantum mechanics. Assume we are given a Hamilton operator H, written for convenience using creation and destruction operators $a^+ = (a_1^+ \dots a_N^+)$, $a = (a_1 \dots a_N)$,

$$H = E_0 + H_0 + \sum_{k=1}^{\infty} \lambda^k H_k \tag{46}$$

where

$$H_0 = \sum_{n=1}^{N} \omega_n a_n^+ a_n \tag{47}$$

and $H_k(a^+, a)$ are polynomials of degree $\leqslant k+2$.

For reasons explained below they need not be homogeneous. Zero-point energies of the unperturbed harmonic oscillators have been absorbed in the constant E_0. Here, we will ignore any ordering problems in quantising a classical normal form. For Hamiltonians of the form kinetic plus potential energy, as in our examples, they do not arise and the theory developed below is independent of the particular classical-quantum association.

Our aim will be to approximate the eigenvalues of H in a simple manner, namely as polynomials in the number operators $N_n = a_n^+ a_n$. So we say a quantum operator is in normal form if it is a function of N_n alone.

Spectra of self-adjoint operators are invariant under unitary transformations U, which may be written as $U = \mathrm{e}^{-\tilde{S}}$ with some anti-self-adjoint operator $\tilde{S}$: $\tilde{S}^+ = -\tilde{S}$.

We have

$$H = U^+ H U \tag{48}$$

$$= \mathrm{e}^{\tilde{S}} H \, \mathrm{e}^{-\tilde{S}} \tag{49}$$

or

$$H' = \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{[\tilde{S}, [\tilde{S}, \dots, [\tilde{S}}_{n\ \text{times}}, H]]. \tag{50}$$

Note the similarity to equation (13). Much of the discussion in § 3, including the 'superconvergent' method, can immediately be carried over, if one reads commutators instead of Poisson brackets. In particular, the existence of the operator S is linked to range questions of the operator $\mathscr{D} = [\cdot, H_0]$. We have

$$\mathscr{D} = \sum_{n=1}^{N} \omega_n [\cdot, a_n^+ a_n] \tag{51}$$

and

$$\mathscr{D}(a^{+l} a^m) = \omega(l-m) a^{+l} a^m. \tag{52}$$

So the kernel of $\mathscr{D}$ again contains normal forms with $l = m$ and, for the case of resonance, resonant monomials and powers thereof.

There is one difference, though, arising from the fact that a^+ and a do not commute. For definiteness we will require that in each monomial all creation operators are commuted to the left of destruction operators (normal ordering). In the final result, however, normal forms are more conveniently expressed in powers of number operators, e.g. $a^{+4}a^4 = N^4 - 6N^3 + 11N^2 - 6N$, etc.

We define a quantisation operator Q by the rule

$$Q(z^{*l}z^m) = a^{+l}a^m. \tag{53}$$

Then

$$[a^{+l}a^m, a^{+l'}a^{m'}] = Q(\{z^{*l}z^m, z^{*l'}z^{m'}\}) + \text{monomials of order } L-2,\ L-4,\ L-6, \ldots 1 \text{ (or 0)} \tag{54}$$

where $L = l + l' + m + m' - 2$.

Because of this, it is not possible, as it is in the classical case, to transform increasing orders of homogeneous polynomials to normal form. Instead, it is more like a perturbative expansion in λ: coefficients of increasing powers of λ are transformed to normal form. However, within these coefficients, the leading power of the number operator is given correctly by the classical series, if quantised using (53).

In the case of resonance, other monomials besides powers of number operators survive the transformation. As shown by Robnik (1984) degenerate quantum perturbation theory can then be applied.

We conclude this section with a discussion of the quantum version of the examples of § 3. Because of the similarity, we will be rather brief.

Example 1. One-dimensional oscillator

The quantum Hamiltonian for equation (31) is

$$H = \hbar(a^+a + \tfrac{1}{2}) + \tfrac{3}{2}\lambda\hbar^2(2a^+aa^+a + 2a^+a + 1) + \tfrac{1}{4}\lambda\hbar^2(a^{+4} + 4a^{+3}a + 4a^+a^3 + a^4 + 6a^{+2} + 6a^2). \tag{55}$$

With the anti-self-adjoint operator

$$S = \tfrac{1}{4}\lambda\hbar(\tfrac{1}{4}a^{+4} + 2a^{+3}a - 2a^+a^3 - \tfrac{1}{4}a^4 + 3a^{+2} - 3a^2) \tag{56}$$

and an expansion like (36) we obtain

$$H = \hbar(N + \tfrac{1}{2}) + \tfrac{3}{2}\lambda\hbar^2(N^2 + N + \tfrac{1}{2}) - \tfrac{1}{8}\lambda^2\hbar^3(34N^3 + 51N^2 + 59N + 21) + \tfrac{1}{16}\lambda^3\hbar^4(375N^4 + 750N^3 + 1416N^2 + 1329N + 333) + O(\lambda^4) \tag{57}$$

where $N = a^+a$.

Example 2. Two-dimensional oscillator

The quantum analogue of model (42) is

$$H = \tfrac{1}{2}\hbar(\omega_1 + \omega_2) + H_0 + H_{2N} + H_{2R} \tag{58}$$

with

$$H_0 = \hbar\omega_1 a^+a + \hbar\omega_2 b^+b \tag{59a}$$

$$H_{2N} = \tfrac{3}{2}\alpha\hbar^2(a^+aa^+a + a^+a + b^+bb^+b + b^+b + 1) + \gamma\hbar^2(a^+a + \tfrac{1}{2})(b^+b + \tfrac{1}{2}) \tag{59b}$$

$$H_{2R} = \tfrac{1}{4}\alpha\hbar^2(a^{+4} + 4a^{+3}a + 6a^{+2} + b^{+4} + 4b^{+3}b + 6b^{+2}) + \tfrac{1}{4}\gamma\hbar^2(a^{+2}b^{+2} + a^{+2}b^2 + a^{+2} + b^{+2} + 2a^+ab^{+2} + 2a^{+2}b^+b) + \text{Hermitian conjugate.} \tag{59c}$$

The anti-self-adjoint operator S is

$$S=\frac{\alpha\hbar}{4}\left(\frac{1}{4\omega_1}a^{+4}+\frac{2}{\omega_1}a^{+3}a+\frac{3}{\omega_1}a^{+2}+\frac{1}{4\omega_2}b^{+4}+\frac{2}{\omega_2}b^{+3}b+\frac{3}{\omega_2}b^{+2}\right)$$
$$+\frac{\gamma\hbar}{4}\left(\frac{a^{+2}b^{+2}}{2(\omega_1+\omega_2)}+\frac{a^{+2}b^2}{2(\omega_1-\omega_2)}+\frac{1}{2\omega_1}a^{+2}+\frac{1}{2\omega_2}b^{+2}+\frac{a^+a}{\omega_2}b^{+2}+\frac{b^+b}{\omega_1}a^{+2}\right)$$
$$+\text{Hermitian conjugate.} \tag{60}$$

The final result is

$$H=\hbar\omega_1(N+\tfrac{1}{2})+\hbar\omega_2(M+\tfrac{1}{2})+\tfrac{3}{2}\alpha\hbar(N^2+M^2+N+M+1)$$
$$+\gamma\hbar^2(N+\tfrac{1}{2})(M+\tfrac{1}{2})$$
$$-\frac{\alpha^2\hbar^2}{8}[34(N^3+M^3)+51(N^2+M^2)+59(N+M)+42]$$
$$-\frac{3\alpha\gamma\hbar^2}{8\omega_1}(2M+1)[(2N+1)^2+1]-\frac{3\alpha\gamma\hbar^2}{8\omega_2}(2N+1)[(2M+1)^2+1]$$
$$-\frac{\gamma^2\hbar^2}{32}\left(\frac{2}{\omega_2}-\frac{\omega_2}{\omega_1^2-\omega_2^2}\right)(2N+1)^2(2M+1)-\frac{3\gamma^2\hbar^2}{32(\omega_1+\omega_2)}(N+M-1)$$
$$-\frac{\gamma^2\hbar^2}{32}\left(\frac{2}{\omega_2}+\frac{\omega_1}{\omega_1^2-\omega_2^2}\right)(2N+1)(2M+1)^2-\frac{3\gamma^2\hbar^2}{32(\omega_1-\omega_2)}(N-M) \tag{61}$$

where $N=a^+a$ and $M=b^+b$.

5. Convergence of the series

We begin with a discussion of the classical transformation. As is well known, convergence of the transformation to the Birkhoff–Gustavson normal form implies integrability. This is a non-generic property of Hamiltonian systems, so we typically expect the series to diverge and to be asymptotic at best. If convergent, it may have a natural radius of convergence because of changes in the topology of motion (bounded–unbounded), as seen in the first example in § 3. Similar things might happen for the two-dimensional oscillator, which is integrable for specific values of the parameters (Bountis *et al* 1982). However, even if the system is not integrable the series can be sensible to this change in topology. For sufficiently small λ and energy, it can show the usual asymptotic behaviour, yielding reasonable approximations with a few terms before diverging, whereas for larger values it can be strongly divergent. This behaviour has apparently been observed by Williams and Koonin (1982).

In quantum mechanics, the situation is more uniform: independent of the degrees of freedom the series is at best asymptotic. Roughly speaking, the series is dominated by the highest powers of operators and thus a change of sign in λ can bring about qualitative changes in the spectrum from discrete to continuous. The following scaling argument for one-dimensional anharmonic oscillators can perhaps be generalised to cover normal forms as well (see Simon 1970, 1982, Reed and Simon 1978).

Let

$$H=\tfrac{1}{2}(p^2+\alpha x^2)+\lambda x^4 \tag{62}$$

and let $E_n(\alpha, \lambda)$ denote the nth eigenvalue. Following Symanzik we set $p' = \lambda^{1/6} p$ and $x' = \lambda^{-1/6} x$ so that

$$H(\alpha, \lambda) = \lambda^{1/3} H(\alpha\lambda^{-2/3}, 1) \tag{63}$$

and

$$E_n(1, \lambda) = \lambda^{1/3} E_n(\lambda^{-2/3}, 1). \tag{64}$$

Now x^4 is a singular perturbation to $p^2 + x^2$, so regular perturbation theory in the sense of Kato-Rellich does not apply, but x^2 is small compared with $p^2 + x^4$, so the strong coupling expansion

$$E_n(1, \lambda) = \lambda^{1/3}(a_0 + a_1\lambda^{-2/3} + a_2\lambda^{-4/3} + \ldots) \tag{65}$$

exists. From this one can deduce (Simon 1970) that $E_n(1, \lambda)$ is not analytic at $\lambda = 0$.

The normal forms derived in § 4 are formal power series around this singularity, so they cannot be convergent. For this particular case and a few others, one can prove they are asymptotic. Moreover, Borel summation and Padé approximants allow to recover the exact eigenvalues (for a recent review see Simon 1982).

To see the relationship to Rayleigh-Schrödinger perturbation theory, recall that asymptotic expansions are unique. So, if the unitary perturbation theory developed in § 4 and standard perturbation theory are both asymptotic, they have to be identical since they approximate the same object.

6. Summary and concluding remarks

In this paper, we have established a very close connection between the Lie algebra version of the transformation to the Birkhoff-Gustavson normal form and unitary perturbation theory in quantum mechanics. Compared to the standard procedure used to achieve normal form in classical mechanics, this one is simpler and should be easier to program on computers capable of symbolic algebraic manipulations: the generator for the canonical transformation can be read off directly (no matrix inversion required) and the new Hamiltonian is obtained from a sequence of Poisson brackets. Moreover, a 'superconvergent', i.e. fast, procedure can be given.

A quantum analogue is obtained if Poisson brackets are replaced by commutators. The resulting quantum normal form is equivalent to Rayleigh-Schrödinger perturbation theory. In each order of the parameter λ, the coefficient of the leading power of the number operator is given correctly by the classical normal form if quantised using, for example, equation (53).

Robnik (1984) suggested using

$$Q((z^*z)^l) = (a^+a + \tfrac{1}{2})^l \tag{66}$$

for normal monomials. For the examples given above, this rule produces the quantum perturbative result up to $O(\lambda^2)$ and up to a constant.

Note that because of equation (59) problems with small denominators also appear in quantum mechanics. Results for anharmonic oscillators (Bender and Wu 1969) show, however, that they might be absorbed in the divergence of lower powers of number operators (for instance, the constants at each order of λ diverge like $CA^n n!$ with constants C and A).

We finally note an open problem. The quantum perturbation series can be resummed to yield the exact eigenvalue. It is not known whether the classical Birkhoff–Gustavson normal form can be resummed and to what it should converge. Preliminary data of Shirts and Reinhardt (1982) indicate that Padé approximants to the constant of motion derived from the normal form behave nicely. There might exist a connection to cantori as discussed by MacKay *et al* (1984), but it still needs to be worked out.

Acknowledgment

This work was supported by Forschungsschwerpunkt 'Dynamische Systeme'.

References

Abraham R and Marsden J E 1981 *Foundations of Mechanics* (Reading, MA: Benjamin/Cummings) 2nd edn
Ali M K 1985 *J. Math. Phys.* **26** 2565
Arnold V I 1978 *Mathematical Methods of Classical Mechanics* (Berlin: Springer)
Bender C M and Wu T T 1969 *Phys. Rev.* **184** 1231
Berry M V 1983 *Chaotic Behaviour in Deterministic Systems* ed G Iooss, R H G Helleman and R Stora (Amsterdam: North-Holland) p 171
Berry M V and Tabor M 1976 *Proc. R. Soc.* A **349** 101
Birkhoff G D 1927 *Dynamical Systems* (Providence, RI: Am. Math. Soc.)
Bjorken S D and Drell S D 1954 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)
Bountis T, Segur H and Vivaldi F 1982 *Phys. Rev.* A **25** 1257
Dragt A J and Finn J M 1976 *J. Math. Phys.* **17** 2215
Duistermaat J J 1984 *Bifurcation Theory and Applications* ed L Salvadori (*Lecture Notes in Math.* **1057**) (Berlin: Springer)
Gustavson F G 1966 *Astron. J.* **71** 670
Jaffé C and Reinhardt W P 1982 *J. Chem. Phys.* **77** 5191
Lichtenberg A J and Liebermann M A 1983 *Regular and Stochastic Motion* (Berlin: Springer)
MacKay R S, Meiss J D and Percival I C 1984 *Physica* **13D** 55
Percival I C 1977 *Adv. Chem. Phys.* **36** 1
Reed M and Simon B 1978 *Methods of Modern Mathematical Physics* vol IV *Analysis of Operators* (New York: Academic)
Rice S A 1981 *Adv. Chem. Phys.* **47** 117
Robnik M 1984 *J. Phys. A: Math. Gen.* **17** 109
Shirts R B and Reinhardt W P 1982 *J. Chem. Phys.* **77** 5204 (erratum **79** 3173)
Siegel C L and Moser J K 1971 *Lectures on Celestial Mechanics* (Berlin: Springer)
Simon B 1970 *Ann. Phys., NY* **58** 76
—— 1982 *Int. J. Quantum Chem.* **21** 3
Swimm R T and Delos J B 1979 *J. Chem. Phys.* **71** 1706 (erratum **78** 4795)
Thirring W 1978 *A Course in Mathematical Physics* vol 1 *Classical Dynamical Systems* (Berlin: Springer)
Williams R D and Koonin S E 1982 *Nucl. Phys.* A **391** 72