THE WEYL REPRESENTATION IN CLASSICAL AND QUANTUM MECHANICS

Alfredo M. OZORIO DE ALMEIDA[a,b]

[a] *Centro Brasileiro de Pesquisas Físicas – CBPF Rua Dr. Xavier Sigaud, 150 22290-180, Rio de Janeiro, RJ, Brazil*

[b] *Institut Henri Poincaré, 11 Rue P. et M. Curie, 75005 Paris, France*

ELSEVIER

AMSTERDAM – LAUSANNE – NEW YORK – OXFORD – SHANNON – TOKYO

ELSEVIER

Physics Reports 295 (1998) 265–342

PHYSICS REPORTS

The Weyl representation in classical and quantum mechanics

Alfredo M. Ozorio de Almeida[a,b]

[a] *Centro Brasileiro de Pesquisas Físicas – CBPF Rua Dr. Xavier Sigaud, 150 22290-180, Rio de Janeiro, RJ, Brazil*

[b] *Institut Henri Poincaré, 11 rue P. et M. Curie, 75005 Paris, France*

Received May 1997; editor: E. Brezin

Contents

Abstract

The position representation of the evolution operator in quantum mechanics is analogous to the generating function formalism of classical mechanics. Similarly, the Weyl representation is connected to new generating functions described by chords and centres in phase space. Both classical and quantal theories relie on the group of translations and reflections through a point in phase space. The composition of small time evolutions leads to new versions of the classical variational principle and to path integrals in quantum mechanics. The strong resemblance between the two theories allows a clear derivation of the semiclassical limit in which observables evolve classically in the Weyl representation. The restriction of the motion to the energy shell in classical mechanics is the basis for a full review of the semiclassical Wigner function and the theory of scars of periodic orbits. By embedding the theory of scars in a fully uniform approximation, it is shown that the region in which the scar contribution is oscillatory is separated from a decaying region by a caustic that touches the shell along the periodic orbit and widens quadratically within the energy shell. © 1998 Elsevier Science B.V.

PACS: 03.; 03.65.-w; 03.65.Sq

Keywords: Variational principle; Path integral; Wigner function; Scars

PII S0370-1573(97)00070-7

0. Introduction

The Wigner function was originally devised with the purpose of reconciling quantum statistical mechanics with classical phase space [1–3]. An apparent contradiction with the uncertainty principle tends to surround this Weyl representation of the density operator with a certain mystery. As a result, the Wigner function can be negative-valued, disallowing its direct correspondence with the classical Liouville probability density. However, its projections correspond to probabilities in positions or momenta, and its smoothing, known as the Husimi function [4, 5], is also positive definite.

The study of the semiclassical limit of pure quantum states has also directed attention to the Wigner function [6,7]. The important point is that the features that distinguish integrable from chaotic motion in classical mechanics manifest themselves most clearly in the full phase-space picture. Thus, it is expected that their emergence in the semiclassical limit should be optimally illuminated by the Weyl representation. Indeed, Berry [7] showed that the amplitude peak of the Wigner function for the pure state of an autonomous system with a single freedom lies close to the energy shell and Ozorio de Almeida and Hannay [38] generalized this picture for the invariant tori of classical integrable systems.

For a point x within the torus, Berry obtained the phase of the oscillating semiclassical Wigner function by an appealing "chord construction": the phase is proportional to the symplectic area (or action) bounded by the torus and the torus chord that is centred on x. Later, Marinov [8] found that an analogous chord construction determined the semiclassical limit of the Weyl propagator, i.e. the Weyl transform of the evolution operator. The difference is that in this case the tips of the chord centred on x must lie on the same classical orbit. Marinov [9] showed that the phase of the Weyl propagator satisfies a new version of the classical Hamilton–Jacobi equation, thus bringing the chord construction into classical mechanics itself.

The derivation of a path integral for the Weyl propagator introduces the chord construction into the core of quantum mechanics. Even though the original presentation by Berezin and Marinov [8, 10] obscured this point by doubling the number of variables, the present author [11] obtained a formula that is indeed analogous to the classical variational principle for the centre generating function [12].

These new presentations of path integrals and classical variational principles are connected with the traditional versions through the notion of double phase space. Since the dynamical state of a system is completely determined by a point in phase space, its evolution after a time t is described by an ordered pair of such points, or, alternatively, by a point in double phase space. For a Hamiltonian system with L freedoms, phase space has $(2L)$ dimensions and double phase space has $(4L)$ dimensions, within which the classical transformation is represented by a $(2L)$-dimensional surface. The uncertainty principle only permits complete knowledge of half the coordinates of both the initial and the final phase space, usually chosen to be positions, though the momenta are equally legitimate. The Weyl representation relies on the alternative choice of the centre of the vector joining each pair of initial and final phase space points.

For this reason I shall often refer to the Weyl representation as the centre representation, specially when considering the complementary chord representation. The latter arises from the perception that the family of affine vectors that we could choose for a given centre are complementary variables to the centres themselves, in analogy to the familiar relation between positions and momenta: In both cases we obtain a classical action by integrating one variable with respect to the other. The centre representation of the unitary evolution operator relies on half the coordinates of double phase

space, just as the usual position representation. In the latter case we forego the specification of the momenta, whereas in the former it is chords that remain undetermined.

Though classical mechanics allows us to specify complementary variables simultaneously, we can mimic the quantum view through the formalism of generating functions. Determining a canonical transformation via a generating function that depends only on positions, the complementary initial and final momenta result from appropriate derivatives. We can also specify a canonical transformation by its centre generating function, from which we obtain the corresponding chords by differentiation. It is in this sense that we may speak of a Weyl representation in classical mechanics. The Legendre transform of the centre generating function coincides with the chord generating function and vice versa, whereas symmetrized Legendre transforms connect these with the familiar position or momentum generating functions. All these transformations of representation are a consequence of the Lagrangian property of the $(2L)$-dimensional surface in double phase space that defines a canonical transformation, that is, it is a null surface for the appropriate action (or two-form).

The relationship between each of the centre, chord, momentum or position representations in quantum mechanics bears exact analogy with that connecting the corresponding classical generating functions if we merely substitute Fourier for Legendre transforms. I shall not attempt to provide a fully comprehensive survey, but will reveal here instead the novel aspects of this conceptual web.

The unifying thread is provided by the fundamental operations of translation and reflection in phase space. By reflection, I shall always mean reflection through a point, synonymous to inversion or half-turn [13]. It is well known that the translations form a group, but the fact that this is contained in a wider group which includes the reflections about each point in phase space [13] has not been widely appreciated in dynamics. In quantum mechanics, the unitary reflection operators are defined as Fourier transforms of the familiar translation operators [57, 58].

Consider the uniform translation, of the points $x = (p_1, \ldots, p_L, q_1, \ldots, q_L)$ in $(2L)$-dimensional phase space, by a vector α,

$$x_+ = T_\alpha(x_-) = x_- + \alpha\,,$$

and the corresponding quantum operator $\hat{T}_\alpha$. The reflection through the point a is defined as

$$x_+ = R_a(x_-) = -x_- + 2a\,,$$

corresponding to the quantum operator $\hat{R}_a$. The Weyl (centre) representation of an arbitrary linear operator $\hat{B}$ is the function $B(x)$ defined by the trace of $\hat{R}_x\hat{B}$. Similarly, the trace of $\hat{T}_{-\xi}\hat{B}$ defines its chord symbol $B(\xi)$.

In classical mechanics, the full canonical transformation $C: x_+ = C(x_-)$ is either obtained by expressing $\xi = x_+ - x_-$ in terms of $x = (x_+ + x_-)/2$, or vice versa. In the former case we start with the centre generating function $S(x)$ and ξ is determined by its derivatives, whereas, in the latter, we differentiate the chord generating functions $S(\xi)$. We can also determine ξ directly, given x, by noting that $x_- = x - \xi/2$ is the fixed point of the combined transformation $R_x \circ C$. Likewise, if we start with a given displacement ξ, we obtain x_- as the fixed point of $T_{-\xi} \circ C$.

Thus, we again find a close analogy between the chord or the centre representations in quantum and in classical mechanics. In both cases we combine the operation to be described with either a translation or a reflection. Then we take the trace in quantum mechanics or find the classical fixed point.

There are definite advantages of the centre representation over the one in terms of chords. The former is ideally suited for the description of transformations close to the identity and, therefore, continuous time evolution, for which the chord representation becomes singular. Another advantage is that the Weyl representation of a Hamiltonian operator is necessarily real and, of course, there are the many desirable properties of the Wigner function. However, it turns out that there is a great increase of analytic flexibility, simplifying the development of the theory if we develop the chord along with the centre representation.

The first three sections of this review are dedicated to classical mechanics. Linearizing the flow in the neighbourhood of a classical orbit leads to the Cayley parametrization of symplectic matrices and hence to the quadratic centre and chord generating functions for linear canonical transformations. Section 1 also develops the relationship between these classical representations and the fundamental group of translations and reflections in phase space, a normal subgroup of the full affine symplectic group.

Section 2 is concerned with the composition of canonical transformations. These are described by polygons in phase space, specified by the centres of their sides in the case of the Weyl representation. Dividing a finite time flow into an arbitrarily large number of steps, we obtain a new derivation of the variational principle: Among all the infinite-sided polygons where the centre of one of its sides is kept fixed, the symplectic area is only stationary when all the other sides together coincide with a classical trajectory. There follows a derivation of the more usual formulations of the variational principle and of Marinov's version of the Hamiltonian–Jacobi equation. Further mathematical results concerning the geometry of polygons in phase space are provided in Appendix A.

Section 3 deals with the variational principle for fixed energy. The construction of the corresponding orbits that lie on the energy shell, such that their tips are centred on a given point in phase space, leads us to the definition of the centre map. The desired orbits determine fixed points of this variation of more familiar Poincaré maps. For the latter, the fixed points determine periodic orbits, so we obtain a family of solutions of the centre variational principle for each periodic orbit. The centre representation breaks down along centre caustics, corresponding to the caustics of the semiclassical Wigner function [7]. It is shown that in the case of more than one freedom there is one fold caustic for each period orbit, along which the caustic touches the energy shell. Inside the shell the caustic widens smoothy. Traversal of such caustics involves a passage through the chord representation.

From then on the presentation is concerned with quantum mechanics. The definition of the operator for translations in phase space in Section 4 is chosen so that compositions of translations generate polygons analogous to those in classical mechanics. Together with the reflections obtained as their Fourier transform, they reproduce the classical group. To end this section we review the representations of quantum states, introducing the Husimi representation in terms of the diagonal matrix elements in the basis of coherent states.

Section 5 presents the centre and chord representations for quantum operators, connecting them with the usual representations. We study a few examples of the Wigner function, from which we obtain the Husimi representation as a Gaussian smoothing of the Weyl representation.

The chord representation of the products of operators leads to a generalized formula for the Weyl representation of an arbitrary number of operators, which is a semiclassical expansion in powers of Planck's constant in the case of observables, generalizing the Groenewold rule [32]. The rest of Section 6 is dedicated to the path integral for the Weyl propagator, obtained from the integral

formula for the centre representation of the product of operators. The evolution is expressed as a superposition of phases defined as the symplectic area of all the infinite-sided polygonal lines with ends centred at a given phase-space point. The trace of this operator is defined by the set of closed polygonal lines. The deduction is quite analogous to that of the classical variational principle, so that the semiclassical limit becomes specially transparent. Relying on the group of metaplectic operators, corresponding to classical linear transformations, we find that observables evolve classically in the semiclassical limit of the Weyl representation.

In the final section, we take the time Fourier transform of the Weyl propagator to obtain the spectral Wigner function and hence discuss the nature of stationary states and their semiclassical limit. Building on the classical theory presented in Section 3, we obtain a uniform theory for the semiclassical limit as the evolution point passes into the energy shell. This includes a curvature correction to the Berry theory for the scars of periodic orbits [14] which damps the contribution of far away orbits. There results a new obstacle to be surmounted by the resummation of the periodic orbits for the spectral Wigner functions.

The geometrical constructions underlying the results in Sections 3 and 7 are more complicated than those necessary for the time-dependent theory. Thus, the reader that is not primarily interested in the energy shell structure may profitably skip directly from Section 2 to 4.

The basic purpose of this report is to present a self-consistent formalism for nonrelativistic classical and quantum mechanics based entirely on centres and chords. We can certainly translate all the quantal results into the familiar position representation of wave mechanics, just as we can also derive all of classical mechanics from a Lagrangian formulation, but the Weyl representation stands on its own as an alternative. In spite of the length of this text, it has been pruned so as to present strictly the main line of reasoning. Thus, there are many important subjects, such as Weyl ordering, for which the reader will depend on the basic references provided.

Is there any merrit in presenting yet another reformulation of mechanics? Though it is early to predict the possible uses of the present theory, I believe that the fundamental point concerns the invariance of the theory with respect to different groups of transformations: classical mechanics is invariant with respect to the full group of canonical transformations, whereas quantum mechanics is not. The generating functions of classical mechanics also lack this flexibility, though each type of generating function is invariant with respect to a particular subgroup of canonical transformations. In particular, the position generating functions are invariant with respect to point transformation, i.e. those where the positions transform among themselves without mixing with the momenta. Obtained from coordinate transformations of the Lagrangian, we can also bring this invariance into wave mechanics. The prevalence of simple Hamiltonians where the momentum dependence is separate and quadratic accounts for the enormous success of wave mechanics.

Quantum mechanics is also invariant with respect to metaplectic transformations, corresponding to linear canonical transformations. This fact becomes self-evident in the Weyl representation and is reflected in the invariance of the chord and centre generating functions with respect to symplectic transformations. The basis of the semiclassical approximation is the linearization of the motion along a classical orbit, so that the semiclassical approximation becomes exact for linear motion. It is natural that the deepest insight into the semiclassical limit should be obtained in the representation that is symplectically invariant.

Work is in progress along some of the many directions that are left open by the following presentation. The most obvious need is for a complete theory of Morse and Maslov indices for the

semiclassical approximation, since these are only determined explicitly in the text for small times. Another challenge that is being tackled is to derive a resummed formula for individual Wigner functions. All the same, I felt that the remarkable evolution of both the extent and the global coherence of the present theory since my last papers on this subject [11, 12] warranted the publication of this report resulting from a series of lectures presented in CBPF (Rio), IF-USP (São Paulo) and finally the Institut Henri Poincaré (Paris), whose hospitality I thank. I am grateful to all the discussions, with numerous colleagues and students, in particular, S. Fishman, J. Keating, R. Prange, A. Rivas, P. Rios, M. Saraceno, A. Voros and W. Wrezinsky throughout the past year, that have continuously clarified my views on the subject and greatly refined the presentation of this review.

1. Linear transformations in classical mechanics

Dynamical systems obtained as the classical limit of quantum mechanical systems are characterized by a *Hamiltonian function* $H(x,t)$, where x is a point in an even-dimensional *phase space*. Usually, the coordinates of this $(2L)$-dimensional space are separated into L momenta and L positions, so that $x=(p_1,\ldots,p_L,q_1,\ldots,q_L)$. In any case, Darboux's theorem [15] guarantees that this coordinatization is always possible. The dynamical system is defined by *Hamilton's equations*,

$$\dot{p}=-\partial H/\partial q\,, \qquad \dot{q}=\partial H/\partial p\,, \tag{1.1}$$

which may be compactified into the form

$$\dot{x}=\mathfrak{J}\,\partial H/\partial x\,, \tag{1.2}$$

with the definition of the $(2L\times 2L)$-dimensional matrix

$$\mathfrak{J}=\left[\begin{array}{c|c} 0 & -1 \\ \hline 1 & 0 \end{array}\right]. \tag{1.3}$$

It is important to note that the transpose $\mathfrak{J}'=\mathfrak{J}^{-1}=-\mathfrak{J}$.

Since the equations of motion are of first-order, there exists exactly one orbit passing through each point in phase space. Let us choose the origin to be on the orbit that we wish to study; then we obtain the neighbouring orbits for a short time by expanding the Hamiltonian in a Taylor series:

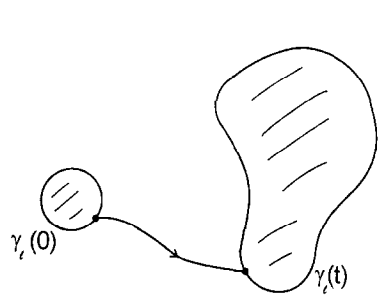
$$H(x)=H(0)+\partial H/\partial x|_0\cdot x+\tfrac{1}{2}x\mathscr{H}_0x+\cdots\,, \tag{1.4}$$

where $\mathscr{H}_0$ is the *Hessian matrix* $\partial^2H/\partial x^2$ evaluated at the origin.
The velocities near the origin become, for $t=0$,

$$\dot{x}(x)=\mathfrak{J}\,\partial H/\partial x|_0+\mathfrak{J}\mathscr{H}_0x+\cdots\,, \tag{1.5}$$

so that

$$\delta\dot{x}=\dot{x}(x)-\dot{x}(0)=\mathfrak{J}\mathscr{H}_0x=\mathfrak{J}\mathscr{H}_0\delta x \tag{1.6}$$



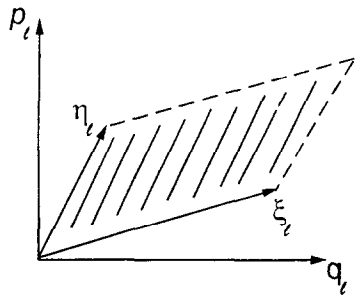


Fig. 1.1. A closed curve γ in phase-space projects onto L closed curves γ_ℓ in the L conjugate planes. In general, the area within each of these circuits evolves, leaving invariant the full symplectic area (1.9).

Fig. 1.2. The parallelogram formed by the vectors ξ and η in phase-space projects into parallelograms in each of the conjugate planes.

to first order in δx. Thus, the short time motion surrounding a chosen orbit is determined by a linear dynamical system. I have not included the time dependence explicitly because it will only affect the matrix $\mathscr{H}_0$ and not the form of the equations.

After an infinitesimal time δt, we may shift the origin to $x = \delta t \dot{x}(0)$ and there obtain a new expansion. Iterating this procedure, we find that the flow $x(0) \to x(t)$ in the neighbourhood of a given orbit $x(0,0) \to x(t,0)$ may be approximated by a linear flow $\delta x(0) \to \delta x(t)$ resulting from the time-dependent linear Hamiltonian system

$$\delta \dot{x} = \mathfrak{J}\mathscr{H}(x(t),t)\delta x\,. \tag{1.7}$$

Since the product of (infinitesimal) linear transformations is necessarily linear, we can define the matrix $\mathscr{M}_t$ such that

$$\delta x_t = \mathscr{M}_t \delta x_0\,. \tag{1.8}$$

It should be noted that for the critical points of the Hamiltonian, where $\partial H/\partial x = 0$, the orbit of x_0 reduces to an equilibrium point, simplifying the foregoing theory.

The study of the possible matrices $\mathscr{M}_t$, or linear maps, that can arise in Hamiltonian systems is of fundamental importance. The essential property is that they preserve the *symplectic area* (or *action*) of any closed circuit γ in phase space:

$$S = \sum_{\ell=1}^{L} S_\ell = \sum_{\ell=1}^{L} \oint_{\gamma_\ell} p_\ell \cdot \mathrm{d}q_\ell = \oint_\gamma p \cdot \mathrm{d}q\,, \tag{1.9}$$

where the γ_ℓ are projections of γ onto the L conjugate planes (p_ℓ, q_ℓ) as displayed in Fig. 1.1.

This property results from the conservation of the symplectic area of the parallelogram formed by any pair of vectors ξ and η:

$$\sum_{\ell=1}^{L} (\xi_{p_\ell}\eta_{q_\ell} - \xi_{q_\ell}\eta_{p_\ell}) = (\mathfrak{J}\xi)\cdot\eta \equiv \xi \wedge \eta\,, \tag{1.10}$$

where the last identity defines the *skew product* of ξ and η. Note that the projections onto the L conjugate planes are also parallelograms, as shown in Fig. 1.2. Since

$$\frac{\mathrm{d}}{\mathrm{d}t}(\xi \wedge \eta) = (\mathfrak{J}\dot{\xi}) \cdot \eta + (\mathfrak{J}\xi) \cdot \dot{\eta}$$

$$= (\mathfrak{J}\mathfrak{J}\mathscr{H}\xi) \cdot \eta + (\mathfrak{J}\xi) \cdot \mathfrak{J}\mathscr{H}\eta = -(\mathscr{H}\xi) \cdot \eta + \xi \cdot (\mathscr{H}\eta) = 0\,, \tag{1.11}$$

the total change in this symplectic area is zero. Using $\xi_t = \mathscr{M}_t\xi$ and $\eta_t = \mathscr{M}_t\eta$, we must then impose that

$$(\mathfrak{J}\mathscr{M}_t\xi) \cdot (\mathscr{M}_t\eta) = \mathfrak{J}\xi \cdot \eta\,, \tag{1.12}$$

which reduces to

$$\mathscr{M}_t'\mathfrak{J}\mathscr{M}_t = \mathfrak{J}\,, \tag{1.13}$$

the definition of a *symplectic matrix*. Thus, the linearized flow in the neighbourhood of any orbit in a Hamiltonian system is determined by a symplectic matrix.

An immediate consequence of the symplectic property is that

$$\det \mathscr{M}_t = 1\,. \tag{1.14}$$

Taking determinants of the product in (1.13), we obtain the unit modulus. Then continuity, with the fact that $\mathscr{M}_0 = 1$, the unit matrix, determines the sign. Thus linear flows preserve phase space volumes.

The properties that we have found for linear systems are extendable to nonlinear flows. Any circuit may be divided into an arbitrary number of parallelograms for which symplectic area is preserved in the limit of smallness, so the full symplectic area is invariant. Likewise, any phase space volume may be indefinitely subdivided into hypercubes for which volume conservation holds.

It is easy to see that the product of symplectic matrices is also symplectic, that is, symplectic matrices form a group. It follows that similarity transformations between symplectic matrices are symplectic. We will now show that symplectic matrices that are diagonalized or taken to Jordan normal form are also symplectic and hence this is also a property of the matrices that reduce them, though they may be complex.

To see this we show that if γ is an eigenvalue of $\mathscr{M}$, so is γ^{-1}. Consider the following manipulations on the characteristic polynomial

$$P(\gamma) = \det[\mathscr{M} - \gamma 1] = \det[\mathfrak{J}^{-1}\mathscr{M}'^{-1}\mathfrak{J} - \gamma 1] = \det[\mathscr{M}'^{-1} - \gamma 1]$$

$$= \frac{(-\gamma)^{2L}}{\det(\mathscr{M}')} \det[\mathscr{M}' - \gamma^{-1}1] = \frac{\gamma^{2L}}{\det(\mathscr{M})} \det[\mathscr{M} - \gamma^{-1}1]\,. \tag{1.15}$$

Thus,

$$P(\gamma) = \pm\gamma^{2L}P(\gamma^{-1}) \tag{1.16}$$

and if γ_0 is a root of $P(\gamma)$, so is γ_0^{-1}.

It is immediately verified that the matrix

$$\mathscr{M} = \left[\begin{array}{cc|cc} \gamma_1 & & & \\ & \gamma_2 & 0 & \\ & & \ddots & \\ \hline & & \gamma_1^{-1} & \\ 0 & & & \gamma_2^{-1} \\ & & & & \ddots \end{array}\right] \tag{1.17}$$

is symplectic, so it was diagonalized by a symplectic transformation. However, this transformation will only be real if all the eigenvalues are real.

Exponentiation permits us to generate symplectic matrices from symmetric matrices. Indeed, if $\frac{1}{2}x\mathscr{H}x$ is a time-independent quadratic Hamiltonian, the linear flow after a time t is simply

$$x(t) = \exp[t\mathfrak{J}\mathscr{H}]x(0) = \mathscr{M}_t x(0)\,. \tag{1.18}$$

Though this is an explicit construction of the symplectic matrix, it relies on an infinite series and it does not include all the possible matrices that could result from a quadratic time-independent Hamiltonian, such as

$$\mathscr{M} = \begin{pmatrix} -\gamma & 0 \\ 0 & -\gamma^{-1} \end{pmatrix}. \tag{1.19}$$

In both respects there is a definite advantage in using the *Cayley parametrization* [16]:

$$\mathscr{M} = [1 - \mathfrak{J}\mathscr{B}][1 + \mathfrak{J}\mathscr{B}]^{-1} = [1 + \mathfrak{J}\mathscr{B}]^{-1}[1 - \mathfrak{J}\mathscr{B}]\,. \tag{1.20}$$

Direct insertion of (1.20) into the definition (1.13) shows that this is necessarily a symplectic matrix, unless $\mathfrak{J}\mathscr{B}$ has the eigenvalue -1. The inverse is immediately obtained as

$$\mathfrak{J}\mathscr{B} = [1 + \mathscr{M}]^{-1}[1 - \mathscr{M}] = [1 - \mathscr{M}][1 + \mathscr{M}]^{-1}\,. \tag{1.21}$$

Thus, it is only symplectic matrices with eigenvalue -1 for which there is no Cayley parametrization. In the same way as with the exponential parametrization, changing the sign of $\mathscr{B}$ generates $\mathscr{M}^{-1}$:

$$[1 + \mathfrak{J}\mathscr{B}][1 - \mathfrak{J}\mathscr{B}]^{-1}[1 - \mathfrak{J}\mathscr{B}][1 + \mathfrak{J}\mathscr{B}^{-1}] = 1\,. \tag{1.22}$$

The significance of the Cayley parametrization is emphasized by interpreting

$$S(x) = x\mathscr{B}x \tag{1.23}$$

as a function that implicitly generates the transformation $x_- \to x_+$, where

$$x_\pm = x \pm \tfrac{1}{2}\xi \tag{1.24}$$

and

$$\xi = -\mathfrak{J}(\partial S/\partial x)(= -2\mathfrak{J}\mathscr{B}x)\,. \tag{1.25}$$

We will henceforth refer to x as the *centre* of the transformation and ξ as the *chord*. Evidently, the matrix for the linear transformation (1.24) generated by the quadratic form (1.23) is exactly (1.20), so the transformation is symplectic. Rewriting

$$S(x) = \frac{1}{2}\frac{\partial S}{\partial x}\cdot x = -\frac{1}{2}\mathfrak{J}\frac{\partial S}{\partial x}\wedge x = \frac{1}{2}\xi\wedge x\,, \tag{1.26}$$

we can interpret $S(x)$ as the area of the triangle with corners at $0, x_-$ and x_+, as shown in Fig. 1.3.

If we expand a general nonlinear function $S(x)$ in a Taylor series:

$$S(x) = S(x_0+\delta x) = S(x_0) + \alpha_0\wedge\delta x + \delta x\mathscr{B}_0\delta x + 0(\delta x^3) \tag{1.27}$$

while keeping the implicit definition of the transformation $x_- \to x_+$ given by (1.24) and (1.25), we obtain

$$\xi(x) = \alpha_0 - 2\mathfrak{J}\mathscr{B}_0\delta x + 0(\delta x^2)\,. \tag{1.28}$$

Therefore, the Jacobian of the transformation is the matrix (1.20), i.e. even the nonquadratic $S(x)$ is the *centre generating function* of a canonical transformation. We shall usually refer to it as the *centre function*, for short. It is easily seen that, if $\mathscr{M}_0$ is the symplectic matrix corresponding to $\mathscr{B}_0$, then the transformation generated by (1.27) up to second order is

$$x_+ = \mathscr{M}_0(x_- + \alpha_0/2) + \alpha_0/2\,, \tag{1.29}$$

i.e. we sandwich the linear transformation $\mathscr{M}_0$ between two translations by $\alpha_0/2$.

For small times $t=\varepsilon$, we may identify

$$S_\varepsilon(x) = -\varepsilon H(x) + \mathcal{O}(\varepsilon^3)\,, \tag{1.30}$$

since the flow is simply

$$x_+ = x_- + \xi = x_- + \varepsilon\mathfrak{J}(\partial H/\partial x) + 0(\varepsilon^3)\,, \tag{1.31}$$

i.e. $\mathscr{B}_0$ is just the Hessian $-(\varepsilon/2)\mathscr{H}_0$. In general, exchanging $t\to -t$ implies $\xi\to -\xi$, so $S_t(x)$ must be an odd function of t. The third-order correction to (1.30) is derived in Appendix B.

Even for finite times, the chord ξ will be tangent to the surface $S(x)=$ constant, just as shown in Fig. 1.3, because $\dot{x}$ is tangent to $H(x)=E$. In particular, the critical points of $S(x)$ correspond to fixed points of the canonical transformation. These will coincide with the equilibria of $H(x)$ in the case of a Hamiltonian flow. In a way, $-S(x)$ is a finite time Hamiltonian, for which we obtain a single canonical transformation, specified by $\xi(x)$, instead of a group of transformations by integrating $\dot{x}(x)$. For this reason we cannot simply add the generating functions so as to compose transformations. In the following section we shall study the delicate geometrical patterns involved in the composition of centre generating functions.

The simplest kind of motion that we can consider is that generated by a linear Hamiltonian $H(x) = -a\wedge x$. Hamilton's equations will then just be $\dot{x}=a$, so that immediate integration renders the flow after the time t as $x_+ = ta + x_-$. In short, the canonical transformations representing the flow will be merely the uniform *translations* of phase space $x_+ = T_\alpha(x_-)$, where $\alpha = ta$.

We shall find that translations play a fundamental role in the following theory. Evidently, the composition of translations, $T_{\alpha_2}\circ T_{\alpha_1}(x_-) = T_{\alpha_1+\alpha_2}(x_-)$ forms a continuous group. The centre function

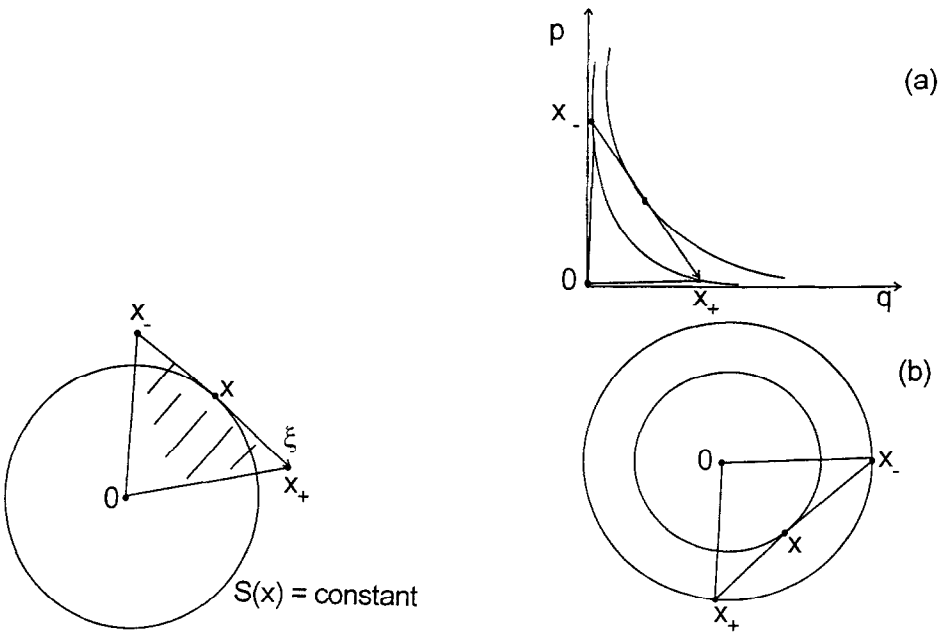


Fig. 1.3. The chord ξ is always tangent to the surface $S(x)=$ constant. If $S(x)$ is quadratic, the centre function equals at each point the symplectic area of the triangle that $\xi(x)$ subtends with the origin.

Fig. 1.4. The level curves of the centre functions for the flow derived from the Hamiltonian (a) $H(x)=\gamma pq$ and (b) $H(x)=\frac{1}{2}w(p^2+q^2)$. They are the same families of hyperbolae and circles as for the Hamiltonian themselves.

for T_α is just the linear function $S(x)=\alpha\wedge x$, such that the chord $\xi=\alpha$, a constant. This is the only case where the short time relation (1.30) between $H(x)$ and $S(x)$ can be extrapolated for all time. It is pleasing that the definition of the centre function makes sense even when its quadratic part (which motivated the definition) cancels.

The next instance where we may hope to obtain the explicit form of the centre function is for the homogeneous linear transformations from which we started. Combining (1.18) with (1.21), we obtain, for $H(x)=\frac{1}{2}x\mathcal{H}x$,

$$\mathcal{J}\mathcal{B}=[1-\exp(t\mathcal{J}\mathcal{H})][1+\exp(t\mathcal{J}\mathcal{H})]^{-1}\,. \tag{1.32}$$

Obviously, the job of inverting and exponentiating matrices will be simplified if they are first diagonalized. The easiest case is when the eigenvalues of $\mathcal{J}\mathcal{H}$ are real. In the case of one freedom, then $H=\gamma pq$, so that

$$\mathcal{J}\mathcal{H}=\begin{bmatrix}-\gamma & 0\\ 0 & \gamma\end{bmatrix}$$

and $\mathcal{J}\mathcal{B}$ will also have the diagonal form

$$\begin{bmatrix}\tanh\frac{t\gamma}{2} & 0\\ 0 & -\tanh\frac{t\gamma}{2}\end{bmatrix}.$$

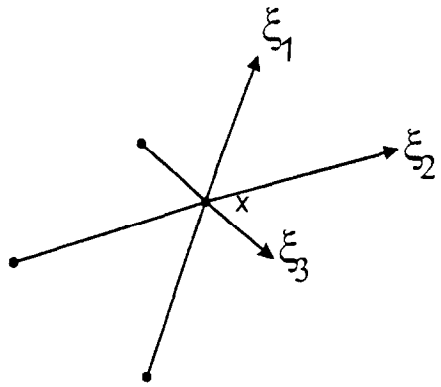


Fig. 1.5. The reflection R_x about a given point x in phase space defines a degenerate chord construction where all chords pass through the same point.

It follows that $S(x) = -2\tanh(t\gamma/2)pq$. So, in this particular case, $\mathscr{H}(x)$ and $S(x)$ have the same level curves, as shown in Fig. 1.4(a). We also notice that the Cayley parametrization is never singular.

Let us now examine the more familiar case of the harmonic oscillator $H(x) = \frac{1}{2}w(p^2+q^2)$. Then we cannot diagonalize $\mathfrak{J}\mathscr{H}$ by a real symplectic transformation, but it is easy to verify that the equation of motion has the well known solution

$$\begin{pmatrix} p_+ \\ q_+ \end{pmatrix} = \begin{pmatrix} \cos wt & -\sin wt \\ \sin wt & \cos wt \end{pmatrix} \begin{pmatrix} p_- \\ q_- \end{pmatrix}.$$

Since the movement is merely a rotation of the phase plane, we may rely on elementary geometry to observe that the chord ξ, between x_- and x_+, is tangent at its midpoint x, to a circle of the same concentric family where $H(x)$ is constant. We can easily calculate the area of the triangle whose base is ξ and height $|x| = \sqrt{q^2+p^2}$, so that using (1.26), $S(x) = -\tan\frac{1}{2}wt(p^2+q^2)$.

Again we find a close family resemblance to the Hamiltonian, as shown in Fig. 1.4(b), but now the centre representation breaks down when $wt = \pi$. At this instant the flow of the harmonic oscillator reduces to a reflection through the origin: $x_+ = R_0(x_-) = -x_-$. Indeed, the centre function becomes singular whenever the map reduces locally to a *reflection* R_x about any point in phase space, i.e. the map is such that $(x_+ - x) = -(x_- - x)$, represented in Fig. 1.5. Evidently, we obtain a degenerate chord structure for this point x.

This problem is neatly dealt with by defining the complementary *chord generating* function $S(\xi)$ for the transformation $x_- \to x_+$, or *chord function*, for short. We still use (1.24), but now we obtain the centre from

$$x = \mathfrak{J}\,\partial S/\partial\xi\,. \tag{1.33}$$

Expanding this generating function in a Taylor series

$$S(\xi) = S(\xi_0) + \delta\xi \wedge a_0 + \tfrac{1}{4}\delta\xi\beta\delta\xi + \cdots, \tag{1.34}$$

we obtain the Jacobian of the transformation in the form

$$\mathscr{M} = -(1+\mathfrak{J}\beta)(1-\mathfrak{J}\beta)^{-1}, \tag{1.35}$$

which is a complementary parametrization of symplectic matrices, unless $\mathfrak{J}\beta$ has the eigenvalue 1. Observing its inverse

$$\mathfrak{J}\beta = (\mathscr{M} - 1)^{-1}(\mathscr{M} + 1), \tag{1.36}$$

we see that it is only the symplectic matrices with eigenvalue 1 that cannot be parametrized in this form, including, of course, the identity matrix.

The simplest example of a chord function is just linear: $S(\xi) = \xi \wedge a$. We then obtain that a is the centre for all chords ξ. In short, this is the generating function for the reflection R_a. So, if a representation in terms of centres is complementary to that of chords, in a similar sense, translations T_ξ are complementary to reflections through a point: R_x. This theme will be developed in the following theory.

Let us now pursue our simple quadratic examples. For $H(x) = \gamma pq$, $\mathfrak{J}\beta$ will again be diagonal,

$$\begin{bmatrix} -\coth\dfrac{t\gamma}{2} & 0 \\ 0 & \coth\dfrac{t\gamma}{2} \end{bmatrix},$$

so that $S(\xi) = \frac{1}{2}\coth(t\gamma/2)\xi_p\xi_q$. We thus see that the chord function becomes singular as $t \to 0$. This is a general feature for all Hamiltonian flows.

For the harmonic oscillator, $H(x) = \frac{1}{2}w(p^2 + q^2)$, it is now easy to verify that the chord function for the flow after a time t is just $S(\xi) = \frac{1}{2}\cot(wt/2)(\xi_p^2 + \xi_q^2)$. Again, this is singular as $t \to 0$. However, we now have a generating function that is perfect for dealing with the singular region of the centre function, since at $wt = \pi$ we just have $S(\xi) = 0$, which is the generating function for the reflection at the origin, R_0.

The general role of reflections and translations in the definition of the centre and the chord generating functions is revealed as we enquire into the existence of a chord for a given centre, or vice versa. When a canonical transformation $C : x_- \to x_+$ is described by a centre function $S(x)$, the existence of a chord $\xi(x)$ is tantamount to the existence of a point $x_-(x)$, such that $C(x_-) = R_x(x_-)$. Since $R_x \circ R_x = 1$, the identity transformation, it follows that the point x_- is determined as the fixed point of C composed with R_x:

$$x_- = R_x \circ C(x_-), \tag{1.37}$$

as shown in Fig. 1.6.

Conversely, if we describe the transformation C by a chord function $S(\xi)$, the centre of the given chord $x(\xi)$ is defined when we can find a point $x_-(\xi)$, for which $C(x_-) = T_\xi(x_-)$, as shown in Fig. 1.6. In this case the point x_- is defined as the fixed point of C composed with $T_{-\xi}$:

$$x_- = T_{-\xi} \circ C(x_-). \tag{1.38}$$

Thus, it is fair to describe the centre function as "viewing the transformation" as a reflection, whereas the chord function "views it" is a translation. The feasibility of using either description will depend on the existence of the fixed point for the appropriate compound transformation. To this end we return to the first-order expansions of the generating functions (1.27) and (1.34). We have seen that $S(x) = \alpha \wedge x$ generates a translation, T_α, while $S(\xi) = \xi \wedge a$ generates the reflection, R_a. Therefore the existence of the required fixed point depends on the products of translations and reflections.

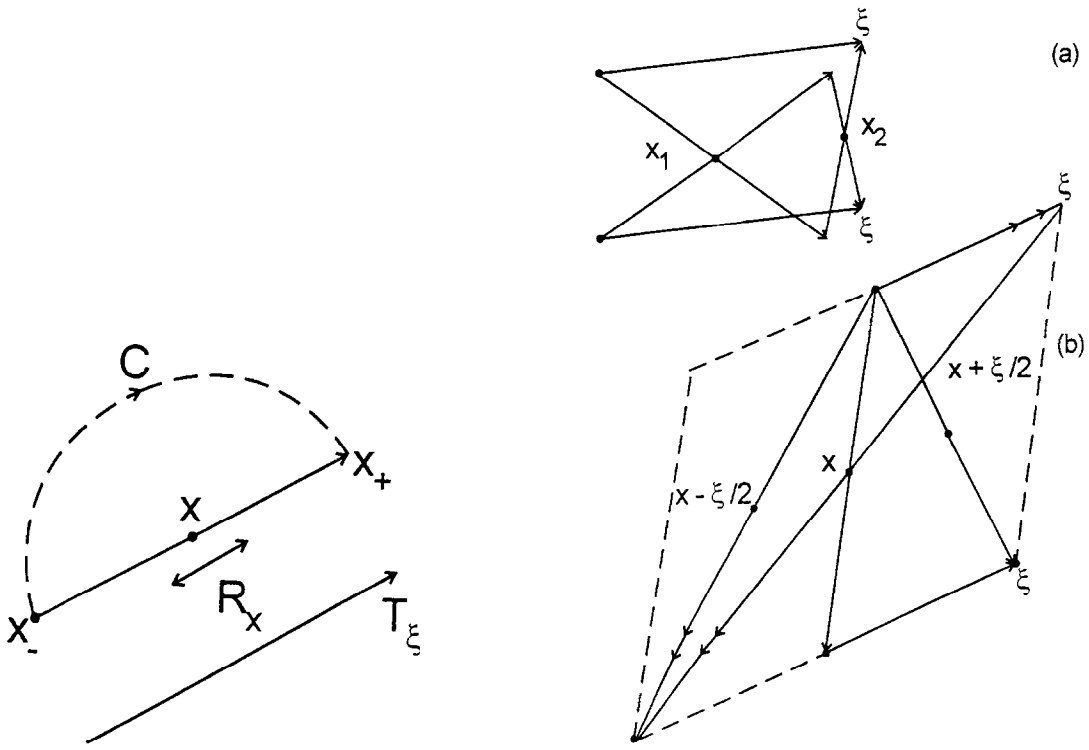


Fig. 1.6. The centre representation describes the canonical transformation $C: x_- \to x_+$ as a reflection R_x about the point x, so that x_- is the fixed point of the composed map $R_x \circ C$. The chord function represents the same canonical transformation as the translation T_ξ by a given vector ξ, so that x_- is the fixed point of the map $T_{-\xi} \circ C$.

Fig. 1.7. (a) The composition of the two reflections $R_{x_2} \circ R_{x_1}$ is an uniform translation T_ξ, where $\xi = 2(x_2 - x_1)$. (b) The composition of a reflection R_x with a translation T_ξ is a reflection about the point $x \pm \xi/2$ (depending on the order of the transformations).

Though we have seen that the set of all translations form a group,

$$T_{\xi_2} \circ T_{\xi_1} = T_{\xi_1+\xi_2}\,, \tag{1.39}$$

this is not true of the reflections. Indeed, we easily verify with reference to Fig. 1.7(a) that the product of two reflections is always an uniform translation:

$$R_{x_2} \circ R_{x_1} = T_{2(x_2-x_1)}\,. \tag{1.40}$$

However, since the product of a reflection and a translation is itself a reflection, such that according to Fig. 1.7(b),

$$T_\xi \circ R_x = R_{x+\xi/2}\,, \tag{1.41}$$

$$R_x \circ T_\xi = R_{x-\xi/2}\,, \tag{1.42}$$

we see that, together, the set of all reflections and translations does form a group. This is the product of the translation group with the group of reflections at the origin R_0, which is identical

to the group $\mathbb{Z}_2$ of the ordinary product of 1 and -1. Indeed, according to (1.41) and (1.42), R_x equals the product of $T_{\pm 2x}$ with R_0, depending on the order, or $R_x = T_x R_0 T_{-x}$.

With the aid of these results, we see immediately that, if C reduces to a reflection, the fixed point of (1.38) exists, but not that of (1.37). Conversely, if C is a uniform translation, it is only $R_x \circ C$ that has an unique fixed point. In other words, the chord function is ideal for considering near reflections viewed as translations, whereas the centre function is constructed to view near translations as reflections. The great advantage of the centre function is that we are mostly concerned with the arbitrarily small translations that are generated by the Hamiltonian flow.

The general problem of determining all the chords through all the centres x in phase space reduces to that of finding the fixed points of the $(2L)$-parameter family of maps $K_x \equiv R_x \circ C$. Likewise, the problem of obtaining the centres for each chord ξ depends on the $(2L)$-parameter family $K_\xi = T_{-\xi} \circ C$.

Let us, therefore, establish the continuity of the fixed points $y(\lambda)$ of a one-parameter family of maps, $K_\lambda : x_- \to x_+ = K_\lambda(x_-)$, with respect to the parameter. Differentiating $y = K_\lambda(y)$ with respect to λ,

$$\frac{\mathrm{d}y}{\mathrm{d}\lambda} = \frac{\partial K_\lambda}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}\lambda} + \frac{\partial K_\lambda}{\partial \lambda}, \tag{1.43}$$

we note that $\partial K_\lambda / \partial y = \mathscr{M}_\lambda$, the symplectic matrix for the linearized transformation near the fixed point. Thus, its rate of change with the parameter is just

$$\mathrm{d}y/\mathrm{d}\lambda = (1 - \mathscr{M}_\lambda)^{-1} \partial K_\lambda / \partial \lambda . \tag{1.44}$$

Therefore, the implicit function theorem allows us to define the family of fixed points $y(\lambda)$, except at the *resonant parameters*, where

$$\det(1 - \mathscr{M}_\lambda) = 0 . \tag{1.45}$$

What goes wrong at these parameter events? We can ascribe them to coalescing fixed points. The deduction of (1.44) presupposed that we follow an isolated fixed point, so this formula breaks down when they come together. Meyer [17] proved that the generic resonance of an area-preserving map is a *bifurcation* where two fixed points merge and disappear. In this case there exists a canonical coordinate system such that the centre generating function for the map takes on the local form [17]

$$S_\lambda(x) = \tfrac{1}{2}q^2 + \lambda p + \tfrac{1}{3}p^3 , \tag{1.46}$$

in which $p(\lambda) = (\pm\sqrt{-\lambda}, 0)$. It is easy to verify that one of these fixed points is unstable (real eigenvalues e^{γ} and $\mathrm{e}^{-\gamma}$) and the other is stable (eigenvalues $\mathrm{e}^{\pm \mathrm{i}w}$). At $\lambda = 0$, the eigenvalue is unity. The bifurcation diagram is shown in Fig. 1.8.

The simplest application of this theory should be to the Hamiltonian flow. The parameter of the family of maps is then just the time. However, all the points of a periodic orbit with a given period will be fixed points of the flow for this time. So, we cannot apply directly the above theory, which presupposes an isolated fixed point. Instead, the theory of the bifurcation of periodic orbits relies on sections through the orbits that we shall study in Section 3.

Returning to the chord problem, we seek the fixed points of the mapping $K_x = R_x \circ C$, so x itself is now the $(2L)$-dimensional parameter of the matrix $\mathscr{M}_x$ in (1.45). This equation defines the surfaces,

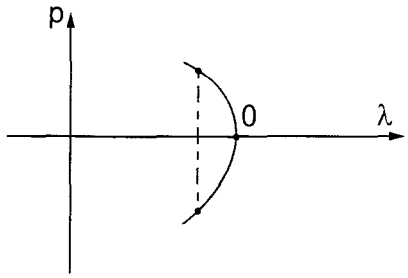


Fig. 1.8. Generic bifurcation diagram for a one-parameter family of symplectic maps.

centre caustics, where pairs of chords coalesce. In other words, they are the loci where the mapping $x \to \xi$ becomes singular. Since ξ is determined by the gradient of $S(x)$, we obtain the alternative equation for the centre caustic as

$$\det(\partial^2 S/\partial x^2)(x) = \infty . \tag{1.47}$$

However, we easily find that the mapping $\xi \to x$ is nonsingular along the centre caustic and that, indeed

$$\det(\partial^2 S/\partial \xi^2)(\xi) = 0 \tag{1.48}$$

along this surface. Of course, there will be *chord caustics* where $\xi \to x$ is singular, but these will be nicely described by the centre function. The great advantage of the latter is that $S(x) \to 0$ for a small time flow, so that (1.47) is entirely avoided. Conversely $S(\xi)$ is globally singular as $t \to 0$.

The chord function and the centre function are reciprocally related as Legendre transforms. Indeed, according to (1.25) we may write the differential of $S(x)$ as

$$\mathrm{d}S = \xi(x) \wedge \mathrm{d}x . \tag{1.49}$$

Thus, if we define

$$F(x, \xi) = \xi \wedge x - S(x) \tag{1.50}$$

and

$$S(\xi) = F(x(\xi), \xi) , \tag{1.51}$$

with $x(\xi)$ prescribed by $\partial F/\partial x = 0$, we obtain precisely Eq. (1.33).

Let us consider x and ξ as alternative coordinates for the *double phase space* ($4L$ dimensions) of initial positions x_- and final positions x_+. Then the fact that (1.48) is an exact differential implies that there exists a ($2L$)-dimensional surface in which

$$\oint \xi \wedge \mathrm{d}x = \oint [\xi_p \cdot \mathrm{d}q - \xi_q \cdot \mathrm{d}p] = 0 \tag{1.52}$$

for any reducible circuit. Therefore, we can furnish double phase space with a symplectic structure of the same form as ordinary phase space by considering that now ξ_{pi} is the conjugate variable to q_i, whereas $-\xi_{qi}$ becomes the conjugate variable p_i. In other words, the canonical coordinates of

double phase space are $(x, \mathfrak{J}\xi)$ or $(-\mathfrak{J}x, \xi)$. It follows that (1.52) defines a *Lagrangian surface* in double phase space, along which Stoke's Theorem [15] leads to

$$d\xi_p \wedge dq - d\xi_q \wedge dp = 0 . \tag{1.53}$$

By making more exchanges of variable we arrive at the traditional generating functions of classical mechanics. Indeed, recalling that $q = (q_- + q_+)/2$ and defining

$$F\left(\frac{q_+ + q_-}{2}, p\right) = S\left(\frac{q_+ + q_-}{2}\right) + p(q_+ - q_-) , \tag{1.54}$$

we find that the function

$$S(q_+, q_-) = F((q_+ + q_-)/2, p(q_+ - q_-)) , \tag{1.55}$$

with $p(q_+ - q_-)$ given by $\partial F/\partial p = 0$, is the generating function of the canonical transformation determined implicitly by

$$p_+ = \partial S/\partial q_+ , \qquad p_- = -\partial S/\partial q_- . \tag{1.56}$$

Therefore, the surface in double phase space that determines the canonical transformation has zero action

$$\oint p_+ . dq_+ - \oint p_- . dq_- = 0 , \tag{1.57}$$

which can be obtained directly from Stoke's theorem and (1.52), since

$$\begin{aligned} &(dp_+ - dp_-) \wedge \left(\frac{dq_+ + dq_-}{2}\right) - (dq_+ - dq_-) \wedge \left(\frac{dq_+ + dp_-}{2}\right) \\ &\quad = dp_+ \wedge dq_+ - dp_- \wedge dq_- = 0 . \end{aligned} \tag{1.58}$$

Evidently, we have returned to the conservation of symplectic area for canonical transformations, with which we started. However, in the context of double phase space, we now find that (1.53) and (1.58) are equivalent properties of the same (Lagrangian) surface that determines the canonical transformation. We always use half the coordinates of double phase space when generating the canonical transformation, in such a way that the number of old coordinates equals that of the new; which we choose, is a matter of convenience. Further discussion of double phase space is found in Refs. [49, 12, 48].

One important criterion selects those generating functions that do not become singular in the neighbourhood of the identity map. These are $S(p_+, q_-), S(p_-, q_+)$ and the centre function $S(x)$; they also share the property that their critical points are fixed points of the transformation. However, another criterion is the ease of composing transformations. Here $S(q_+, q_-)$ has a distinct advantage. Indeed, we compose transformations using $S(p_+, q_-)$ by taking their Legendre transforms, obtaining $S(q_+, q_-)$. We shall study the beautiful patterns resulting from the composition of the centre functions in the next section and hence derive the variational principle.

2. Classical variational principles

How do we obtain the centre function for the composition of two canonical transformations, $S_1(x)$ and $S_2(x)$, respectively. The requirement is that the chord η_1 through x_1 in the first transformation join onto the chord η_2 through x_2, to form the chord through some centre x for the composition. We will refer to the resulting triangle formed by η_1, η_2 and η as the *circumscribed triangle* to the midpoints x_1, x_1, and x, as opposed to the *inscribed triangle* with corner at these points.

This simple geometry is displayed in Fig. 2.1. It is important to note that a circumscribed polygon in the full phase space projects into L polygons in the conjugate planes that are circumscribed around the projections of the midpoints. Hence, the following symplectic geometry is reduced to plane geometry. The symplectic area of the circumscribed triangle is the sum of the areas of each projection:

$$\Delta_3(x,x_1,x_2) = 2(x_1 - x) \wedge (x_2 - x) = 2[x_1 \wedge x_2 + x_2 \wedge x + x \wedge x_1], \tag{2.1}$$

from which we obtain

$$\partial \Delta_3/\partial x_1 = -2\mathfrak{J}(x_2 - x) = -\mathfrak{J}\eta_1, \qquad \partial \Delta_3/\partial x_2 = 2\mathfrak{J}(x_1 - x) = -\mathfrak{J}\eta_2, \tag{2.2}$$

where η_j is the jth side of Δ_3, in the clockwise direction.

For the desired composition of the two canonical transformations, we demand that the chords for each transformation satisfy

$$\xi_1 \equiv -\mathfrak{J}\,\partial S_1/\partial x_1 = \eta_1 \quad \text{and} \quad \xi_2 \equiv -\mathfrak{J}\,\partial S_2/\partial x_2 = \eta_2, \tag{2.3}$$

which is equivalent to imposing zero derivatives for

$$S(x,x_1,x_2) = S_1(x_1) + S_2(x_2) + \Delta_3(x,x_1,x_2). \tag{2.4}$$

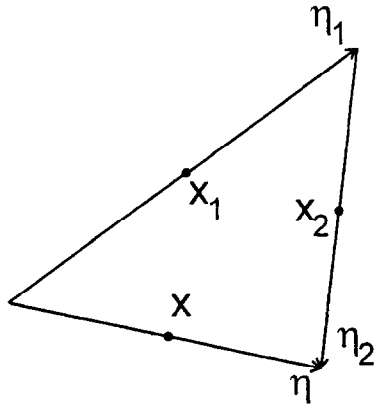


Fig. 2.1. The composition of two canonical transformations requires that the resulting chord η close a triangle whose sides are centred on the three centres x_1, x_2 and x.

Indeed, we verify that

$$\frac{\partial S}{\partial x_1} = \frac{\partial S_1}{\partial x_1} + \frac{\partial \Delta_3}{\partial x_1} = \mathfrak{J}\xi_1 - 2\mathfrak{J}(x_2 - x) = 0\,,$$
$$\frac{\partial S}{\partial x_2} = \frac{\partial S_2}{\partial x_2} + \frac{\partial \Delta_3}{\partial x_2} = \mathfrak{J}\xi_2 + 2\mathfrak{J}(x_1 - x) = 0\,. \tag{2.5}$$

Consider now the 'multiple Legendre transform' $S(x,x_1(x), x_2(x))$, obtained by substituting the solutions of (2.5) into (2.4). Then the chord for this generating function is just given by $-\mathfrak{J}\,\mathrm{d}S/\mathrm{d}x$, where

$$\frac{\mathrm{d}S}{\mathrm{d}x} = \frac{\partial S}{\partial x_1}\,\frac{\partial x_1}{\partial x} + \frac{\partial S}{\partial x_2}\,\frac{\partial x_2}{\partial x} + \frac{\partial S}{\partial x}\,. \tag{2.6}$$

But, because of (2.5),

$$\mathrm{d}S/\mathrm{d}x = \partial S/\partial x = \partial \Delta_3/\partial x = 2\mathfrak{J}(x_2 - x_1) = \mathfrak{J}(\xi_1 + \xi_2)\,. \tag{2.7}$$

Therefore, $S(x,x_1(x),\ x_2(x))$ is the desired generating function.

As a simple example, we consider the composition of homogeneous translations by α_1, and α_2 in phase space. The centre functions are simply $S_j(x) = \alpha_j \wedge x$ and conditions (2.5) then become

$$x_1 = -\tfrac{1}{2}\alpha_2 + x\,, \qquad x_2 = \tfrac{1}{2}\alpha_1 + x\,, \tag{2.8}$$

so that

$$S(x) = \alpha_1 \wedge (x - \tfrac{1}{2}\alpha_2) + \alpha_2 \wedge (x + \tfrac{1}{2}\alpha_1) + 2(-\tfrac{1}{2}\alpha_2) \wedge (\tfrac{1}{2}\alpha_1)$$
$$= (\alpha_2 + \alpha_2) \wedge x + \text{constant}\,, \tag{2.9}$$

as expected. For linear transformations $S_j(x) = x\mathfrak{B}_j x$, conditions (2.5) become a system of $4L$ linear equations. The solutions are linear in x, so that (2.4) is again a quadratic form in x. The geometrical interpretation of $S(x)$ as the symplectic area of a triangle fits nicely with the composition rule. Consider Fig. 2.2; as drawn, $S_j = \frac{1}{2}\xi_j \wedge x_j$ have negative areas (the skew product obeys the left-hand rule), whereas Δ_3 has positive area. It follows that the three terms in (2.4) combine to give the (negative) area of the triangle joining ξ to the origin, as required.

The important point that we must consider in order to generalize the composition rule (2.4) is that, according to (2.2), the derivatives of the symplectic area of the triangle Δ_3 are specified by the corresponding sides. The side η does not depend on the centre x, since it may be considered as the translation resulting from the reflections at x_1 and x_2. All compositions of these two reflections generate the same translation by $\eta = 2(x_2 - x_1)$ according to (1.40), so we may place the centre x anywhere without changing η.

If we now add two more reflections at arbitrary points x_3 and x_4, we again obtain a specific uniform translation by $2(x_3 - x_4)$. Since it also has a free centre, we can always fit these together to obtain $\eta = 2(x_4 - x_3) + 2(x_2 - x_1)$ as shown in Fig. 2.3. Evidently, we may repeat this procedure

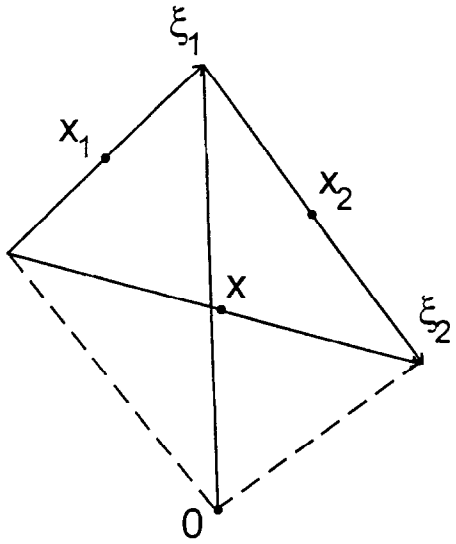


Fig. 2.2. The interpretation of the quadratic centre function as the area of a triangle is fully consistent with the composition of such canonical transformations.

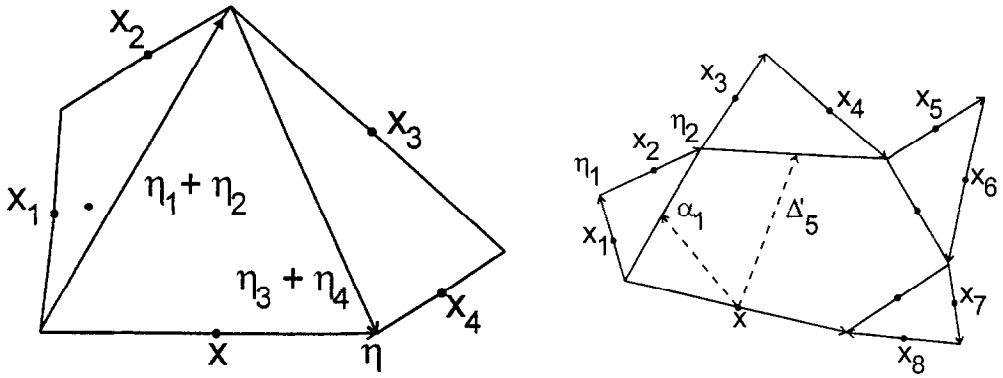


Fig. 2.3. The composition of four canonical transformations defines a pentagon with midpoints at x_1, x_2, x_3 and x. The triangle Δ_3' formed by $\eta_1+\eta_2, \eta_3+\eta_4$ and η depends on the positions of $x_1 \ldots x_4$ but not on x.

Fig. 2.4. The vectors α_j joining x to the centre of the vectors $\eta_{2j}+\eta_{2j-1}$ are invariant with respect to x. The area of the enneagon is decomposed into the pentagon Δ_5' and four triangles.

for any even number of reflections to obtain the remaining side of an odd-sided polygon as

$$\eta = 2\sum_{j=1}^{2n}(-1)^j x_j\,, \tag{2.10}$$

which is independent of x, just as for the triangle.

Now, we may calculate the total symplectic area of the polygon Δ_{2n+1} by subdividing it into an internal polygon with symplectic area Δ_{n+1}' and n triangles. Any change of x translates the internal

polygon rigidly, hence Δ'_{n+1} is independent of x. Therefore, if we lable the fixed vector between x and the centre of the jth side of Δ'_{n+1} as α_j (see Fig. 2.4) we obtain

$$\Delta_{2n+1} = \Delta'_{n+1} + \sum_{j=1}^{n} 2(x_{2j} - x_{2j-1}) \wedge (x + \alpha_j - x_{2j}) = \text{constant} + \eta \wedge x \,. \tag{2.11}$$

We have defined the vector η as the sum of the other sides of the polygon η_j, so it has the opposite orientation to theirs. For each of the other sides we can therefore generalize (2.2) to obtain

$$\partial \Delta_{2n+1} / \partial x_k = -\mathfrak{J}\eta_k \,. \tag{2.12}$$

Here η_k is the kth side of the polygon, so the condition that this coincides with the chord ξ_k of the kth canonical transformation is equivalent to the constraint that the derivative of

$$S(x) = S_1(x_1) + \cdots + S_2(x_{2n}) + \Delta_{2n+1}(x, x_1, \ldots, x_{2n}) \tag{2.13}$$

with respect to x_k be zero. Substitution of all the centres x_k as a function of x, obtained by the zero gradient condition, defines the centre function of the composed transformation.

Consider now the flow generated by the Hamiltonian $H(x)$. It was shown in Section 1 that for a small time $t = \varepsilon$, the corresponding centre function is $S_\varepsilon(x) = -\varepsilon H(x)$, to third order in ε. We can now use the composition rule to estimate the correction, that is, choosing $\varepsilon = t/2$, we obtain the time t flow as the composition of two ε-flows. Thus, distinguishing the phase-space velocity $\dot{x}_1$ for the first interval and $\dot{x}_2$ for the second, determines

$$\Delta_3 = \tfrac{1}{2}(\dot{x}_1 \varepsilon) \wedge (\dot{x}_2 \varepsilon) \,, \tag{2.14}$$

according to Fig. 2.5. Introducing the 'acceleration',

$$\ddot{x} \equiv [\dot{x} \cdot \partial / \partial x]\dot{x} = \mathfrak{J}\mathscr{H}\dot{x} \,, \tag{2.15}$$

so that $\dot{x}_2 \simeq \dot{x}_1 + \ddot{x}_1 \varepsilon$, we obtain

$$\Delta_3 = \tfrac{1}{2}\varepsilon^3 \dot{x} \wedge \ddot{x} = \tfrac{1}{2}\varepsilon^3 \dot{x} \mathscr{H} \dot{x} \,, \tag{2.16}$$

according to Fig. 2.5. Here, the Hessian matrix of the Hamiltonian can be evaluated at the point x, to lowest order in ε. We must now evaluate $-\varepsilon H(x_1)$ and $-\varepsilon H(x_2)$ to calculate the centre function for the composition (2.4). Expanding

$$H(x \pm \varepsilon \dot{x}) = H(x) + \tfrac{1}{2}\varepsilon^2 \dot{x} \mathscr{H} \dot{x} + \cdots \,, \tag{2.17}$$

we obtain, to lowest order in ε,

$$H(x_1) = H(x_2) = H(x) + \tfrac{1}{2}\varepsilon^2 \dot{x} \mathscr{H} \dot{x} \,. \tag{2.18}$$

Therefore, the composition of two short time flows is approximately

$$S_t(x) = -tH(x) - \tfrac{1}{2}(t/2)^3 \dot{x} \mathscr{H} \dot{x} \,. \tag{2.19}$$

Comparing this result with (B.10) in Appendix B, we find that we have overestimated here the nonlinear term: We shall show that S_t depends on the area between the smooth trajectory and the chord, rather than the triangle Δ_3.

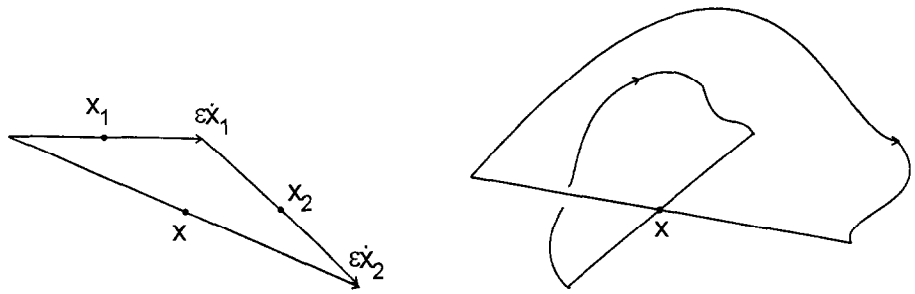


Fig. 2.5. To evaluate the correction of the centre function for short times t, we compose two flows for $\varepsilon = t/2$.

Fig. 2.6. Two possible paths, whose actions are compared by the centre variational principle.

It is now possible to obtain higher order corrections to the centre function for a time t flow by composing four elementary generating functions of the form $-\varepsilon H(x)$, with $\varepsilon = t/N$, in the guise of two transformations generated by (2.19). However, our immediate objective is to note that composing an even number N of flows for time $\varepsilon = t/N$ we obtain the centre function

$$S_t(x) = \sum_{n=1}^{N} S_{t/N}(x_n) + \Delta_{N+1}(x, x_1, \ldots, x_N)$$

$$= -\frac{t}{N} \sum_{n=1}^{N} H(x_n) + \Delta_{N+1}(x, x_1, \ldots, x_N) + \mathrm{O}(t^3/N^2)\,. \tag{2.20}$$

Increasing the number of subdivisions of the interval to infinite, we thus obtain

$$S_t(x) = \lim_{N\to\infty} \left\{ -\frac{t}{N} \sum_{n=1}^{N} H(x_n) + \Delta_{N+1}(x, x_1, \ldots, x_N) \right\}. \tag{2.21}$$

The polygon Δ_{N+1} has one large side ξ passing through the centre x and N small chords

$$\eta_j = \xi_j \underset{N\to\infty}{\longrightarrow} (t/N)\mathfrak{J}\,\partial H/\partial x_j\,, \tag{2.22}$$

i.e. they are tangents to the orbit as $N \to \infty$. Throughout the limiting process, we guarantee that $\partial S_t/\partial x_j = 0$, so that we arrive at the *centre variational principle*: The *centre action*

$$S_t(x) = \oint_x p \cdot \mathrm{d}q - \int H(x(t))\,\mathrm{d}t \tag{2.23}$$

is stationary along the classical trajectory. The paths to be compared always have their endpoints centred on the point x. The second integral is evaluated along this path, whereas the symplectic area defined by the first integral is closed off by the chord centred on x, as displayed in Fig. 2.6. Note that $S_t(x)$ is necessarily an odd function of t.

This particular version of the variational principle dispenses with any specification of appropriate topologies for sophisticated spaces of paths. For any large but finite N, the path becomes a polygonal line, uniquely specified by the centres of each of its sides and the point x. In the large N limit, the sides of the stationary polygonal line shrink to form a smooth curve that coincides with the

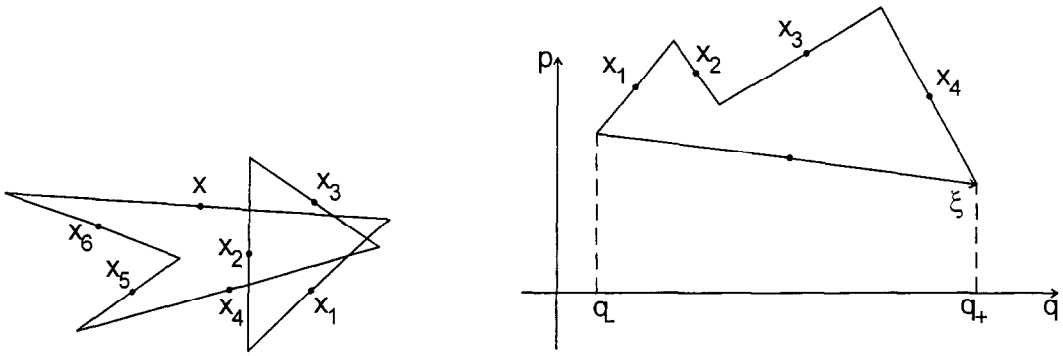


Fig. 2.7. A typical polygonal line with self-crossings for $N=6$.

Fig. 2.8. The symmetrized Legendre transform of the centre action in the variational principle restricts the q component of ξ to equal $(q_+ - q_-)$. The new action now includes all the symplectic area between the polygonal line and the q-plane.

trajectory. What is the result of moving one centre x_j from the polygonal approximation for the trajectory? We have seen that the side ξ_j is unaffected, whereas, according to (2.10), all the other sides will be altered by $\delta\xi_k = \pm 2\delta x_j$. Therefore, the neighbouring paths are also polygonal lines with small sides of order δx_j. Evidently, there will be collective variations of centres such that the variations of the sides of the polygonal line remain bounded and therefore smooth in the limit $N \to \infty$: The centre action will still be stationary among this subset of paths. Conversely, no matter how large we take N, general paths defined by an arbitrary choice of centres will be very jagged polygons which may self-intersect many times! This is illustrated in Fig. 2.7.

We can guarantee that there exists an unique solution to the centre variational problem for sufficiently small times. To see this, recall that the tip of the chord x_- is determined by (1.37), where now the canonical transformation C is just G_t, the Hamiltonian flow. If t is small enough, the stationary chord through x is just $\xi = t\dot{x}$. Expanding the Hamiltonian about the orbit to second order (see Section 1), we obtain the linear transformation, $\delta x_+ = \mathcal{M}_t \delta x_-$, such that $\mathcal{M}_t$ has the Cayley parametrization (1.20). Since the centre function is proportional to $H(x)$ for small times, we obtain

$$\mathcal{M}_t = [1 + \tfrac{1}{2}t\mathfrak{J}\mathcal{H}][1 - \tfrac{1}{2}t\mathfrak{J}\mathcal{H}]^{-1}, \tag{2.24}$$

where $\mathcal{H}$ is the Hessian matrix for $H(x)$. For small times $\mathcal{M}_t$ is close to the unit matrix and therefore G_t is far from being locally a reflection. It is only when t is large enough for G_t to become locally a reflection that a bifurcation will occur, beyond which there will be more than one chord for each centre.

Evidently, the action $S_t(x)$ that is minimized by the variational principle is the centre function that generates the Hamiltonian flow for the time t. In the special case of linear transformations studied in Section 1, we identified the centre function with the area of the triangle in Fig. 1.3. If we extend the composition of an arbitrarily large number of small time evolutions, such as shown in Fig. 2.2, we now identify the first integral in (2.23) with the area between the orbit and the chord. If the quadratic Hamiltonian is time-independent, the second integral is just Et, the area of the circular sector subtended between x_+ and x_- in Fig. 1.4(b), hence the difference is the area of the triangle. Indeed, we can generally equate $S_t(x)$ with the area of a curvilinear triangle between

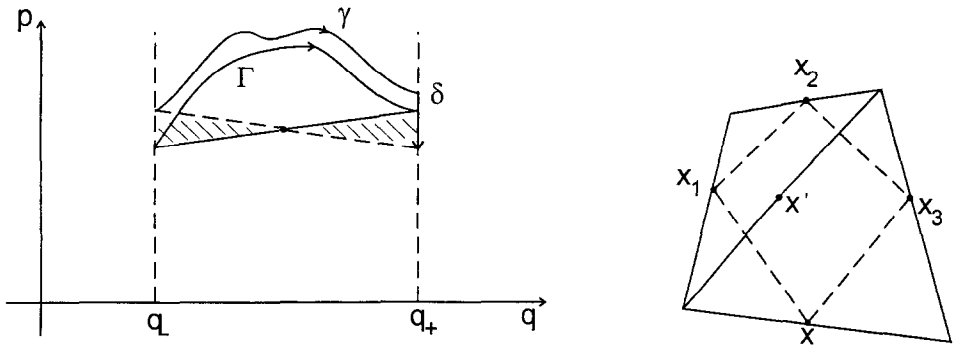


Fig. 2.9. Neighbouring paths γ with the same fixed endpoints q_+ and q_- as the trajectory Γ may also be considered as neighbouring paths with the same centre, by adding a vertical segment δ. Stationarity of the action for the latter implies that of the former.

Fig. 2.10. The composition of three centre functions requires that two triangles with centres x_1, x_2, x' and x', x_3, x share a common side. This imposes the restriction that x_1, x_2, x_3, x, must form a parallelogram.

ξ and the origin, but this is not often useful, because it requires the knowledge of time evolution of orbits with different energies, whereas (2.23) depends on a single orbit.

To derive the usual variational principle with fixed positions, we merely perform the symmetrized Legendre transform (1.54), (1.55) on $S(x)$. Distinguishing the various actions or generating functions only by their arguments, we obtain

$$S(q_+, q_-) = S\left(p, \frac{q_+ + q_-}{2}\right) + p \cdot (q_+ - q_-), \tag{2.25}$$

with p determined as a function of $(q_+ - q_-)$ by the condition that

$$\xi_q = \frac{\partial}{\partial p} S\left(p, \frac{q_+ + q_-}{2}\right) = -(q_+ - q_-). \tag{2.26}$$

In other words, we are merely demanding that the chord ξ fit precisely between the two planes $q = q_-$ and $q = q_+$, as shown in Fig. 2.8. Since the symplectic area between ξ and the q axes is equal to the last term in (2.25), we obtain

$$S(q_+, q_-) = \int_{q_-}^{q_+} p \cdot \mathrm{d}q - \int H \, \mathrm{d}t, \tag{2.27}$$

in which the first integral is the entire area under the path.

It remains to show that the action (2.27) of the trajectory Γ, joining the endpoints q_+ and q_- in position space is stationary with respect to neighbouring paths with the same endpoints and the same interval of time. But reference to Fig. 2.9 shows that all such paths γ may also be considered as paths with the same centre that are closed by an extra vertical segment, δ. The centre variational principle guarantees that the action $S_\Gamma(p, q)$ is stationary among such paths. Since the two shaded triangles in Fig. 2.9 have the same symplectic area, the difference in action between $S_\gamma(p, q)$ and $S(q_+, q_-)$ for all the paths that are being compared is the constant $p \cdot (q_+ - q_-)$. Therefore $S(q_+, q_-)$ is also

stationary among all the neighbouring paths γ. The deduction of the centre variational principle from the one with fixed endpoints is presented in Ref. [12].

It is important to note that discretization of the variational principle with fixed endpoints is not favoured by the cancellation of the second order term as in (2.19). However, it is simple to compose generating functions $S(q_+, q_-)$ by using the rule that

$$S(q_-, q_+) = S_1(q_-, q_1) + S_2(q_1, q_2) + \cdots + S_n(q_{n-1}, q_+) \tag{2.28}$$

with

$$\partial S/\partial q_j = 0\,. \tag{2.29}$$

This condition implies that the final momentum for the jth transformation,

$$p_{jj} = \partial S_j/\partial q_j = \partial S_{j+1}/\partial q_j = p_{jj+1}\,, \tag{2.30}$$

where p_{jj+1} is the initial momentum for the $(j+1)$th transformation, for all j, so that we obtain an unique path. However, the small time decompositions $S_j(q_j, q_{j+1})$ are not obtained directly from the Hamiltonian, forcing us to use a Lagrangian formulation which is awkward for quantization.

If we enquire into the time variation of the centre function (2.23) defined along the trajectory, we find that, to first order, $\delta S_t = -H(x_+)\delta t$, because of the variational principle. Combining this with (1.24) and (1.25), we obtain a partial differential equation for the action,

$$\partial S_t/\partial t + H(x - \tfrac{1}{2}\Im\,\partial S_t/\partial x) = 0\,, \tag{2.31}$$

which is Marinov's version of the Hamilton–Jacobi equation [9]. Actually, the deduction from the variational principle shows that we could also use $H(x + \frac{1}{2}\Im\partial S_t/\partial x)$ or any weighted average of these two arguments of the Hamiltonian.

To conclude this section we should discuss the composition of an odd number of canonical transformations. This includes the important case of viewing a canonical transformation C in new coordinates: $C \to C' = G^{-1}CG$. First we note that the general rule (2.13) could have been obtained iteratively by joining successive triangles. For example, the composition of four transformations can be viewed as $C_1 \circ C_2$ composed with $C_3 \circ C_4$. Thus, the pentagon that determines the transformation is considered as the sum of three triangles in Fig. 2.3. In general, we then view Δ_{2n+1} as the sum of a $(2n-1)$-sided polygon with two triangles.

Evidently, we can then derive the composition of an odd number of transformations by adding a single triangle to an odd-sided polygon. There results an even-sided polygon – a circumscribed quadrilateral, in the case of three transformations. The problem is that we are not free to place the centres of this even polygon just anywhere we wish. In the case of the quadrilateral, there is a simple theorem that the inscribed quadrilateral must be a parallelogram as shown in Fig. 2.10. In the general case, we recall that an odd number of reflections about pre-assigned centres is again a reflection, rather than a translation. So the centre of this reflection, which must coincide with the centre of the remaining side of the even polygon, is uniquely determined by the other centres. Taken together, we therefore have $2n$ reflections adding up to the identity. Considering the rule for composing reflections (1.40), we obtain the compatibility condition as

$$\sum_{n=1}^{2n} (-1)^k x_k = 0\,, \tag{2.32}$$

where we include the centre $x = x_{2n}$.

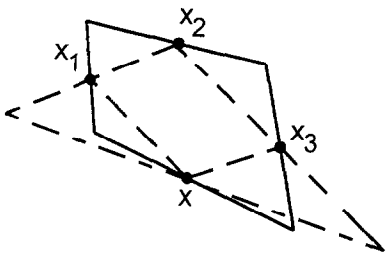


Fig. 2.11. There are an infinite number of circumscribed quadrilaterals to the centres x, x_1, x_2 and x_3 if they form a parallelogram, but they all have the same symplectic area. In particular, one of the sides (ξ_2 in this case) may shrink to zero to form a triangle.

Conversely, once we fix the remaining centre, so that we describe the polygon as $2n$ reflections, we can choose an arbitrary initial corner for the circumscribed polygon, since the compatibility condition is just that these reflections add up to the identity transformation. In the case of the quadrilateral it is obvious that the symplectic area of all these quadrilaterals equals twice the area of the fixed inscribed quadrilateral, as shown in Fig. 2.11. It is not hard to show that in general the symplectic area of the circumscribed even polygon depends uniquely on the centres of its sides. The proof is presented in Appendix A.

We can use the freedom in the choice of the circumscribed quadrilateral to fit three of its sides to chords obtained from the generating function being composed. Thus defining

$$S(x) = S_1(x_1) + S_2(x_1 + x_3 - x) + S_3(x_3) + \Delta_4 , \tag{2.33}$$

where we used (2.33) to eliminate x_2, and then eliminating x_1 and x_3 through the conditions

$$\partial S/\partial x_1 = \partial S/\partial x_3 = 0 , \tag{2.34}$$

we obtain the centre function for the full transformation. It should be noted that conditions (2.34) will be satisfied only when all sides of Δ_4 coincide with the chords through the centres of its sides.

We will not be concerned with the rules for comparing chord transformations. The reader will find a presentation in Ref. [12]. They involve the description of phase-space polygons in terms of the vectors composing their sides, rather than the centres. The relation of this rather more natural geometry to the specification in terms of centres is discussed in Appendix A, which also presents general formulae for the symplectic area of polygons and their variation with respect to arbitrary displacements of the centres.

3. The energy shell, sections and maps

So far we have considered paths that take the same time, rather than polygonal lines whose centres have a fixed energy. However, for an autonomous system we know that the $N \to \infty$ limit of the action, i.e. the solution of the variational problem is a trajectory with constant energy. Thus, we can restrict the variations to other paths where all the centres have the same energy, without violating

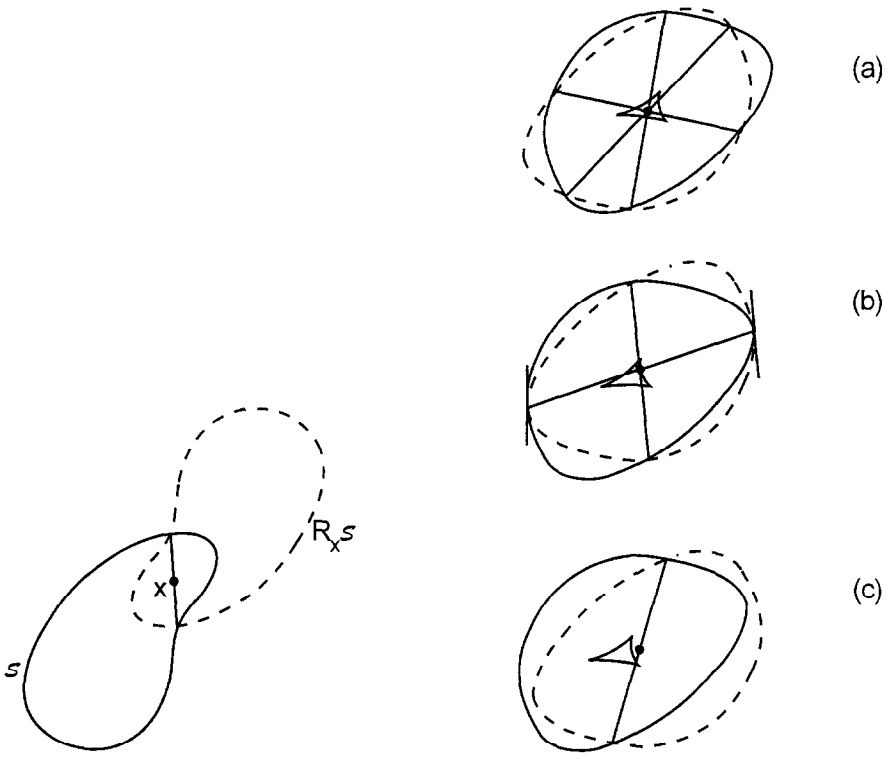


Fig. 3.1. The chord through x is determined by the intersections of the energy curve $\mathcal{S}$ with its reflection through $x: R_x\mathcal{S}$. Moving x by δx translates R_xS rigidly by $2\delta x$.

Fig. 3.2. The Wigner caustic is the small cusped curve within the energy shell. (a) Within the caustic there are three chords. (b) On the caustic two chords coalesce. The tangents at the tips of the degenerate chord are parallel, so they are obtained as a reflection through the centre. (c) Outside of the Wigner caustic there is only one chord.

the variational principle. Within this restricted class, for which $H(x_j)=E$, if we still demand that all paths take the same time, we find that $\delta S_t=\delta\Delta_{N+1}$. But the symplectic area of this polygon is not itself a function of time, so we arrive at the *energy (centre) variational principle*: The symplectic area $\oint_x p\cdot dq=S_E$ between all the paths on the energy shell and all the chords centred on x is stationary for a classical orbit. Evidently, we obtain the usual energy variational principle, defined in terms of the end positions, by taking the symmetrized Legendre transform of the centre action, in the same way as for the time variational principle.

Are there solutions to the energy variational problem? First, let us examine the simplest case of a single freedom ($L=1$), where the energy shell, $\mathcal{S}$, defined by $H(x)=E$, is just a compact curve. The tips of the chord through x result from the reflection, R_x, through this point. We therefore obtain the chord by reflecting the entire shell and taking the intersection of $\mathcal{S}$ with $R_x\mathcal{S}$. This is displayed in Fig. 3.1.

Note that displacements, δx, merely translate the reflected curve rigidly by $2\delta x$, because of (1.40) and (1.41). If $\mathcal{S}$ is convex, there will be no intersection of $\mathcal{S}$ with $R_x\mathcal{S}$ when x lies outside of $\mathcal{S}$. If x is inside and close to $\mathcal{S}$, there will be an unique chord. Indeed, $S_E(x)$ is just the area

sandwitched between the chord and the shell, whereas $\xi(x)=\Im\partial S/\partial x$. Thus, $S_E(x)$ grows as x is brought further into the interior of $\mathscr{S}$.

Deep in the interior of $\mathscr{S}$ we run into a bifurcation, beyond which there are two new chords. The locus of these points is known as the *Wigner caustic*, because it was first studied by Berry [7] in the context of the semiclassical Wigner function. Within this closed curve there are three chords and hence three functions $S_E(x)$ [7]. As x is moved onto the Wigner caustic, two chords coalesce and vanish, so that their corresponding centre functions become singular, according to (1.47), while the third centre function moves regularly through the caustic. This process is depicted in Fig. 3.2. Notice that the tangents to the shell at the tips of the degenerate chord are parallel. The implication is that, if the curvature were ignored, there would be a continuum of chords in the shell centred about the same point x, i.e., related by a reflection through x. Further discussion of the Wigner caustic for $L=1$ is provided by Berry [7], where it is shown that there are always three cusps if $\mathscr{S}$ is convex.

To obtain the chords that solve the variational problem for a given centre x, when $L>1$, we again reflect the energy shell $\mathscr{S}$ and examine the intersection of $\mathscr{S}$ with $R_x\mathscr{S}$. Since both of these surfaces have $(2L-1)$ dimensions, their intersection defines the *centre section*. This $(2L-2)$-dimensional surface is the locus of the tips of all the chords on the energy shell that are centred on x, the problem now being reduced to finding which chord tips belong to the same orbit.

Let us now consider the orbits in $\mathscr{S}$ through a given neighbourhood of $\mathscr{S}\cap R_x\mathscr{S}$. If $\mathscr{S}$ is compact, these orbits will eventually reintersect the section, defining the *centre map*, C_x, of the section onto itself (as a consequence of the Poincaré recurrence theorem [18, 19]). Evidently, the fixed points of the composition $R_x\circ C_x$ define the solutions of the variational problem.

If $\mathscr{S}$ is convex (as well as compact) and x lies outside of $\mathscr{S}$, then $R_x\mathscr{S}$ does not intersect $\mathscr{S}$ and there are no solutions to the energy variational problem. As x moves onto $\mathscr{S}$, $R_x\mathscr{S}$ is translated rigidly until we obtain a tangency of $\mathscr{S}$ with $R_x\mathscr{S}$. The chord then coincides with the locally straight orbit through x and $S_E(x)$ is still zero. Moving x into $\mathscr{S}$ determines a smooth family of maps $R_x\circ C_x$ with $(2L)$ parameters, the components of x. The fixed points of these maps will be smooth functions of these parameters, unless there are bifurcations. Therefore, if x is close enough to the shell, we can guarantee the existence of unique short orbit solutions to the variational problem.

However, these are not the only possible short chord solutions even if x is close to $\mathscr{S}$. Consider again the case of $L=1$, where the section is just a pair of points. We may consider the unique chord through x as pertaining to the short orbit or the long orbit, as shown in Fig. 3.3, or to different multiples of these combined, so that we wind many times around this periodic orbit while travelling from one tip of ξ to the other. Exactly the same happens for a point x at the centre of a periodic orbit chord if $L>1$. Each of these different windings will introduce a chord corresponding to a different centre map for which it is the fixed point. Since all these maps are smooth with respect to variations of x, we also obtain solutions for nearby centres x, even though the segments no longer constitute a periodic orbit. Thus, close to the shell, the chords can be classified as belonging to short orbits or to smooth continuations of chords belonging to nearby periodic orbits.

It will be important to understand the geometric structure of the centre section. If x lies close to the shell, this geometry is revealed by expanding $H(x+X)$ in powers of X:

$$H(x+X)=H(x)+(\partial H/\partial x)\cdot X+\tfrac{1}{2}X\mathscr{H}_xX+\cdots, \tag{3.1}$$

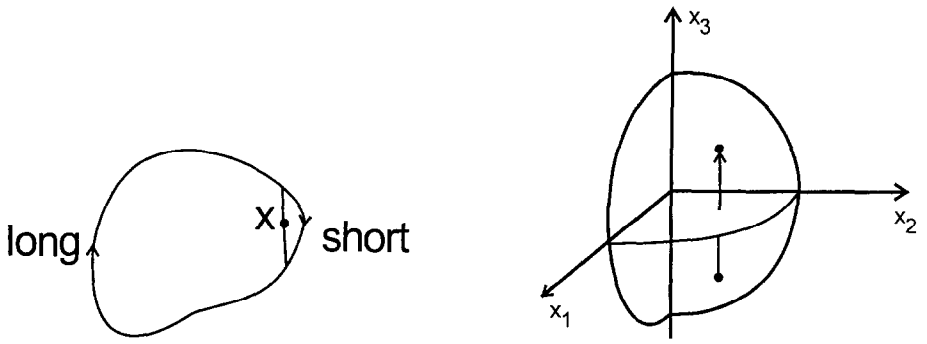


Fig. 3.3. The chord through x determines two alternative centre actions $S(x)$ related to the short and the long segments of the periodic orbit that constitutes the shell in the case $L=1$.

Fig. 3.4. The centre section has the topology of a sphere. If x is sufficiently close to the shell, the short orbits will be approximately straight. The (short) centre map injects points from the lower into the upper hemisphere.

where $\mathscr{H}_x$ is the Hessian matrix at x, i.e. at $X=0$. The centre section is defined by the equations $H(X)=E$ and $H(X)=H(-X)$, which becomes

$$(\partial H/\partial x)\cdot X=0 \quad \text{or} \quad \Im(\partial H/\partial x)\wedge X=0 \tag{3.2}$$

to second order in X. Thus, the section of $\mathscr{S}_E$ by $R_x\mathscr{S}_E$ coincides in this limit with the section by the plane tangent to the shell through the point x. This is the more familiar construction for a *Poincaré section* usually chosen to be a plane cutting the energy shell.

The form of the section is now revealed by including (3.2) in the expansion (3.1), so that

$$E-H(x)=\tfrac{1}{2}X\mathscr{H}_xX\,. \tag{3.3}$$

Adopting coordinates such that the plane (3.2) becomes $X_{2L}=0$, the $(2L-2)$-dimensional section becomes the ellipsoid (3.3) with $X_{2L}=0$. Therefore, the topology of the centre section is that of a $(2L-2)$-sphere in the $(2L-1)$-dimensional energy shell. As x approaches the shell, this reduces to an ellipsoid, but, even when x is far from the shell, the section remains invariant with respect to reflections about x.

Taking coordinates such that $X_{2L}=0$ is the tangent plane implies that locally H is only a function of X_{2L}. Therefore, the nearly straight parallel orbits intersecting this small ellipsoid will have the direction of the conjugate coordinate, which we shall call X_{2L-1}. Taking this as the vertical direction transverse to the plane $X_{2L-1}=0$, we can use the remaining coordinates $(X_1,\ldots,X_{2L-2})$ as coordinates for the two halves of the ellipsoid, as shown in Fig. 3.4 for $L=2$. The usual definition of a *Poincaré map* involves only one sense of the vertical traversal of the trajectory in or out of the sphere; since we work directly with the coordinates in the $(X_1,\ldots,X_{2L-2})$ plane, the Poincaré section is considered as a disk. The centre map C_x concerns chords centred on x (i.e. $X=0$), so we must consider its full spherical structure.

Evidently, the spherical topology is preserved for large spheres centred on x far inside $\mathscr{S}$. The orbits that intersect such a sphere will divide it into an incoming and an outgoing hemisphere, which correspond, respectively, to the lower and the upper hemispheres in the previous coordinates. In the

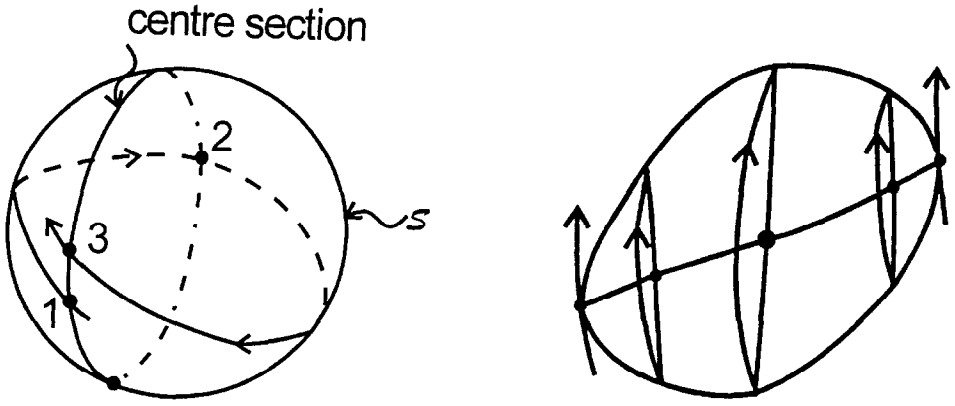


Fig. 3.5. Reducing by one the dimension of the shell $\mathscr{S}$, it would become a two-dimensional spherical surface and the centre section ($\mathscr{S} \cap R_x \mathscr{S}$) would be represented by a closed curve. The orbits define the Poincaré map ($1 \rightarrow 3$) as the composition of the in-map ($1 \rightarrow 2$) with the out-map ($2 \rightarrow 3$).

Fig. 3.6. The equatorial surface is the locus of midpoints of the chords joining the outgoing and the incoming regions of the centre section.

simplest case, $L = 2$, the two hemispheres are divided by a line (the equator) where the orbits are tangent to the sphere.

The centre map, when viewed on the sphere, carries points from the incoming hemisphere to the outgoing hemisphere. It should be born in mind that the definition of inside and outside of the centre section is only natural when x is close to the shell. Since $\mathscr{S}$ is compact, the two regions into which it is sectioned are also compact. So if we take x continuously from one side of the shell to the other, we exchange the 'natural' inside and outside of the sphere. In this traversal there may be further bifurcations analogous to those for $L = 1$, but these have not been analysed so far.

Once we choose which side is labeled incoming and which is outgoing, we see that the usual Poincaré section is the composition of two maps between the hemispheres. First we travel 'inside' the sphere, then return to the original hemisphere following an outer orbit, as sketched in Fig. 3.5. This is analogous to the way that Bogomolny obtains his sections in quantum mechanics [20].

In the case of ordinary Poincaré sections by an arbitrary $(2L-1)$-dimensional plane, such as $x_{2L} = 0$, we merely describe the appropriate hemisphere by means of the coordinates $x' = (x_1, \ldots, x_{2L-2})$. That is why the section appears as a disc. For the description of the centre map it is natural to define the *equatorial surface* of the sphere. This is the locus of midpoints for the chords of all orbits that intersect the spherical surface as shown in Fig. 3.6. We have verified that, as $x \rightarrow \mathscr{S}$, the orbits collapse onto their straight chords. Thus, in this limit, the equatorial surface becomes indistinguishable from the plane used for the Poincaré section.

Let us now consider the restriction of the action $S_E(x)$ to the equatorial surface. The chords generated by differentiating S_E as a centre function, will be of the form $\xi(x') = (\xi', \xi_{2L-1}, \xi_{2L})$, where ξ' is a vector within the tangent plane to the equatorial surface, at the point x'. Since the centre function generates canonical transformations, (1.53) holds, i.e.

$$\mathrm{d}\xi \wedge \mathrm{d}x|_{x'} = \mathrm{d}\xi' \wedge \mathrm{d}x' = 0 \tag{3.4}$$

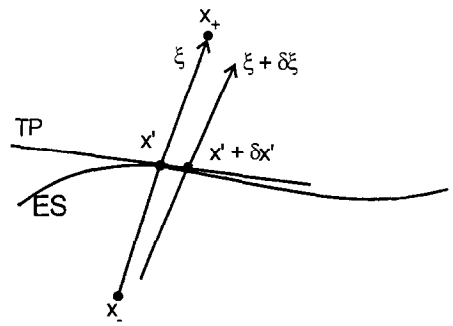


Fig. 3.7. The point x' is constrained to the equatorial surface (ES). The variation $\delta\xi'$ is the projection of $\delta\xi$ onto the tangent plane (TP).

(the geometrical basis for this equation is displayed in Fig. 3.7). This shows that the projection of the transformation between the two hemispheres is canonical when restricted to the tangent plane of the equatorial surface. In the limit of very small spheres, this surface tends to a plane, along which the transformation will be canonical. However, when the curvature of the equatorial surface must be taken into account, there will be no single plane along which the transformation induced on the section will be seen as canonical, even though the full symplectic action is preserved for any section. This is consequence of the Poincaré–Cartan theorem [15], which may be directly obtained from the classical variational principles [19].

The situation is simpler for the more familiar position action

$$S_E(q_+, q_-) = \int_{q_-}^{q_+} p \cdot \mathrm{d}q \tag{3.5}$$

(where $p(q)$ is an orbit on the energy shell) which is obtained by taking the Legendre transform (1.54), (1.55) of $S_E(x)$. Obviously, we obtain the corresponding momenta $p_\pm$ by taking the derivatives (1.56). Therefore the restriction to particular planes, $q_{+L} = a_+, q_{-L} = a_-$, implies that this action, $S_E(q_+, q_-)$, becomes the generating function of the canonical transformation $x'_- = (p_1, \ldots, p_{L-1}, q_1, \ldots, q_{L-1}) \to x'_+$, through the equations

$$p'_+ = (\partial S_E/\partial q'_+)(q'_+, a_+, q'_-, a_-)\,, \qquad p'_- = -(\partial S_E/\partial q'_-)(q'_+, a_+, q'_-, a_-)\,. \tag{3.6}$$

In the particular case where $a_+ = a_- = a$, we obtain the generating function for the Poincaré maps.

Sections are ideal instruments for the study of periodic orbits. The Poincaré section can often be chosen so as to include almost all the orbits in the shell as periodic points. Conversely, by bringing a centre within the shell as close as we like to a periodic orbit, we can make the spherical section as small as we like, thus excluding nearby periodic orbits. It is true that, since these are dense, there will always be very long periodic orbits crossing the section, but these will not appear as a fixed point in the first iteration of the map.

Let us consider a tube of trajectories that crosses the centre sphere and includes a periodic orbit. This orbit will return to the section and it will bring back with it a neighbourhood of non-periodic

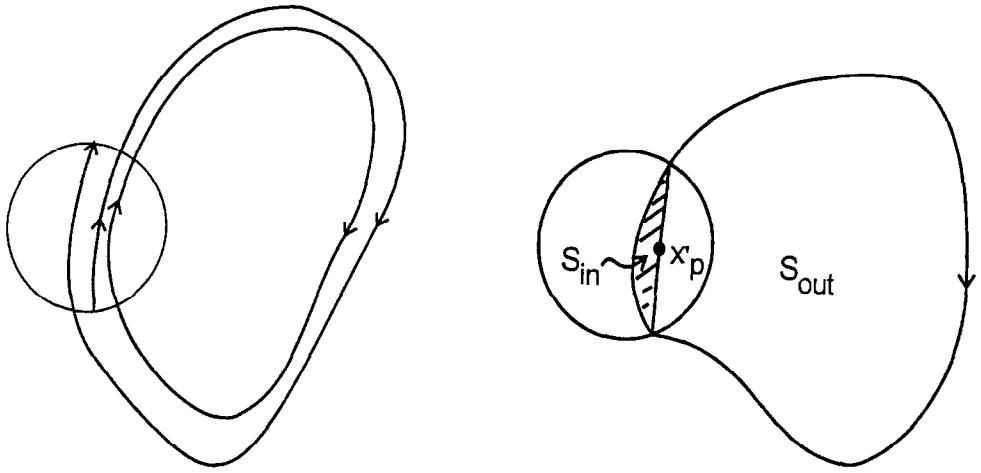


Fig. 3.8. Generally, the orbits in a narrow neighbourhood of a periodic orbit that intersects a centre section are not themselves periodic.

Fig. 3.9. The total symplectic area $\mathscr{S}$ for a periodic orbit that crosses the centre section is split by the chord through x'_p into the components $S_{\rm in}$ and $S_{\rm out}$.

orbits as shown in Fig. 3.8. Eventually, nearly all the orbits will return, because of the Poincaré recurrence theorem [18, 19], but they may explore entirely different regions of the energy shell. Indeed, if the system is ergodic, they will come arbitrarily close to any point in the shell.

The periodic orbit closes in a sequence of two mappings between the ingoing and the outgoing hemispheres. We have seen that the *in-map* is trivial for all the orbits, in the limit of small spheres. It is then approximately the identity map in the coordinates of the equatorial plane. The action $S_{\rm in}(x')$ will thus be small for all the orbits, so that later we shall use the results of Appendix B to estimate it. We can clearly decompose the action of the periodic orbit itself as

$$\mathscr{S} = \oint p \cdot {\rm d}q = S_{\rm in}(x'_p) + S_{\rm out}(x'_p)\,, \tag{3.7}$$

where x'_p is the centre of the chord for the periodic orbit (lying in the equatorial surface) and $S_{\rm out}(x'_p)$ is the action for the *out-map*. Thus there are two chords $\xi_{\rm out}(x'_p) = -\xi_{\rm in}(x'_p)$ shown in Fig. 3.9, corresponding to two different solutions of the variational problem, centred on x'_p.

These two solutions have very different actions, but, by adding two in-traversals to the out-orbit, as shown in Fig. 3.10, we obtain a new solution with action $S_{\rm long}(x')$, for which again $\xi_{\rm long}(x') = -\xi_{\rm out}(x') = \xi_{\rm in}(x')$. Evidently, the long solution has nearly the same action as the out solution, since

$$S_{\rm long}(x'_p) = \mathscr{S} + S_{\rm in}(x'_p)\,, \tag{3.8}$$

whereas $S_{\rm out}(x'_p) = \mathscr{S} - S_{\rm in}(x'_p)$. There is an evident extension for the case of multiple windings of the periodic orbit.

So far we have considered only chords lying on a periodic orbit. However, what we are really interested in is the chord whose centre coincides with the centre, x, of the sphere. Its tips need not

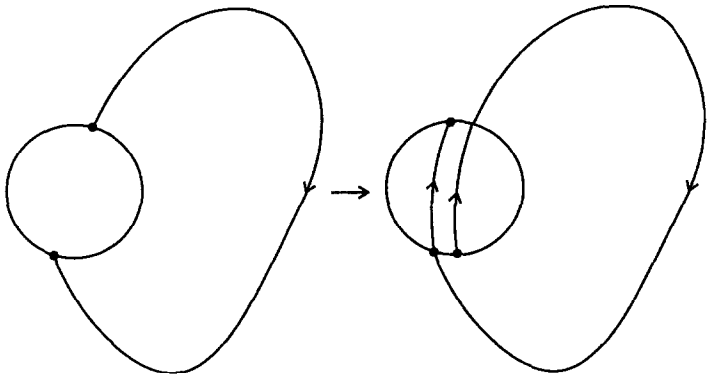


Fig. 3.10. The long orbit is obtained by adding in-segments to the tips of an out-orbit. In the case of periodic orbits, the in-segments overlap.

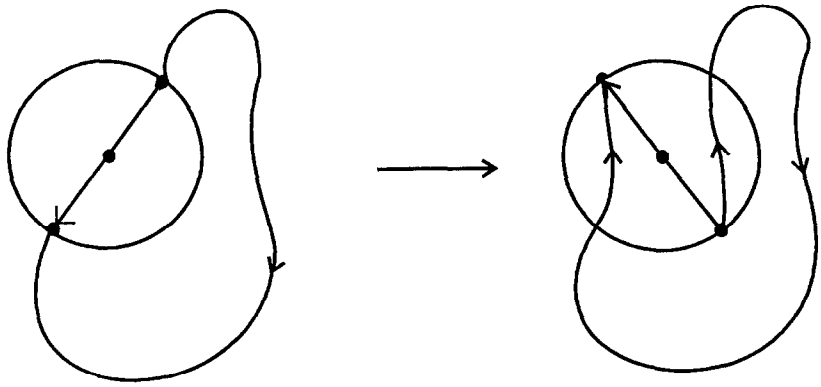


Fig. 3.11. The chords passing through the centre of the section do not generally belong to a periodic orbit, but, again, we obtain the long orbit by adding two in-segments to an out orbit.

lie on a periodic orbit. For any solution $\xi_{\rm out}(x)$, we will also find a solution $\xi_{\rm long}(x)$, obtained by adding two in-mappings to the out-map, as shown in Fig. 3.11. Let us define x'_+ to be the centre of the chord for the in-map that extends the out-map in the forward direction and x'_- as the centre for the backwards extension. Then clearly

$$S_{\rm long}(x) = S_{\rm out}(x) + S_{\rm in}(x'_-) + S_{\rm in}(x'_+)\,. \tag{3.9}$$

For the chords centred on x, we will have $x'_+ = -x'_-$, but the corresponding centre actions need not be equal. It is important to note that the relation (3.9) between the long action and the out action is quite general for chords of orbits traversing a small section. It is therefore entirely independent of our ability of expressing either of these actions in terms of the action $\mathscr{S}$ of a nearby periodic orbit.

Let us now suppose that indeed there is a periodic orbit traversing the section. Then x'_p will be a fixed point of the out-map, but $x'_+ = -x'_- \neq 0$ if $x'_p \neq 0$. (If the periodic orbit determines periodic points of the Poincaré map, then we must compose further in and out maps so as to obtain a fixed point. In any case, reduction of the section can eliminate all but one of the periodic points).

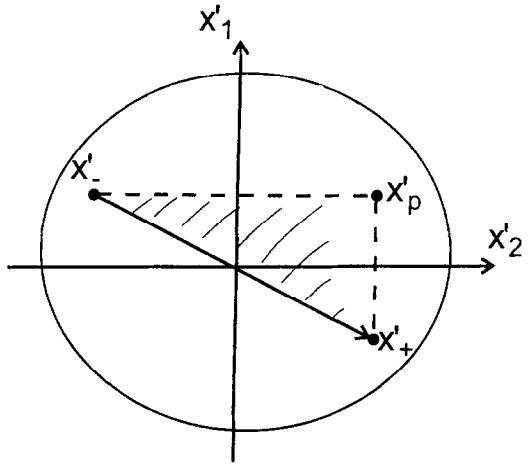


Fig. 3.12. The shaded area of the triangle determines the centre function for the canonical transformation $x_- \to x_+$, linearized around x'_p. In fact, this is just the out-map, so the symplectic area of the triangle is $\delta S_{\rm out}$. The chord for the long map projects onto the same vector between x'_- and x'_+. When these two points touch the outer curve that limits the classical motion in the equatorial plane (i.e. the x' plane), the full chords $\xi_{\rm out}$ and $\xi_{\rm long}$ coalesce in a Wigner caustic.

Evidently, the action $S_{\rm out}$ at the centre of the sphere can be obtained from the action of the periodic orbit:

$$S_{\rm out}(x) = S_{\rm out}(x'_p) + \delta S_{\rm out} = \mathscr{S} - S_{\rm in}(x'_p) + \delta S_{\rm out}\,, \tag{3.10}$$

$$S_{\rm long}(x) = \mathscr{S} - S_{\rm in}(x'_p) + S_{\rm in}(x'_+) + S_{\rm in}(x'_-) + \delta S_{\rm out}\,. \tag{3.11}$$

We may say that $\delta S_{\rm out}$ takes care of "transverse" action differences with respect to the periodic orbit of both $S_{\rm out}$ and $S_{\rm long}$, since the "longitudinal" differences between these is already taken into account by the $S_{\rm in}$ terms.

If we now linearize $\delta S_{\rm out}$ around the fixed point x'_p, we find that, according to (1.26), $\delta S_{\rm out}$ can be identified with the area of the triangle formed by (x'_p, x'_-, x'_+), displayed in Fig. 3.12. Thus,

$$\delta S_{\rm out} \simeq (x - x'_p)\mathfrak{B}_p(x - x'_p) = X'_p \mathfrak{B}_p X'_p\,, \tag{3.12}$$

where

$$\mathfrak{B}_p = [1 - m_p][1 + m_p]^{-1}\,. \tag{3.13}$$

For a small enough sphere, we can identify m_p, the Jacobian matrix for the fixed point of the out-map, with the stability matrix for the full return of the periodic orbit.

It only remains to estimate $S_{\rm in}(x')$ for the traversal of a small spherical section. The symplectic area of a chord for a short interval is obtained to third order in time in (B.13) of Appendix B. This is just $S_{\rm in}(x')$ for a point in the equatorial plane:

$$S_{\rm in}(x') \simeq \tfrac{1}{12} t^3 \dot{x}' \wedge \ddot{x}' = \tfrac{1}{12} t^3 \dot{x}' \mathscr{H}' \dot{x}'\,, \tag{3.14}$$

where $\dot{x}'$ and $\ddot{x}'$ are the velocity and the acceleration at the point x' and $\mathscr{H}'$ is the Hessian of the Hamiltonian at this point. The "time of flight" for the in-map can now be eliminated by using (2.17) or (B.11):

$$E - H(x') = \tfrac{1}{8} t^2 \dot{x}' \mathscr{H}' \dot{x}' , \tag{3.15}$$

so that

$$S_{\rm in}(x') \simeq \tfrac{4}{3} 2^{1/2} [(E - H(x'))^{3/2}] / [\dot{x}' \mathscr{H}' \dot{x}']^{1/2} . \tag{3.16}$$

Thus, the action of a small chord is obtained by placing $x' = 0$ (at the centre of the section), whereas $S_{\rm long}$ and $S_{\rm out}$ result from the evaluation of (3.16) for $x'_{\pm}$ and x'_p, where appropriate.

If we are merely calculating $S_{\rm in}$ itself, it is clear that this action is positive when the shell is convex and tends to zero as the centre approaches the shell. Outside the shell, we would obtain an imaginary action, according to (3.16). For the evaluation of $S_{\rm out}$ and $S_{\rm long}$, we must then calculate $S_{\rm in}$ at points where it is smaller than at the centre of the section. Indeed, as $x'_{\pm}$ move to the boundary of the equatorial plane shown in Fig. 3.12, $H(x') \to E$ and we obtain $S_{\rm out} = S_{\rm long}$, according to (3.9). This is a Wigner caustic, beyond which we *again have imaginary action, even though the reflection centre is inside the shell.* The possibility of expressing the actions in terms of that of a periodic orbit also clearly depends on this orbit crossing the section, since otherwise (3.16) becomes imaginary for the orbit. In the original Berry theory [14], the various contributions of the in-action were all evaluated at the reflection centre. This is an excellent approximation to the present theory, as long as x'_p and $x'_{\pm}$ all lie well within the spherical section, but it breaks down as the sphere is shrunk to zero by bringing the reflection centre onto the shell away from a particular periodic orbit.

We can now resolve $x'_{\pm}$ in terms of x'_p, within the linear approximation, using $X'_- = -X'_+$ and

$$X'_+ - X'_p = m_p (X'_- - X'_p) , \tag{3.17}$$

where $X'_j = x'_j - x$ and all vectors are restricted to the equatorial plane. It follows that

$$X'_{\pm} = \pm \mathfrak{J} B_p X'_p \tag{3.18}$$

so that close to the shell

$$S_{\rm out}(x) = \mathscr{S} + X'_p B_p X'_p - \tfrac{4}{3} 2^{1/2} \{[E - H(x'_p)]^{3/3}\} / [\dot{X}'_p \mathscr{H} \dot{X}'_p]^{1/2} , \tag{3.19}$$

whereas

$$S_{\rm long}(x) = S_{\rm out}(x) + \tfrac{8}{3} 2^{1/2} [E - H(x'_+)]^{3/2} / [\dot{X}'_+ \mathscr{H} \dot{X}'_+]^{1/2} \tag{3.20}$$

with $x_+ = x + \mathfrak{J} B_p X'_p$. Note that the symmetric matrices $\mathscr{H}_j$ and B_p have distinct status. The former represents local structure of the Hamiltonian near the point x, whereas the latter depends on the motion along the full extent of the periodic orbit.

Finally, we derive the local structure of the Wigner caustic where the second term cancels in (3.20). Expanding $H(x'_+)$ about the point x, we obtain the energy difference between the point $x_{\rm c}$ on the caustic and the energy shell as

$$\delta E(x_{\rm c}) = \tfrac{1}{2} X'_+ \mathscr{H}_x X'_+ = -\tfrac{1}{2} X'_p B_p \mathfrak{J} \mathscr{H}_x \mathfrak{J} B_p X'_p , \tag{3.21}$$

i.e. we have a quadratic form in $x_{\rm c} - x_p$. For X'_p within the closed curve (3.21) there are two chords; these coalesce on this ellipse and cease to contribute as $x_{\rm c}$ is pushed away from the periodic orbit

leaving δE constant. Conversely, if we start on the same point, but move onto the shell keeping X_p' constant, we also cross the caustic into the classicaly forbidden region, because $\delta E \to 0$ in (3.21).

We are now in a position to appreciate the multiplicity of chords and caustics for a reflection centre that is not too close to the shell. It will then have several periodic orbits appearing as fixed points of the out-map. Its present position could have been obtained by bringing it smoothly from the shell, starting on a point of any of these periodic orbits. In each case this displacement would bring with it a pair of chords that would evolve smoothly. In each of these one parameter families of maps there must be points where the pair of chords ascribed to the other periodic orbits enter the section. Therefore, in each of these evolutions the reflection centre crosses a sequence of Wigner caustics.

4. Quantum operators: representations

The formalism of quantum mechanics has become more familiar to physicists than the more elementary structure of classical mechanics. Indeed, the ease with which we manipulate the symbols of the compact Dirac notation hides the greater difficulties of deriving rigorous results within the quantum theory. As we are not here concerned with rigor, it will be sufficient to make a quick review of the general formalism, so as to establish the notation and highlight the differences with classical mechanics.

The *state of a dynamical system* is represented by a vector $|\psi\rangle$ in Hilbert space (or Banach space) when it is described by quantum mechanics, as opposed to a point in classical phase space. The evolution of all the states for a finite time, analogous to the canonical transformations that we have been studying, results from the action of a linear *unitary operator* $\hat{U}_t$:

$$|\psi_t\rangle = \hat{U}_t|\psi_0\rangle\,, \tag{4.1}$$

with the property that the adjoint $\hat{U}_t^\dagger$ equals the inverse $\hat{U}_t^{-1}$. The one-parameter family of evolution operators $\hat{U}_t$ are the solutions of *Schrödinger's equation*

$$\mathrm{i}\hbar(\partial/\partial t)\hat{U}_t = \hat{H}\hat{U}_t\,, \tag{4.2}$$

where the *Hamiltonian operator* $\hat{H}$ is *Hermitian*, or *self-adjoint*: $\hat{H}^\dagger = \hat{H}$.

If the Hamiltonian is independent of time (autonomous), we can integrate (4.2) immediately, to obtain

$$\hat{U}_t = \exp(-\mathrm{i}\hbar^{-1}t\hat{H})\,, \tag{4.3}$$

where we define the exponential by its Taylor series. Conversely, we may define the Hamiltonian as

$$\hat{H} = \mathrm{i}\hbar \lim_{t\to 0}(\hat{U}_t - \hat{1})/t\,, \tag{4.4}$$

where $\hat{1}$ is the *identity operator*. Thus, any Hermitian operator can be considered as the "infinitesimal generator" of a possible evolution in Hilbert space, just as any real function in classical phase space is capable of generating a Hamiltonian flow.

The eigenvalues of Hermitian operators, being real, allow us to identify them with the results of ideal measurements. These quantum equivalents of classical variables are thus referred to as

observables. For an arbitrary state $|\psi\rangle$ the average value predicted for the measurement of the observable $\hat{A}$ is

$$\langle A\rangle = \langle\psi|\hat{A}|\psi\rangle\,. \tag{4.5}$$

Consider now the "evolution" of a classical variable $A(x)$ under the flow of the Hamiltonian $H(x)$. For any initial position x_0, we obtain $A_t(x_0) = A(x(x_0,t))$, so that A_t results from considering the classical states (i.e., phase space points) as fixed and attributing the change experienced, because of movement of the state, to A itself. If we now assume this same point of view for the quantum evolution (the Heisenberg "picture") we find that, attributing the change of $\langle A\rangle_t$ to $\hat{A}$ rather than to $|\psi\rangle$ in (4.5) implies that

$$\hat{A}_t = \hat{U}_t^{\dagger}\hat{A}_0\hat{U}_t\,. \tag{4.6}$$

Thus, to first order in t,

$$\hat{A}_t = (\hat{1} + (\mathrm{i}/\hbar)t\hat{H})\hat{A}_0(1 - (\mathrm{i}/\hbar)t\hat{H}) = \hat{A}_0 - (\mathrm{i}/\hbar)t[\hat{A}_0,\hat{H}]\,, \tag{4.7}$$

using the definition of the *commutator*,

$$[\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}\,. \tag{4.8}$$

So, the evolution of the observable $\hat{A}$ is governed by *Heisenberg's equation*,

$$(\mathrm{d}/\mathrm{d}t)\hat{A} = -(\mathrm{i}/\hbar)[\hat{A},\hat{H}]\,. \tag{4.9}$$

It is well known that (4.9) corresponds precisely to the classical equation $\dot{A} = \{A,H\}$, defined in terms of Poisson brackets [21,23]. In particular, if A is just a component of position or momentum, we recover one of Hamilton's equations.

Just as in classical mechanics, we are primarily interested in the system's position q_0, described as the eigenvalue of the operator $\hat{q}$ for the eigenstate $|q_0\rangle$:

$$\hat{q}|q_0\rangle = q_0|q_0\rangle\,, \tag{4.10}$$

and likewise the system's momentum:

$$\hat{p}|p_0\rangle = p_0|p_0\rangle\,. \tag{4.11}$$

The fundamental postulate is that these observables do not commute, indeed the components $\hat{q}_k$ and $\hat{p}_j$ satisfy

$$[\hat{q}_k,\hat{p}_j] = \mathrm{i}\hbar\delta_{kj}\hat{1}\,, \tag{4.12}$$

where $\hbar$ is Planck's constant.

The fact that the commutators (4.12) are a multiple of the identity operator (with the dimension of action) has far reaching consequences. If we construct the family of unitary operators

$$\hat{T}_q = \exp(-\mathrm{i}\hbar^{-1}q\cdot\hat{p})\,, \tag{4.13}$$

where the L parameters q_k have the dimension of distance in position space, we can transform any observable $\hat{A}\to\hat{A}' = \hat{A}(q)$ according to (4.6). Obviously, $\hat{T}_q$ leaves the observable $\hat{p}$ invariant, but,

if we take $\hat{A}=\hat{q}_j$ and associate any of the parameter components q_k with the time in (4.9), we obtain the derivative of $\hat{A}(q)$ as

$$(\partial/\partial q_k)\hat{q}_j=(-\mathrm{i}/\hbar)[\hat{q}_j,\hat{p}_k]=\delta_{kj}\hat{1}\,. \tag{4.14}$$

Integrating with respect to the parameter components q_k, we then obtain the action of $\hat{T}_{q_a}$ on $\hat{q}$, for a particular vector parameter q_a, to be

$$\hat{q}'_a=\hat{T}_{-q_a}\hat{q}\hat{T}_{q_a}=\hat{q}+q_a\hat{1}\,, \tag{4.15}$$

so that

$$\hat{q}'_a|q_b\rangle=(q_b+q_a)|q_b\rangle\,. \tag{4.16}$$

Evidently, we are using the "Heisenberg picture". In the "Schrödinger picture", the states are transformed rather than the operators; so, to obtain the new eigenvalue we must have

$$\hat{T}_{q_a}|q_b\rangle=|q_a+q_b\rangle\,. \tag{4.17}$$

In the same manner, we can form the operators

$$\hat{T}_p=\exp(\mathrm{i}\hbar^{-1}p\cdot\hat{q})\,, \tag{4.18}$$

obtaining from the fundamental commutator (4.12) that

$$\hat{p}'=\hat{T}_{-p_a}\hat{p}\hat{T}_{p_a}=\hat{p}+p_a\hat{1} \tag{4.19}$$

and

$$\hat{T}_{p_a}|p_b\rangle=|p_a+p_b\rangle\,. \tag{4.20}$$

We can now compose $\hat{T}_p$ with $\hat{T}_q$ to obtain operator equivalents of general translations in phase space. The fact that these two families of operators do not commute is not a major problem because of the fundamental theorem [21]: If both $\hat{A}$ and $\hat{B}$ commute with their commutator, then

$$\mathrm{e}^{\hat{A}+\hat{B}}=\mathrm{e}^{\hat{A}}\mathrm{e}^{\hat{B}}\mathrm{e}^{-(1/2)[\hat{A},\hat{B}]}\,. \tag{4.21}$$

Since the condition is satisfied by $\hat{p}$ and $\hat{q}$, we obtain the translation operator by $\xi=(p,q)$ as

$$\begin{aligned}\hat{T}_\xi&\equiv\exp((\mathrm{i}/\hbar)\xi\wedge\hat{x})\equiv\exp[(\mathrm{i}/\hbar)(p\cdot\hat{q}-q\cdot\hat{p})]\\&=\hat{T}_p\hat{T}_q\exp[-(\mathrm{i}/2\hbar)p\cdot q]=\hat{T}_q\hat{T}_p\exp[(\mathrm{i}/2\hbar)p\cdot q]\,,\end{aligned} \tag{4.22}$$

where naturally $\hat{x}=(\hat{p},\hat{q})$. In other words, the order of $\hat{T}_p$ and $\hat{T}_q$ affects only the overall phase of the product, allowing us to define the translation as above since unitary operators act in joint pairs to transform other operators in (4.6). However, we shall find that the detailed form of these phase factors plays an important role in the following theory. $\hat{T}_\xi$ is also known as a *Heisenberg operator*.

The group property is maintained within a phase factor:

$$\hat{T}_{\xi_2}\hat{T}_{\xi_1}=\hat{T}_{\xi_1+\xi_2}\exp[(-\mathrm{i}/2\hbar)\xi_1\wedge\xi_2]=\hat{T}_{\xi_1+\xi_2}\exp[(-\mathrm{i}/\hbar)D_3(\xi_1,\xi_2)]\,, \tag{4.23}$$

where D_3 is the symplectic area of the triangle determined by two of its sides. Evidently, the inverse of the unitary operator $\hat{T}_\xi^{-1} = \hat{T}_\xi^\dagger = \hat{T}_{-\xi}$ and we can generalize (4.23):

$$\hat{T}_{\xi_n} \cdots \hat{T}_{\xi_2} \hat{T}_{\xi_1} = \hat{T}_{\xi_1+\cdots+\xi_n} \exp[(-\mathrm{i}/\hbar) D_{n+1}(\xi_1, \ldots, \xi_n)] , \tag{4.24}$$

where D_{n+1} is the symplectic area of the resulting $(n+1)$-sided polygon defined in Appendix A. Thus, it can be said that the phase stores the memory of the sequence of translations of the Heisenberg group. This is known as a *cocycle* [22] in the theory of representations of groups.

Let us now define a new family of linear operators as proportional to the Fourier transform of the family of translations $\hat{T}_\xi$:

$$\hat{R}_x \equiv (4\pi\hbar)^{-L} \int \mathrm{d}\xi \exp[(\mathrm{i}/\hbar) x \wedge \xi] \hat{T}_\xi = (4\pi\hbar)^{-L} \int \mathrm{d}\xi \exp[(\mathrm{i}/\hbar) \xi \wedge (\hat{x} - x)] . \tag{4.25}$$

Evidently, this relation can be inverted, such that

$$\begin{aligned} (\pi\hbar)^{-L} \int \mathrm{d}x \exp[-(\mathrm{i}/\hbar) x \wedge \xi] \hat{R}_x &= (2\pi\hbar)^{-2L} \int \mathrm{d}x \, \mathrm{d}\xi' \exp[-(\mathrm{i}/\hbar) x \wedge (\xi' - \xi)] \hat{T}_{\xi'} \\ &= \int \mathrm{d}\xi' \ \delta(\xi - \xi') \hat{T}_{\xi'} = \hat{T}_\xi . \end{aligned} \tag{4.26}$$

Comparing the last integrand in (4.25) with the middle integrand in (4.26), we see that, in a way, the $\hat{R}_x$ corresponds to an operator version of the Dirac δ-function. However, the following investigation of the properties of this family of operators through their combinations with $\hat{T}_\xi$ will reveal their true nature. First we verify that

$$\begin{aligned} &\hat{R}_x \hat{T}_\xi = (4\pi\hbar)^{-L} \int \mathrm{d}\xi' \exp[(\mathrm{i}/\hbar) x \wedge \xi'] \hat{T}_{\xi'} \hat{T}_\xi = (4\pi\hbar)^{-L} \int \mathrm{d}\xi' \exp\left[\frac{\mathrm{i}}{\hbar}\left(x - \frac{\xi}{2}\right) \wedge \xi'\right] \hat{T}_{\xi+\xi'} \\ &(4\pi\hbar)^{-L} \int \mathrm{d}\eta \exp\left[\frac{\mathrm{i}}{\hbar}\left(x - \frac{\xi}{2}\right) \wedge (\eta - \xi)\right] \hat{T}_\eta = \exp[-(\mathrm{i}/\hbar) x \wedge \xi] \hat{R}_{x-\xi/2} , \end{aligned} \tag{4.27}$$

whereas, similarly

$$\hat{T}_\xi \hat{R}_x = \exp[-(\mathrm{i}/\hbar) x \wedge \xi] \hat{R}_{x+\xi/2} . \tag{4.28}$$

These products may now be used to simplify the evaluation of

$$\begin{aligned} \hat{R}_{x_2} \hat{R}_{x_1} &= (4\pi\hbar)^{-L} \int \mathrm{d}\xi \exp[(\mathrm{i}/\hbar) x_2 \wedge \xi] \hat{T}_\xi \hat{R}_{x_1} = (4\pi\hbar)^{-L} \int \mathrm{d}\xi \exp[(\mathrm{i}/\hbar)(x_2 - x_1) \wedge \xi] \hat{R}_{x_1+\xi/2} \\ &= (\pi\hbar)^{-L} \int \mathrm{d}\eta \exp[(\mathrm{i}/\hbar) 2(x_2 - x_1) \wedge (\eta - x_1)] \hat{R}_\eta = \exp[(\mathrm{i}/\hbar) x_1 \wedge x_2] \hat{T}_{2(x_2-x_1)} . \end{aligned} \tag{4.29}$$

Comparing (4.24) and (4.27)–(4.29) with (1.39)–(1.42) we verify that the operators $\hat{R}_x$ and $\hat{T}_x$ form a group that is analogous to the classical reflections and translations in phase space. From (4.29) we see that

$$\hat{R}_x^2 = \hat{1} , \tag{4.30}$$

so $\hat{R}_x$ definitely does not correspond to a δ-function operator. Also, since it is obvious that $\hat{R}_x^\dagger = \hat{R}_x$, we find that $\hat{R}_x$ has been defined as an unitary operator. Therefore, the full group of $\hat{R}_x$ and $\hat{T}_\xi$ is an unitary quantal representation of the classical group of translations and reflections [57, 58].

These properties will be of fundamental importance when we come to study the Weyl representation. But now we will start with the more familiar representations of quantum mechanics. The eigenfunctions of position and momentum form a complete set, so we can expand

$$|\psi\rangle = \int \mathrm{d}q|q\rangle\langle q|\psi\rangle = \int \mathrm{d}p|p\rangle\langle p|\psi\rangle\,, \tag{4.31}$$

where $|\langle q|\psi\rangle|^2$ and $|\langle p|\psi\rangle|^2$ are, respectively, the probability density that the state $|\psi\rangle$ has the position q and the momentum p. Evidently,

$$\langle q'|q\rangle = \delta(q-q') \quad \text{and} \quad \langle p'|p\rangle = \delta(p-p')\,, \tag{4.32}$$

as a consequence of (4.31), whereas it is well known that

$$\langle p|q\rangle = (2\pi\hbar)^{-L/2}\,\exp[-(\mathrm{i}/\hbar)p\cdot q]\,. \tag{4.33}$$

Therefore, the state $|q_0\rangle$ has a specific position q_0, but an uniform probability density for its momentum, while $|p_0\rangle$ specifies the momentum rather than the position.

These representations depart radically from the classical concept of a state as being represented by a point in phase space where both the momentum and the position are precisely determined. It is therefore natural to seek representations in terms of states such that both the position and the momentum are as narrowly determined as possible. Recalling the definition of the mean value (4.5) of the observable $\hat{A}$ for the state $|\psi\rangle$, we relate the uncertainty in A to the *variance* $(\Delta A)^2 = \langle(\hat{A} - \langle A\rangle)^2\rangle$. Then, following standard manipulations [21, 23], one obtains the *uncertainty principle*:

$$(\Delta p)^2(\Delta q)^2 \geq \tfrac{1}{4}\hbar^2\,. \tag{4.34}$$

The states of minimum uncertainty, for which the equality holds in (4.34), are known as *coherent states*. Most of their basic properties (reviewed in Refs. [24, 25] and in complement G_v of the standard reference [23]) are derived from the fact that coherent states are eigenstates of

$$\hat{a} = (2\hbar)^{-1/2}(w^{1/2}\hat{q} + \mathrm{i}w^{-1/2}\hat{p})\,, \tag{4.35}$$

the *lowering operator* for the harmonic oscillator:

$$\hat{H} = \tfrac{1}{2}(\hat{p}^2 + w^2\hat{q}^2) = w\hbar(w\hat{q}^2 + w^{-1}\hat{p}^2)/2\hbar = w\hbar[\hat{a}^+\hat{a} + \tfrac{1}{2}]\,, \tag{4.36}$$

where the adjoint,

$$\hat{a}^\dagger = (2\hbar)^{-1/2}(w^{1/2}\hat{q} - \mathrm{i}w^{-1/2}\hat{p})\,, \tag{4.37}$$

is known as *the raising operator*. Indeed, recalling that the eigenstates of $\hat{H}$ have a regularly spaced spectrum,

$$\hat{H}|n\rangle = w\hbar(n+\tfrac{1}{2})|n\rangle\,, \tag{4.38}$$

it is shown [21, 23] that

$$\hat{a}^\dagger|n\rangle = (n+1)^{1/2}|n+1\rangle \quad \text{and} \quad \hat{a}|n\rangle = n^{1/2}|n-1\rangle\,. \tag{4.39}$$

(Note that, for simplicity, we are now restricting consideration to the case $L=1$.)

The eigenvalues of the lowering operator (4.35) need not be real, since it is not self-adjoint. Let us solve the eigenvalue equation

$$\hat{a}|z\rangle = z|z\rangle\,, \tag{4.40}$$

where we will express the eigenvalue z in terms of two real parameters, P and Q:

$$z = (2\hbar)^{-1/2}(w^{1/2}Q + \mathrm{i}w^{-1/2}P) \tag{4.41}$$

in a natural way. In the position representation $\hat{p} = -\mathrm{i}\hbar\partial/\partial q$, so that (4.40) becomes the differential equation

$$(w^{1/2}q + \hbar w^{-1/2}\partial/\partial q)\langle q|z\rangle = z\sqrt{2\hbar}\langle q|z\rangle\,, \tag{4.42}$$

with the normalized solution

$$\langle q|z\rangle = (w/\pi\hbar)^{1/4}\exp[-(w/2\hbar)(q-Q)^2 + (\mathrm{i}/\hbar)Pq]\,. \tag{4.43}$$

Since the Fourier transform of a Gaussian has the same form, we obtain the momentum representation of the coherent state as

$$\langle p|z\rangle = (1/\pi\hbar w)^{1/4}\exp[-(1/2\hbar w)(p-P)^2 - (\mathrm{i}/\hbar)Qp]\,. \tag{4.44}$$

It is evident from the form of (4.43) and (4.44) that the average values satisfy $\langle q\rangle_z = Q$ and $\langle p\rangle_z = P$. Furthermore, we can verify that the product $\Delta p\Delta q$ is independent of w and of the eigenvalue z, providing the equality in (4.34). In the limit $w \to \infty$, except for the need to redefine the normalization, we obtain $\langle q|z\rangle \to \delta(q-Q)$, i.e. $|z\rangle$ becomes a position eigenstate. Alternatively, when $w \to 0$, $\langle q|z\rangle \to \exp[-(\mathrm{i}/\hbar)Pq]$, which is a momentum eigenstate. By choosing $w \sim 1$, we localize both the momentum and the position within an $\hbar$-area of the expected value (P,Q).

The coherent states would thus form an ideal "phase-space basis", if it were not for their overcompleteness. There are many ways of showing this and of determining a complete subset of coherent states. Let us just note that in the limit $w \to \infty$, the subset of states with eigenvalue $P = 0$ becomes the set $|q\rangle$, which is complete. Indeed, any function $P(q)$ specifies a complete subset of these "extreme" coherent states. All the same, we shall verify that we still obtain useful expressions by treating the decompositions into the full set of coherent states as if they were merely complete.

Another natural representation for the coherent states is the harmonic oscillator basis $|n\rangle$ defined by (4.38) (with the same choice of frequency w). Expanding

$$|z\rangle = \sum_n c_n(z)|n\rangle\,, \tag{4.45}$$

we obtain

$$\hat{a}|z\rangle = \sum_n c_n(z)\sqrt{n}|n-1\rangle = \sum_n zc_n(z)|n\rangle\,, \tag{4.46}$$

using (4.39) and (4.40). Thus, normalizing the resulting recurrence relation, we obtain

$$|z\rangle = \mathrm{e}^{-|z|^2/2}\sum_n \frac{z^n}{\sqrt{n!}}|n\rangle\,. \tag{4.47}$$

We can now use this decomposition to represent an arbitrary state $|\psi\rangle$ in the coherent state representation. Indeed, the expansion coefficient is

$$\langle z|\psi\rangle = \mathrm{e}^{-|z|^2/2}F(z^*)\,, \tag{4.48}$$

where

$$F(z)=\sum_{n=0}^{\infty}\frac{z^n}{(n!)^{1/2}}\langle n|\psi\rangle \tag{4.49}$$

is an entire function, since the sum $\langle n|\psi\rangle$ is normalized to unity. It is then easy to show that the scalar product can be decomposed in the usual form implied by the Dirac notation,

$$\langle\phi|\psi\rangle=(1/\pi)\int \mathrm{d}z\langle\phi|z\rangle\langle z|\psi\rangle=\int \mathrm{d}P\mathrm{d}Q/2\pi\hbar\langle\phi|z\rangle\langle z|\psi\rangle\,, \tag{4.50}$$

even though this representation is overcomplete. It follows that

$$\langle z|\hat{A}|\psi\rangle=(1/\pi)\int \mathrm{d}z'\langle z|\hat{A}|z'\rangle\langle z'|\psi\rangle\,, \tag{4.51}$$

where $\langle z|\hat{A}|z'\rangle$ is the coherent state representation of the operator $\hat{A}$, and we can use the matrix product rule

$$\langle z|\hat{A}_2\hat{A}_1|z'\rangle=(1/\pi)\int \mathrm{d}z''\langle z|\hat{A}_2|z''\rangle\langle z''|\hat{A}_1|z'\rangle\,. \tag{4.52}$$

Alternatively, we can employ the analyticity of the coherent state representation to resolve its overcompleteness. It is known that if an entire function of z and z' is zero along $z'=z^*$, then it cancels for all z and z' [26]. Since

$$\mathrm{e}^{(1/2)|z|^2+(1/2)|z'|^2}\langle z|\hat{A}|z'\rangle=\sum_n\frac{|z^*|^n|z'|^m}{(n!m!)^{1/2}}\langle n|\hat{A}|m\rangle \tag{4.53}$$

is an entire function of z^* and z', then it is uniquely determined by the diagonal elements $\langle z|\hat{A}|z\rangle$. The diagonal coherent state representation is known as the *Husimi representation* [4, 5] usually considered as a function of the real phase space variables (P,Q), rather than of the complex number z:

$$A_{\mathrm{H}}(P,Q)=\left\langle\frac{w^{1/2}Q+\mathrm{i}w^{-1/2}P}{(2\hbar)^{1/2}}|\hat{A}|\frac{w^{1/2}Q+\mathrm{i}w^{-1/2}P}{(2\hbar)^{1/2}}\right\rangle\,. \tag{4.54}$$

The essential analytic part of the Husimi representation is known as the *Bargmann representation* [54, 55]. However, the present emphasis on the real phase-space variables justifies the former term.

We obtain a more intuitive grasp of the Husimi representation by rewriting

$$A_{\mathrm{H}}(P,Q)=\mathrm{Tr}(|z\rangle\langle z|\hat{A})\,. \tag{4.55}$$

The projection operator $|z\rangle\langle z|$ onto the coherent state $|z\rangle$ is identified with the density operator ρ_z for the pure state $|z\rangle$. Recalling the definition of averages for operators, given the density ρ [21,23], we find that $A_{\mathrm{H}}(P,Q)=\langle\hat{A}\rangle_z$, the average value of $\hat{A}$ for a system that is in the coherent state $|z\rangle$.

As well as considering $|z\rangle\langle z|$ as a density operator, it is important to study the Husimi representation of an arbitrary pure state density operator $\rho = |\psi\rangle\langle\psi|$:

$$\rho_{\mathrm{H}}(P,Q) = \left| \left\langle \frac{w^{1/2}Q + \mathrm{i}w^{-1/2}P}{(2\hbar)^{1/2}} \middle| \psi \right\rangle \right|^2 . \tag{4.56}$$

We see immediately that this *Husimi function* represents the density operator as a real non-negative function in phase space.

As an example, let us consider the Husimi function for the eigenstates of the harmonic oscillator. Since the coherent state itself can be decomposed into eigenstates of the harmonic oscillator (4.47), we obtain

$$\langle z|m\rangle = \mathrm{e}^{-|z|^2/2} z^{*m}/\sqrt{m!} . \tag{4.57}$$

Therefore, the Husimi function will be

$$|\langle z|m\rangle|^2 = (1/2\hbar m!)(wQ^2 + P^2/w)^{2m} \exp[-(1/2\hbar)(wQ^2 + P^2/w)] , \tag{4.58}$$

which is constant along the elliptical level curves of the classical Hamiltonian. It should be recalled that $(P,Q) = (\langle p\rangle, \langle q\rangle)$, i.e. the mean values for the coherent state $|z\rangle$.

It is important to note that the simplicity of (4.58) depends on the choice of the free parameter w' for the coherent state as equal to w, the frequency of the harmonic oscillator. If we chose $w' \gg w$, the Husimi function would have oscillations resembling those of $|\langle q|m\rangle|^2$ instead of being smooth. In the case where $w = 1$, the elliptical level curves of (4.58) become circles of radius $|z|$. The maximum intensity of the Husimi function is found by taking

$$(\mathrm{d}/\mathrm{d}|z|^2)|\langle z|m\rangle|^2 = [-1 + m/|z|^2]|\langle z|m\rangle|^2 = 0 , \tag{4.59}$$

i.e. at $|z|^2 = m$. Recalling that the classical Hamiltonian is just $H_{\mathrm{c}}(p,q) = \frac{1}{2}(p^2 + q^2) = \hbar|z|^2$, we find that the maximum intensity of the Husimi function is obtained at

$$H_{\mathrm{c}}(p,q) = m\hbar , \tag{4.60}$$

which is just inside the quantized circle $\frac{1}{2}(p^2 + q^2) = (m + \frac{1}{2})\hbar$, obtained as the classical orbit with the energy of the mth eigenstate.

This example is not untypical of the marvellous way that the Husimi function highlights the classical structure underlying a quantum state. We see, however, that this depends on an appropriate choice of the free parameter. The fact that the phase space is complexified is not a major problem. Indeed, the classical theory of normal forms around stable equilibria [15, 19] relies on an entirely analogous procedure.

Retrieval of the phase of the coherent state representation of a given state from its (Husimi) modulus involves analytic continuation. This can be neatly obtained from the knowledge of the zeros of the Husimi function as a consequence of Hadamard's factorization theorem [27]. Thus, it is true to say that the zeros of the Husimi function encapsule a minimal determination of a quantum state. Leboeuf and Voros [28, 29] have verified that the nature of the distribution of zeros in the phase plane is correlated to the chaotic or regular character of the corrresponding classical motion. The remarkable feature is that these zeros do not generally lie in the region of classical motion. We

shall see in the next section that the previous simple example is quite typical, in that the Husimi distribution has a smooth maximum on the classically allowed region, thus expelling the zeros from this neighbourhood.

The formulae beyond Eq. (4.34) in this section have been restricted to a single degree of freedom so as not to encumber the notation. Thus, for instance, the representation of a coherent state in the harmonic oscillator basis (4.47) generalizes to

$$|z_1 \cdots z_L\rangle = \exp\left\{-\frac{1}{2}[|z_1|^2 + \cdots + |z_L|]^2\right\} \sum_{n_1 \cdots n_L} \frac{z_1^{n_1} \cdots z_1^{n_L}}{\sqrt{n_1! \cdots n_L!}} |n_1 \cdots n_L\rangle \,. \tag{4.61}$$

Here, each of the oscillators is free to have a different frequency: $w_1 \cdots w_L$, so that the position representation will be a $(2L)$-dimensional Gaussian with L different widths. The only difficulty lies in the analytic continuation required to obtain the phase of the coherent state representation from the Husimi function, involving the analytic continuation for several complex variables. In particular, the zeroes of entire functions will not be restricted to points in phase space.

5. Centres and chords in quantum mechanics

By expressing the Husimi symbol of an operator as the trace (4.55) of its product with another operator, we attain a different level of representation in quantum mechanics. Indeed, the trace of an operator is invariant with respect to any unitary transformation of the basis with which we represent it, so we may omit this basis from consideration. We shall now develop this theme, while substituting different choices of unitary operators $\hat{U}$ for the projection operator $|z\rangle\langle z|$ onto a coherent state. First, let us note that $\mathrm{Tr}\,\hat{U}$ corresponds intuitively to a measure of the fixed points of a classical mapping, since $|\langle\psi|\hat{U}|\psi\rangle|^2$ is the probability that the *quantum map*,

$$|\psi'\rangle = \hat{U}|\psi\rangle \,, \tag{5.1}$$

brings the states $|\psi\rangle$ back onto itself.

Let us examine specifically the translation and reflection (unitary) operators defined in the previous section. Starting with the translations,

$$\begin{aligned}\mathrm{Tr}\,\hat{T}_\xi &= \int \langle q|T_\xi|q\rangle \,\mathrm{d}q = \int \mathrm{d}q \exp\left[\frac{1}{2\hbar}\xi_p \cdot \xi_q\right] \langle q|\hat{T}_{\xi_q}\hat{T}_{\xi_p}|q\rangle \\ &= \exp\left(\frac{\mathrm{i}}{2\hbar}\xi_p \cdot \xi_q\right) \int \mathrm{d}q \exp\left[\frac{\mathrm{i}}{\hbar}\xi_p \cdot q\right] \langle q|q + \xi_q\rangle = (2\pi\hbar)^L \delta(\xi_p)\delta(\xi_q) \\ &= (2\pi\hbar)^L \delta(\xi) \,,\end{aligned} \tag{5.2}$$

and then taking the Fourier transform,

$$\mathrm{Tr}\,\hat{\mathbb{R}}_x = \mathrm{Tr}\, 2^L \hat{R}_x = (2\pi\hbar)^{-L} \int \mathrm{d}\xi \exp\left[(\mathrm{i}/\hbar)x \wedge \xi\right] \mathrm{Tr}\,\hat{T}_\xi = 1 \,, \tag{5.3}$$

where, it is now also convenient to define the exact Fourier transform $\hat{\mathbb{R}}_x$ of $\hat{T}_\xi$. We recall that the classical transformation R_x has a single fixed point (x itself), whereas T_ξ has fixed points only if

$\xi = 0$, when all points are fixed. These results are in general agreement with our intuition as to the classical correspondence of the traces of unitary operators.

We can now consider the possibility of expressing any operator $\hat{A}$ as a linear superposition of elementary translation operators:

$$\hat{A} = \int (\mathrm{d}\xi/(2\pi\hbar)^L) A(\xi)\hat{T}_\xi \,. \tag{5.4}$$

The confirmation results from

$$\mathrm{Tr}(\hat{T}_{-\xi}\hat{A}) = \mathrm{Tr}\int \frac{\mathrm{d}\xi'}{(2\pi\hbar)^L}\, A(\xi')\hat{T}_{-\xi}\hat{T}_{\xi'} = \int \frac{\mathrm{d}\xi'}{(2\pi\hbar)^L} A(\xi') \exp\left[\frac{\mathrm{i}}{2\hbar}\xi' \wedge \xi\right] \mathrm{Tr}\, \hat{T}_{\xi'-\xi} = A(\xi)\,, \tag{5.5}$$

which furnishes explicitly the expansion coefficient. Comparison with (1.38) reveals that, when $\hat{A}$ is an unitary operator, equation (5.5) represents the unitary map, in close analogy to the way that the chord generating function represents a canonical transformation. In the latter case, we are given the chord ξ and determine the missing centre, x, by finding the fixed point of the transformation combined with $\hat{T}_{-\xi}$. In quantum mechanics we now take the trace of the combined operators to dermine the *chord representation* $A(\xi)$ that uniquely specifies $\hat{A}$ according to (5.4). Of course, we cannot determine the centre x simultaneously with ξ because of the uncertainty principle. However, we can equally represent any operator $\hat{A}$ as a superposition of reflections:

$$\hat{A} = \int \frac{\mathrm{d}x}{(2\pi\hbar)^L} A(x)\hat{\mathbb{R}}_x = \int \frac{\mathrm{d}x}{(\pi\hbar)^L} A(x)\hat{R}_x \,. \tag{5.6}$$

Again we obtain the expansion coefficient by calculating

$$\begin{aligned}\mathrm{Tr}(\hat{\mathbb{R}}_x\hat{A}) &= \mathrm{Tr}\int \frac{\mathrm{d}x'}{(2\pi\hbar)^L} A(x')\hat{\mathbb{R}}_x\hat{\mathbb{R}}_{x'} = \int \frac{\mathrm{d}x'}{(2\pi\hbar)^L} A(x')2^{2L} \exp\left[\frac{\mathrm{i}2}{\hbar}x' \wedge x\right] \mathrm{Tr}\, \hat{T}_{2(x-x')} \\ &= \int \frac{\mathrm{d}x'}{(2\pi\hbar)^L} A(x')2^{2L} \exp\left[\frac{\mathrm{i}2}{\hbar}x' \wedge x\right] (2\pi\hbar)^L \frac{\delta(x-x')}{2^{2L}} = A(x)\,.\end{aligned} \tag{5.7}$$

Notice that comparison of (5.4) and (5.6) with (4.25) and (4.26) yields

$$R_x(\xi) = 2^{-L} \exp\left[(\mathrm{i}/\hbar)x \wedge \xi\right] \quad \text{and} \quad T_\xi(x) = \exp\left[-(\mathrm{i}/\hbar)x \wedge \xi\right] \,.$$

In analogy with our previous result, we may refer to $A(x)$ as the *centre representation* of the operator $\hat{A}$, but the historic term is the *Weyl representation*. If $\hat{A}$ is an unitary operator, we are now specifying the centre x, rather than the chord (now unknown), and we determine the Weyl transform (5.7) in close analogy to (1.37). Evidently, the origin of the analogy between (1.37) and (1.38) with (5.7) and (5.5) is the achievement of defining the fundamental group of translations and reflections in both classical and quantum mechanics.

By defining different orderings for the operator $\hat{T}_p$ and $\hat{T}_q$ in (4.22), we arrive at alternative representations that are discussed in the review by Balazs and Jennings [30], on which this section is largely based. It is also important to note that we are not able to express an arbitrary operator as a superposition of $|z\rangle\langle z|$, so as to invert the Husimi representation in the manner of $\hat{T}_\xi$ or $\hat{\mathbb{R}}_x$, since $\mathrm{Tr}(|z\rangle\langle z|z'\rangle\langle z'|) = |\langle z|z'\rangle|^2$ is not a δ-function.

Our ability to exchange chord and centre generating functions by means of Legendre transforms in classical mechanics is exactly matched by the Fourier transform between the chord and the centre representations. Inserting (4.25) into (5.6),

$$\hat{A} = (4\pi\hbar)^{-L} \int \frac{\mathrm{d}x}{(\pi\hbar)^L} \mathrm{d}\xi A(x) \exp\left[\frac{\mathrm{i}}{\hbar} x \wedge \xi\right] \hat{T}_\xi$$
$$= \int \frac{\mathrm{d}\xi}{(2\pi\hbar)^L} \left\{ (2\pi\hbar)^{-L} \int \mathrm{d}x\, A(x) \exp\left[\frac{\mathrm{i}}{\hbar} x \wedge \xi\right] \right\} \hat{T}_\xi , \tag{5.8}$$

so that comparison with (5.4) reveals that

$$A(\xi) = (2\pi\hbar)^{-L} \int \mathrm{d}x\, A(x) \exp\left[(\mathrm{i}/\hbar) x \wedge \xi\right] . \tag{5.9}$$

An analogous treatment of (4.26) and (5.4) compared to (5.6) reveals that

$$A(x) = (2\pi\hbar)^{-L} \int \mathrm{d}\xi\, A(\xi) \exp\left[-(\mathrm{i}/\hbar) x \wedge \xi\right] . \tag{5.10}$$

Consider now the coordinate representation of the operator $\hat{A}$. Using (5.4), we have

$$\langle q_+ | \hat{A} | q_- \rangle = \int \frac{\mathrm{d}\xi}{(2\pi\hbar)^L} A(\xi) \langle q_+ | \hat{T}_\xi | q_- \rangle$$
$$= \int \frac{\mathrm{d}\xi}{(2\pi\hbar)^L} A(\xi) \delta(q_+ - q_- - \xi_q) \exp\left[\frac{\mathrm{i}}{\hbar}\, \xi_p \cdot \left(q_- + \frac{\xi_q}{2}\right)\right]$$
$$= \int \frac{\mathrm{d}\xi_p}{(2\pi\hbar)^L} A(\xi_p, q_+ - q_-) \exp\left[\frac{\mathrm{i}}{\hbar}\, \xi_p \cdot \frac{q_- + q_+}{2}\right] , \tag{5.11}$$

which is just a (symmetrized) Fourier transform, whose inverse is

$$A(\xi) = \int \mathrm{d}\bar{q} \langle \bar{q} + \tfrac{1}{2}\xi_q | \hat{A} | \bar{q} - \tfrac{1}{2}\xi_q \rangle \exp[-(\mathrm{i}/\hbar)\, \xi_p \cdot \bar{q}] . \tag{5.12}$$

If we now take the full Fourier transform of $A(\xi)$, we also obtain the Weyl representation of $\hat{A}$ as a Fourier transform of $\langle q_+ | \hat{A} | q_- \rangle$, i.e.

$$A(x) = \int \mathrm{d}\xi_q \langle q + \tfrac{1}{2}\xi_q | \hat{A} | q - \tfrac{1}{2}\xi_q \rangle \exp[-(\mathrm{i}/\hbar) p \cdot \xi_q] , \tag{5.13}$$

with its inverse

$$\langle q_+ | \hat{A} | q_- \rangle = \int \frac{\mathrm{d}p}{(2\pi\hbar)^L} A\left(p, \frac{q_+ + q_-}{2}\right) \exp\left[\frac{\mathrm{i}}{\hbar}\, p \cdot (q_+ - q_-)\right] . \tag{5.14}$$

Comparison of these last equations with (1.49)–(1.55) reveals that the Legendre transforms of classical mechanics between generating functions are isomorphic to the corresponding Fourier transforms in quantum mechanics.

The reciprocity of the relations (5.4) with (5.5) and of (5.6) with (5.7) would have allowed for an arbitrary real factor in the definition of the centre and the chord representations. As well as agreeing with various previous definitions of the Weyl transformation, the choice adopted guarantees that

$$\mathrm{Tr}\,\hat{A} = A(\xi = 0) = (2\pi\hbar)^{-L} \int \mathrm{d}x A(x)\,. \tag{5.15}$$

Also, we obtain a pleasing simplicity in the representation of the identity operator $\hat{1}$:

$$1(\xi) = (2\pi\hbar)^{L}\delta(\xi) \quad \text{and} \quad 1(x) = 1\,, \tag{5.16}$$

within a phase factor. Substitution of these expressions in the previous formulae furnishes a good check for stray constants.

The advantage of the centre representation lies in its ability to deal with motion for short intervals of time, just as we found in classical mechanics. The corresponding operator will be very close to the identity, hence it will be smooth and close to unity in the centre representation, whereas its chord representation will be a sharply peaked function at the origin. This short time motion is generated by Hermitian operators (4.4). Since the respective representations of the adjoint operator, $\hat{A}^{\dagger}$, are

$$A^{\dagger}(\xi) = [A(-\xi)]^{*} \quad \text{and} \quad A^{\dagger}(x) = [A(x)]^{*}\,, \tag{5.17}$$

we see that the centre representation of a Hermitian operator is always a real function, while the restriction on its chord representation is not so satisfying.

We find immediately from (5.13) that any function $f(\hat{q})$ has $f(q)$ as its Weyl transform, whereas the Fourier transform of (5.13),

$$A(x) = \int \mathrm{d}\xi_p \, \langle p + \tfrac{1}{2}\xi_q | \hat{A} | p - \tfrac{1}{2}\xi_p \rangle \exp[(\mathrm{i}/\hbar) q \cdot \xi_q]\,, \tag{5.18}$$

shows that the centre representation of $f(\hat{p})$ is $f(p)$. The linearity of the Weyl transform then implies that the Weyl representation of the important class of Hamiltonians

$$\hat{H} \equiv (1/2m)\hat{p}^2 + V(\hat{q}) \underset{\text{Weyl}}{\longrightarrow} H(x) = (1/2m)\ p^2 + V(q)\,, \tag{5.19}$$

i.e. the Weyl representation equals the classical Hamiltonian $H_c(x)$. This is not true for general Hamiltonians or Hermitian operators, but the Weyl representation will be a smooth real function close to the classical Hamiltonian and tending to it in the limit $\hbar \to 0$. The chord representation of these Hermitian operators, will therefore be close to the Fourier transform of the corresponding classical functions.

The advantage of the centre representation for Hermitian operators has obscured its basic reciprocity with the chord basis. It seems mysterious to find a representation of quantum mechanics where observables are represented by real functions that are at least close to the corresponding classical variables. The situation becomes clearer if we focus on unitary operators which transform states. These correspond classically to transformations in double phase space, as presented in Section 1. The generating functions of these canonical transformations are defined in terms of only half the variables of double phase space, but the other half can be obtained explicitly as derivatives. The corresponding representations of the unitary transformations in quantum mechanics are also determined by half the variables in double phase phase, but we cannot obtain the other half explicitly

because of the uncertainty principle. If we now consider the observables as generators of unitary transformations, we find that, though the centre representation is just one of many options related by Fourier transforms, it is remarkably appropriate.

Even though we can migrate among the several representations of operators via Fourier transforms, there is an important distinction concerning the centre and the chord representations, namely that they are not directly concerned with the states. We may consider that the position, the momentum and even the coherent state representations represent operators as a consequence of the way they represent the states. Indeed, we may choose the Schrödinger picture with static operators and moving states, or the Heisenberg picture that reverses the attribute of mobility. The centre and the chord representations concern only the operators and are therefore bound to the Heisenberg picture.

Only the density operators, $\hat{\rho}$, will represent states in this view, be they pure or statistical mixtures. Historically, the centre representation of the density operator

$$W(x) \equiv (2\pi\hbar)^{-L}\rho(x) \tag{5.20}$$

is named the *Wigner function*. In the case of a pure state, $\hat{\rho} = |\psi\rangle\langle\psi|$, we see from (5.13) that the Wigner function can be considered as the Fourier transform of the spatial correlation of the wave function [1]:

$$W(x) = (1/2\pi\hbar)^{L} \int \mathrm{d}q' \langle q + \tfrac{1}{2}q'|\psi\rangle\langle\psi|q - \tfrac{1}{2}q'\rangle \exp[-(\mathrm{i}/\hbar)p \cdot q'] . \tag{5.21}$$

Projecting this back onto the q-plane, we find that

$$\int \mathrm{d}p\, W(p,q) = |\langle q|\psi\rangle|^2 , \tag{5.22}$$

whereas a similar treatment of the momentum representation reveals that

$$\int \mathrm{d}q\, W(p,q) = |\langle p|\psi\rangle|^2 . \tag{5.23}$$

Integrating again either (5.22) or (5.23), we find that the Wigner function of a normalized state satisfies

$$\int \mathrm{d}p\, \mathrm{d}q W(p,q) = 1 = \mathrm{Tr}\,\hat{\rho} . \tag{5.24}$$

We therefore find that the Wigner function represents the state $|\psi\rangle$ as a real function in phase space; even though $W(x)$ may be somewhere negative, its projectors onto the position and the momentum planes correspond respectively to the correct probability densities.

Unlike the Weyl transforms of observables that are smooth functions in phase space, we expect a function that projects into correct probability densitites to be sharply peaked in the regions of classical motion, as we found with the Husimi function. Let us consider some examples. The first is a travelling wave in the interval $(-l/2, l/2)$, with periodic boundary conditions. The wave function is

$$\langle q|\psi\rangle = (l)^{-1/2} \exp[(\mathrm{i}/\hbar)p_n q] , \quad p_n = 2\pi n\hbar/l . \tag{5.25}$$

Inserting the corresponding density matrix into (5.13), we obtain

$$W(p,q) = \frac{1}{\pi l} \frac{\sin[l\hbar^{-1}(p - p_n)]}{p - p_n} \xrightarrow[\hbar\to 0]{} \frac{1}{l}\, \delta(p - p_n) . \tag{5.26}$$

The Wigner function is independent of position and it is sharply peaked around the classical motion as was found for the Husimi function.

Replacing the boundary condition by hard walls at $\pm\frac{1}{2}l$ leads to the wave function

$$\langle q|\psi\rangle = (2/l)^{1/2}\cos[(1/\hbar)p_n q]\,. \tag{5.27}$$

Care must now be taken to cancel (5.27) whenever one of the two wave functions lies outside the walls, i.e. if $q>0$, then the limit of integration is $l/2-q$. Therefore,

$$\begin{aligned} W(p,q) &= \frac{1}{\pi\hbar l}\int_{-l/2+q}^{l/2-q} \mathrm{d}q \cos\left[\frac{p_n}{\hbar}\left(q+\frac{q'}{2}\right)\right]\cos\left[\frac{p_n}{\hbar}\left(q-\frac{q'}{2}\right)\right]\exp\left[-\frac{\mathrm{i}p}{\hbar}\,q'\right] \\ &= \frac{1}{2\pi l}\left\{\frac{\sin[\hbar^{-1}2(l/2-q)(p-p_n)}{(p-p_n)} + \frac{\sin[\hbar^{-1}2(l/2-q)(p+p_n)]}{p+p_n}\right. \\ &\quad \left. + \frac{2\cos(\hbar^{-1}p_n q)\sin[\hbar^{-1}2(l/2-q)p]}{p}\right\}. \end{aligned} \tag{5.28}$$

In the limit $\hbar\to 0$, the first two terms condense onto the classical manifold $p=\pm p_n$, in the same way as for the travelling wave. But now there is also a nonclassical term. This interference term oscillates *along* the q-axis itself and it is essential for the correct projection of $|\langle q|\psi\rangle|^2$ onto the q-axis:

$$\int \mathrm{d}p\, W(p,q) = l^{-1}[1+\cos(2\hbar^{-1}p_n q)] = 2l^{-1}\cos^2(\hbar^{-1}p_n q)\,. \tag{5.29}$$

In the classical limit, the oscillations of $|\langle q|\psi\rangle|^2$ become infinitely rapid, so we can measure only its average $|\langle q|\psi\rangle|^2 = 1/l$. By the same token, the interference term of the Wigner function can be considered to vanish as a consequence of its infinitely rapid positive and negative oscillations. The smoothed Wigner function corresponds to the classical probability density. It is interesting to note that the Wigner function broadens nonclassically, close to the wall. This feature is not present in the case of two periodic waves travelling in opposite directions, treated in previous presentations of this example [19, 31].

As another important example, we now derive the Wigner function for a coherent state $|z\rangle$ with mean values $\langle p\rangle = P$ and $\langle q\rangle = Q$:

$$\begin{aligned} &W_z(x) \\ &= \left(\frac{1}{2\pi\hbar}\right)^L\left(\frac{w}{\pi\hbar}\right)^{L/2}\int \mathrm{d}q' \exp\left\{-\frac{w}{2\hbar}\left[q-Q+\frac{q'}{2}\right]^2 - \frac{w}{2\hbar}\left[q-Q-\frac{q'}{2}\right]^2 + \frac{\mathrm{i}}{\hbar}\,(P-p)q'\right\} \\ &= (\hbar\pi)^{-L}\exp\left\{-\frac{w}{\hbar}(q-Q)^2 - \frac{1}{w\hbar}(p-P)^2\right\}. \end{aligned} \tag{5.30}$$

This minimum uncertainty state is thus a non-oscillating Gaussian centred on (P,Q). In the case where $(P,Q)=0$, we obtain the Wigner function for the ground state of the harmonic oscillator.

We have used a multidimensional version of (4.43) in deriving the above formula, with the same frequency for each freedom, but this restriction is not necessary in (5.30).

We can now use (5.30) to derive the relation between the Weyl and the Husimi representation. Developing

$$A_{\mathrm{H}}(P,Q)=\langle z|\hat{A}|z\rangle=\int \mathrm{d}q'\,\mathrm{d}q''\langle z|q''\rangle\langle q''|\hat{A}|q'\rangle\langle q'|z\rangle \tag{5.31}$$

and taking $q=(q'+q'')/2$ and $q'''=q''-q'$, we obtain

$$\langle z|\hat{A}|z\rangle=\int \mathrm{d}q\,\mathrm{d}q'''\left\langle q-\frac{q'''}{2}\middle|z\right\rangle\left\langle z\middle|q+\frac{q'''}{2}\right\rangle\int\frac{\mathrm{d}p}{(2\pi\hbar)^L}\,A(p,q)\exp\left[\frac{\mathrm{i}}{\hbar}p\cdot q'''\right]$$

$$=\int\frac{\mathrm{d}x}{(2\pi\hbar)^L}\,W_z(x)A(x)\,. \tag{5.32}$$

In the case where $\hat{A}=\hat{\rho}_\psi$, the density operator for the state $|\psi\rangle$, we obtain

$$\rho_{\mathrm{H}}(P,Q)=\int \mathrm{d}x\,W_z(x)W_\psi(x)=|\langle z|\psi\rangle|^2\,. \tag{5.33}$$

So we find that the Husimi representation can be viewed as a Gaussian smoothing of the Weyl representation. Because of the analyticity of coherent states, it is possible though nontrivial to retrieve the information masked by the smoothing. Conversely, we see from (5.33) that a Gaussian smoothing of the Wigner function necessarily produces a positive definite distribution. Indeed, we may consider the projections (5.22) and (5.23) as particular instances of (5.33), when we choose $w\to\infty$ or $w\to 0$, for which $w_z(x)\to\delta(q-Q)$ or $\delta(p-P)$, respectively. Since the Husimi function must be positive (or zero) for any w, we deduce that the Wigner function can only be negative in the context of narrow oscillations.

Returning to the example of a particle in a box with hard walls, we see immediately that a convenient choice of w will produce a smoothing that effectively cancels oscillations along $p=0$. We thus obtain a representation which underlines the classical structure. However, we cannot recover the wave intensity by a mere projection, once the Husimi smoothing has been performed. (See Ref. [31] for further discussion of smoothings of the Wigner function.)

A similar situation arises with the eigenstates of the harmonic oscillator. According to Groenewold [32] the Wigner functions in the case of $w=1$ are

$$W_n(x)=((-1)^n/\pi\hbar)\exp(-x^2/\hbar^2)L_n(2x^2/\hbar^2)\,, \tag{5.34}$$

were $L_n(z)$ is the nth Laguerre polynomial normalized to unity at the origin. We recognize here the same oscillatory structure that is familiar in the position representation. This must be so, if we are to obtain the correct projections. These oscillations are wiped out by the Gaussian smoothing which leaves the single maximum along the classical manifold that we found for the Husimi function (4.58). Berry [7] showed that generally the semiclassical approximation of the Wigner function for a pure state has a maximum along the energy shell in the case that $L=1$. This is the border of Airy function fringes inside the shell (se also [19]) that are wiped out by the Husimi smoothing. Therefore, the zeros of the Husimi function will be forced away from the classical region.

6. Products of operators and path integrals

It may be surprising that the derivation of the composition rule for successive unitary transformations is actually simpler than the classical counterpart studied in Section 2. We shall see that the corresponding geometrical structures arise as a mere consequence of the algebraic rules that we have defined in the two previous sections.

Starting with the chord representation, we have, for the product $\hat{A}_n\hat{A}_{n-1}\cdots\hat{A}_1$,

$$A_n\cdot A_{n-1}\cdots A_1(\xi)=\mathrm{Tr}\left(\frac{1}{2\pi\hbar}\right)^{nL}\int \mathrm{d}\xi_n\cdots\mathrm{d}\xi_1 A_n(\xi_n)\cdots A_1(\xi_1)\hat{T}_{-\xi}\hat{T}_{\xi_n}\hat{T}_{\xi_{n-1}}\cdots\hat{T}_{\xi_1}$$

$$=\left(\frac{1}{2\pi\hbar}\right)^{n-1}\int \mathrm{d}\xi_n\cdots\mathrm{d}\xi_1\, A_n(\xi_n)\cdots A_1(\xi_1)\mathrm{Tr}[\hat{T}_{\xi_1+\cdots+\xi_n-\xi}]$$

$$\times\exp\left[-\frac{\mathrm{i}}{\hbar}D_{n+2}(\xi_1,\ldots,\xi_n,-\xi)\right]$$

$$=\left(\frac{1}{2\pi\hbar}\right)^{L(n-1)}\int \mathrm{d}\xi_n\cdots\mathrm{d}\xi_1 A_n(\xi_n)\cdots A_1(\xi_1)\delta(\xi_1+\cdots\xi_n-\xi)$$

$$\times\exp\left[-\frac{\mathrm{i}}{\hbar}D_{n+1}(\xi_1,\ldots,\xi_n)\right]\,, \tag{6.1}$$

where we note that the Dirac δ-function has reduced the $(n+2)$-sided polygon with symplectic area D_{n+2} to an $(n+1)$-sided polygon, with n free sides. Evidently, we can now use the δ-function to remove one of the variables in the integral, but (6.1) is in its most symmetric form.

To obtain the composition rule of operators in the centre representation, we can proceed in several ways. The simplest is just to take the Fourier transform of (6.1):

$$A_n\cdot A_{n-1}\cdots A_1(x)=(1/2\pi\hbar)^{Ln}\int \mathrm{d}\xi_n\cdots\mathrm{d}\xi_1\,\mathrm{d}\xi\, A_n(\xi_n)\cdots A_1(\xi_1)\delta(\xi_1+\cdots\xi_n-\xi)$$

$$\times\exp\{-(\mathrm{i}/\hbar)[D_{n+1}(\xi_1,\ldots,\xi_n)-x\wedge\xi]\}\,. \tag{6.2}$$

Recalling that $\xi=\xi_1+\cdots+\xi_n$, we now define the multivariable function

$${}'A_n\cdots A_1'(x_1,\ldots,x_n)=\left(\frac{1}{2\pi\hbar}\right)^{Ln}\int \mathrm{d}\xi_1\cdots\mathrm{d}\xi_n\exp\left[-\frac{\mathrm{i}}{\hbar}D_{n+1}(\xi_1,\ldots,\xi_n)\right]$$

$$\times\prod_{j=1}^{n}A_j(\xi_j)\exp\left[-\frac{\mathrm{i}}{\hbar}x_j\wedge\xi_j\right]\,, \tag{6.3}$$

which takes on the special value

$${}'A_n\cdots A_1'(x,x,\ldots,x)=A_n\cdots A_1(x)\,. \tag{6.4}$$

This multivariable function is now evaluated by expanding the exponential with the polygonal area D_{n+1}, so as to yield a series of integrals. For the zeroth order, the integrals decouple, so that

$$'A_n \cdots A_1'(x_1,\ldots,x_n)_0 = A_1(x_1)\cdots A_n(x_n)\,. \tag{6.5}$$

The next term is given by

$$(2\pi\hbar)^{Ln}\ \ 'A_n \cdots A_1'(x_1,\ldots,x_n)_1 = -\frac{\mathrm{i}}{\hbar}\int \mathrm{d}\xi_1 \cdots \mathrm{d}\xi_n D_{n+1}(\xi_1,\ldots,\xi_n) \times \prod_{j=1}^{n} A_j(\xi_j)\exp\left[-\frac{\mathrm{i}}{\hbar}x_j \wedge \xi_j\right], \tag{6.6}$$

so, if we notice that $D_{n+1}(\xi_1,\ldots,\xi_n)$ is a bilinear function of its arguments and that

$$\left[\frac{\partial}{\partial x_j}\wedge\frac{\partial}{\partial x_k}\right]\exp\left\{-\frac{\mathrm{i}}{\hbar}[x_j\wedge\xi_j + x_k\wedge\xi_k] = \frac{1}{\hbar^2}[\xi_j\wedge\xi_k] \times \exp\left\{-\frac{\mathrm{i}}{\hbar}[x_j\wedge\xi_j + x_k\wedge\xi_k]\right\}\right\}, \tag{6.7}$$

we obtain

$$'A_n \cdots A_1'(x_1,\ldots,x_n)_1 = -\mathrm{i}\hbar D_{n+1}(\partial/\partial x_1,\ldots,\partial/\partial x_n)A_1(x_1)\cdots A_n(x_n)\,. \tag{6.8}$$

But we can easily generalize (6.8) to higher-order derivatives, allowing us to resum the exponential expansion in the symbolic form

$$'A_n \cdots A_1'(x_1,\ldots,x_n) = \exp\{-\mathrm{i}\hbar D_{n+1}(\partial/\partial x_1,\ldots,\partial/\partial x_n)\}A_1(x_1)\cdots A_n(x_n)\,. \tag{6.9}$$

For the product of only two operators, (6.9) reduces to the *Groenewold rule* [32]:

$$A_2\cdot A_1(x) = \exp\left\{-\frac{\mathrm{i}\hbar}{2}\ \frac{\partial}{\partial x_1}\wedge\frac{\partial}{\partial x_2}\right\}A_1(x_1)|_{x_1=x}A_2(x_2)|_{x_2=x}\,. \tag{6.10}$$

Since the Weyl representation of Hermitian operators is of zero-order in Planck's constant, (6.10) and its generalization (6.9) are useful starting points for semiclassical expansions. In particular, we see that

$$A_2\cdot A_1(x) = A_1(x)A_2(x) - \frac{\mathrm{i}\hbar}{2}\frac{\partial A_1}{\partial x}\wedge\frac{\partial A_2}{\partial x} + \mathrm{O}(\hbar^2)\,, \tag{6.11}$$

where we recognize the *Poisson bracket* $\{A_1,A_2\}$ in the term that is of first order in $\hbar$. Hence, the commutator has the Weyl transform

$$[A_1,A_2](x) = \mathrm{i}\hbar\{A_1(x),A_2(x)\} + \mathrm{O}(\hbar^2)\,, \tag{6.12}$$

neatly reflecting the correspondence between commutators and Poisson brackets which was the starting point of quantum mechanics. For the symmetrized product,

$$\tfrac{1}{2}(A_1\cdot A_2 + A_2\cdot A_1)(x) = A_1(x)A_2(x) + \mathrm{O}(\hbar^2)\,. \tag{6.13}$$

If $\hat{A}_1$ and $\hat{A}_2$ are Hermitian, so is their symmetrized product, implying that the first term holds to second order in $\hbar$, because the Weyl transform must be real. Using (6.9) we also find that the power of an Hermitian operator $\hat{H}$ is represented by

$$H^n(x) = [H(x)]^n + \mathrm{O}(\hbar^2)\,. \tag{6.14}$$

In particular, we may expand the square of an observable as

$$H^2(x) = [H(x)]^2 + \tfrac{1}{8}\hbar^2 \mathrm{Tr}(\mathfrak{J}\mathscr{H})^2 + \mathrm{O}(\hbar^4)\,, \tag{6.15}$$

where $\mathscr{H}$ is the Hessian of $H(x)$ and $\mathfrak{J}$ is defined by (1.3).

We shall also need integral formulae for the product of operators. The result depends crucially on the parity of the number of operators so we will start with the simplest case where $n=2$. Proceeding from the definition (5.6), we obtain

$$\begin{aligned} A_2 \cdot A_1(x) &= \mathrm{Tr}\left(\frac{1}{2\pi\hbar}\right)^{2L} \int \mathrm{d}x_2\, \mathrm{d}x_1\, A_2(x_2) A_1(x_1)\, \hat{\mathbb{R}}_x\, \hat{\mathbb{R}}_{x_2}\, \hat{\mathbb{R}}_{x_1} \\ &= \left(\frac{1}{\pi\hbar}\right)^{2L} 2^L \int \mathrm{d}x_2\, \mathrm{d}x_1\, A_2(x_2) A_1(x_1) \exp\left[\frac{\mathrm{i}}{\hbar} x_1 \wedge x_2\right] \mathrm{T}_2\left[\hat{R}_x \hat{T}_{(x_2 - x_1)}\right] \\ &= \left(\frac{1}{\pi\hbar}\right)^{2L} 2^L \int \mathrm{d}x_2\, \mathrm{d}x_1\, A_2(x_2) A_1(x_1) \exp\left[\frac{\mathrm{i}2}{\hbar} x_1 \wedge x_2 - \frac{\mathrm{i}}{\hbar} x \wedge 2(x_2 - x_1)\right] \mathrm{Tr}\, \hat{R}_{x-(x_2-x_1)} \\ &= \left(\frac{1}{\pi\hbar}\right)^{2L} \int \mathrm{d}x_2\, \mathrm{d}x_1\, A_2(x_2) A_1(x_1) \exp\left[\frac{\mathrm{i}}{\hbar} A_3(x, x_1, x_2)\right]. \end{aligned} \tag{6.16}$$

The extension to $(2n)$ operators can be derived by dividing the polygon of $(2n+1)$ sides η_j into a set of n triangles that result from pairing the sides (as shown in Figs. 2.3 and 2.4) and an internal polygon Δ'_{n+1}. Its symplectic area is

$$\Delta'_{n+1} = \tfrac{1}{2}\,\alpha_1 \wedge (\eta_1 + \eta_2) + \cdots + \tfrac{1}{2}\,\alpha_n \wedge (\eta_{2n-1} + \eta_{2n})\,, \tag{6.17}$$

where α_j is the vector joining x to the centre of the jth side of Δ'_{n+1}. Thus, we can insert (5.6) for each operator into (5.7), to obtain

$$\begin{aligned} A_{2n} \cdots A_1(x) &= \mathrm{Tr}\left(\frac{1}{2\pi\hbar}\right)^{2nL} \int \mathrm{d}x_{2n} \cdots \mathrm{d}x_1\, A_{2n}(x_{2n}) \cdots A_1(x_1)\, \hat{\mathbb{R}}_x\, \hat{\mathbb{R}}_{x_{2n}} \cdots \hat{\mathbb{R}}_1 \\ &= \left(\frac{1}{\pi\hbar}\right)^{2nL} \int \mathrm{d}x_{2n} \cdots \mathrm{d}x_1\, A_{2n}(x_{2n}) \cdots A_1(x_1) \\ &\quad \times \exp\left[\frac{\mathrm{i}2}{\hbar}(x_1 \wedge x_2 + \cdots + x_{2n-1} \wedge x_{2n})\right] \mathrm{Tr}\left[\hat{R}_x \hat{T}_{2(x_{2n} - x_{2n-1})} \cdots \hat{T}_{2(x_2 - x_1)}\right] \\ &= \left(\frac{1}{\pi\hbar}\right)^{2nL} 2^L \int \mathrm{d}x_{2n} \cdots \mathrm{d}x_1 A_{2n}(x_{2n}) \cdots A_1(x_1) \exp\left[\frac{\mathrm{i}2}{\hbar}(x_1 \wedge x_2 + \cdots + x_{2n-1} \wedge x_{2n})\right. \\ &\quad \left. - \frac{\mathrm{i}}{\hbar}\Delta'_{n+1} - \frac{\mathrm{i}}{\hbar} x \wedge \xi\right] \mathrm{Tr}\, \hat{R}_{x - \xi/2} \\ &= \left(\frac{1}{\pi\hbar}\right)^{2nL} \int \mathrm{d}x_{2n} \cdots \mathrm{d}x_1 A_{2n}(x_{2n}) \cdots A_1(x_1) \end{aligned}$$

$$\times \exp\left\{\frac{\mathrm{i}}{\hbar}[\Delta_3(x+\alpha_1,x_1,x_2)+\cdots+\Delta_3(x+\alpha_{2n-1},x_{2n-1},x_{2n})+\Delta'_{n+1}]\right\}$$

$$=\left(\frac{1}{\pi\hbar}\right)^{2nL}\int \mathrm{d}x_{2n}\cdots\mathrm{d}x_1\,A_{2n}(x_{2n})\cdots A_1(x_1)\exp\left\{\frac{i}{\hbar}\Delta_{2n+1}(x,x_1,\ldots,x_{2n})\right\}. \quad (6.18)$$

Here the symplectic area Δ_{2n+1} corresponds to the $(2n+1)$-sided polygon circumscribed around the centres $(x,x_1,\ldots,x_{2n})$. In Ref. [11] this result was derived by induction.

The simplest way to derive the composition law for an odd number of operators is merely to particularize $\hat{A}_{2n}=\hat{1}$. In this case we immediately obtain

$$A_{2n-1}\cdots A_1(x)=\left(\frac{1}{\pi\hbar}\right)^{2nL}\int \mathrm{d}x_{2n-1}\cdots\mathrm{d}x_1\,A_{2n-1}(x_{2n-1})\cdots A_1(x_1)$$

$$\times\int \mathrm{d}x_{2n}\exp\left\{\frac{\mathrm{i}}{\hbar}\Delta_{2n+1}(x,x_1,\ldots,x_{2n})\right\}. \quad (6.19)$$

But, according to (2.11), we have

$$\int \mathrm{d}x\exp\left\{\frac{\mathrm{i}}{\hbar}\Delta_{2n+1}(x,x_1,\ldots,x_{2n})\right\}=(2\pi\hbar)^{2L}\delta(\xi)=(2\pi\hbar)^{2L}\delta\left(2\sum_{k=1}^{2n}(-1)^k x_k\right), \quad (6.20)$$

where ξ is the side centred on x. Since this becomes zero, the polygon loses one side and

$$A_{2n-1}\cdots A_1(x)=(1/\pi\hbar)^{(2n-1)L}\int \mathrm{d}x_{2n-1}\cdots\mathrm{d}x_1\,A_{2n-1}(x_{2n-1})\cdots A_1(x_1)$$

$$\times\delta[x_{2n-1}-\cdots+x_1-x]\exp\{(\mathrm{i}/\hbar)\Delta_{2n}(x,x_1,\ldots,x_{2n-1})\}. \quad (6.21)$$

Taking one of the sides of Δ_{2n+1} as zero has created an even-sided polygon Δ_{2n}. This will only be properly defined by its midpoints if the argument of the δ-function cancels. In the important case of three operators, we obtain

$$A_3\cdot A_2\cdot A_1(x)=(1/\pi\hbar)^{3L}\int \mathrm{d}x_3\,\mathrm{d}x_2\,\mathrm{d}x_1\,A_3(x_3)A_2(x_2)A_1(x)\delta[x_3-x_2+x_1-x]$$

$$\times\exp\{(\mathrm{i}/\hbar)\Delta_4(x,x_1,x_2,x_3)\}. \quad (6.22)$$

The δ-function here forces the inscribed quadrilateral to be a parallelogram, which is necessary for the existence of the circumscribed quadrilateral, as we saw in Section 2 (Fig. 2.10). Evidently, we can use the δ-function to eliminate one of the variables, say x_2 as shown in Fig. 2.11, so that

$$A_3\cdot A_2\cdot A_1(x)=(1/\pi\hbar)^{3L}\int \mathrm{d}x_3\,\mathrm{d}x_1\,A_3(x_3)A_2(x_3+x_1-x)A_1(x_1)$$

$$\times\exp\{(\mathrm{i}/\hbar)\Delta_3(x,x_1,x_3)\}, \quad (6.23)$$

where we use the fact that all the quadrangles circumscribed on a parallelogram have the same area, as proved in Appendix A. The triangle adopted here is just a particular case where the side $\xi_2=0$.

We will now discuss a few important results concerning the phase space representation of products of operators. Even though the chord and the centre representations of the product of two operators

are not equal to the product of the functions that represent them, this interchange does hold for the trace:

$$\mathrm{Tr}\,\hat{A}_2\hat{A}_1 = \int (\mathrm{d}\xi/(2\pi\hbar)^L)A_2(\xi)A_1(\xi) = \int (\mathrm{d}x/(2\pi\hbar)^L)A_2(x)A_1(x)\,. \tag{6.24}$$

These results are easily obtained by applying (5.15) to (6.1) and (6.16). Clearly, they reflect the fact that the trace of the product of two operators commutes. Since the trace of the product of more than two operators does not necessarily commute, we cannot expect to obtain such a remarkable simplification of the trace in general. In Section 7 we shall discuss the trace of the product of an arbitrary even number of unitary operators.

A direct consequence of (6.24) is that, for any observable $\hat{A}$,

$$\langle\hat{A}\rangle = \int \mathrm{d}x\, A(x)W(x)\,. \tag{6.25}$$

This remarkable formula has been for a long time one of the main attractions of the Weyl representation; indeed, it is the motivation for Wigner's original paper [1]. We see that even though the Wigner function need not be positive definite, it allows us to calculate mean values of the observables in the same manner as with the classical Liouville probability distribution. This is specially appealing in the cases where the Weyl representation of the observable coincides exactly with the classical variable. This feat cannot be emmulated by the positive-definite Husimi function.

The product rules can be used to study unitary operators. First, let us establish that we can define a group of unitary operators by their Weyl representation of the form

$$U(x) = c\exp[\mathrm{i}\hbar^{-1}x\mathfrak{B}x] \tag{6.26}$$

for any symmetric matrix $\mathfrak{B}$. To do this we use (6.16) to show that the corresponding operator satisfies $\hat{U}\hat{U}^\dagger = \hat{1}$, for an appropriate choice of the constant c. Thus, we define

$$\begin{aligned} I(\mathfrak{B}) &= \frac{1}{(\pi\hbar)^{2L}}\int \mathrm{d}x_1\,\mathrm{d}x_2 U(x_1)[U(x_2)]^* \exp\left[\frac{\mathrm{i}}{\hbar}\Delta_3(x,x_1,x_2)\right] \\ &= \frac{|c|^2}{(\pi\hbar)^{2L}}\int \mathrm{d}x_1\,\mathrm{d}x_2 \exp\left\{\frac{\mathrm{i}}{\hbar}[x_1\mathfrak{B}x_1 - x_2\mathfrak{B}x_2 - 2(x_1-x)\mathfrak{J}(x_2-x)]\right\} \\ &= \frac{|c|^2}{(\pi\hbar)^{2L}}\int \mathrm{d}x_1\,\mathrm{d}x_2 \exp\left\{\frac{\mathrm{i}}{\hbar}[x_1\mathfrak{B}x_1 - x_2\mathfrak{B}x_2 - 2x_1\mathfrak{J}x_2]\right\}\,, \end{aligned} \tag{6.27}$$

where in the last equation we have taken the origin to the stationary point of the quadratic phase; since $x_1 = x_2$ there, the x dependence cancels. We easily ascertain that $I(\mathfrak{B})$ is real and that $I(\mathfrak{B}) = I(-\mathfrak{B})$.

To calculate the resulting multiple Gaussian integral, it is convenient to define the double vector $X = (x_1, x_2)$ and the $(4L \times 4L)$ symmetric matrices

$$\mathbb{B} = \left[\begin{array}{c|c} \mathfrak{B} & 0 \\ \hline 0 & -\mathfrak{B} \end{array}\right] \quad \text{and} \quad \mathbb{J} = \left[\begin{array}{c|c} 0 & \mathfrak{J} \\ \hline \mathfrak{J} & 0 \end{array}\right] \tag{6.28}$$

so that

$$I(\mathfrak{B}) = \frac{|c|^2}{(\pi\hbar)^{2L}} \int \mathrm{d}X \exp\left\{\frac{\mathrm{i}}{\hbar} X[\mathbb{B} - \mathbb{J}]X\right\} = \frac{|c|^2}{(\pi\hbar)^{2L}} \left[\frac{(\pi\hbar)^{4L}}{|\det[\mathbb{B} - \mathbb{J}]|}\right]^{1/2}. \tag{6.29}$$

Therefore, we need to define

$$c = |\det[\mathbb{B} - \mathbb{J}]|^{1/4} \mathrm{e}^{\mathrm{i}\theta} \tag{6.30}$$

for (6.26) to represent an unitary operator. To simplify this $(4L \times 4L)$ determinant, we first note that

$$|\det[\mathbb{B} - \mathbb{J}]| = |\det[1 - \mathbb{J}\mathbb{B}]| = \det\left|\left[\begin{array}{c|c} 1 & \mathfrak{JB} \\ \hline \mathfrak{JB} & 1 \end{array}\right]\right|, \tag{6.31}$$

since $\mathbb{J}^2 = 1$. Now we use the fact that $I(\mathfrak{B}) = I(-\mathfrak{B})$ to write

$$\begin{aligned} |\det[\mathbb{B} - \mathbb{J}]| &= |\det[1 + \mathbb{J}\mathbb{B}][1 - \mathbb{J}\mathbb{B}]|^{1/2} \\ &= \left|\det\left[\begin{array}{c|c} 1 - (\mathfrak{JB})^2 & 0 \\ \hline 0 & 1 - (\mathfrak{JB})^2 \end{array}\right]\right|^{1/2} = |\det[1 - (\mathfrak{JB})^2]|, \end{aligned} \tag{6.32}$$

which has reduced to a $(2L \times 2L)$ determinant. This can be further simplified by recalling the relation (1.20) between the symmetric matrix $\mathfrak{B}$ and the sympletic matrix $\mathscr{M}$, such that $\det \mathscr{M} = 1$. The result is that

$$|\det[1 - \mathfrak{JB}]|^{1/2} = |\det[1 + \mathfrak{JB}]|^{1/2} = |\det[1 - (\mathfrak{JB})^2]|^{1/4}. \tag{6.33}$$

Thus, we obtain alternative forms for the unitary operators related to the symmetric matrix $\mathfrak{B}$ as

$$U(x) = |\det[1 \pm \mathfrak{JB}]|^{1/2} \exp[\mathrm{i}\hbar^{-1} x\mathfrak{B}x + \mathrm{i}\theta] = 2^L |\det(1 + \mathscr{M})|^{-1/2} \exp[\mathrm{i}\hbar^{-1} x\mathfrak{B}x + \mathrm{i}\theta]. \tag{6.34}$$

The phase θ is unimportant for the action of the unitary operator on another operator, since it cancels in (4.6). However, we have seen that the definition of the Weyl representation relies on the linear superposition of unitary operators, for which it is essential to take account of the phase. Moreover, we can also obtain the position representation $\langle q'|U|q\rangle$, through a symmetrized Fourier transform and hence transform wave functions $\langle q|\psi\rangle$. Though these are only defined within a phase factor, the latter must be accounted for in the phenomena of interference between waves. The determination of these phases in the context of the semiclassical approximations will be the subject of future work.

How do these unitary operators act on other operators? Using (6.23) to obtain the Weyl representation of $\hat{A}' = \hat{U}^\dagger \hat{A} \hat{U}$, we have

$$\begin{aligned} A'(x) &= \int \frac{\mathrm{d}x_1\, \mathrm{d}x_2}{(\pi\hbar)^{2L}} [U(x_1)]^* A(x_2 + x_1 - x) U(x_2) \exp\left[\frac{\mathrm{i}}{\hbar} \Delta_3(x, x_1, x_2)\right] \\ &= |\det[1 - \mathfrak{JB}]| \int \frac{\mathrm{d}x_1\, \mathrm{d}x_2}{(\pi\hbar)^{2L}} \exp\left\{\frac{\mathrm{i}}{\hbar}[-x_1\mathfrak{B}x_1 + x_2\mathfrak{B}x_2 - 2(x_1 - x)\mathfrak{J}(x_2 - x)]\right\} \\ &\quad \times A(x_2 + x_1 - x) \end{aligned} \tag{6.35}$$

with the change of variables $x_2 + x_1 - x = X$, $x_2 - x_1 = Y$ (which has Jacobian 2^{2L}), the integral becomes

$$A'(x) = |\det[1 - \mathfrak{J}\mathfrak{B}]| \int \frac{\mathrm{d}XA(X)}{(2\pi\hbar)^{2L}} \int \mathrm{d}Y \exp\left\{\frac{\mathrm{i}}{\hbar}[Y\mathfrak{B}(X+x) + Y\mathfrak{J}(X-x)]\right\}$$

$$= |\det[1 - \mathfrak{J}\mathfrak{B}]| \int \mathrm{d}X\ A(X)\delta(\mathfrak{B}(X+x) + \mathfrak{J}(X-x)). \tag{6.36}$$

Thus, recalling again (1.20), we obtain

$$A'(x) = A(\mathscr{M}x), \tag{6.37}$$

where $\mathscr{M}$ is the symplectic matrix corresponding to the centre function $S(x) = x\mathfrak{B}x + \theta$.

Actually, we can add a linear term, to obtain $S(x) = \alpha \wedge x + x\mathfrak{B}x$. This is the centre function for a non-homogeneous (linear) symplectic transformation in classical mechanics. The only effect will be to change the origin of the preceding Gaussian integrals. Thus, we find that there is a one-to-one correspondence between this restricted group of classical canonical transformations generated by the centre functions $S(x)$ with the subgroup of quantum unitary operators, given in the Weyl representation by

$$U(x) = 2^L|\det(1 + \mathscr{M})|^{-1/2} \exp[-\mathrm{i}\hbar^{-1}(\alpha \wedge x + x\mathfrak{B}x) + \mathrm{i}\theta]$$

$$= |\det[1 \pm \mathfrak{J}\partial^2 S/\partial x^2]|^{1/2} \exp[\mathrm{i}\hbar^{-1}S(x)]. \tag{6.38}$$

Taking the Fourier transform of this expression, we find that the corresponding chord representation is

$$U(\xi) = 2^L|\det(1 - \mathscr{M})|^{-1/2}\ \exp[(\mathrm{i}/\hbar)(a \wedge \xi + \tfrac{1}{4}\xi\mathfrak{B}\xi) + \mathrm{i}\theta'], \tag{6.39}$$

where now the symmetric matrix $\mathfrak{B}$ is given by $\mathscr{M}$ in (1.35). The corresponding position representation of the evolution operator, resulting from the symmetrized Fourier transform (5.14), is

$$\langle q'|U|q\rangle = (2\pi\hbar)^{-1/2}|\det(\partial^2 S/\partial q\partial q')|^{1/2} \exp[(\mathrm{i}/\hbar)S(q',q)], \tag{6.40}$$

where $S(q',q)$, the symmetrized Legendre transform (1.54) of $S(x)$, will again be quadratic. In these representations, the undetermined phase θ, can be interpreted as the arbitrary additive constant in the definition of the corresponding generating function. The unitary operators that we have been discussing form the *inhomogeneous metaplectic group* [41,47]. The latter reference includes some discussion of the overall phase θ.

We do not have simple closed forms for the general unitary transformations generated by nonlinear Hamiltonians. Just as in classical mechanics, it is then important to develop approximations that are valid for small intervals of time. By expanding the exponential (4.3), we obtain the Weyl transform term by term:

$$U_t(x) = 1 - \mathrm{i}(t/\hbar)H(x) - \tfrac{1}{2}(t/\hbar)^2H^2(x) + \cdots + (1/n!)(-\mathrm{i}t/\hbar)^n H^n(x) + \cdots. \tag{6.41}$$

We have seen that $H(x)$ will be within $\mathrm{O}(\hbar)$ of the classical Hamiltonian $H_c(x)$ and (6.14) shows that the centre representation of $\hat{H}^n$ is $\mathrm{O}(\hbar^2)$ of $[H(x)]^n$. Therefore, the above series is uniformly convergent, which allows us to rearrange the terms in the form

$$U_t(x) = \exp[-\mathrm{i}(t/\hbar)H(x)] + \mathrm{O}((t/\hbar)^2). \tag{6.42}$$

For sufficiently small times, we may thus use the exponential of $H(x)$ as an approximation for the Weyl propagator. This is in close analogy to the theory in Section 1, where the centre function was found to have $-tH_c(x)$ as a limit. However, the classical approximation holds to $O(t^3)$, whereas the range of validity of (6.42) is exceedingly small in the semiclassical limit. Futhermore, contrary to the classical theory where $-tH_c(x)$ can always be used to generate a transformation that is guaranteed to be canonical, we cannot be sure that (6.42) represents an unitary operator.

To improve our approximation, we can include the first correction (6.15):

$$U_t(x) = \exp[-\mathrm{i}(t/\hbar)H(x)] - \tfrac{1}{2}t^2/\hbar^2[(\hbar^2/8)\mathrm{Tr}(\mathfrak{J}\mathscr{H})^2 + O(\hbar^4)] + O((t/\hbar)^3)$$
$$\times \exp[-\mathrm{i}(t/\hbar)H(x)]\{1 - (t^2/16)\mathrm{Tr}(\mathfrak{J}\mathscr{H})^2\} + O(t^2\hbar^2) + O((t/\hbar)^3)\,, \tag{6.43}$$

where $\mathscr{H}$ is the Hessian matrix evaluated at the point x. Even though this is still not as good as the achievement in classical mechanics, the first correction becomes small specially in the semiclassical limit, while the second correction is now of third order. The small time propagator can now be rewritten as

$$U_t(x) = |\det[1 - (\mathfrak{J}(t/2)\mathscr{H})^2]|^{1/4} \exp[-(\mathrm{i}t/\hbar)H(x)] + O(t^2\hbar^2) + O((t/\hbar)^3)\,. \tag{6.44}$$

The advantage of this form is that comparison with (6.33) and (6.34) demonstrates it to represent an unitary operator if $H(x)$ is quadratic. Indeed, $-(t/2)\mathscr{H}$ will then just be the constant matrix $\mathfrak{B}$ in accordance with the theory in Section 1. Since we can always expand $H^n(x)$ to second order terms, we find that the simple rearrangement between (6.43) and (6.44) really accounts for a resummation of higher-order terms in the expansion of $H^n(x)$.

We can now extend the range of the Weyl propagator for the Hamiltonian flow by composing the unitary operators corresponding to small periods:

$$U_t(x) = \int \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_N}{(\pi\hbar)^{NL}} \prod_{n=1}^{N} U_{t/N}(x_n) \exp\left\{\frac{\mathrm{i}}{\hbar}[\Delta_{N+1}(x, x_1, \ldots, x_N)]\right\}. \tag{6.45}$$

This formula is exact, but we can only insert (6.44) for $U_{t/N}(x)$ in the limit as $N \to \infty$. We can then ignore the amplitude, since

$$\prod_{n=1}^{N} \det\left[1 - \left(\mathfrak{J}\frac{t}{2N}\mathscr{H}_n\right)^2\right] \underset{N\to\infty}{\longrightarrow} 1\,, \tag{6.46}$$

if the Hessian $\mathscr{H}_n$ remains bounded for each centre x_n. Thus, we obtain the *path integral*

$$U_t(x) = \lim_{N\to\infty} \int \frac{\mathrm{d}x_1 \cdots \mathrm{d}x_N}{(\pi\hbar)^{NL}} \exp\left\{\frac{\mathrm{i}}{\hbar}\left[\Delta_{N+1}(x, x_1, \ldots, x_N)) - \frac{t}{N}\sum_{n=1}^{N} H(x_n)\right]\right\}. \tag{6.47}$$

We immediately recognize that the phase of this integral coincides with the action of the variational principle (2.23) for the polygonal path with endpoints centred on x and whose kth side is centred on x_k. Just as with the variational principle, we need not worry about the definition of a "path space" since (6.47) is an ordinary multiple integral. The only cost is that there will be some very jagged polygonal paths (see Fig. 2.7) as well as smooth paths, such as the classical trajectory that solves the variational problem.

It is interesting to note that the geometry of $U_t(x)$, represented in terms of the polygon Δ_{N+1}, is reminiscent of that for the Wigner function $W_n(x)$ obtained in [19]. This will result in the next section from the Fourier transform of $U_t(x)$. The original derivation relied on the iteration of the pure state condition $\hat{\rho}_n^2 = \hat{\rho}_n$, instead of the group property for $\hat{U}_t$.

The clearest way to derive the semiclassical approximation for the Weyl propagator is to return to (6.45) with the individual small time propagators specified by (6.44). The latter are already in their semiclassical form, in the limit of small intervals. The semiclassical limit for the full propagator now involves evaluating the multiple integral (6.45) by stationary phase. This depends only on the quadratic expansion of the phase and, since Δ_{N+1} is already quadratic, we need only expand the Hamiltonian about each stationary point x_n:

$$H(x_n + X_n) = H(x_n) + \partial H/\partial x_n \cdot X_n + X_n \mathscr{H}_n X_n + \cdots . \tag{6.48}$$

But to this order, each of the propagators $U_{t/N}(x_n)$ has exactly the Gaussian form (6.38) with the symmetric matrix $\mathfrak{B} = -(t/2N)\mathscr{H}_n$. We have seen that these propagators correspond to the group of classical (linear) symplectic transformations. Therefore, the composition of these metaplectic operators is isomorphic to the composition of the corresponding classical transformations. In other words, the stationary phase evaluation of (6.18) at a given point x must have the same value as (6.38), with $\mathscr{M}$ being the symplectic matrix for the full linear transformation obtained by composing the linearized transformations about each stationary centre x_n.

The stationary points lie on the classical trajectory which solves the variational problem. Thus, with this choice of x_n, we cancel the sum of linear terms in the phase of (6.45), whereas the constant part of the phase about which we expand is just $\hbar^{-1}S_t(x)$ as given by (2.23). It follows that the semiclassical approximation is just

$$U_t(x)_{\mathrm{SC}} = \exp\left[\frac{\mathrm{i}}{\hbar}S_t(x)\right] \lim_{N\to\infty} \int \frac{\mathrm{d}X_1 \cdots \mathrm{d}X_N}{(\pi\hbar)^{NL}} \prod_{n=1}^{N} U_{t/N}^n(X_n)$$
$$\times \exp\{(\mathrm{i}/\hbar)\Delta_{N+1}(0, X_1, \ldots, X_N)\} , \tag{6.49}$$

where $X_n = x_n - x_{n_0}$, x_{n_0} being the stationary point on the classical orbit, so that

$$U_{t/N}^n(X_n) = 2^L |(1 + \mathscr{M}_n)|^{-1/2} \exp[-\mathrm{i}(t/\hbar N) X_n \mathscr{H}_n X_n] \tag{6.50}$$

is the metaplectic propagator for the linearized (t/N)-flow about x_{n_0}. It is important to note that, for small times (t/N) we can determine the phase θ in (6.36) to be zero, since we use $1(x) = 1$. The decomposition of the polygonal area in the above formula results from (A.13) in Appendix A. The explicit result of this infinite composition of linear transformations is finally obtained as

$$U_t(x)_{\mathrm{SC}} = 2^L |(1 + \mathscr{M})|^{-1/2} \exp[(\mathrm{i}/\hbar) S_t(x)] , \tag{6.51}$$

where $\mathscr{M}$ is the symplectic matrix for the linearized transformation between the neighbourhood of the tips of the chord $\xi(x)$ generated by $S_t(x)$ as a centre function.

This result is only valid for sufficiently short times such that the variational problem has an unique solution. Eventually, there will be bifurcations producing more chords whose number increases with time. So, generally we will have

$$U_t(x)_{\mathrm{SC}} \sim \sum_j 2^L |\det(1 + \mathscr{M}_j)|^{-1/2} \exp\{\mathrm{i}\hbar^{-1} S_{tj}(x) + \mathrm{i}\gamma_j\} , \tag{6.52}$$

where the index runs over all the contributing classical orbits. In the case of a single orbit, the corresponding *Morse index* $\gamma_j = 0$. The linearized motion around each orbit can also be used to approximate the propagation of wave packets as developed by Heller [52, 47].

Derived in this way, we understand that the semiclassical approximation has more meaning than a mere expansion in powers of $\hbar$ to some arbitrary order. It is exact in the case of linear transformations and it would be exact if we could approximate the nonlinear evolution by a sequence of linear transformations. If we keep $\mathscr{M}$ constant, while expanding the action $S_t(x)$ to second order in x, we obtain the Weyl representation of an operator that is exactly unitary. On the other hand, we easily verify that the semiclassical Weyl propagator corresponding to a nonlinear transformation is self-consistently unitary within the stationary phase approximation to the first integral in (6.27). Indeed, this is completely determined by the quadratic approximation of the phase, so that $I(\mathfrak{B})$ reduces to its metaplectic approximation in the case where there is only one chord. Should there be more chords, we insert (6.50) into (6.27) to obtain an integral that decomposes into terms of the form

$$I_{jj'} = 2^{2L}|\det(1+\mathscr{M}_j)\det(1+\mathscr{M}_{j'})|^{-1}\int \mathrm{d}x_1\mathrm{d}x_2\exp\{(\mathrm{i}/\hbar)[S_{tj}(x_1) - S_{tj'}(x_2) - 2(x_1-x)\mathfrak{J}(x_2-x)] + \mathrm{i}(\gamma_j-\gamma_{j'})\}\,. \tag{6.53}$$

The stationary-phase condition is precisely that $x_1(x)$ and $x_2(x)$ be chosen so the canonical transformation generated by $S_{tj}(x)$ be combined with that generated by $-S_{tj'}(x_2)$ to result in a new canonical transformation, as explained at the beginning of Section 2. In other words, the endpoint x_{1+} for the first evolution must coincide with x_{2-} for the second, as shown in Fig. 6.1(a). For the diagonal terms ($j=j'$), $-S_{tj}(x)$ generates the inverse canonical transformation to $S_{tj}(x)$, so the return and the outgoing orbit coincide. There may be many chords through the point x_1 in Fig. 6.1(b), but it is clearly the chord for which $x_{1-}=x$ that determines the integral with the stationary point. In this case the other centre $x_2=x_1$, whereas the chord ξ and the triangle Δ_3 collapse. There can be no stationary points for the nondiagonal terms ($j\pm j'$), because there is only one orbit reaching the corner of Δ_3 opposite to x in Fig. 6.1(a): The only backwards transformation matching onto that generated by S_{tj} is $-S_{tj}$ with $j'=j$. The principle behind this simplification is that there may be many chords through each centre for long times, but there is still an unique canonical transformation. Therefore, there is only one orbit through each endpoint, $x_\pm$, travelling forward or backward in time.

The property that $S_t(x)$ is always an odd function of t guarantees that $U_{-t}(x)=[U_t(x)]^*$ in (6.52). This is a necessary, but not a sufficient condition for $U_t(x)$ to represent an unitary operator. Still we find that there is a correspondence between classical canonical transformations generated by the Hamiltonian and semiclassical Weyl propagators (6.52) that are unitary within the stationary-phase approximation. This equivalence is not restricted to the Weyl representation, having been systematically discussed by Miller [50]. It is remarkable that the approximate unitarity of the semiclassical propagator (6.52) is not affected by the choice of phases γ_j, though these will have to be fully determined for the construction of a theory for the semiclassical evolution of Wigner functions [52]. A much simpler task is to derive the semiclassical evolution of observables in the Weyl representation. This is obtained by inserting (6.52) into (6.35) and evaluating the resulting integrals by stationary phase. The fact that the function $A(x)$ representing the observable is smooth, rather than highly oscillatory, reduces the condition of stationary phase to that of the integrals $I_{jj'}$ in (6.53). Therefore, there is only one stationary point for each of the diagonal terms. The full evaluation now

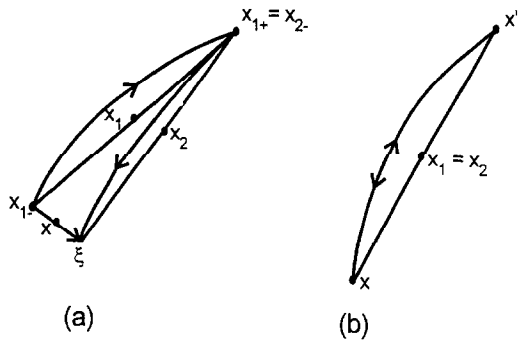


Fig. 6.1. (a) The stationary points $x_1(x)$ and $x_2(x)$ for the composition of two unitary operators in the semiclassical approximation must be such that $x_{1+}=x_{2-}$. (b) In the case that $U_2(x)=U_1^*(x)$, the only solution collapses the triangle Δ_3 into a single chord. The corner opposite x becomes its image under the Hamiltonian flow.

proceeds exactly as in (6.35) and (6.36), hence the result is that

$$A'(x)_{\mathrm{SC}} = A(x'(x)) , \tag{6.54}$$

where $x'(x)$ is the point to which x evolves in the Hamiltonian flow.

Thus, we find that observables evolve classically in the semiclassical limit of the Weyl representation. Again, this result is invariant with respect to the choice of phases in the semiclassical propagator. There will be no problems with crossing caustics in this simple limit of the Heisenberg picture.

Our discussion of the variational principle for fixed time in Section 2 revealed that there will be an unique solution for a sufficiently short interval. This is in sharp contrast to the full quantum path integral which includes all polygonal paths joining the tips of all possible chords centred on x. We obtain a somewhat intermediary situation for the Husimi propagator. Combining (5.32) with (6.47), we obtain

$$U_{\mathrm{H}}(P,Q) = \lim_{N\to\infty} \int \frac{\mathrm{d}x_N \cdots \mathrm{d}x_1}{(\pi\hbar)^{NL}} \exp\left\{ -\mathrm{i}\frac{t}{\hbar N} \sum_{n=1}^{N} H(x_n) \right\}$$

$$\times \frac{1}{(\hbar\pi)^L} \int \mathrm{d}x \exp\left\{ -\frac{w}{\hbar}(q-Q)^2 - \frac{1}{w\hbar}(p-P)^2 + \frac{\mathrm{i}}{\hbar}\Delta_{N+1}(x,x_1,\ldots,x_N) \right\}. \tag{6.55}$$

If we now recall the linear relation (2.11) of Δ_{N+1} with x, (6.55) can be integrated to yield

$$U_{\mathrm{H}}(P,Q) = \lim_{N\to\infty} \int \frac{\mathrm{d}x_n \cdots \mathrm{d}x_1}{(\pi\hbar)^{NL}} \exp\left\{ \frac{\mathrm{i}}{\hbar}\Delta_{N+1}(X,x_1,\ldots,x_N) - \frac{\mathrm{i}t}{\hbar N} \sum_{n=1}^{N} H(x_n) \right\}$$

$$\times \exp\left\{ -\frac{w}{4\hbar}\xi_q^2 - \frac{1}{4w\hbar}\xi_p^2 \right\}, \tag{6.56}$$

where the chord ξ passing through $X=(P,Q)$ depends only on the other centres $x_1,\ldots,x_N$, according to (2.10). This expression again highlights the complementarity between chords and centres. We

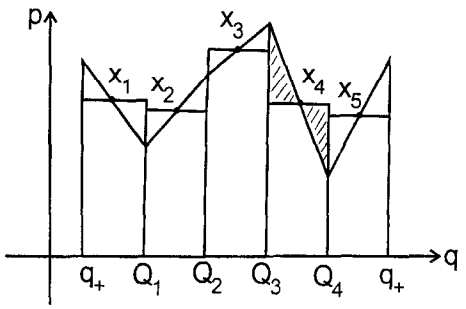


Fig. 6.2. The area between the polygonal line and the q-axis is the same as that of the succession of strips with width $Q_n - Q_{n-1}$ and height p_n, i.e. the p-coordinate of x_n.

derive the Husimi propagator by taking a local average over centres; the consequence is a Gaussian cutoff on the length of the chords. The full semiclassical limit of the Husimi propagator is the subject of ongoing work.

We have discussed the fact that the centre and the chord representations of any operator transform classically as the result of a metaplectic transformation. This property is also explicitly verified for a general Weyl propagator since the Hamiltonian and the symplectic areas determining the phase of the path integral are invariant under a canonical transformation, while the chord construction is preserved by its linearity. However, a general symplectic transformation will alter the Gaussian amplitude in the path integral for the Husimi propagator (6.56) so that this representation will not usually transform classically.

We can now derive the more familiar path integrals by taking the symmetrized Fourier transform (5.14) of (6.47). This matches the Legendre transform between variational principles that we discussed in Section 2:

$$\begin{aligned}\langle q_+|\hat{U}_t|q_-\rangle &= (2\pi\hbar)^{-L}\int \mathrm{d}p\, U_t\left(p, \frac{1}{2}(q_+ + q_-)\right)\exp\left[\frac{\mathrm{i}}{\hbar}p\cdot(q_+ - q_-)\right]\\ &= (2\pi\hbar)^{-L}\lim_{N\to\infty}\int \mathrm{d}p\frac{\mathrm{d}x_N\cdots\mathrm{d}x_1}{(\pi\hbar)^{NL}}\exp\left\{\frac{\mathrm{i}}{\hbar}\left[p\cdot(q_+ - q_-) - \frac{t}{N}\sum_{n=1}^{N}H(x_n)\right.\right.\\ &\quad \left.\left. + \Delta_{N+1}\left(\left(p, \frac{1}{2}(q_+ + q_-)\right), x_1, \ldots, x_n\right)\right]\right\}\\ &= \lim_{N\to\infty}\int \frac{\mathrm{d}x_N\cdots\mathrm{d}x_1}{(\pi\hbar)^{NL}}\delta(q_+ - q_- - \xi_q)\exp\left[\frac{\mathrm{i}}{\hbar}S'_N(x_1,\ldots,x_N)\right]. \end{aligned} \tag{6.57}$$

S' is now the symplectic area between the polygonal line (with centres at $x_1,\ldots,x_N$) and the $p=0$ plane. The δ-function ensures that the tips of this polygonal line project precisely onto q_+ and q_-. To eliminate the δ-function we redefine the coordinates of the integration variables as $q_1 = \frac{1}{2}(Q_1 + q_-),\ldots,q_n = \frac{1}{2}(Q_n + Q_{n-1}),\ldots,q_N = \frac{1}{2}(Q_N + Q_{N-1})$. Since the area under the polygonal line is the same as that of the succession of strips $p_n\cdot(Q_n - Q_{n-1})$, as shown in Fig. 6.2, we can

rewrite

$$S'_N = \sum_n p_n \cdot (Q_n - Q_{n-1}) - \frac{t}{N}\left\{H\left(p_1, \frac{q_- + Q_1}{2}\right) + \cdots \right.$$
$$\left. + H\left(p_n, \frac{Q_n + Q_{n-1}}{2}\right) + \cdots + H\left(p_n, \frac{Q_{N-1} + Q_N}{2}\right)\right\}. \tag{6.58}$$

The argument of the δ-function then reduces to $Q_N - q_+$, so that the phase for each path of steps at Q_1, $Q_2 \ldots$ is just

$$S'_N = S_N(q_+, q_-) = \sum_n p_n(Q_n - Q_{n-1}) - \frac{t}{N}\left\{H\left(p_1, \frac{q_- + Q_1}{2}\right) + \cdots \right.$$
$$\left. + H\left(p_n, \frac{Q_n + Q_{n-1}}{2}\right) + \cdots + H\left(p_n, \frac{Q_{N-1} + q_+}{2}\right)\right\}, \tag{6.59}$$

i.e. the generating function (2.25). Since the Jacobian of the coordinate transformation is 2^{NL}, we have

$$\langle q_+|\hat{U}|q_-\rangle = \lim_{N\to\infty} \int \frac{\mathrm{d}p_1 \cdots \mathrm{d}p_N \, \mathrm{d}Q_1 \cdots \mathrm{d}Q_{N-1}}{(2\pi\hbar)^{NL}} \exp\left\{\frac{\mathrm{i}}{\hbar} S_N(q_+, q_-)\right\}, \tag{6.60}$$

the usual definition of the path integral in phase space [8].

The advantage of the present derivation of path integrals is that there is no reliance on special properties of the Hamiltonian. This is explicitly realized as the Weyl representation of the quantum operator, which will equal the classical Hamiltonian only in simple cases. However, we generalize the usual path integrals to arbitrary Hamiltonians by writing the actions in terms of the Weyl transforms rather than the classical Hamiltonians. On the other hand, the equivalence between the centre path integral and the one defined by (q_-, q_+) allows us to incorporate in the former the difficult discussions about the convergence of the latter. Of course, it will be difficult to verify convergence for the complete generality of path integrals that can be defined in the Weyl representation, though this question is now seen in a clearer light than in the original presentation [11].

7. Stationary states

One of the outstanding problems of quantum mechanics is the development of a full theory for the semiclassical limit of stationary states. We have presented a simple derivation of such a limit for the evolution operator for an arbitrary quantum Hamiltonian, but the traditional theory for energy eigenstates is restricted to classically integrable systems. These have L independent constants of the motion commuting with the Hamiltonian. Classically, this restricts the motion to invariant L-dimensional tori in phase space. Arbitrary small perturbations will destroy this perfect foliation of phase space even if many tori are preserved. For stronger perturbations most of phase space becomes chaotic and we must completely abandon our reliance on invariant surfaces. Arnold's book [15] is the standard reference on the perturbation of classical integrable systems. For the problems concerning their semiclassical limit see [19, 33–35].

The best course is to rely on the semiclassical theory that we have obtained for the evolution operator, based on its spectral decomposition,

$$\hat{U}_t = \exp\{-(\mathrm{i}t/\hbar)\hat{H}\} = \sum_n |n\rangle\langle n| \exp\{-(i/\hbar)E_n t\}\,, \tag{7.1}$$

where $|n\rangle$ are now taken to be the eigenstates of the Hamiltonian. Taking appropriate transforms of this formula we can extract information about the eigenstates and the energy spectrum. In particular, we have the *spectral operator*

$$\begin{aligned} -\frac{1}{\pi}\,\mathrm{Im}(E+\mathrm{i}\varepsilon-\hat{H})^{-1} &= \frac{1}{\pi\hbar}\,\mathrm{Re}\int_0^\infty \mathrm{d}t \exp\left\{\frac{\mathrm{i}}{\hbar}(E-\hat{H})t - \frac{\varepsilon t}{\hbar}\right\} \\ &= \frac{1}{\pi}\,\frac{\varepsilon}{(E-\hat{H})^2+\varepsilon^2} \equiv \delta_\varepsilon(E-\hat{H}) = \sum_n |n\rangle\delta_\varepsilon(E-E_n)\langle n|\,, \end{aligned} \tag{7.2}$$

where $\delta_\varepsilon(x)$ is a normalized function whose width ε can be taken to be arbitrarily small, so that δ_ε tends to the Dirac δ-function as $\varepsilon \to 0$. Notice though that in this limit we require knowledge of the evolution operator for all time.

By taking the trace of (7.2), we obtain the smoothed density of states:

$$\sum_n \delta_\varepsilon(E-E_n) = -(1/\pi)\,\mathrm{Im}\,\mathrm{Tr}(E+\mathrm{i}\varepsilon-\hat{H})^{-1}\,. \tag{7.3}$$

We can also sum the wave intensities over a narrow energy range:

$$-(1/\pi)\langle q_+|\mathrm{Im}(E+\mathrm{i}\varepsilon-\hat{H})^{-1}|q_-\rangle = \sum_n \delta_\varepsilon(E-E_n)\langle q_+|n\rangle\langle n|q_-\rangle\,. \tag{7.4}$$

Finally, by taking the Weyl transform of (7.2), we define, as Berry [14], the *spectral Wigner function*,

$$W(x;E,\varepsilon) = (2\pi\hbar)^L \sum_n \delta_\varepsilon(E-E_n)W_n(x) = (1/\pi\hbar)\,\mathrm{Re}\int_0^\infty \mathrm{d}t \exp\{(\mathrm{i}/\hbar)t(E+\mathrm{i}\varepsilon)\}U_t(x)\,. \tag{7.5}$$

Let us first consider the energy spectrum. Since the trace is merely the integral of the Weyl representation (5.15),

$$\sum_n \delta_\varepsilon(E-E_n) = (1/\pi\hbar)\,\mathrm{Re}\int_0^\infty \mathrm{d}t \exp\{(\mathrm{i}/\hbar)t(E+\mathrm{i}\varepsilon)\}\int \mathrm{d}x(2\pi\hbar)^{-L}U_t(x)\,. \tag{7.6}$$

Inserting the path integral (6.47) into (7.6) and recalling that the x dependence of $U_t(x)$ is just a phase equal to $\xi \wedge x$, where ξ is the side of the polygon Δ_{N+1} centred on x, we obtain, using (6.20),

$$\begin{aligned} \mathrm{Tr}\,\hat{U}_t = \int \frac{\mathrm{d}x U_t}{(2\pi\hbar)^L}(x) &= 2^L\left(\frac{1}{\pi\hbar}\right)^{(N-1)L} \lim_{N\to\infty}\int \mathrm{d}x_N \cdots \mathrm{d}x_1\,\mathrm{d}x\,\delta(\xi) \\ &\quad\times \exp\left\{\frac{i}{\hbar}\Delta_N(x_1,\ldots,x_N) - \frac{t}{N}\sum_{n=1}^N H(x_n)\right\}\,. \end{aligned} \tag{7.7}$$

The arbitrary side of the polygon has thus collapsed. In the semiclassical limit, all the other remaining sides will be small and parallel to a classical trajectory, which must therefore be periodic. If we now perform the time integration in (7.6), we find that the density of states will depend on all the periodic orbits with periods below that allowed by the cutoff factor.

Before deriving the semiclassical amplitudes of the contributions of the periodic orbits to the density of states, let us view the general features of the spectral Wigner function. Inserting (6.47) into (7.5), we obtain

$$(2\pi\hbar)^L \sum_n \delta_\varepsilon(E-E_n)W_n(x) = 2\mathrm{Re}\lim_{N\to\infty}\int \frac{\mathrm{d}x_N\cdots\mathrm{d}x_1}{(\pi\hbar)^{NL}} \times \exp\left\{\frac{i}{\hbar}\Delta_{N+1}(x,x_1,\ldots,x_N)\right\}\delta_\varepsilon\left(E-\frac{1}{N}\sum_{n=1}^{N}H(x_n)\right). \tag{7.8}$$

If we neglected the width of the δ-function in this expression, it would seem that all the centres of the polygon, except x itself, would be forced onto the energy shell $\mathscr{S}$ of energy E. Among all these possible polygons, we would obtain a stationary action Δ_{N+1} for the one that coincides with the classical trajectory that solves the energy variational problem studied in Section 3. We recall that these trajectory solutions arise in continuous families as the centre x is varied and that, as x approaches the shell, the trajectory either becomes very short, or it approaches a periodic orbit. It is important to note that, contrary to the formula for the density of states, these semiclassical *scars* of periodic orbits, that affect the spectral Wigner function close to $\mathscr{S}$, involve only the reduced subset of periodic orbits that intersect the small centre section studied in Section 3.

However, the δ-function in (7.8) does not really require that all the centres lie on the same energy shell. Indeed, it is only the energy average that is restricted. In the semiclassical limit this is sufficient, since the stationary trajectory must lie entirely within a single shell if the Hamiltonian is autonomous. Even so, there is a real need to relax the constraint on the allowed paths. Though there is not yet a full theory of tunnelling in the Weyl representation, we cannot hope to deal with the Wigner function in a double well potential without allowing chords to join the separate energy shells in phase space.

Let us now evaluate the semiclassical approximation of the spectral Wigner function, i.e., we insert the semiclassical approximation of the Weyl propagator (6.52) into (7.5):

$$W(x;E,\varepsilon) = \frac{2^L}{\pi\hbar}\sum_j \mathrm{Re}\int_0^\infty \mathrm{d}t \frac{\mathrm{e}^{-\varepsilon t/\hbar}}{|\det[1+\mathscr{M}_j(x)]|^{1/2}} \exp\left\{\frac{\mathrm{i}}{\hbar}[S_{tj}(x)+Et]+\mathrm{i}\lambda_j\right\}. \tag{7.9}$$

This integral will be dominated by its points of stationary phase, such that

$$(\mathrm{d}/\mathrm{d}t)[S_{tj}(x)+Et]=0. \tag{7.10}$$

This determines the $t_j(E)$ for which

$$S_{tj}(x)+Et_j(E)=S_{Ej}(x) \tag{7.11}$$

is the energy-dependent action defined in Section 3, where we found that, if x lies inside the energy shell $\mathscr{S}$, there will be at least one chord in the centre section uniquely determined by the point x.

If all the chords centred on x and having tips on the energy shell $\mathscr{S}$ are sufficiently separated, we may evaluate (7.9) by the method of stationary phase, which substitutes the neighbourhood of each stationary point by a Gaussian integral. Taking

$$\mathrm{d}^2 S_{tj}/\mathrm{d}t^2 = \mathrm{d}E_j/\mathrm{d}t = (\mathrm{d}t_j/\mathrm{d}E)^{-1}\,, \tag{7.12}$$

we obtain the semiclassical Wigner functions as

$$W(x;E,\varepsilon) = \frac{2^{(L+1)}}{(2\pi\hbar)^{\frac{1}{2}}} \sum_j \{(\mathrm{d}t_j/\mathrm{d}E)|\det[1+\mathscr{M}_j]|^{-1}\}^{1/2} \mathrm{e}^{-\varepsilon t_j/\hbar} \cos[\hbar^{-1} S_{E_j}(x) + \lambda_j']\,. \tag{7.13}$$

Before discussing this formula, let us simplify the amplitude of each of the classical contributions. Recall that we can choose coordinates near the orbit γ_j, such that one coordinate is the energy E and the conjugate coordinate is the time along the orbit, t. In Section 3, these were, respectively, called X_{2L} and X_{2L-1} close to the shell. In the transformation $\delta x_+ = \mathscr{M}_j \delta x_-$, we evidently obtain $\delta E_+ = \delta E_-$ and $\delta t_+ = \delta t_-$. So, using the coordinates $x = (x', \delta t, \delta E)$ we have

$$\mathscr{M}_j = \left(\begin{array}{c|cc} m_j & 0 & \\ \hline 0 & 1 & 0 \\ & 0 & 1 \end{array}\right) \tag{7.14}$$

and so

$$\det[1+\mathscr{M}_j] = 4\det[1+m_j]\,, \tag{7.15}$$

where m_j is now the $(2L-1)\times(2L-1)$ symplectic matrix for the centre map determined by the jth orbit or chord. Hence, we finally obtain

$$W(x;E,\xi) = \frac{2^{(L)}}{(2\pi\hbar)^{1/2}} \sum_j \left\{\frac{\mathrm{d}t_j}{\mathrm{d}E} \det[1+m_j]^{-1}\right\}^{1/2} \mathrm{e}^{-\varepsilon t_j/\hbar} \cos[\hbar^{-1} S_{Ej}(x) + \lambda_j']\,. \tag{7.16}$$

To understand this semiclassical expression for the spectral Wigner function, we must return to the discussion of the chord solutions of the centre map. If the energy shell is closed and convex, there are no chords when x lies outside of $\mathscr{S}$. Once x is inside $\mathscr{S}$ there will be solutions, but with small chords if x is close to $\mathscr{S}$. For one freedom there will be a single chord, though there will be many possible windings around the periodic orbit, which in this case we must identify with the shell. Berry [7] showed that in this case we can infer the correct quantization condition from the identification of the Wigner function for all the different windings. His work dealt directly with a single Wigner function, rather than the spectral Wigner function; this can be sampled in (7.16) by taking $\varepsilon \to 0$.

When $L \geq 2$, we can no longer identify the energy shell with a single periodic orbit. If x lies close to the shell, we will always find a small chord connecting the tips of a short orbit for the centre section. We will also have short chords connecting the tips of orbits that wind very closely around periodic orbits. In the limit $\varepsilon \to 0$, there will always be periodic orbits traversing the section, no matter how small. However, for any finite ε, the period of most of these orbits will be too long for them to contribute to (7.16). Overall, the majority of periodic orbits never enter a small section.

The simple semiclassical approximation to the spectral Wigner function breaks down when x is taken very close to the shell, for then the time t_0 of the short orbit approaches the limit of integration.

Instead of dealing directly with this problem, it is preferable to double the range of integration and to use a smooth cutoff function at the origin. Thus, we substitute (7.9) by

$$W(x;E,\varepsilon)\simeq\frac{2^L}{2\pi\hbar}\int_{-\infty}^{\infty}\mathrm{d}t\frac{\mathrm{e}^{-\varepsilon^2t^2/\hbar^2}}{|\det[1+\mathscr{M}(x)]|^{1/2}}\exp\left\{\frac{\mathrm{i}}{\hbar}[S_t(x)+Et]\right\}, \tag{7.17}$$

so that, instead of interacting with the origin, the stationary point t_0, interacts with its time-reversal pair at $-t_0$. This is typical of the way that the stationary points of the Airy integral

$$\int_{-\infty}^{\infty}\exp\left\{\mathrm{i}\left[\alpha\frac{t^3}{3}-\beta t\right]\right\}\mathrm{d}t=|\alpha|^{-1/3}2\pi Ai[-(\alpha)^{-1/3}\beta] \tag{7.18}$$

interact as the free parameters α and β are changed. Hence, we can map the integral (7.17) onto (7.18) using the method of uniform approximation [36, 37], choosing the parameters in (7.18) so that the stationary phases of both integrals coincide, i.e.,

$$\beta=-(3/2\hbar t_0)S_E(x)\quad\text{and}\quad\alpha=\beta/t_0^2, \tag{7.19}$$

where t_0 is the (positive) stationary point for the short orbit and $S_E(x)$ is the corresponding action (7.11). The result is the uniform expression for the spectral Wigner function

$$W(x;E,\varepsilon)=\frac{2^{L-1}}{\hbar}\frac{\mathrm{e}^{-\varepsilon^2t_0^2/\hbar^2}}{|\det[1+m_0]|^{1/2}}\left(\frac{2\hbar}{3S_E(x)}\right)^{1/3}t_0Ai\left[-\left(\frac{3S_E(x)}{2\hbar}\right)^{2/3}\right]. \tag{7.20}$$

Strictly this last formula only holds if the cutoff parameter ε is large enough to cancel all orbits except the shortest. Otherwise, we must add the contributions of the other orbits already obtained in (7.16). If the evaluation point x is brought away from the shell, we may expand the Airy function asymptotically as a cosine and thus retrieve the previous semiclassical approximation. More interesting is the limit where x approaches $\mathscr{S}$. This is just the situation analysed at the end of Section 3, where we derived that the action $S_E(x)\to S_{\mathrm{in}}(x)$ for the orbit of the in-map within the small centre section, where

$$S_{\mathrm{in}}(x)\simeq\tfrac{1}{12}t_0^3\dot{x}\mathscr{H}\dot{x}=\tfrac{4}{3}2^{1/2}(E-H(x))^{3/2}/[\dot{x}\mathscr{H}\dot{x}]^{1/2}. \tag{7.21}$$

Also m_0 becomes the identity map, so $\det[1+m_0]=2^{(2L-2)}$.

We thus resolve the indeterminacy in (7.20), where both $S_E(x)$ and t_0 tend to zero, as

$$W(x;E,\varepsilon)\underset{x\to\mathscr{S}}{\longrightarrow}2(\hbar^2\dot{x}\mathscr{H}\dot{x})^{-1/3}Ai\left[-2\frac{E-H(x)}{(\hbar^2\dot{x}\mathscr{H}\dot{x})^{1/3}}\right]. \tag{7.22}$$

We therefore find that the spectral Wigner function oscillates inside the energy shell, arrives at a peak just inside the shell and decays exponentially outside. This behaviour was deduced by Berry [56], having derived it for the Wigner function itself [7] when $L=1$. The caustic structure and the structure of the Wigner function close to a quantized torus of an integrable system are discussed in Refs. [38, 39]. If we convolute (7.22) with a Gaussian window, we find that the spectral Husimi function is appreciable only very close to the energy shell. This semiclassical result generalizes our previous deduction for the harmonic oscillator to any convex energy shell for an arbitrary number degrees of freedom.

If instead of integrating (7.22) with a Gaussian window, we integrate it straight, we will obtain the trace of the spectral operator (7.3). Semiclassically the integral will be dominated by the region close to the shell. Therefore, if we keep $(\dot{x}\mathscr{H}\dot{x})$ constant and recall that the integral of the Airy function is unity, we obtain

$$\sum_n \delta_\varepsilon(E - E_n) \simeq (2\pi\hbar)^{-L}\, \mathrm{d}V/\mathrm{d}E\,, \tag{7.23}$$

where V is the phase-space volume of the energy shell $\mathscr{S}$. Integrating over the energy, we obtain the approximate number of eigenstates below the energy E as

$$N(E) = V(E)/(2\pi\hbar)^L\,. \tag{7.24}$$

This is the well known *Weyl rule* for the density of states, that the number of states below the energy E is the phase-space volume of the shell divided by the minimum uncertainty volume. Evidently, this is a smoothed approximation, not only because (7.24) should represent a sequence of unit steps, but also we recall that it is only valid in the limit where the smoothing ε is large enough to cancel the contribution of all other orbits. This is only the first term in an asymptotic series in powers of $\hbar$ for the smooth approximation of the density of states. (See [40] for a derivation based on the Weyl representation.)

If we reduce the smooothing, there will be new contributions to the spectral Wigner function. If x is inside the shell and very close to it, we must take into account pairs of chords for orbits that wind very close to the periodic orbits that intersect the centre section. These pairs of chords subtend actions very close to that of the periodic orbit $\mathscr{P}_E$ as was derived in Eqs. (3.10) and (3.11). The periods of this pair of contributions also approximate the periods of the periodic orbit itself. In the limit as x touches the periodic orbit in the shell, or in the limit when the contributing chords enter the centre section at the border of the equatorial plane, both the chords will determine orbits with the same period. We therefore have to deal with Airy-like contributions to the spectral Wigner function, very similar to those already deduced for the short orbit.

Close to the energy shell we can assume that both chords ξ_{long} and ξ_{out} project onto the same end points $(x \pm X'_+)$ in the equatorial surface, defined in Section 3. Furthermore, we can identify the corresponding mappings m_j. Therefore, the contribution of this pair of chords near the jth periodic orbit is

$$W_j(x,E,\varepsilon) = \frac{2^{L-1}}{2\pi\hbar}\,\frac{\mathrm{e}^{-\varepsilon\tau_j/\hbar}}{|\det[1+m_j]|^{1/2}}\cos\left[\frac{\bar{S}_j}{\hbar}+\lambda_j\right] \times \int_{-\infty}^{\infty}\mathrm{d}t\,\exp\left[\frac{\mathrm{i}}{\hbar}\,(\delta S_t + Et)\right]\,, \tag{7.25}$$

where $\bar{S}_j = \frac{1}{2}(S_{\mathrm{out}} + S_{\mathrm{long}})$ and $\delta S_t + Et = \frac{1}{2}(S_{\mathrm{long}} - S_{\mathrm{out}})$ for this pair of orbits at the stationary time. Thus, the time origin has been brought over to the average time τ_j. Integrating, we obtain

$$W_j(x;E,\varepsilon) = \frac{2^{L-1}}{\hbar}\,\frac{\mathrm{e}^{-\varepsilon\tau_j/\hbar}}{|\det[1+m_j]|^{1/2}}\cos\left\{\frac{\bar{S}_j}{\hbar}+\lambda_j\right\} \times t_{\mathrm{in}}\left(\frac{2\hbar}{3\delta S_j}\right)^{1/3} Ai\left[-\left(\frac{3\delta S_j}{2\hbar}\right)^{2/3}\right]\,, \tag{7.26}$$

corresponding to (7.20). Here $t_{\rm in}$ is the time for one of the tips of the out-orbit to complete another in-traversal of the centre section.

It is worth noting that, so far, we have not invoked the periodic orbit, i.e., the structure of (7.26) arises merely from the need to deal with a pair of nearly degenerate actions. However, we can now insert the approximations (3.19) and (3.20) for $S_{\rm out}$ and $S_{\rm long}$ into (7.26), obtaining

$$W_j(x;E,\varepsilon) = \frac{2^L}{|\det[1+m_j]|^{1/2}} \mathrm{e}^{-\varepsilon\tau_j/\hbar} \cos\left\{ \mathscr{S}_j + X'_p \mathfrak{B}_j X'_p - \frac{4}{3}\, 2^{1/2} \frac{[E-H(x'_p)]^{3/2}}{[\dot{X}'_p \mathscr{H} \dot{X}'_p]^{1/2}} \right\}$$
$$\times (\hbar^2 \dot{X}'_+ \mathscr{H} \dot{X}'_+)^{-1/3} Ai\left\{ -2 \frac{E-H(x'_+)}{[\hbar^2 \dot{X}'_+ \mathscr{H} \dot{X}'_+]^{1/3}} \right\}. \tag{7.27}$$

Now the time τ_j refers to the periodic orbit crossing the centre section, $\mathscr{S}_j$ is its action (with the multiplicity appropriate to the map considered) and $x'_p = x + X'_p$ is the point where it crosses the approximately plane equatorial surface. Thus, the cosine term exhibits a small correction with respect to the Berry theory [14]. Even so the general agreement with that result is emphasised by recalling that $\mathfrak{B}_j$ is the symmetric matrix that parametrizes the symplectic matrix m_j according to (1.20) and that $\dot{X}\mathscr{H}\dot{X} = \ddot{X} \wedge \dot{X}$.

The essential improvement of the present theory involves the subtle exchange of

$$x \to x'_+ = x + \mathfrak{J}\mathfrak{B}_p X'_p \tag{7.28}$$

in the argument of the Hamiltonian in the Airy function. Because of this difference, the movement of x away from the periodic orbit will lead it through a caustic even if it remains within the energy shell. Beyond the caustic the contribution of this pair of chords becomes evanescent.

Thus, we find that there are both important similarities and differences between the contributions of short orbits and pairs of orbits close to periodic orbits. In both cases, a fold caustic separates oscillatory contributions to the Wigner function from evanescent regions, obtaining maximum amplitude close to the caustic. However, the caustic for the short orbits is identified with the energy shell and there are no phase oscillations along the caustic. In contrast, the periodic orbit caustic only touches the shell along the orbit itself, then it receeds into the shell, widening smoothly according to (3.21). Moroever, the cosine term in (7.27) leads to phase oscillations along the caustic.

Let us now consider the spectral Husimi function. Taking the view that this is just the Gaussian smoothing of the spectral Wigner function, we see that the oscillations inside the shell will be cancelled even along the caustic, but not the scars of the periodic orbits, very near the shell. Indeed, we have seen that the Husimi propagator will be a path integral only over small chords. In the semiclassical limit of the spectral Husimi function, these small chords must connect the tips of trajectories on the same energy shell. No matter how much we reduce the energy smoothing, we never obtain an appreciable amplitude far from the energy shell. We can thus justify a weakened version of the hypothesis of Voros and Berry [41, 42]. In its original form it amounted to the conjecture that the Wigner function would be uniformly peaked along the energy shell. Now we understand that it is the Husimi function which should be so peaked. However, the amplitude is not uniform because of the contributions of the periodic orbits.

Returning to the density of states, we can now deduce the oscillatory corrections to (7.23) by integrating over the contributions of the scars in the energy shell. Of course, these scar amplitudes are only evaluated locally and we have seen that they become exponentially damped rather than

oscillatory far from the periodic orbit. However, in both cases we only obtain an appreciable contribution to the integral close to the periodic orbit, so we shall extrapolate the validity of these amplitudes for all x.

By reverting to the special coordinates such that $x = (x', t, E)$ and taking X' as the difference in coordinates with respect to the periodic orbit, we obtain the contribution of each scar to the trace of the spectral operator as approximately

$$\rho_j = \int \frac{\mathrm{d}x}{(2\pi\hbar)^L} W_j(x; E, \varepsilon)$$
$$= \frac{2^L}{(\pi\hbar)^L} \frac{\mathrm{e}^{-\varepsilon\tau_j/\hbar}\tau_{jp}}{|\det[1 + m_j]|^{1/2}} \cos\left(\frac{1}{\hbar}\mathscr{S}_{E_j} + \lambda'_j\right) \int \mathrm{d}X' \exp\left[-\frac{1}{\hbar} X' B_j X'\right] . \tag{7.29}$$

Here we have used the fact that the Airy function integrates to unity and we neglected the energy dependence of all the other terms. The integral over t is just τ_{jp}, the period of the *primitive periodic orbit*, no matter how many windings we may attribute to the jth periodic orbit. This is because we are performing a spatial integral in which we have merely used the time along the orbit as an useful coordinate; integrating over all the phase space requires only a single winding around the orbit. To complete the evaluation of the periodic orbit contribution, we have merely to evaluate the Gaussian integral. Recalling from Section 1 the Cayley parametrization of symplectic matrices, we obtain immediately

$$\det \mathfrak{B}_j = \det[m_j - 1]/\det[m_j + 1] , \tag{7.30}$$

so that

$$\rho_j = \frac{\tau_{jp}\mathrm{e}^{-\varepsilon\tau_j/\hbar}}{\det[1 - m_j]^{1/2}} \cos\left\{\frac{1}{\hbar}\ \mathscr{S}_j(E) + \lambda'_j\right\} . \tag{7.31}$$

Adding the contributions of all the periodic orbits to the smooth Weyl term, we obtain the celebrated *Gutzwiller trace formula* [34, 43]

$$\sum_n \delta_\varepsilon(E - E_n) \approx \left(\frac{1}{2\pi\hbar}\right)^L \frac{\mathrm{d}V}{\mathrm{d}E} + \sum_j \rho_j , \tag{7.32}$$

where the sum over the periodic orbits includes all the repetitions of all the primitive periodic orbits in the energy shell $\mathscr{S}$ of energy E. There is no problem in using (7.31) to obtain accurate estimates of the smoothed density of states if the damping factor excludes all but a few periodic orbits with short periods. In this limit the density will be a smooth function which does not allow us to sample individual states.

The problem with attempting to go beyond this is that we must then reduce the smoothing to allow many periodic orbits. For chaotic systems the number of periodic orbits increases exponentially with period, whereas the amplitudes of the contributions decrease exponentially with period. Therefore, (7.32) is divergent, or at best conditionally convergent in the limit $\varepsilon \to 0$.

There are many ways that we can deal with this difficulty of obtaining individual energy levels. The simplest is to use Gaussian smoothing as in (7.20), which will override any exponential divergence in (7.32) for arbitrarily small ε. Berry and Keating [44] have advanced more sophisticated resummation methods which allow us to obtain good estimates of individual levels from finite sums

of periodic orbits. The number of terms increases in the semiclassical limit, because the average spacing of levels (the inverse of (7.23)) decreases as $\hbar \to 0$. (The energy width, ε, for the cutoff needs to be of the order of the average spacing $\overline{\delta E}$). It is natural to interpret these results as meaning that the periodic orbits up to the time $\hbar/\varepsilon$ contain the information that is necessary to determine the individual energy level.

It is tempting to extrapolate these results so as to resum the periodic orbit contributions for the individual Wigner functions, $W_n(x)$, in (7.4). This is more problematic, because we do not know a priori the energy E_n. Indeed, we should obtain a sharp peak in $W(x; E, \varepsilon)$ whenever $E = E_n$ for all x! Evidently, it will be very difficult for our semiclassical theory to achieve this result for points that are outside of the energy shell, $\mathscr{S}_{E_n}$, since the contributions of the Weyl term (short orbits) and the periodic orbits are all exponentially damped. But even inside and close to the energy shell, the great majority of the periodic orbits will also be exponentially damped. It is only the subset of periodic orbits that penetrate the small centre section that will make undamped contributions. If we allow the period of these orbits to be arbitrarily large, it is possible that they will behave like a typical sample, providing an energy peak at exactly the same energy as the full set of periodic orbits, but this places a new restriction on the relation between the value of $\hbar$ and the minimum volume of the centre section that we can allow. Hence, it is not possible to resum the orbits that are not part of the local subset within the present theory as attempted by Agam and Fishman [45]. It would be necessary to restrict the sum to those orbits that penetrate the centre section, though it is hard to verify that this produces the same poles.

We could try to avoid this problem by moving the evaluation point x well into the energy shell. Certainly, we will then have a much more representative subset of periodic orbits intersecting the centre section. However, we then have to face a new problem since the orbits that contribute with sizeable chords to (7.13) will no longer be periodic. It is true that we obtain each of the contributions as continuous families that include periodic orbits or short orbits as we bring x onto different points of the shell (as explained in Section 3) but the actions of each contribution will differ by many multiples of $\hbar$ from that of $\mathscr{S}_{E_j}$. It would then be necessary to show that the sum (7.13), over contributions with very different phases from those of the periodic orbit sum, have dominant contributions when the energy shell has the correct eigenenergy. Though this is not impossible and indeed it was shown to be true by Berry for the case of a single freedom [7], it is hard to demonstrate when $L \geq 2$.

It may seem that we cannot use the existing semiclassical theory as a basis for obtaining local knowledge of individual eigenstates. We have seen that there is an exact path integral for the trace of the propagator and that its semiclassical limit depends on all the periodic orbits with the same period. Thus we could obtain the Gutzwiller trace formula by taking the transform of this semiclassical trace, even if we could not define the semiclassical limit of the propagator itself. Ozorio de Almeida and da Luz [46] have produced an example where the operations of taking the trace and taking the semiclassical limit do not commute for the baker's map. The present analysis of the semiclassical limit of the spectral operator and its trace suggests that commutativity could break down for features that depend on a sufficiently long time of propagation.

Alternatively, it may turn out that the program of orbit resummation does indeed overcome the challenge of providing the same poles for the Green's function at all points inside the energy shell. In so doing it should reveal a few of the secrets of its subtle nature. Present work indicates that the Fredholm method [53] can be adapted so as to encompass, the intricate geometrical constructions of the centre section, providing a resummed theory for the scars of individual Wigner functions.

Appendix A. Polygons in phase space

A sequence of n translations in phase space $T_{\xi_1}, T_{\xi_2}, \ldots, T_{\xi_n}$ define an $(n+1)$-sided polygon. Its symplectic area is obtained from Fig. A.1 as

$$D_{n+1}(\xi_1, \ldots, \xi_n) = \tfrac{1}{2}[(\xi_1 \wedge \xi_2) + (\xi_1 + \xi_2) \wedge \xi_3 + \cdots (\xi_1 + \cdots + \xi_{n-1}) \wedge \xi_n], \tag{A.1}$$

a bilinear quadratic form involving all the components of each of the chords ξ_j. As usual, we can define this in terms of a symmetric matrix acting on the $(2L)^n$-dimensional vectors $(\xi_1, \xi_2, \ldots, \xi_n)$. To this end, we construct the matrix with n diagonal blocks of J:

$$\mathscr{J} = \begin{pmatrix} \mathfrak{J} & 0 & 0\cdots \\ 0 & \mathfrak{J} & 0\cdots \\ 0 & 0 & \mathfrak{J} \\ \vdots & \vdots & \vdots \end{pmatrix} \tag{A.2}$$

and another $L^n \times L^n$ matrix

$$\mathit{\Pi}_n = \begin{pmatrix} 0 & -1 & -1\cdots \\ 1 & 0 & -1\cdots \\ 1 & 1 & 0 \\ \vdots & \vdots & \vdots \end{pmatrix}, \tag{A.3}$$

so that

$$D_{n+1} = \tfrac{1}{4}(\xi_1, \ldots, \xi_n)\mathscr{J}\mathit{\Pi}_n(\xi_1, \ldots, \xi_n). \tag{A.4}$$

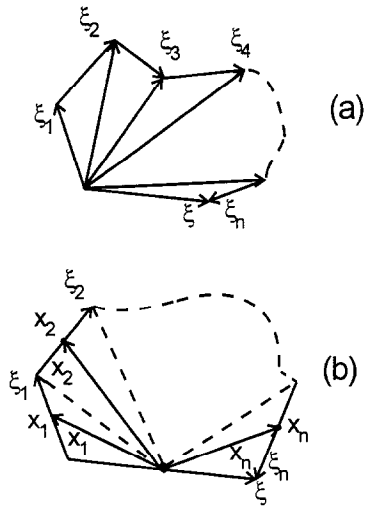


Fig. A.1. A set of n vectors $\xi_1, \ldots, \xi_n$ determines an unique $(n+1)$-sided polygon. (a) Its symplectic area is obtained by adding the successive triangles as each new vector is added to the sequence. (b) An alternative tecelation of the polygon into triangles with the common vertex x leads to the expression (A.8) for the symplectic area, where the X_j are the positions of the centres x_j, taking x as the origin.

Since $\mathcal{J}$ and Π_n are antisymmetric matrices,

$$\mathcal{J}\Pi_n = \Pi_n\mathcal{J}_n = \left(\begin{array}{c|c|c} 0 & -\mathfrak{I} & -\mathfrak{I}\cdots \\ \hline \mathfrak{I} & 0 & -\mathfrak{I}\cdots \\ \hline \mathfrak{I} & \mathfrak{I} & 0 \\ \vdots & \vdots & \end{array}\right), \tag{A.5}$$

is a symmetric matrix.

The matrix Π_n is easily seen to specify the positions of the centres of a polygon, $x_j = x + x_j$, as a function of the sides.

$$(x_1,\ldots,x_n) = \tfrac{1}{2}\Pi_n(\xi_1,\ldots,\xi_n), \tag{A.6}$$

as can be seen in Fig. A.1(b). Hence, according to the discussion at the end of Section 2, Π_n can only be inverted for even n, i.e. in the case of an odd-sided polygon. Using (2.10) for each of the sides, while remembering that the special side ξ has the opposite orientation to all the others, yields

$$\Pi_n^{-1} = \left(\begin{array}{c|c|c|c} 0 & 1 & -1 & 1\cdots \\ \hline -1 & 0 & 1 & -1\cdots \\ \hline 1 & -1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \end{array}\right) \quad (n \text{ even}). \tag{A.7}$$

Combining now (A.4) with (A.6), we have

$$D_{n+1} = \tfrac{1}{2}(\xi_1,\ldots,\xi_n)\mathcal{J}_n(X_1,\ldots,X_n), \tag{A.8}$$

which has the obvious interpretation as the sum of the triangles that each of the sides ξ_j determines with x, as shown in Fig. A.1(b). Finally, by eliminating the ξ_j's in (A.7), we obtain

$$D_{n+1} = \Delta_{n+1}(0,X_1,\ldots,X_n) = (X_1,\ldots,X_n)\mathcal{J}_n\Pi_n^{-1}(X_1,\ldots,X_n), \tag{A.9}$$

where

$$\mathcal{J}_n\Pi_n^{-1} = \Pi_n^{-1}\mathcal{J}_n = \left(\begin{array}{c|c|c} 0 & \mathfrak{I} & -\mathfrak{I}\cdots \\ \hline -\mathfrak{I} & 0 & -\mathfrak{I}\cdots \\ \hline 1 & -\mathfrak{I} & 0 \\ \vdots & \vdots & \vdots \end{array}\right). \tag{A.10}$$

Therefore, in the case of odd-sided polygons, their area is a bilinear quadratic form of the centres of their sides, as of the vectors that make up the sides themselves. Though the simplest formula is (A.8), it is redundant to determine both the $\xi_{j'S}$ and $X_j = x_j - x$.

Adding arbitrary displacements to x_j, we will alter the area of the polygon:

$$\begin{aligned}\Delta_{n+1}(0,X_1+\delta X_1,\ldots,X_n+\delta X_n) = \Delta_{n+1}(0,X_1,\ldots,X_n) + 2(\delta X_1,\ldots,\delta X_n)\mathcal{J}_n\Pi_n^{-1}(X_1,\ldots,X_n) \\ + (\delta X_1,\ldots,\delta X_n)\mathcal{J}_n\Pi_n^{-1}(\delta X_1,\ldots,\delta X_n).\end{aligned} \tag{A.11}$$

By inspection of (A.10), we find that the coefficient of each term that is linear in δX_j does not dependent on X_j itself. This is also true of ξ_j, i.e. we found in (2.10) that, if we determine a polygon

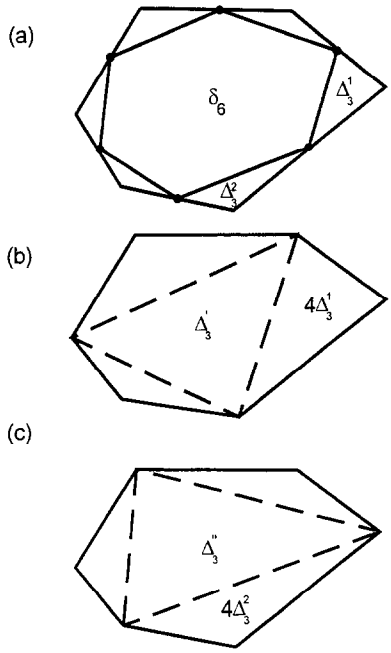


Fig. A.2. Decompositions of a hexagon: (a) Subtracting the inscribed hexagon δ_6, the remaining symplectic area is subdivided into six triangles, (b) and (c) represent both alternative ways of subtracting an inner triangle. The three remaining triangles have four times the symplectic area of those in (a).

by its centres, the side with a given centre actually does not depend on this particular centre. We can thus use (A.6) to reduce (A.11) to

$$\Delta_{n+1}(0,X_1+\delta X_1,\ldots,X_n+\delta X_n)=\Delta_{n+1}(0,X_1,\ldots,X_n)+(\delta X_1,\ldots,\delta X_n)\mathcal{J}_n(\xi_1,\ldots,\xi_n)+\Delta_{n+1}(0,\delta X_1,\ldots\delta X_n)\,, \tag{A.12}$$

which generalizes (2.11) for arbitrary changes of all the sides. For a polygon where $x\neq 0$, we obtain

$$\Delta_{n+1}(x_1+\delta x_1,\ldots,x_n+\delta x_n)=\Delta_{n+1}(x,x_1,\ldots,x_n)+(\delta x_1,\ldots,\delta x_n)\mathcal{J}_n(\xi_1,\ldots,\xi_n)+\Delta_{n+1}(0,\delta x_1,\ldots,\delta x_n)\,. \tag{A.13}$$

To prove that the area of a circumscribed even polygon depends only on the midpoints of its sides, even though the overall shape is not unique, we first express this area as the sum of the area of the inscribed polygon δ_{2n} with that of the $2n$ triangles at the corners of the circumscribed polygon

$$\Delta_{2n}=\delta_{2n}+\Delta_3^1+\Delta_3^2+\cdots+\Delta_3^{2n}\,. \tag{A.14}$$

This is the subdivision in Fig. A.2(a). However, we note that the two other adjoining subdivisions are also possible. The sides of Δ_n' and Δ_n'' are obtained by doubling alternate sides of δ_{2n}, so that their area is independent of the arbitrary shape of the circumscribed polygon. Adding the areas in the second and third figures determines

$$2\Delta_{2n}=\Delta_n'+\Delta_n''+4(\Delta_3^1+\Delta_3^2+\cdots+\Delta_3^n)\,, \tag{A.15}$$

which, combined with (A.14), yields

$$\Delta_{2n} = 2\delta_{2n} - \tfrac{1}{2}(\Delta'_n + \Delta''_n)\,, \tag{A.16}$$

which is invariant with respect to the arbitrary corner of the circumscribed polygon.

The reader is reminded that the plane figures used to illustrate each result may be considered as projections of polygons in a phase space of arbitrary even dimension onto one of the conjugate planes.

Appendix B. Centre generating function for short times

The approximation (1.30) for the centre function only holds for times that are short enough for us to approximate the trajectory by a straight segment. To extend the approximation, we must include the curvature which depends on the quadratic part of the Hamiltonian, so that locally, near the arbitrary point 0,

$$H(x) = h \cdot x + \tfrac{1}{2} x \mathscr{H}_0 x\,, \tag{B.1}$$

to third order in x. A change of origin to $\gamma = -\mathscr{H}^{-1} h$ (the "centre of curvature") turns the Hamiltonian into an homogeneous quadratic

$$H(x) = \tfrac{1}{2}(x-\gamma)\mathscr{H}_0(x-\gamma) + C\,, \tag{B.2}$$

corresponding to the linear flow

$$(x_+ - \gamma) = \mathscr{M}_t(x_- - \gamma)\,, \tag{B.3}$$

where

$$\mathscr{M}_t = \mathrm{e}^{t\mathfrak{J}\mathscr{H}_0} = 1 + t\mathfrak{J}\mathscr{H}_0 + \tfrac{1}{2}t^2(\mathfrak{J}\mathscr{H}_0)^2 + \tfrac{1}{6}t^3(\mathfrak{J}\mathscr{H}_0)^3 + \cdots\,. \tag{B.4}$$

To derive the function that generates (B.3) we merely determine the coefficients in the expansion

$$\mathfrak{J}\mathscr{B}_+ = b_1 t + b_2 t^2 + b_3 t^3 + \cdots\,, \tag{B.5}$$

from the Cayley parametrization

$$\mathscr{M}_t = [1 - \mathfrak{J}\mathscr{B}_t][1 + \mathfrak{J}\mathscr{B}_t]^{-1} = 1 - 2\mathfrak{B}_t + 2(\mathfrak{J}\mathscr{B}_t)^2 - 2(\mathfrak{J}\mathscr{B}_t)^3 + \cdots\,. \tag{B.6}$$

The result is

$$\mathscr{B}_t = -\tfrac{1}{2}t\,\mathscr{H}_0 + \tfrac{1}{24}t^3\,\mathscr{H}_0\mathfrak{J}\mathscr{H}_0 + \mathrm{O}(t^5)\,, \tag{B.7}$$

so that the quadratic approximation to the centre function becomes

$$S_t(x) = -\tfrac{1}{2}t(x-\gamma)\mathscr{H}_0(x-\gamma) + = \tfrac{1}{24}t^3(x-\gamma)\mathscr{H}_0\mathfrak{J}\mathscr{H}_0\mathfrak{J}\mathscr{H}_0(x-\gamma)\,. \tag{B.8}$$

If we now notice that

$$\dot{x} = \mathfrak{J}\,\partial H/\partial x \simeq \mathfrak{J}\mathscr{H}_0(x-\gamma)\,, \tag{B.9}$$

we obtain the short time approximation to the centre function as

$$S_t(x) = -tH(x) - \tfrac{1}{24}t^3\,\dot{x}\mathscr{H}_x\dot{x} + \mathrm{O}(t^5)\,, \tag{B.10}$$

where $\mathscr{H}_x$ is the Hessian matrix for the Hamiltonian evaluated at the point x.

Evidently, we can continue this expansion for arbitrarily long times in the case of quadratic Hamiltonians. This can be obtained explicitly by expanding tan and tanh in the simple examples of Section 1.

We can now express S_t in terms of the energy of the orbit by noting that, to lowest order in time,

$$E = H(x_+) \simeq H(x \pm \tfrac{1}{2}t\,\dot{x}) = H(x) \pm \tfrac{1}{2}(\tfrac{1}{2}t)^2\dot{x}\mathscr{H}\dot{x}\,, \tag{B.11}$$

so that

$$S_t(x) = -tE(t) + \tfrac{1}{12}t^3\,\dot{x}\mathscr{H}_x\dot{x}\,. \tag{B.12}$$

Thus, the energy action can be identified as

$$S_E(x) = \tfrac{1}{12}t^3\,\dot{x}\mathscr{H}_x\dot{x}\,. \tag{B.13}$$

This is generally a better approximation than (2.16), being the exact third-order expansion in the case of quadratic Hamiltonians.

References

[1] E.P. Wigner, Phys. Rev. 40 (1932) 749.
[2] S.R. De Groot, L.G. Suttorp, Foundations of Electrodynamics, North-Holland, Amsterdam, 1972.
[3] R.P. Feynman, Statistical Mechanics, Benjamin, Reading, MA, 1972.
[4] K. Husimi, Proc. Phys. Math. Soc. Japan 22 (1940) 264.
[5] K. Takashi, J. Phys. Soc. Japan 55 (1986) 762.
[6] F.A. Berezin, M.A. Shubin, in: Colloquia Mathematica Societatis Janos Bolyiai, North-Holland, Amsterdam, 1972, p. 21.
[7] M.V. Berry, Phil. Trans. R. Soc. 287 (1977) 237.
[8] M.S. Marinov, Phys. Rep. 60 (1980) 1.
[9] M.S. Marinov, J. Phys. A 12 (1979) 31.
[10] F.A. Berezin, M.S. Marinov, Ann. Phys. (NY) 104 (1977) 336.
[11] A.M. Ozorio de Almeida, Proc. R. Soc. Lond. A 439 (1992) 139.
[12] A.M. Ozorio de Almeida, Proc. R. Soc. Lond. A 431 (1990) 403.
[13] H.S.M. Coxeter, Introduction to Geometry, Wiley, New York, 1969.
[14] M.V. Berry, Proc. R. Soc. Lond. A 423 (1989) 219.
[15] Arnold, Mathematical Methods of Classical Mechanics, Springer, New York, 1978.
[16] V.I. Arnold, A.B. Givental, in: V.I. Arnold, S.P. Novikov (Eds.), Dynamical Systems IV, Encyclopedia of Mathematical Sciences IV, Springer, Berlin, 1990.
[17] K.R. Meyer, Trans. Am. Math. Soc. 149 (1970) 95.
[18] V.I. Arnold, A. Avez, Ergodic Problems of Classical Mechanics, Benjamin, Reading, MA, 1968.
[19] A.M. Ozorio de Almeida, Hamiltonian in Systems: Chaos and Quantization, Cambridge University Press, Cambridge, 1988.
[20] E. Bogomolny, Nonlinearity 5 (1990) 1055.
[21] A. Messiah, Quantum Mechanics, vol. 1, North-Holland, Amsterdam, 1972.

[22] R. Jackiw, Comments Nucl. Parts Phys. 15 (1985) 99.
[23] C. Cohen Tannoudji, B. Diu, F. Laeoe, Quantum Mechanics, Wiley, New York, 1977.
[24] J.R. Klauder, B. Skagerstam, Coherent States, World Scientfic, Singapore, 1985.
[25] A. Perelomov, Generalized Coherent States and their Applications, Springer, New York, 1986.
[26] S. Bochner, W.T. Martin, Several Complex Variables, Princeton University Press, Princeton, 1948.
[27] E.C. Titchmarish, The Theory of Functions, 2nd ed, Oxford University Press, Oxford.
[28] P. Leboeuf, A. Voros, J. Phys. A 23 (1990) 1765.
[29] P. Lebocuf, A. Voros, in: G. Gasati and B. Chirikov (Eds.), Quantum Chaos, Cambridge University Press, Cambridge, 1995, p. 507.
[30] N.L. Balazs, B.K. Jennings, Phys. Rep. 104 (1984) 347.
[31] M.A.M. de Aguiar, A.M. Ozorio de Almeida, J. Phys. A 23 (1990) L 1025.
[32] H.J. Groenewold, Physica 12 (1946) 405.
[33] B. Eckhardt, Phys. Rep. 163 (1988) 205.
[34] M.C. Gutzwiller, Chaos in Classical and Quantum Physics, Springer, New York, 1990.
[35] M.-J. Giannoni, A. Voros, J. Zinn-Justin (Eds.), Chaos and Quantum Physics (Session LII, Les Houches), North-Holland, Amsterdam, 1991.
[36] S.C. Miller, R.H. Good Jr., Phys. Rev. 91 (1953) 174.
[37] M.V. Berry, Adv. Phys. 25 (1976) 1.
[38] A.M. Ozorio de Almeida, J.H. Hannay, Ann. Phys. (NY) 138 (1982) 115.
[39] A.M. Ozorio de Almeida, Ann. Phys. (NY) 145 (1983) 100.
[40] B. Grammatics, A. Voros, Ann. Phys. (NY) 123 (1979) 359.
[41] A. Voros, Ann. Inst. H. Poincaré 24A (1976) 31.
[42] M.V. Berry, J. Phys. A 10 (1977) 2083.
[43] M.C. Gutzwiller, J. Math. Phys. 12 (1971) 343.
[44] M.V. Berry, J.P. Keating, Proc. R. Soc. Lond. A 437 (1992) 151.
[45] O. Agam, S. Fishman, J. Phys. A 26 (1993) 2113.
[46] M.G.E. da Luz, A.M. Ozorio de Almeida, Nonlinearity 8 (1995) 43.
[47] R.G. Littlejohn, Phys. Rep. 138 (1986) 193.
[48] R.G. Littlejohnm, in: G. Casati and B. Chirikov (Eds.), Quantum Chaos, Cambridge University Press, Cambridge, 1995, p. 343.
[49] R. Abraham, J. Marsden, Foundations of Mechanics, Benjamin, Reading, MA, 1978.
[50] W.H. Miller, Adv. Chem. Phys. 25 (1974) 69.
[51] E.J. Heller, in: Chaos and Quantum Physics (see [35]) 1989, p. 547.
[52] E.J. Heller, J. Chem. Phys. 65 (1976) 1289.
[53] S. Fishman, B. Georgeot, R.G. Prange, J. Phys. A 29 (1996) 919.
[54] V. Bargmann, Comm. Pure Appl. Math. 14 (1961) 187.
[55] V. Bargmann, Comm. Pure Appl. Math. 20 (1967) 1.
[56] M.V. Berry, Proc. R. Soc. Lond. A 424 (1989) 279.
[57] A. Royer, Phys. Rev. A 15 (1977) 449.
[58] A. Grossmann, P. Huguenin, Helv. Phys. Acta 51 (1978) 252.