Existence of the Hannay angle for single-frequency systems

1988 J. Phys. A: Math. Gen. 21 4535

(http://iopscience.iop.org/0305-4470/21/24/009)

J. Phys. A: Math. Gen. **21** (1988) 4535–4547. Printed in the UK

Existence of the Hannay angle for single-frequency systems

Simon Golin†

Fachbereich Mathematik, Technische Universität Berlin, Sekr MA 7-2, D-1000 Berlin 12, Federal Republic of Germany

Received 22 April 1988

Abstract. The Hannay angle was introduced by Hannay as a means of measuring an anholonomy effect in classical mechanics closely corresponding to Berry's phase in quantum mechanics. Such classical adiabatic angle shifts, the Hannay angles, arise when an integrable classical Hamiltonian involves time-dependent parameters which undergo a closed adiabatic excursion. In this paper two proofs of an averaging type of theorem for a single-frequency dynamical system are given. As a consequence one can establish the existence of the Hannay angle for a class of smooth classical Hamiltonian systems with one degree of freedom. Moreover, a review of the rotated rotator illustrates the usefulness of our averaging theorem.

1. Introduction

Classical mechanics is concerned with either solving or, whenever this goal cannot explicitly be achieved, exhibiting the qualitative features of the solutions of the equations of motion for the dynamical system under consideration. In particular, one is often looking for first integrals of motions, i.e. those functions of the motion which are preserved under time evolution. For instance, in the case of a conservative Hamiltonian system, i.e. when the Hamiltonian of the dynamical system is time independent, one such integral is provided by the energy itself. Moreover, if the system is integrable one can find a canonical transformation to angle–action variables [1] and the actions are independent constants of motion. This is not so for time-dependent Hamiltonians. However, a simple intuitive argument (cf, e.g., [2]) based on two physical principles, angle averaging and Liouville's theorem, leads to the assertion of the existence of adiabatic invariants. Here 'adiabatic' denotes slow time dependence. The adiabatic theorem of classical mechanics [3, 4] claims, roughly speaking, that if $H(q, p, X(t))$ is an integrable classical Hamiltonian depending on some exterior parameters X, which are supposed to depend on time only in a slow fashion, then the actions are adiabatic invariants, i.e. are almost constant over a long timescale in the limit of slow time dependence of the Hamiltonian. A more precise statement of this fact will be given below.

Hannay [2] and Berry [5] investigated the fate of the angle variables under an adiabatic excursion in the space of classically integrable Hamiltonians. It turns out that when the Hamiltonian is taken around a closed loop then this results in an extra angle, the Hannay angle, in addition to the time integral over the instantaneous frequency. This anholonomy effect is in correspondence to Berry's phase in quantum

† Supported by the Deutsche Forschungsgemeinschaft (DFG).

mechanics and, in fact, a semiclassical relation between these two quantities was proposed by Berry [5].

It is the aim of this paper to analyse the existence of the Hannay angle. We do so by establishing an appropriate averaging theorem which will immediately imply the existence of the Hannay angle for smooth classical Hamiltonians. However, the analysis of the present paper is restricted to the case of one degree of freedom. Averaging results for systems with several frequencies are much more intricate due to the unavoidable occurrence of resonances.

It seems natural to embark on the averaging theorem below by means of the averaging methods of classical perturbation theory [1, 3] and, in fact, we will give a proof of the theorem along these lines. Moreover, an alternative proof is given applying an integration by parts method. This procedure, which is non-standard in classical perturbation theory, was motivated by the search for generalisation to the case of more degrees of freedom. (In a future paper [6] we will deal with the multi-frequency case and also give a geometrical interpretation of the Hannay angle.) A direct illustration of the averaging theorem will then be given for the example of the rotated rotator.

The plan of the paper is as follows. In § 2 the averaging theorem needed for showing the existence of the Hannay angle is stated and proved by means of averaging methods of classical perturbation theory. Then some physics related to the Hannay angle is explained, and its existence is shown in § 3. Section 4 is devoted to the alternative proof in terms of integration by parts. The example of the rotated rotator is discussed in § 5, and § 6 contains the conclusion.

2. The averaging theorem

My purpose in this section is to formulate the general framework of the averaging method in classical perturbation theory for single-frequency systems, and to prove the kind of averaging theorem that will be needed for application to the Hannay angle.

The setting of averaging comprises a hierarchy of dynamical systems whose time evolutions are to be compared in a suitable (the adiabatic) limit. The *unperturbed* system is given by the set of differential equations

$$\dot{\theta} = \omega(I) \tag{1a}$$

$$\dot{I} = 0 \tag{1b}$$

with $\theta \in \Pi^1 \equiv S^1$, $I \in K \subset \mathbb{R}^n$ and $\omega : K \to \mathbb{R}$. In what follows it will be essential that the angle θ varies only on the one-dimensional torus Π^1; this is, by definition, the single-frequency case. The action I, however, will be allowed to assume values in (a subset of) $\mathbb{R}^n$. As a matter of fact, this generality will be made use of in the next section (for $n = 2$). The *perturbed* system to be considered is governed by

$$\dot{\theta} = \omega(I) + \varepsilon f(\theta, I) \tag{2a}$$

$$\dot{I} = \varepsilon g(\theta, I) \tag{2b}$$

where f and g are functions 2π-periodic in θ, i.e. defined on $\Pi^1 \times K$, and $\varepsilon \geq 0$ is a small parameter. More generally, we could have admitted an additional dependence of f and g on the parameter ε without any resulting essential change in the proofs below. By Taylor expansion, one just gets higher-order perturbations. For simplicity of presentation we will omit such a dependence.

In the following $\theta = \theta(t)$, $I = I(t)$ will always denote a solution of the perturbed system, and it is the time evolution of these functions that we are interested in for small ε and on a timescale of order $1/\varepsilon$, i.e. $0 \leqslant t \leqslant 1/\varepsilon$. Averaging theorems are concerned with comparing this evolution with the one coming from the *averaged equation*

$$\dot{J} = \varepsilon g^{\mathrm{av}}(J) \tag{3}$$

where $g^{\mathrm{av}}(J)$ is the average of $g(\theta, J)$:

$$g^{\mathrm{av}}(J) := \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} g(\theta, J). \tag{4}$$

To avoid misunderstandings, it is worth pointing out that the solutions of (2) and (3) depend on the parameter ε, i.e. $\theta_\varepsilon = \theta_\varepsilon(t)$, $I_\varepsilon = I_\varepsilon(t)$ and $J_\varepsilon = J_\varepsilon(t)$. In accordance with common usage in averaging theory we drop the subscript throughout the paper.

Typically, averaging theorems state [1] under which conditions the averaged solution $J = J(t)$ approximates the solution $I = I(t)$ of the perturbed equation (2). Before giving some precise results of this approximation let us fix the assumptions. In the following $\|\cdot\|_0$ will denote the supremum norm of a function and $\|\cdot\|_1$ will be the supremum of the moduli of the function and its first derivatives.

(*i*) *Assumptions on phase space.* The phase space is $M := \Pi^1 \times K$, where K is a convex and compact subset of $\mathbb{R}^n$.

(ii) Assumptions on the angular frequency. We assume the angular frequency $\omega \in C^1(K)$ to be nowhere vanishing in K. (Note that the compactness of K implies $\|\omega^{-1}\|_0 < \infty$.)

(iii) Assumptions on f and g. f and g are in $C^1(M)$.

(iv) Assumptions on J(t). Fix the initial condition $J(0) = I_0 \in K$ and assume that, for $\varepsilon = 1$, $J(t)$ does not go out to ∂K (the boundary of K) for $0 \leqslant t \leqslant 1$. (For arbitrary $\varepsilon > 0$ this immediately implies that $J(t)$ does not cross ∂K for $0 \leqslant t \leqslant 1/\varepsilon$, since $J_\varepsilon(t) = J_1(\varepsilon t)$.)

It is worth pointing out that the essential assumption is $\omega \neq 0$, which corresponds to a non-resonance condition.

The usual averaging theorem on the evolution of the actions $I(t)$ on a timescale of order $1/\varepsilon$ can then be proven (cf, e.g., [3]).

Theorem 1. For any initial angle $\theta(0) = \theta_0 \in \Pi^1$, initial actions $I(0) = I_0$ $(=J(0))$ and under assumptions (i)–(iv) the perturbed system (2) has a unique solution up to time $1/\varepsilon$ satisfying

$$(\exists c > 0)(\exists \varepsilon_0 > 0)(\forall \varepsilon \leqslant \varepsilon_0) \sup_{0 \leqslant t \leqslant 1/\varepsilon} |I(t) - J(t)| \leqslant c\varepsilon. \tag{5}$$

This is the standard averaging theorem. We will now establish another averaging theorem, and in this section its proof will rely on the idea of angle averaging. It should be stressed that the next theorem will essentially deal with the behaviour of the angle variable (this will become clear from the discussion in § 3) whereas theorem 1 considers the variation of the actions. However, the adiabatic invariance of the actions will be a necessary ingredient in theorem 2.

Let $a \in C^1(\Pi^1 \times K)$ and suppose that $\|a\|_1 \leqslant c_1$ for some $c_1 > 0$. As before, an additional ε dependence is possible but omitted for simplicity. Define

$$A(t) := \varepsilon \int_0^t \mathrm{d}u\, a(\theta, I) \tag{6a}$$

$$A^{\mathrm{av}}(t) := \varepsilon \int_0^t \mathrm{d}u\, a^{\mathrm{av}}(J) \tag{6b}$$

where, again, $a^{\mathrm{av}}(J) := \int_0^{2\pi} a(\theta, J)\, \mathrm{d}\theta / 2\pi$. We have chosen an explicit factor of ε in these definitions so as to keep the difference of A and A^{av} of the order of ε for times of order $1/\varepsilon$.

Theorem 2. Under the same assumption as before

$$(\exists \tilde{c} > 0)(\forall \varepsilon \leqslant \varepsilon_0) \sup_{0 \leqslant t \leqslant 1/\varepsilon} |A(t) - A^{\mathrm{av}}(t)| \leqslant \varepsilon \tilde{c}. \tag{7}$$

Here ε_0 is taken from theorem 1.

Proof. Let $\Lambda(t)$ be given by $\Lambda := A + \varepsilon\lambda(\theta, I)$, where the function λ (2π-periodic in θ) is to be suitably defined in such a way that $\Lambda(t)$ will be close to both $A(t)$ and $A^{\mathrm{av}}(t)$. Rather than working with Λ itself, consider its time derivative

$$\begin{aligned} \dot{\Lambda} &= \dot{A} + \varepsilon(\dot{\theta}\partial_\theta\lambda + \dot{I}\nabla_I\lambda) \qquad \text{with } \dot{I}\nabla_I\lambda := \sum_{i=1}^n \dot{I}_i \partial_{I_i}\lambda \\ &= \varepsilon(a + \omega\partial_\theta\lambda) + R_1 \end{aligned} \tag{8}$$

where $R_1 := \varepsilon^2(f\partial_\theta\lambda + g\nabla_I\lambda)$ (again abbreviating $g\nabla_I\lambda := \Sigma_{i=1}^n g_i\partial_{I_i}\lambda$). Since we are in the single-frequency case and, by fiat, $\omega \neq 0$, the homological equation $a + \omega\partial_\theta\lambda = a^{\mathrm{av}}$ admits a solution for λ, namely

$$\lambda(\theta, I) = -\frac{1}{\omega(I)}\left(\int_{\theta_0}^{\theta} \mathrm{d}\vartheta\, a(\vartheta, I) - (\theta - \theta_0)a^{\mathrm{av}}(I)\right) \tag{9}$$

and obviously λ is 2π-periodic in θ. Now we obtain

$$\dot{\Lambda} = \varepsilon a^{\mathrm{av}}(I) + R_1 = \varepsilon a^{\mathrm{av}}(J) + R_2 \tag{10}$$

and $R_2 := R_1 + \varepsilon[a^{\mathrm{av}}(I) - a^{\mathrm{av}}(J)]$.

Since $|R_1| \leqslant \varepsilon^2(\|f\|_0 + \|g\|_0)\|\lambda\|_1$, this entails

$$|R_2| \leqslant \varepsilon^2 c_2 \qquad \varepsilon \leqslant \varepsilon_0 \tag{11}$$

where $c_2 := (\|f\|_0 + \|g\|_0)\|\lambda\|_1 + cc_1$. Here ε_0 and c are taken from theorem 1. Thus

$$|(\Lambda - A^{\mathrm{av}})^{\cdot}| \leqslant \varepsilon^2 c_2 \tag{12a}$$

$$|\Lambda(t) - A^{\mathrm{av}}(t)| \leqslant \varepsilon c_2 \qquad \varepsilon \leqslant \varepsilon_0 \text{ and } 0 \leqslant t \leqslant 1/\varepsilon. \tag{12b}$$

By the very definition of Λ, it is close to A and hence

$$|A(t) - A^{\mathrm{av}}(t)| \leqslant |A(t) - \Lambda(t)| + |\Lambda(t) - A^{\mathrm{av}}(t)| \leqslant \varepsilon\tilde{c} \tag{13}$$

where $\tilde{c} := \|\lambda\|_0 + c_2$.

3. The Hannay angle

Originally Hannay [2] was led to the investigation of classical adiabatic angle shifts by the discovery [7] of a quantum mechanical quantity, Berry's phase, for which he wanted to give a classical analogue. In this section we will briefly review the setting of the Hannay angle. Then we will show how theorem 2 can be applied to prove the existence of the Hannay angle for single-frequency systems and to establish a particular simple 'averaged' form of it. Throughout the following we will assume that the systems under consideration have only one degree of freedom.

Let an integrable classical Hamiltonian $H(q, p, X)$ be given. Here q and p denote the position variable and the momentum variable respectively, and X is some exterior parameter. Since, by fiat, the Hamiltonian is integrable for fixed values of X there exists, for all X, a canonical transformation to angle–action variables θ and I so that the Hamiltonian becomes cyclic when expressed in the new variables, i.e. it does not depend on the angle θ. We call $h_0(I, X)$ the Hamiltonian in the new variables.

Now suppose that the Hamiltonian describes a time-dependent dynamical system and the time dependence enters by changing the parameters X as time evolves. Being interested in the adiabatic limit, i.e. the limit of slow time dependence, we consider $X = X(\varepsilon t)$, where $\varepsilon \geqslant 0$ is assumed to be small. The evolution of the angle–action variables in the non-conservative case is not governed by $h_0(I, X)$ alone, but one has to take into account the time derivative of the generating function $S(q, I, X)$ which describes the change of variables from (q, p) to (θ, I). More precisely, the generating function is generally multi-valued and one has to consider all the branches; we will neglect this for simplicity of presentation. Expressing the generating function completely in the new variables by

$$\mathcal{S}(\theta, I, X) := S(q(\theta, I, X), I, X) \tag{14}$$

the dynamics in angle–action variables is given by the Hamiltonian

$$h = h_0(I, X) + \varepsilon X' h_1(\theta, I, X)$$

where

$$h_1(\theta, I, X) := (\nabla_X \mathcal{S} - p \nabla_X q)(\theta, I, X).$$

Thus the canonical equations are

$$\dot{\theta} = \omega(I, X) + \varepsilon X' \partial_I h_1(\theta, I, X) \tag{15a}$$

$$\dot{I} = -\varepsilon X' \partial_\theta h_1(\theta, I, X). \tag{15b}$$

Here $\omega(I, X) := \partial_I h_0(I, X)$ is the instantaneous angular frequency.

Until recently, mainly the behaviour of the action variables $I(t)$ was considered. The adiabatic thorem of classical mechanics states that for sufficiently slow time dependence, i.e. ε small, the variation of the action is small on a timescale of order $1/\varepsilon$. In other words, the action is an adiabatic invariant.

Hannay [2] and Berry [5], however, considered the time evolution of the angle variable θ. The Hannay angle $\Delta\theta$ is defined by comparing the evolution when the exterior parameters in the Hamiltonian undergo a closed adiabatic loop with the evolution due to the instantaneous frequency. Since the definition of angle variables involves a certain degree of arbitrariness related to the choice of origin ($S(q, I, X) \to S(q, I, X) + \hat{\theta} I$, with $\hat{\theta} = \text{constant}$) one cannot, generally speaking, compare angles belonging to distinct values of X, and it is for this reason that the excursion of the parameter X is required to be closed. Let us assume that $X(0)=X(1)$. Then the Hannay angle is defined by

$$\Delta\theta := \lim_{\varepsilon\to 0}\left(\theta(1/\varepsilon)-\theta_0-\int_0^{1/\varepsilon} \mathrm{d}t\,\omega(I(t), X(\varepsilon t))\right) \tag{16}$$

provided this limit exists.

We shall now show that theorem 2 may be applied to prove that the Hannay angle is well defined in the single-frequency case, i.e. when the underlying Hamiltonian describes a system of one degree of freedom. Since theorems 1 and 2 apply to autonomous systems we must get rid of the explicit time dependence in (15). This can be achieved by introducing a new 'action' $\tau(t) := \varepsilon t$. Then (15) is equivalent to the autonomous system

$$\dot\theta = \omega(I, X(\tau)) + \varepsilon X'(\tau)\partial_I h_1(\theta, I, X(\tau)) \tag{17a}$$

$$\dot I = -\varepsilon X'(\tau)\partial_\theta h_1(\theta, I, X(\tau)) \tag{17b}$$

$$\dot\tau = \varepsilon. \tag{17c}$$

Here the corresponding averaged equations are particularly simple:

$$\dot J = 0 \tag{18a}$$

$$\dot\tau = \varepsilon. \tag{18b}$$

Since (17) now has the required structure (the number of actions being two) theorem 2 implies the existence of the Hannay angle, provided the assumptions of § 2 are satisfied.

Theorem 3. $\Delta\theta$ is well defined by (16) and moreover

$$\Delta\theta = -\partial_I \oint_{\mathscr{C}} \mathrm{d}X(p\nabla_X q)^{\mathrm{av}}(I_0, X). \tag{19}$$

Here $\mathscr{C}$ denotes the contour of $X = X(\varepsilon t)$ in parameter space.

Proof. By the equation of motion (17a) and theorem 2:

$$\begin{aligned}
\theta(t)-\theta_0-\int_0^{1/\varepsilon}\mathrm{d}t\,\omega(I, X(\tau)) &= \varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\,X'(\tau)\partial_I h_1(\theta, I, X(\tau))\\
&= \varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\,X'(\tau)(\partial_I h_1)^{\mathrm{av}}(I_0, X(\tau)) + \mathrm{O}(\varepsilon)\\
&= \partial_I\int_0^1 \mathrm{d}t\,\dot X(t)(\nabla_X\mathscr{S} - p\nabla_X q)^{\mathrm{av}}(I_0, X(t)) + \mathrm{O}(\varepsilon)\\
&= -\partial_I\oint_{\mathscr{C}}\mathrm{d}X(p\nabla_X q)^{\mathrm{av}}(I_0, X) + \mathrm{O}(\varepsilon).
\end{aligned} \tag{20}$$

The term with the generating function $\mathscr{S}$ has vanished since it is a gradient and integration is performed around the closed contour $\mathscr{C}$.

As one can see from (20), the Hannay angle depends only on the initial action I_0 and the contour $\mathscr{C}$, $\Delta\theta = \Delta\theta(I_0, \mathscr{C})$, but is independent of θ_0. The independence of the angle shift on θ_0 is, of course, desirable in view of the arbitrariness in the choice of the origin of the angle variable. Moreover, (20) makes it obvious that a non-vanishing Hannay angle can only arise if the loop is performed in a parameter space which is either at least two dimensional or homeomorphic to the circle S^1. A closed loop in $\mathbb{R}^1$ has to retrace itself and thus yields a zero contour integral.

For completeness let us note that by Stokes' theorem the Hannay angle can be given a different representation, namely as an integral over a 2-form [2, 5]. Let $\mathscr{F}$ be a surface in parameter space with boundary $\mathscr{C}$ and denote by $\bar{d}$ the exterior differential in parameter space; then

$$\Delta\theta = \partial_I \int_{\mathscr{F}} W(I_0, X) \tag{21a}$$

$$W(I, X) := (\bar{d}q \wedge \bar{d}p)^{\mathrm{av}}(I, X) = \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} (\bar{d}q \wedge \bar{d}p)(\theta, I, X). \tag{21b}$$

Note that W is a 2-form in parameter space, called the *angle 2-form.*

As for the situation with more than one frequency, the reader is referred to [6]. Moreover, one knows that Berry's phase is formulated irrespective of whether the underlying classical system is integrable, ergodic or in between. Since, semiclasically, Berry's phase is related to the Hannay angle one wonders what the correspondence might be in classically non-integrable situations. A geometric interpretation of the Hannay angles (in accordance with Simon's treatment [8] of Berry's phase) answering this question will follow in [6].

Another case where a Hannay angle can easily be computed is a (not necessarily completely integrable) classical Hamiltonian system whose Hamiltonian admits one and the same symmetry for all values of the parameters X. Roughly speaking, the symmetry allows, by Noether's theorem, for a constant of the motion for each fixed X, and then one can perform a canonical transformation on phase space to make the Hamiltonian cyclic. The shift for the corresponding variable, the one which is absent in the Hamiltonian, can be computed and turns out to be zero. This is proved in [9].

4. Integration by parts argument

The proof of theorem 2 relied on standard methods in the theory of averaging. I will now give a second proof based on a rather unconventional approach to classical perturbation theory, namely via integration by parts. The main motivation for looking for a different proof of theorem 2 was the desire for generalisation to systems with several frequencies. To indicate the idea of how the following proof could be adapted to the multi-frequency case consider 'well behaved' initial values (θ_0, I_0), i.e. those for which $I(t)$ enters the resonant zones at most for short times. Then (29) might be used by splitting integration over 'resonant' and 'non-resonant' time intervals and giving an upper bound to the resonant one according to the assumption on the character of the initial value.

The general intention here is to construct a proof of the existence of the Hannay angle by making use as much as possible of already existing results in averaging. So the following proof will also make use of the usual averaging theorem, i.e. theorem 1. We shall now take for granted that the previous assumptions of § 2 are satisfied.

Alternative proof of theorem 2. Consider

$$A(t)-A^{\mathrm{av}}(t)=\varepsilon\int_0^t \mathrm{d}u[a(\theta,I)-a^{\mathrm{av}}(J)]$$
$$=\varepsilon\int_0^t \mathrm{d}u[a(\theta,J)-a^{\mathrm{av}}(J)]+\varepsilon\int_0^t \mathrm{d}u[a(\theta,I)-a(\theta,J)]. \tag{22}$$

According to the assumptions on the function a the modulus of the last integral may be bounded by $c\|a\|_1$ for $0\leqslant t\leqslant 1/\varepsilon$; this follows from the mean value theorem. Thus

$$A(t)-A^{\mathrm{av}}(t)=\varepsilon\int_0^t \mathrm{d}u[a(\theta,J)-a^{\mathrm{av}}(J)]+\mathrm{O}(\varepsilon) \tag{23}$$

where here and in the following $\mathrm{O}(\varepsilon)$ is always meant uniformly for $0\leqslant t\leqslant 1/\varepsilon$.

Being left with controlling the first term in (23) we introduce the Fourier expansion of a

$$a(\theta,I)=\sum_{k=-\infty}^{\infty}\hat{a}_k(I)\,\mathrm{e}^{\mathrm{i}k\theta} \tag{24a}$$

where the Fourier coefficients are given by

$$\hat{a}_k(I)=\frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\theta\,\mathrm{e}^{-\mathrm{i}k\theta}a(\theta,I). \tag{24b}$$

Since, by fiat, $a(\cdot,I)$ is a continuously differentiable function, the Fourier series (24a) converges for each I uniformly in $\theta\in\Pi^1$. In fact, the compactness of K implies that the convergence in (24a) is uniform on $\Pi^1\times K$. This can be seen by adapting the usual proof of uniform convergence in θ (cf, e.g., [10]) to the additional I dependence. Upon averaging with respect to θ we may therefore exchange the order of integration and summation so that $a^{\mathrm{av}}=\hat{a}_0$ and

$$A(t)-A^{\mathrm{av}}(t)=\varepsilon\int_0^t \mathrm{d}u\sum_{k=0}\hat{a}_k(J)\,\mathrm{e}^{\mathrm{i}k\theta}+\mathrm{O}(\varepsilon)$$
$$=\varepsilon\sum_{k\neq 0}\int_0^t \mathrm{d}u\,\hat{a}_k(J)\,\mathrm{e}^{\mathrm{i}k\theta}+\mathrm{O}(\varepsilon). \tag{25}$$

The last step is, again, a consequence of the uniform convergence.

Now let us consider each of the non-zero modes ($k\neq 0$) separately. By the equation of motion (2a)

$$\theta(u)=\theta_0+\int_0^u \mathrm{d}v\,\omega(I)+\varepsilon\int_0^u \mathrm{d}v f(\theta,I). \tag{26}$$

Since $\omega\neq 0$ was assumed we may use the equality

$$\exp\left(\mathrm{i}k\int_0^u \mathrm{d}v\,\omega(I)\right)=\frac{1}{\mathrm{i}k\omega(I)}\partial_u\exp\left(\mathrm{i}k\int_0^u \mathrm{d}v\,\omega(I)\right) \tag{27}$$

so as to carry out a partial integration:

$$\int_0^t \mathrm{d}u\, \hat{a}_k(J)\, \mathrm{e}^{\mathrm{i}k\theta}$$

$$= \frac{\mathrm{e}^{\mathrm{i}k\theta_0}}{\mathrm{i}k} \int_0^t \mathrm{d}u \frac{\hat{a}_k(J)}{\omega(I)} \partial_u \left[\exp\left(\mathrm{i}k \int_0^u \mathrm{d}v\, \omega(I) \right) \right] \exp\left(\mathrm{i}\varepsilon k \int_0^u \mathrm{d}v f(\theta, I) \right)$$

$$= \frac{\hat{a}_k(J)}{\mathrm{i}k\omega(I)} \mathrm{e}^{\mathrm{i}k\theta} \Big|_{u=0}^{t} - \varepsilon \int_0^t \mathrm{d}u \frac{1}{\omega(I)} \Bigg[\hat{a}_k(J) f(\theta, I)$$

$$+ \frac{1}{\mathrm{i}k} \left(g^{\mathrm{av}}(J) \partial_J \hat{a}_k(J) - \hat{a}_k(J) g(\theta, I) \frac{\partial_I \omega(I)}{\omega(I)} \right) \Bigg] \mathrm{e}^{\mathrm{i}k\theta}. \tag{28}$$

Again, the equations of motion (2*a*, *b*) were used here. Note that $\Sigma_{k\neq 0} k^{-1}\, \partial_J \hat{a}_k(J)\, \mathrm{e}^{\mathrm{i}k\theta}$ is uniformly convergent on $\Pi^1 \times K$ since it is the Fourier series of $\mathrm{i}(b(\theta, J) - b^{\mathrm{av}}(J))$, where $b(\theta, J) := \int \mathrm{d}\vartheta\, \partial_J [a(\vartheta, J) - a^{\mathrm{av}}(J)]$, and so the remark following (24*b*) applies. This now entails

$$\sum_{k\neq 0} \int_0^t \mathrm{d}u\, \hat{a}_k(J)\, \mathrm{e}^{\mathrm{i}k\theta}$$

$$= \frac{1}{\mathrm{i}\omega(I)} \sum_{k\neq 0} \frac{\hat{a}_k(J)}{k} \mathrm{e}^{\mathrm{i}k\theta} \Big|_{u=0}^{t} - \varepsilon \int_0^t \mathrm{d}u \frac{1}{\omega(I)} \Bigg(f(\theta, I) \sum_{k\neq 0} \hat{a}_k(J)\, \mathrm{e}^{\mathrm{i}k\theta}$$

$$- \mathrm{i} g^{\mathrm{av}}(J) \sum_{k\neq 0} \frac{\partial_J \hat{a}_k(J)}{k} \mathrm{e}^{\mathrm{i}k\theta} + \mathrm{i} g(\theta, I) \frac{\partial_I \omega(I)}{\omega(I)} \sum_{k\neq 0} \frac{\hat{a}_k(J)}{k} \mathrm{e}^{\mathrm{i}k\theta} \Bigg) \tag{29}$$

and thus this quantity is uniformly bounded for $0 \leqslant t \leqslant 1/\varepsilon$. Therefore, (25) yields $A\ (t) - A^{\mathrm{av}}(t) = \mathrm{O}(\varepsilon)$.

5. Example: rotated rotator

This example is devoted to an illustration of the general theory explained so far. In a different version it was already treated in [2, 5].

By the rotated rotator we mean the free motion of a particle of unit mass on a planar loop (in coordinate space) which is itself rotating through a complete turn. The rotation of the loop is assumed to take place around an origin in its own plane and to leave the loop in this plane.

As a result of the time-dependent holonomic constraint this system has only one degree of freedom and thus lies within the scope of the method of this paper. Let us choose a frame of reference fixed with respect to the loop and denote the position of the particle in this frame by $x = x(s)$, where s is the arc length as measured from some fixed (material) point on the loop. Moreover let Φ be the angle by which this frame of reference differs from the inertial one (cf figure 1). If q is the position vector in the inertial frame of reference then

$$q(s, \Phi) = A(\Phi) x(s) \qquad A(\Phi) := \begin{pmatrix} \cos\Phi & -\sin\Phi \\ \sin\Phi & \cos\Phi \end{pmatrix}. \tag{30}$$

The Lagrangian $L = \frac{1}{2}(\dot{q})^2$ is easily expressed in the coordinates of the rotating frame:

$$L = \tfrac{1}{2}(\dot{s})^2 + \varepsilon\Omega\dot{s}\Delta(s) + \tfrac{1}{2}\varepsilon^2\Omega^2 x^2(s) \tag{31}$$

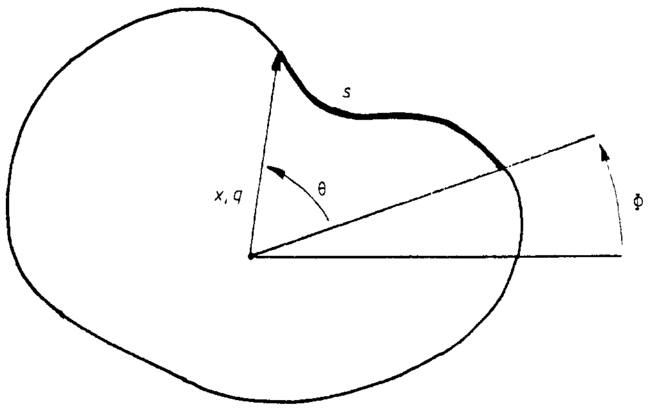

Figure 1.

where $\varepsilon\Omega(\varepsilon t) = (\mathrm{d}/\mathrm{d}t)\Phi(\varepsilon t)$ is the angular frequency of the rotation of the loop and $\Delta(s) := \det(x(s), x'(s))$. From this one may deduce the Hamiltonian

$$h = \tfrac{1}{2}(p - \varepsilon\Omega\Delta(s))^2 - \tfrac{1}{2}\varepsilon^2\Omega^2 x^2(s) \tag{32}$$

where the canonical momentum p conjugate to s is just the ordinary momentum for the inertial observer, $p = \dot{s} + \varepsilon\Omega\Delta(s)$.

It is worth emphasising that this Hamiltonian differs from the standard type $h = h_0 + \varepsilon h_1$ (with h_0 and h_1 both ε independent) as it contains an ε^2 term. This difference comes about because we are actually reducing a two-dimensional system by a holonomic constraint. Strictly speaking, two-dimensional systems are not within the scope for our averaging theorem. Only if the higher-dimensional analogue is understood might one expect to compute the angle shift for the rigid constraint by first solving the system with a soft constraint (i.e. a potential well around the loop) and then taking the limit which implements the rigid constraint.

The Hamiltonian equations of (32) are

$$\dot{s} = p - \varepsilon\Omega\Delta(s) \tag{33a}$$

$$\dot{p} = \varepsilon\Omega(p - \varepsilon\Omega\Delta(s))\Delta'(s) + \varepsilon^2\Omega^2 x(s) \cdot x'(s). \tag{33b}$$

As pointed out before, theorem 2 remains valid if ε^2 perturbations are admitted, so that

$$\Delta s = -\lim_{\varepsilon\to 0} \varepsilon \int_0^{1/\varepsilon} \mathrm{d}t\, \Omega\Delta(s) = -\frac{4\pi\mathscr{A}}{\mathscr{L}} \tag{34}$$

where $\mathscr{L}$ and $\mathscr{A}$ are the length and the enclosed area of the loop respectively. In terms of the polar angle $\theta := 2\pi s/\mathscr{L}$ (with respect to the rotating frame of reference), the Hannay angle is [2, 5]

$$\Delta\theta = -8\pi\mathscr{A}/\mathscr{L}^2 = -2\pi + 2\pi(1 - 4\pi\mathscr{A}/\mathscr{L}^2). \tag{35}$$

The first term expresses the fact that the loop has carried out one complete turn (so that there is a shift in the origin by -2π) whereas the anholonomy effect is contained in the second term, which is non-negative by the isoperimetric inequality. Note that the second term is independent of the particular shape of the loop but only contains $\mathscr{A}/\mathscr{L}^2$. For non-angular loops the Hannay angle therefore allows one to detect rotations (as was noted by Berry [5]).

Another computation allows one to compare the actual evolution of the rotated rotator to the unperturbed one (where the loop is kept fixed)†. By (33*a*)

$$s(1/\varepsilon)-s_0=\int_0^{1/\varepsilon}\mathrm{d}t\,p-\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\,\Omega\Delta=\int_0^{1/\varepsilon}\mathrm{d}t\,p-4\pi\mathscr{A}/\mathscr{L}+\mathrm{O}(\varepsilon). \tag{36}$$

It follows by partial integration that

$$\begin{aligned}\int_0^{1/\varepsilon}\mathrm{d}t\,p&=\frac{1}{\varepsilon}p_0+\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)\dot{p}\\&=\frac{1}{\varepsilon}p_0+\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)[\varepsilon\Omega(p-\varepsilon\Omega\Delta)\Delta'+\varepsilon^2\Omega^2x\cdot x']\\&=\frac{1}{\varepsilon}p_0+\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)\Omega p\Delta'+\mathrm{O}(\varepsilon)\end{aligned} \tag{37}$$

since

$$\frac{\varepsilon^2}{2}\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)\Omega^2(x^2-\Delta^2)'=\frac{\varepsilon^2}{2}\int_0^{1/\varepsilon}\mathrm{d}t\int_0^{1/\varepsilon}\mathrm{d}u\,\Omega^2(x^2-\Delta^2)' \tag{38}$$

and by averaging

$$\varepsilon\int_0^{1/\varepsilon}\mathrm{d}u\,\Omega^2(x^2-\Omega^2)'=\mathrm{O}(\varepsilon). \tag{39}$$

Moreover

$$\begin{aligned}\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)\Omega p\Delta'&=\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)\Omega(\dot{s}+\varepsilon\Omega)\Delta'\\&=\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\left(\frac{1}{\varepsilon}-t\right)\Omega\dot{s}\Delta'+\mathrm{O}(\varepsilon)\end{aligned} \tag{40}$$

by the same argument as in (38) and (39). Thus

$$\begin{aligned}&\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t(1/\varepsilon-t)\Omega p\Delta'\\&\quad=\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t(1/\varepsilon-t)\Omega\frac{\mathrm{d}}{\mathrm{d}t}\Delta+\mathrm{O}(\varepsilon)\\&\quad=-\Omega(0)\Delta(s_0)+\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t\,\Omega\Delta-\varepsilon\int_0^{1/\varepsilon}\mathrm{d}t(1/\varepsilon-t)\frac{\mathrm{d}}{\mathrm{d}t}\Omega(\varepsilon t)\Delta+\mathrm{O}(\varepsilon)\\&\quad=-\Omega(0)\Delta(s_0)+\frac{4\pi\mathscr{A}}{\mathscr{L}}-\varepsilon\frac{2\mathscr{A}}{\mathscr{L}}\int_0^{1/\varepsilon}\mathrm{d}t(1/\varepsilon-t)\frac{\mathrm{d}}{\mathrm{d}t}\Omega(\varepsilon t)+\mathrm{O}(\varepsilon)\\&\quad=-\Omega(0)\Delta(s_0)+\Omega(0)2\mathscr{A}/\mathscr{L}+\mathrm{O}(\varepsilon).\end{aligned} \tag{41}$$

Altogether

$$s(1/\varepsilon)=\frac{1}{\varepsilon}p_0-\frac{4\pi\mathscr{A}}{\mathscr{L}}+\Omega(0)\left(\frac{2\mathscr{A}}{\mathscr{L}}-\Delta(s_0)\right)+\mathrm{O}(\varepsilon) \tag{42}$$

† I am grateful to R Seiler for pointing out this application of theorem 2.

and this corrects formula (A2.15) given in [5]. It is interesting to note that averaging over the initial positions s_0 yields

$$\langle s(1/\varepsilon)-s_0\rangle_{s_0}=\frac{1}{\varepsilon}p_0-\frac{4\pi\mathscr{A}}{\mathscr{L}}+\mathrm{O}(\varepsilon) \tag{43}$$

i.e. the terms on the right-hand side of (42) reduce to the one from free evolution and the Hannay angle.

6. Conclusion

The Hannay angle was found explicitly only recently, although presumably it might have occurred in some physical applications before. Anyway, it has already found application in optics [11] and celestial mechanics [9]. But also from a theoretical standpoint the Hannay angle is an interesting quantity, particularly in the context of semiclassical approximations, where Berry's phase and Hannay's angle might help to shed some light on the semiclassical features of classical properties such as integrability or chaos. So it seemed worthwhile to start an investigation on the mathematical foundation of the Hannay angle.

This paper dealt exclusively with the investigation of the Hannay angle for systems with one frequency. The case of several frequencies requires a more careful treatment as resonances become unavoidable. From a physical point of view this situation is, of course, also very interesting. This work will appear in [6] together with a geometrical interpretation of the Hannay angle.

According to the results by Lenard [12] and Neishtadt [13] (see also the references contained in [10]) the actions are adiabatic invariants to all orders, i.e. $I(t)-I_0=\mathrm{O}(\varepsilon^n)$ for all $n\in\mathbb{N}$. In contrast to the assumptions made in this paper these results needed that the change of the exterior parameters X tends to zero for infinite times (i.e. as $t\to\pm\infty$) whereas we have been considering periodic time dependence. As for the Hannay angle, one cannot expect that the limit in the definition (16) of the Hannay angle involves only terms of infinite order. A simple counterexample is given by $\dot\theta=1$, $a(\theta,I)=\cos\theta$ so that $A(t)-A^{\mathrm{av}}(t)=\varepsilon\sin t$ and, therefore, $A(1/\varepsilon)-A^{\mathrm{av}}(1/\varepsilon)=\varepsilon\sin(1/\varepsilon)\neq\mathrm{O}(\varepsilon^2)$. Of course, perturbation theory could be applied to obtain the expansion of $A(1/\varepsilon)-A^{\mathrm{av}}(1/\varepsilon)$ in successive powers of ε, but being interested in the limit $\varepsilon\to 0$ we did not consider higher-order expansions.

Acknowledgment

I would like to thank R Seiler for helpful discussions.

References

[1] Arnol'd V I, Kozlov V V and Neishtadt A I 1988 *Encyclopaedia of Mathematical Sciences* vol 3, ed V I Arnol'd (Berlin: Springer)
[2] Hannay J H 1985 *J. Phys. A: Math. Gen.* **18** 221–30
[3] Arnol'd V I 1983 *Geometrical Methods in the Theory of Ordinary Differential Equations* (Berlin: Springer)
[4] Dirac P A M 1925 *Proc. R. Soc.* **107** 725–34
[5] Berry M V 1985 *J. Phys. A: Math. Gen.* **18** 15–27

[6] Golin S, Knauf A and Marmi S 1988 *The Hannay angles: geometry, adiabaticity, and an example. Preprint*
[7] Berry M V 1984 *Proc. R. Soc.* A **392** 45-57
[8] Simon B 1983 *Phys. Rev. Lett.* **51** 2167-70
[9] Golin S and Marmi S 1988 *Symmetries, Hannay angles, and precession of orbits. Preprint*
[10] Moise E E 1967 *Calculus* (Reading, MA: Addison-Wesley)
[11] Berry M V 1986 *Fundamental Aspects of Quantum Theory* ed V Gorini and A Frigerio (New York: Plenum) pp 267-78
[12] Lenard A 1959 *Ann. Phys., NY* **6** 261-76
[13] Neishtadt A I 1981 *Prikl. Mat. Mekh.* **45** 80-7 (Engl. transl. 1982 *J. Appl. Math. Mech.* **45** 58-63)