Geometric phases in classical mechanics

Morten Ib Munk-Nielsen, pql692

Bachelor thesis

Advisor: Jens Paaske
Niels Bohr Institute, University of Copenhagen
June 11, 2014

Abstract

Much attention has been given to geometric phases in quantum mechanics and good, concise introductions to the subject exist. Understanding geometric phases in classical mechanics however requires the piecing together of the content of many articles and textbooks while enduring confusingly different notation and inconsistencies. This is problematic since a bridge between the classical and quantum mechanical paradigm might be crucial in future insight in open questions. The purpose of my bachelor thesis is to help bridge the research gap by providing a thorough and coherent introduction to the subject of geometric phases in classical mechanics, introducing all the necessary concepts so the notation is kept consistent. A contribution is given in the solution of the Foucault pendulum by a direct computation of parallel transportation of a vector on a sphere using differential geometry.

Contents

1 Introduction **1**

2 The Foucault pendulum **1**
2.1 Treatment using fictitious forces 1
2.2 Review of differential geometry 4
2.3 The Foucault pendulum revisited 8

3 Action-angle variables and adiabatic invariance **11**
3.1 Short recap of Hamiltonian mechanics 11
3.2 Action-angle coordinates 13
3.3 Adiabatic invariance 15
3.4 Example of a geometric phase using adiabatic invariance 17

4 The Hannay angle **18**

5 Conclusion **19**

References **20**

A Harmonic oscillator with time dependent frequency **21**

1 Introduction

In the eighties a geometric phenomenon was discovered by Berry [1]. It was already known — and known as the adiabatic theorem — that if a quantum mechanical system dependent on some parameters is in a given eigenstate and the parameters are slowly changed, then the system stays in a state with the same quantum numbers but picks up a phase. Berry's insight was to imagine that the parameters were changed slowly in such a way that the value of all parameters end the same as they started. The non-trivial phase picked up because of the variation is often called Berry's phase. As explained by Hannay a few years later [2], it turned out that Berry's phase wasn't an artifact of weird quantum behaviour, but that a strong analogy exists in classical mechanics for systems exhibiting periodic motion. Here the phase is a change in the angle variable associated with the periodic motion and it is sometimes called Hannay's angle. Because of the similarities between the two concepts, they are referred to as being different cases of geometric phases meaning that they are phases picked up by the system after doing an excursion of the parameters the system depends on, irrespective of how long this excursion takes.

It is fairly simple to introduce the notion of geometric phases in quantum mechanics, for example it is done very well by Griffiths [3]. Curiously enough, however, it quickly becomes tricky to introduce geometric phases in classical mechanics as many concepts are required to fully appreciate the subject. This paper sets out to fill the gap in the literature by providing a comprehensive introduction to and treatment of the subject. We will begin with an illustration of a well understood problem and devote some effort towards understanding the geometry in the problem using differential geometry.

With the intuition from this first example, we will continue with a review of adiabatic invariants and the classical adiabatic theorem; a very convenient tool for seperating dynamic and geometric contributions to the phase. Lastly we will end with a short discussion of the relation of the Hannay angle to the Berry phase.

2 The Foucault pendulum

As the first example of a geometric phase we will consider the familiar Foucault pendulum. This is a simple contraption: It is imply a long and heavy pendulum suspended and set in motion. As time goes the plane of oscillations turn and this is taken as a demonstration that earth is rotating, rather than the universe rotating about earth. As it happens, the Foucault pendulum is also an excellent example of a geometric phase. In this section we will start with treating the problem in the manner it is usually done, following the line of reasoning as Fetter and Walecka [4]. This way of solving the problem doesn't offer much insight into why the acquired angle takes the form it does. Therefore we follow up with a review of differential geometry before solving the problem of parallel transport on the sphere, obtaining the result again.

2.1 Treatment using fictitious forces

Consider a pendulum hung at a fixed position on earth and rotating around with it. We will approximate the earth as a perfect sphere. When the sphere revolves, the pendulums plane of oscillation precesses, except when the pendulum is suspended on equator. First a heuristic argument; Imagine suspending the pendulum exactly on the north pole. From an inertial frame the pendulum doesn't move so it is not subject to any fictitious forces. Seen from an observer on earth, though, the pendulum seems to precess clockwise when viewed from above with a period of exactly one day. On the south pole it seems to precess the other way around with the same period. Because of the symmetry of the sphere putting the pendulum on equator doesn't make it turn as no direction of pressesion is preferable to others.

The phenomenon can be attributed to the Coriolis force. Let the pendulum be suspended at a polar angle θ as shown on figure 1. Newton's second law gives

$$m\ddot{\mathbf{r}} = m\mathbf{g} + \mathbf{T} - 2m\omega \times \dot{\mathbf{r}}, \tag{1}$$

where $\mathbf{T}$ is the force on the bob from the tension in the string.

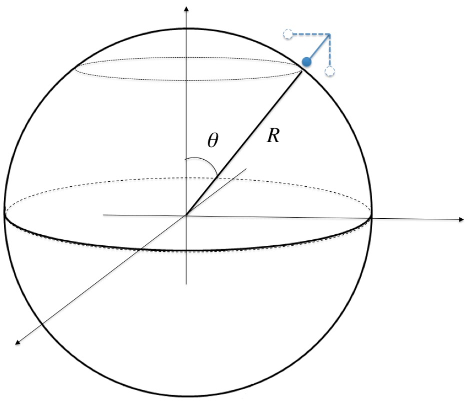

Figure 1: The suspension point of the pendulum is situated at a constant polar angle θ.

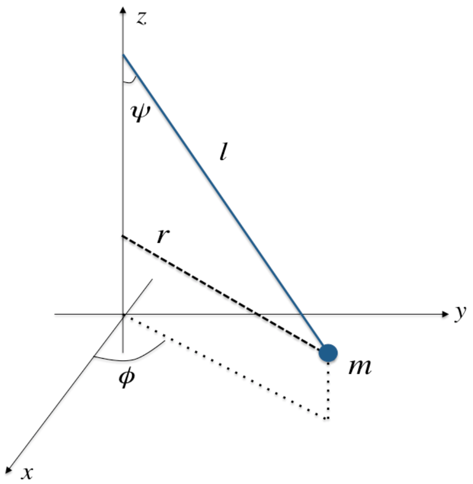

Figure 2: The pendulum in local coordinates.

Now let us invoke a local coordinate system (x, y, z) as shown in figure 2, describing the position of the bob relative to the suspension point of the pendulum which is held fixed in this coordinate system. The xy-plane is the tangent plane to the sphere with the point $(0, 0, 0)$ being the equilibrium point of bob in the pendulum. As shown in the figure we let ϕ denote the azimuthal angle, and let ψ be the angle the pendulum makes with the z-axis and let r denote the horisontal distance from the bob to the z-axis. Let us restrict ourselves to relatively small oscillations. We neglect variations in the distance of the bob to the center of earth so we assume that $\dot{z} \approx 0$. We also assume $\mathbf{T} \approx m\mathbf{g}$ and by using that $\sin\psi = \frac{r}{l}$ we get that (1) becomes

$$m\ddot{x} \approx -T\sin\psi\cos\phi + 2m\omega_\perp \dot{y} = -\frac{mg}{l} r\cos\phi + 2m\omega_\perp \dot{y}$$
$$m\ddot{y} \approx -T\sin\psi\sin\phi - 2m\omega_\perp \dot{x} = -\frac{mg}{l} r\sin\phi - 2m\omega_\perp \dot{y},$$

where $\omega_\perp = \omega\cos\theta$ and the second terms come from the coriolis force. Noticing that $r\cos\phi = x$ and $r\sin\phi = y$ by definition gives

$$\ddot{x} \approx -\frac{g}{l} x\sin\psi\cos\phi + 2\omega_\perp \dot{y} \tag{2}$$
$$\ddot{y} \approx -\frac{g}{l} y\sin\psi\sin\phi - 2\omega_\perp \dot{x}. \tag{3}$$

A convenient way to solve (2) and (3) is to introduce the variable $\zeta(t) \equiv x(t) + iy(t)$ so

$$\ddot{\zeta} = -\frac{g}{l}\zeta - 2i\omega_\perp\dot{\zeta}. \tag{4}$$

The constancy of the coefficients in (4) entails that the solution can be written as an exponential function

$$\zeta(t) = \zeta_0 e^{-i\sigma t}. \tag{5}$$

Substituting (5) into (4) yields

$$\sigma^2 - 2\omega_\perp\sigma - \frac{g}{l} = 0,$$

whereby

$$\sigma_\pm = \omega_\perp \pm \sqrt{\omega_\perp^2 + \frac{g}{l}} \equiv p \pm q.$$

This gives the solution as

$$\zeta(t) = Ae^{-i(p+q)t} + Be^{-i(p-q)t}. \tag{6}$$

Choosing the inital conditions $\zeta(0) = x(0) = a \in \mathbb{R}$ and $\dot{\zeta}(0) = 0$ gives the solution as

$$\zeta(t) = ae^{-ipt}(\cos qt + ipq^{-1}\sin qt) \tag{7}$$

Notice how ζ is always non-zero. This means that it is never a perfect pendulum. The solutions for x and y are then

$$x(t) = a\cos(\omega_\perp t)\cos\left[\left(\omega_\perp^2 + \frac{g}{l}\right)^{1/2} t\right] + a\omega_\perp\left(\omega_\perp^2 + \frac{g}{l}\right)^{-1/2}\sin(\omega_\perp t)\sin\left[\left(\omega_\perp^2 + \frac{g}{l}\right)^{1/2} t\right]$$
$$y(t) = -a\sin(\omega_\perp t)\cos\left[\left(\omega_\perp^2 + \frac{g}{l}\right)^{1/2} t\right] + a\omega_\perp\left(\omega_\perp^2 + \frac{g}{l}\right)^{-1/2}\cos(\omega_\perp t)\sin\left[\left(\omega_\perp^2 + \frac{g}{l}\right)^{1/2} t\right],$$

and if we study the solution in the limit where $\omega_\perp \ll \frac{g}{l}$, which for the simple pendulum would be the frequency of the pendulum, we get

$$x(t) = a\cos(\omega_\perp t)\cos\left[\left(\frac{g}{l}\right)^{1/2} t\right] \tag{8}$$
$$y(t) = -a\sin(\omega_\perp t)\cos\left[\left(\frac{g}{l}\right)^{1/2} t\right]. \tag{9}$$

Now the pendulum goes through the origin and we are lead to the conclusion:

$$\tan\phi = \frac{y(t)}{x(t)} = -\tan\omega_\perp t$$

$$\Leftrightarrow \quad \phi = -\cos(\theta)\omega t = -2\pi\cos(\theta)\frac{t}{T}, \tag{10}$$

where T is the period of the rotation of the earth — that is one day. Now to the point: ϕ seems at first glance to depend on time, but in fact all it depends on is angle the sphere has turned. After a day for instance, with no reference to how long a day is, the pendulum has precessed through an angle $-2\pi\cos\theta$. The pendulum is said to have picked up a geometric phase meaning that the phase is dependent only on the angle which the sphere has turned, ωt, and on θ.

The Foucault pendulum is treated in many textbooks on classical mechanics often as an example of the coriolis force. The strategy is often the same as the one used here, utilizing an accelerated frame and fictitious forces. It is not immediate, however, how the geometric phase relates to the geometry of the problem and why the result takes the form it does. It has a very nice and clear explanation which is fruitful to pursue, but in order to do so we have to venture into the realm of differential geometry which we will do in the next section before returning to the Foucault pendulum.

2.2 Review of differential geometry

In the following section, I will review a few of the important concepts in differential geometry. The central object of concern is that of differential manifolds. The exact, rigorous definition will not be important for the applications in this treatment and for more details I strongly suggest Carroll [5] or for a more mathematical treatment Schlichtkrull [6]. Instead I will follow the typical intuitive approach by physicists and define an m-dimensional manifold, M, to be a space that locally looks like $\mathbb{R}^m$ in a well defined way[1]. Likewise, I will define the tangent space T_pM at a point p on the manifold as the linear span of all tangents to curves on M through p. T_pM is a vector space and an element $V \in T_pM$ is called a vector. We can also define the dual space $T_p^\star M$ at p as the vector space of all linear mappings $T_pM \to \mathbb{R}$. An element ω in the dual space is sometimes called a dual vector.

A basis $\{\hat{e}_{(\mu)}\}$ for $\mu \in (1,\dots,m)$ can be chosen for T_pM as well as a basis $\{\hat{\theta}^{(\mu)}\}$ for $T_p^\star M$ by demanding $\hat{e}_{(\mu)}(\hat{\theta}^{(\nu)}) = \hat{\theta}^{(\nu)}(\hat{e}_{(\mu)}) = \delta^\nu_\mu$, where δ^ν_μ is Kronecker's delta function equal to 1 if $\nu = \mu$ and 0 otherwise. A vector can then be written $V = V^\mu \hat{e}_{(\mu)}$ and a dual vector $\omega = \omega_\mu \hat{\theta}^{(\mu)}$. We now adopt the Einstein summation convention so we always sum over all possible values of an index if it appears both up and down in the same term[2]. With this in mind, the action of vectors and dual vectors on each other is defined, with the summations written out explicitly for clarity:

$$\omega(V) = \omega_\mu\hat{\theta}^{(\mu)}(V^\nu\hat{e}_{(\nu)}) = \omega_\mu V^\nu\hat{\theta}^{(\mu)}(\hat{e}_{(\nu)}) = \omega_\mu V^\nu\delta^\mu_\nu = \sum_{\mu,\nu\in\{1,\dots,m\}}\omega_\mu V^\nu\delta^\mu_\nu = \sum_{\mu\in\{1,\dots,m\}}\omega_\mu V^\mu = \omega_\mu V^\mu,$$

and likewise $V(\omega) = \omega_\mu V^\mu$ so that a vector at point p can be thought of as a linear mapping $T_p^\star \to \mathbb{R}$.

The basis $\{\hat{e}_{(\mu)} = \frac{\partial}{\partial x^\mu}\}_{\mu\in(1,\dots,m)}$ is called the standard basis. Considering basis vectors as partial derivatives may seem odd at first but the idea is that tangent vectors can be thought of as directional derivatives of functions that map from $\mathbb{R}^m$ onto the manifold (this approach has the nice feature that it also makes sense for abstract manifolds).

With tangent- and dual vectors defined, we can now define the more general term of a tensor.

[1] Meaning that there exists a coordinate system at all points $p \in M$ such that the metric (to be defined in a while) is the identity to first order in the coordinates in this coordinate system

[2] This is why there is put a parenthesis on the basis vectors index - these are not to be summed over.

A useful, albeit not particularly deep, definition of a rank (k, l) tensor $T^{\mu_1\mu_2\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_l}$ is that it is an object that transforms in the following way:

$$T^{\mu_1'\mu_2'\cdots\mu_k'}{}_{\nu_1'\nu_2'\cdots\nu_l'} = \frac{\partial x^{\mu_1}}{\partial x^{\mu_1'}}\frac{\partial x^{\mu_2}}{\partial x^{\mu_2'}}\cdots\frac{\partial x^{\mu_k}}{\partial x^{\mu_k'}}\frac{\partial x^{\nu_1'}}{\partial x^{\nu_1}}\frac{\partial x^{\nu_2'}}{\partial x^{\nu_2}}\cdots\frac{\partial x^{\nu_l'}}{\partial x^{\nu_l}}T^{\mu_1\mu_2\cdots\mu_k}{}_{\nu_1\nu_2\cdots\nu_l}. \tag{11}$$

It can be thought of as having k dimensions that behave like tangent vectors and l dimensions that behave like dual vectors. In that sense, a vector is a rank $(1, 0)$ tensor and a dual vector is a rank $(0, 1)$ vector. If we assign a tensor to a particular point p on the manifold, the manifold is a multilinear mapping, $T_p^\star M \times \ldots \times T_p^\star M \times T_pM \times \ldots \times T_pM \to \mathbb{R}$, with $T_p^\star M$ appearing k times and T_pM appearing l times. Differential geometry is a particularly nice tool when working with abstract manifolds but introducing tensors vindicates the use of doing differential geometry on manifolds in $\mathbb{R}^n$ since one can then formulate quantities independently of coordinates since a change of coordinates for a tensor is given by equation (11).
The next thing to do is to introduce an inner product on the manifold. When doing so the enterprise is called Riemannian geometry after the great mathematician, Bernhard Riemann. The inner product is introduced by a rank $(0, 2)$ tensor called the metric tensor which is denoted $g_{\mu\nu}$ and which is symmetric in its indices, that is $g_{\mu\nu} = g_{\nu\mu}$. The metric tensor defines what it means to lower and raise indices by

$$V_\mu \equiv g_{\mu\nu}V^\nu \tag{12}$$
$$\omega^\mu \equiv g^{\mu\nu}\omega_\nu, \tag{13}$$

and noting that this gives the definition of $g^{\mu\nu}$ by

$$g^{\mu\nu} = g^{\mu\rho}g^{\nu\sigma}g_{\rho\sigma}, \qquad g^{\nu\sigma}g_{\rho\sigma} = \delta^\nu_\sigma. \tag{14}$$

so $g^{\mu\nu}$ is the inverse of $g_{\mu\nu}$. With this in mind we now define the inner product $\langle V|U\rangle$ between two vectors V and U in the same tangent space as

$$\langle V|U\rangle \equiv V_\mu U^\mu = g_{\mu\nu}V^\mu U^\nu \quad (= V^\mu U_\mu). \tag{15}$$

The inner product is also sometimes called a contraction because the dependence on one index is removed. If infinitesimal lengths, ds^2, on the manifold are known, then $g_{\mu\nu}$ is also known since

$$ds^2 = \langle dx|dx\rangle = g_{\mu\nu}dx^\mu dx^\nu, \tag{16}$$

so $g_{\mu\nu}$ can be read off as the coefficients. Now that we have the basics and know what a tensor is, we can wonder how we can quantify the change of a tensor. Partial derivatives are very useful in Euclidean space when wanting to compute rates of changes but sadly, partial derivatives of tensors are not themselves tensors because of the Leibniz rule, illustrated here on a vector for clarity:

$$\frac{\partial}{\partial x^\mu}V^\nu = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial}{\partial x^{\mu'}}\Big(\frac{\partial x^\nu}{\partial x^{\nu'}}V^{\nu'}\Big) = \frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial x^\nu}{\partial x^{\nu'}}\frac{\partial}{\partial x^{\mu'}}V^{\nu'} + V^{\nu'}\frac{\partial x^{\mu'}}{\partial x^\mu}\frac{\partial^2 x^\nu}{\partial x^{\mu'}\partial x^{\nu'}}.$$

We would like a linear operation fulfilling Leibniz' product rule generalizing the notion of changes in tensors in the sense that it reduces partial derivative in flat space. This operator is called the covariant derivative and is denoted

$$\nabla_\mu V^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda}V^\lambda, \tag{17}$$

where $\Gamma^{\nu}_{\mu\lambda}$ is called the connection and is chosen so that $\nabla_\mu V^\nu$ and $\nabla_\mu \omega_\nu$ transform according to (11). Note that this means that the connection is not a tensor since $\partial_\mu V^\nu$ is not a tensor. When it is further demanded that the covariant derivative commutes with contractions, that is

$$\nabla_\mu (T^{\lambda}{}_{\lambda\rho}) = (\nabla T)_{\mu}{}^{\lambda}{}_{\lambda\rho}, \tag{18}$$

and demanding for every scalar ϕ that

$$\nabla_\mu \phi = \partial\phi, \tag{19}$$

it can be shown that

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^{\lambda}_{\mu\nu}\omega_\lambda. \tag{20}$$

If it is further assumed that the connection is metric compatible — meaning $\nabla_\mu g_{\mu\nu} = 0$ — and torsion free — meaning $\Gamma^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\nu\mu}$ — then it can be shown that the connection is uniquely given by

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma}\big(\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}\big), \tag{21}$$

which is called the Christoffel connection. Now that we have the covariant derivative and the Christoffel connection we can make sense of comparing vectors in different tangent spaces on the manifold. In flat space we would compare vectors at two different points visually simply by moving them while keeping them constant until they start in the same point. The act of keeping a vector V^μ constant while moving it along a curve $\gamma : I \to M$ with $I \subset R$ amounts to demanding

$$0 = \frac{d}{d\lambda} V^\mu(\lambda) = \frac{dx^\nu}{d\lambda}\partial_\nu V^\mu.$$

We want our condition of keeping a vector constant while moving it about on a general manifold to be independent of coordinates and we now know that the partial derivative is not tensorial and the above condition is therefore dependent on coordinates. We therefore simply replace our partial derivative in the above with a covariant and define that a vector $V(\lambda) \in T_{\gamma(\lambda)}M$ is parallelly transported along a curve $\gamma(\lambda)$ on M if

$$\frac{dx^\nu}{d\lambda}\nabla_\nu V^\mu = 0,$$

or

$$\frac{d}{d\lambda} V^\mu + \Gamma^{\mu}_{\sigma\rho}\frac{dx^\sigma}{d\lambda} V^\rho = 0. \tag{22}$$

Note that parallel transportation preserves the norm if a metric compatible connection is used since if two vectors V and W are parallelly transported, then

$$\frac{dx^\nu}{d\lambda}\nabla_\nu (g_{\mu\nu} V^\mu W^\nu) = V^\mu W^\nu \frac{dx^\nu}{d\lambda}\nabla_\nu g_{\mu\nu} + g_{\mu\nu} W^\nu \frac{dx^\nu}{d\lambda}\nabla_\nu V^\mu + g_{\mu\nu} V^\mu \frac{dx^\nu}{d\lambda}\nabla_\nu W^\nu = 0 \tag{23}$$

The curves on M whose tangent vectors are themselves parallelly transported are called geodesics and one can show that the geodesics of the Christoffel connections uniquely define the shortest routes between points on the manifold. For this reason and since the Christoffel connections are metric compatible, so that parallel transport with respect to them preserves norms, the rest of this project will assume that Christoffel connections are used.

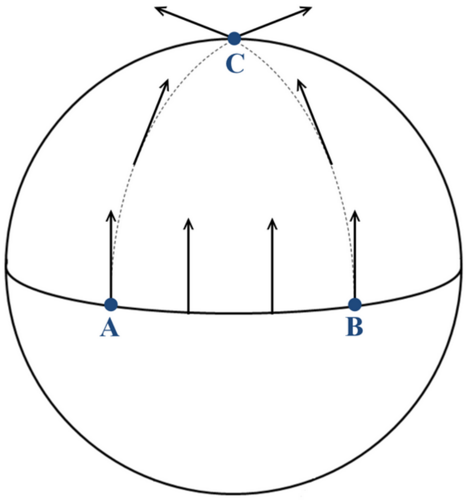

Figure 3: Illustration of a sphere with a vector at point A being parallelly transported along two different curves to the point C.

On figure 3, the notion of parallel transport along geodesics is illustrated. Here the manifold is a sphere so the geodesics can be shown to be great circles. Note that since parallel transportation preserves norms, the angle between a vector and a tangent vectors to a great circles is preserved during parallel transportation along said great circle. On the figure, a vector at point A is transported parallelly two different ways to the north pole. One way is due north along a geodesic and the other is east for some distance and then due north afterwards. Since the angle to the geodesics must be kept constant, the results of transporting the vector are different. This is a good illustration of the fact that it is not obvious how to compare two vectors at different points on the manifold because moving the vector along different curves yields different results.

Equation (22) can be solved exactly, yielding a practical formula as long as the metric and connections for the given system is known. We look for a solution for a given curve $\gamma(\lambda)$ on M where a vector is linearly related with its parallel transported in another tangent space. That is, we are looking for a linear map $T_{x(\lambda_0)} \to T_{x(\lambda)}$ so that $V^\mu(\lambda) \in T_{x^\mu(\lambda)}$ is related to $V^\mu(\lambda_0)$ by

$$V^\mu(\lambda) = \mathbb{P}^\mu{}_\rho(\lambda, \lambda_0) V^\rho(\lambda_0). \tag{24}$$

$\mathbb{P}^\mu{}_\rho$ is called the parallel propagator and the objective now is to find an expression for this that solves the parallel transportation equation.

By defining

$$A^\mu{}_\rho(\lambda) = -\Gamma^\mu_{\sigma\rho} \frac{dx^\sigma}{d\lambda}, \tag{25}$$

equation (22) takes the form

$$\frac{d}{d\lambda} V^\mu = A^\mu{}_\rho V^\rho. \tag{26}$$

Plugging (24) into (26) gives

$$\frac{d}{d\lambda} \mathbb{P}^\mu{}_\rho(\lambda, \lambda_0) A^\rho(\lambda_0) = A^\mu{}_\rho(\lambda) \mathbb{P}^\mu{}_\rho(\lambda, \lambda_0) A^\rho(\lambda_0)$$

$$\Leftrightarrow \quad \frac{d}{d\lambda} \mathbb{P}^\mu{}_\rho(\lambda, \lambda_0) = A^\mu{}_\rho(\lambda) \mathbb{P}^\mu{}_\rho(\lambda, \lambda_0). \tag{27}$$

This can be solved first by integration and then iteration:

$$\mathbb{P}^{\mu}{}_{\rho}(\lambda,\lambda_0) = \delta^{\mu}_{\rho} + \int_{\lambda_0}^{\lambda} A^{\mu}{}_{\sigma}(\eta)\mathbb{P}^{\sigma}{}_{\rho}(\eta,\lambda_0)d\eta \tag{28}$$

$$\Rightarrow \quad \mathbb{P}^{\mu}{}_{\rho}(\lambda,\lambda_0) = \delta^{\mu}_{\rho} + \int_{\lambda_0}^{\lambda} A^{\mu}{}_{\rho}(\eta)d\eta + \int_{\lambda_0}^{\lambda}\int_{\lambda_0}^{\eta_2} A^{\mu}{}_{\sigma}(\eta_2)A^{\sigma}{}_{\rho}(\eta_1)d\eta_1 d\eta_2 + \int_{\lambda_0}^{\lambda}\int_{\lambda_0}^{\eta_3}\int_{\lambda_0}^{\eta_2} A^{\mu}{}_{\sigma}(\eta_3)A^{\sigma}{}_{\nu}(\eta_2)A^{\nu}{}_{\rho}(\eta_1)d\eta_1 d\eta_2 d\eta_3 + \dots \tag{29}$$

In (28) the delta function is introduced so the parallel propagator reduces to the identity when $\lambda = \lambda_0$. It is convenient to introduce the path-ordering operator $\mathcal{P}$ that arranges a product depending on different parameters from largest values of the parameter to the smallest. For instance if $\eta_1 < \eta_2$ then $\mathcal{P}[A(\eta_1)A(\eta_2)] = A(\eta_2)A(\eta_1)$. Using this gives in matrix notation

$$\mathbb{P}(\lambda,\lambda_0) = \mathbb{1} + \sum_{n=1}^{\infty}\frac{1}{n!}\int_{\lambda_0}^{\lambda}\dots\int_{\lambda_0}^{\lambda}\mathcal{P}[A(\eta_n)A(\eta_{n-1})\cdots A(\eta_1)]d\eta_1 d\eta_2\cdots d\eta_n \tag{30}$$

This can be reduced to a more appealing equation by noticing that the above is an exponential series so that

$$\mathbb{P}(\lambda,\lambda_0) = \mathcal{P}\exp\left(\int_{\lambda_0}^{\lambda} A(\eta)d\eta\right),$$

or with indices

$$\mathbb{P}^{\mu}{}_{\nu} = \mathcal{P}\exp\left(-\int_{\lambda_0}^{\lambda}\Gamma^{\mu}_{\sigma\nu}\frac{dx^{\sigma}}{d\eta}d\eta\right). \tag{31}$$

2.3 The Foucault pendulum revisited

We are now ready to return to the Foucault pendulum for a completely geometric approach to deriving (10). The idea is to assume that the pendulum is oscillating at all times in a plane that is free to rotate slowly. This is justified when the pendulum is much smaller than the radius of the sphere and the sphere is rotating with a frequency much slower than the pendulum. In this case the pendulum hardly feels the fictitious forces and the deviation from normal pendulum motion is a small correction.

We will take the pendulum as living on the manifold $\mathcal{S}^2$, which is a sphere, and the orientation of the oscillation on the initial point p on the sphere is represented by a vector $V^{\mu} \in T_p\mathcal{S}^2$, say a vector orthogonal to the plane of oscillations. Instead of considering the earth as rotating with the pendulum on it, we will perceive the situation as the vector V^{μ} being parallelly transported along a curve $\gamma : \mathbb{R} \to \mathcal{S}^2$. If we let the path be given by $\gamma(\lambda) = (\theta, \lambda)$ in spherical coordinates with a constant polar angle θ, then the situation is reminiscent of the one discussed earlier.

In spherical coordinates infinitesimal lengths are given by

$$ds^2 = dR^2 + R^2 d\theta^2 + R^2\sin^2\theta d\phi^2 \tag{32}$$

so on a sphere with constant radius R,

$$ds^2 = R^2 d\theta^2 + R^2\sin^2\theta d\phi^2.$$

This means that the metric tensor is given by

$$g_{ij} = \begin{pmatrix} R^2 & 0 \\ 0 & R^2 \sin\theta \end{pmatrix} \tag{33}$$

and thus

$$g^{ij} = \begin{pmatrix} \frac{1}{R^2} & 0 \\ 0 & \frac{1}{R^2 \sin\theta} \end{pmatrix}. \tag{34}$$

Calculating the Christoffel symbols using equation (21) gives the only non-zero contibutions:

$$\Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta \tag{35}$$

$$\Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} = \cot\theta. \tag{36}$$

On the path γ, we have

$$\frac{dx^\sigma}{d\lambda} = \delta^\sigma_\phi. \tag{37}$$

Therefore, by inserting in (25) we get

$$\begin{aligned} A^{\theta}{}_{\theta} &= -\Gamma^{\theta}_{\sigma\theta}\delta^{\sigma}_{\phi} = 0 \\ A^{\theta}{}_{\phi} &= -\Gamma^{\theta}_{\phi\phi}\delta^{\phi}_{\phi} - \Gamma^{\theta}_{\theta\phi}\delta^{\theta}_{\phi} = \sin\theta\cos\theta \\ A^{\phi}{}_{\phi} &= -\Gamma^{\phi}_{\phi\phi}\delta^{\phi}_{\phi} - \Gamma^{\phi}_{\theta\phi}\delta^{\theta}_{\phi} = 0 \\ A^{\phi}{}_{\theta} &= -\Gamma^{\phi}_{\sigma\theta}\delta^{\sigma}_{\phi} = -\Gamma^{\phi}_{\theta\theta}\delta^{\theta}_{\phi} - \Gamma^{\phi}_{\phi\theta}\delta^{\phi}_{\phi} = -\cot\theta, \end{aligned}$$

or, written more conveniently,

$$A = \begin{pmatrix} 0 & \sin\theta\cos\theta \\ -\cot\theta & 0 \end{pmatrix}. \tag{38}$$

Now we can use equation (30) to find the parallel propagator for the 2-sphere along this particular path. Notice how A is independent of the parameter λ in this case. This means that the path-ordering operator in this case doesn't change anything:

$$\begin{aligned} \mathbb{P}(\lambda,\lambda_0) &= \mathbb{1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\lambda_0}^{\lambda} \int_{\lambda_0}^{\lambda} \cdots \int_{\lambda_0}^{\lambda} d\eta_1 d\eta_2 \ldots d\eta_n \mathcal{P}[A(\eta_1)A(\eta_2)\ldots A(\eta_n)] \\ &= \mathbb{1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\lambda_0}^{\lambda} d^n\eta \mathcal{P}[A^n] = \mathbb{1} + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{\lambda_0}^{\lambda} d^n\eta A^n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{n=1}^{\infty} \frac{1}{n!} (\lambda-\lambda_0)^n A^n \end{aligned}$$

It can be shown by induction that

$$\begin{pmatrix} 0 & a \\ b & 0 \end{pmatrix}^n = \begin{cases} \begin{pmatrix} (ab)^{n/2} & 0 \\ 0 & (ab)^{n/2} \end{pmatrix} & n \text{ even}, \\ \begin{pmatrix} 0 & a^{\frac{n+1}{2}} b^{\frac{n-1}{2}} \\ a^{\frac{n-1}{2}} b^{\frac{n+1}{2}} & 0 \end{pmatrix} & n \text{ odd}. \end{cases} \tag{39}$$

Thus, comparing (38) with (39) we obtain

$$\mathbb{P}(\lambda,\lambda_0) = \sum_{n \text{ even}} \frac{(\lambda-\lambda_0)^n}{n!} \begin{pmatrix} (-\cot\theta\sin\theta\cos\theta)^{n/2} & 0 \\ 0 & (-\cot\theta\sin\theta\cos\theta)^{n/2} \end{pmatrix}$$
$$+ \sum_{n \text{ odd}} \frac{(\lambda-\lambda_0)^n}{n!} \begin{pmatrix} 0 & (-1)^{\frac{n-1}{2}} \frac{\cos^{\frac{n-1}{2}}\theta}{\sin^{\frac{n-1}{2}}\theta} \sin^{\frac{n+1}{2}}\theta\cos^{\frac{n+1}{2}}\theta \\ (-1)^{\frac{n+1}{2}} \frac{\cos^{\frac{n+1}{2}}\theta}{\sin^{\frac{n+1}{2}}\theta} \sin^{\frac{n-1}{2}}\theta\cos^{\frac{n-1}{2}}\theta & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{n \text{ even}} (-1)^{n/2} \frac{(\lambda-\lambda_0)^n}{n!} \cos^{n/2}\theta & 0 \\ 0 & \sum_{n \text{ even}} (-1)^{n/2} \frac{(\lambda-\lambda_0)^n}{n!} \cos^{n/2}\theta \end{pmatrix}$$
$$+ \begin{pmatrix} 0 & -\sum_{n \text{ odd}} (-1)^{\frac{n+1}{2}} \frac{(\lambda-\lambda_0)^n}{n!} \cos^n\theta\sin\theta \\ \sum_{n \text{ odd}} (-1)^{\frac{n+1}{2}} \frac{(\lambda-\lambda_0)^n}{n!} \frac{\cos^n\theta}{\sin\theta} & 0 \end{pmatrix}$$

$$\Rightarrow \quad \mathbb{P}(\lambda,\lambda_0) = \begin{pmatrix} \cos\left[\beta(\lambda-\lambda_0)\right] & \sin\theta\sin\left[\beta(\lambda-\lambda_0)\right] \\ -\sin^{-1}\theta\sin\left[\beta(\lambda-\lambda_0)\right] & \cos\left[\beta(\lambda-\lambda_0)\right] \end{pmatrix} \tag{40}$$

where $\cos\theta \equiv \beta$. Notice that $\mathbb{P}(\lambda_0,\lambda_0) = \mathbb{1}$.
Now that we have the parallel propagator we can compute the angle a vector has rotated when taken around the path. For simplicity let $\lambda_0 = 0$. We then want to compare the vector V_λ with V_0 when

$$V_\lambda \equiv \mathbb{P}(\lambda,0)V_0 = \begin{pmatrix} \cos(\beta\lambda) & \sin\theta\sin(\beta\lambda) \\ -\sin^{-1}\theta\sin(\beta\lambda) & \cos(\beta\lambda) \end{pmatrix} \begin{pmatrix} V_0^\theta \\ V_0^\phi \end{pmatrix}.$$

Let V_0 be normalized. It then follows that V is also normalized since parallel transportation preserves the norm of a vector. The inner product between the two vectors is given by

$$\begin{aligned} g_{\mu\nu}V_0^\mu V_\lambda^\nu &= g_{\mu\nu}V_0^\mu \mathbb{P}(\lambda,0)V_0^\nu \\ &= R^2 V_0^\theta\Big(\cos(\beta\lambda)V_0^\theta + \sin\theta\sin(\beta\lambda)V_0^\phi\Big) + R^2\sin^2\theta V_0^\phi\Big(-\frac{1}{\sin\theta}\sin(\beta\lambda)V_0^\theta + \cos(\beta\lambda)V_0^\phi\Big) \\ &= R^2\cos(\beta\lambda)(V_0^\theta)^2 + R^2\sin^2\theta\cos(\beta\lambda)(V_0^\phi)^2 = g_{\mu\nu}V_0^\mu V_0^\nu\cos(\beta\lambda) = \cos(\beta\lambda), \end{aligned} \tag{41}$$

but the the inner product is in general also given by

$$g_{\mu\nu}V_0^\mu V_\lambda^\nu = \sqrt{\Big(g_{\mu\nu}V_0^\mu V_0^\nu\Big)\Big(g_{\rho\sigma}V_\lambda^\rho V_\lambda^\sigma\Big)}\cos\alpha = \cos\alpha, \tag{42}$$

where α is the sought after angle between V_λ and V_0. Using (41) and (42) we obtain that $\cos(\alpha) = \cos(\beta\lambda)$. This is compatible with

$$\alpha = -\cos(\theta)\lambda, \tag{43}$$

which is identical to (10), proving that the phase obtained by the Foucault pendulum is indeed geometric. This treatment allows the nice intuition that angular momentum is conserved in the system of the earth and the pendulum and when the earth rotates the orientation of the plane of pendulating is parallelly transported. The assumption we used in the first treatment of the Foucault pendulum coincides with the one we used now, being a sort of adiabaticity since we demanded that the pendulum at all times behaved like a regular pendulum but allowed it to act differently over longer time scales.

We shall now turn our attention toward adiabaticity more rigorously. In physics conserved quantities can often make solving problems a lot easier and as it turns out, there exists a conserved quantity when a system exhibiting oscillatory motion is changed slowly. This incidentally offers a nice way of computing geometric phases since geometric phases are unaffected by the change being slow. Finding this conserved quantity, however, relies on some subtle concepts in analytical mechanics which we will review first.

3 Action-angle variables and adiabatic invariance

The purpose of this section is to define the action-angle coordinates and show that the action is an adiabatic invariant. As good as every textbook has its own notation and they even disagree on some points[3]. Therefore I will start by quickly going through the relevant analytical mechanics before defining and discussing the action-angle coordinates. Next I will prove that the action is an adiabatic invariant. The proof is adapted from Wells and Siklos [8] and is important since it gives a stronger result than ones like Landau and Lifshitz [7] or [9] use and to which the literature often refers. The section ends with an example on how to apply the adiabatic theorem to obtain a geometric phase.

We will only work in a flat where the metric is the identity. This means that there is no difference in an index up or down. For clarity I therefore keep all indices down in the following.

3.1 Short recap of Hamiltonian mechanics

One of the advantage of analytical mechanics is that it is formulated using generalized coordinates $\{q_i\}$ and momenta $\{p_i\}$. This gives the theory flexibility since the theory is not bound to regular coordinates. The objective then is usually to determine equations of motions - differential equations of the coordinates - which can then be solved to yield the solution to the problem. These equations of motion are derived from either the Lagrangian or the Hamiltonian. This section is a brief recap of Hamiltonian mechanics. The Hamiltonian H is a Legendre transformation of the Lagrangian $L = T - V$ where T is the kinetic energy and V is the potential energy and is given by

$$H = \sum_i p_i \dot{q}_i - L, \tag{44}$$

where I have adopted the notation $\dot{q} = \frac{dq}{dt}$. In terms of the Lagrangian the generalized momenta are given by $p_i = \frac{\partial L}{\partial \dot{q}_i}$ so for systems where V is independent of $\dot{q}_i$ for all i and T depends quadratically on $\dot{q}_i$ and is independent of q_i for all i, then, using the Einstein summation convention, we have

$$H = \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - T + V = \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} - \dot{q}_i \frac{\partial V}{\partial \dot{q}_i} - T + V = 2T - T + V = T + V = E \tag{45}$$

So H is the total energy of the system in this case.
Once H is known the equations of motion are found using Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \tag{46}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{47}$$

[3]For instance, Goldstein [9] and Fetter-Walecka [4] disagree on whether the action and angle coordinates are cannonically conjugate, though they arrive at the same results.

It is possible to make a change of variables while preserving the form of Hamilton's equations. Such a transformations is called canonical. Denoting the new generalized coordinates $Q_1, Q_2 \ldots, Q_N$ and $P_1, \ldots, P_N$, then the new Hamiltonian $H'(Q_1, \ldots Q_N, P_1 \ldots P_N, t)$ is given by

$$H' = H + \frac{\partial S}{\partial t} \tag{48}$$

for some generating function S depending on either the $\{q_i\}$ or $\{p_i\}$ and either $\{Q_i\}$ or $\{P_i\}$ and time t. The new and old coordinates that S is independent of can be obtained through partial differentiation of S with respect to its coordinates. If for instance $S = S(q_1, \ldots, q_N, P_1 \ldots, P_N, t)$ then it can be shown that for the transformation to be canonical then

$$p_i = \frac{\partial S}{\partial q_i} \tag{49}$$

$$Q_i = \frac{\partial S}{\partial P_i} \tag{50}$$

An example of how to apply the theory of cannonical transformations to solve a seemingly difficult problem is given in appendix A. A particularly nice application of canonical transformations is Hamilton Jacobi theory where S is chosen so $Q \equiv \beta$ and $P \equiv \alpha$ are constants of motion. It is convenient to use a generating function of the above type $S(q_1, \ldots, q_N, P_1 \ldots, P_N, t)$. Noting from (46) and (47) that Q_i and P_i being constants of motion implies that a solution is $H' = 0$, it is seen that the transformation equation (48) is

$$H\big(q_1, \ldots, q_N, \frac{\partial S}{\partial q_i}, \ldots, \frac{\partial S}{\partial q_N}, t\big) + \frac{\partial S}{\partial t} = 0 \tag{51}$$

where I have used (49). This equation is called the Hamilton-Jacobi equation and S is called Hamilton's principal function. The equation can be solved for $S(q_1, \ldots, q_N, \alpha_1, \ldots, \alpha_N)$ (actually the solution would be dependent of an additional additive constant but this is not interesting for our purpose since all we are concerned with are partial derivatives of S) and then $Q_i = \beta_i$ can be found by the equation $\beta_i = \frac{\partial S}{\partial \alpha_i}$.
At this point the original coordinates can be found by (49) and q_i can be found by inverting (50) using $Q_i = \beta_i$ and $P_i = \alpha_i$.

When the Hamiltonian is independent of time then the principal function may be written

$$S(q_1, \ldots, q_N, \alpha_1, \ldots, \alpha_N, t) = W(q_1, \ldots, q_N, \alpha_1, \ldots, \alpha_N) - at. \tag{52}$$

W then is called Hamilton's characteristic function. A formula for W can be obtained by taking the time derivative

$$\frac{dW}{dt} = \frac{\partial W}{\partial q_i} \dot{q}_i \tag{53}$$

Using (49) and (52) it is seen that (53) becomes

$$\frac{dW}{dt} = p_i \dot{q}_i,$$

so

$$W = \int p_i \dot{q}_i dt = \int p_i dq_i \tag{54}$$

With this in mind let us continue discuss the important aciont-angle variables.

3.2 Action-angle coordinates

Consider a system exhibiting periodic motion and which is separable, meaning that the characteristic function can be written[4]

$$W(q_1, \ldots, q_N, \alpha_1, \ldots, \alpha_N) = \sum_{i=1}^{N} W_i(q_i, \alpha_1, \ldots, \alpha_N)$$

If the Hamiltonian does not explicitly depend on time then a particularly nice set of variables called action-angle variables may be chosen. Let the constant Hamiltonian be denoted E. Now define the action variables as

$$I_i = \frac{1}{2\pi} \oint p_i dq_i. \tag{55}$$

By using (49), (52) and separability we get $p_i = \frac{\partial W_i}{\partial q_i}$. Putting this into the equation for I_i gives

$$I_i = \frac{1}{2\pi} \oint \frac{\partial W_i(q_i, \alpha_1, \ldots, \alpha_N)}{\partial q_i} dq_i = I_i(\alpha_1, \ldots, \alpha_N),$$

which can be inverted to yield relations

$$\alpha_i = \alpha_i(I_1, \ldots, I_N). \tag{56}$$

For the Hamilton-Jacobi equation (51) to be satisfied, it must hold that a as defined in (52) is equal to E since W is independent of time. Equation (56) allows me to write W and S as functions of the I's instead of the α's and I will denote the resulting functions with a tilde so as to avoid thinking that the α's are evaluated in the I's, so

$$W\Big(q_1, \ldots, q_N, \alpha_1(I_1, \ldots, I_N), \ldots, \alpha(I_1, \ldots, I_N)\Big) = \widetilde{W}(q_1, \ldots, q_N, I_1, \ldots, I_N) \tag{57}$$

$$\widetilde{S}\Big(q_1, \ldots, q_N, I_1, \ldots, I_N\Big) = \widetilde{W} - a(I_1, \ldots, I_N)t, \tag{58}$$

where the time independence of H ensures that a is only a function of the I's since a dependency on q_i would introduce a the time dependent term $\dfrac{\partial a}{\partial q_i} t$ to H. The functions satisfy the Hamilton-Jacobi equation

$$H\Big(q_1, \ldots, q_2, \frac{\partial \widetilde{S}}{\partial q_1}, \ldots, \frac{\partial \widetilde{S}}{\partial q_N}\Big) + \frac{\partial \widetilde{S}}{\partial t} = 0, \tag{59}$$

with (49) and (50) giving

$$p_i = \frac{\partial \widetilde{S}}{\partial q_i} \tag{60}$$

$$\widetilde{Q}_i = \frac{\partial \widetilde{S}}{\partial I_i}, \tag{61}$$

[4] Note that if $N = 1$ then the system is automatically seperable. In the examples therefore I won't discuss seperability.

and note that $\dfrac{\partial \widetilde{S}}{\partial q_i} = \dfrac{\partial S}{\partial q_i}$ since demanding the constancy of the α's and the I's amount to the same when taking partial derivative with respect to the q's. So it is indeed the same p's as before choosing $\alpha_i(I_1, \ldots, I_N) = I_i$. Since (59) is satisfied

$$\begin{aligned}\widetilde{P}_i &= I_i = \text{constant}\\ \widetilde{Q}_i &\equiv \widetilde{\beta} = \text{constant}.\end{aligned}$$

Now we define the angle variables[5]

$$\theta_i = \frac{\partial \widetilde{W}(q_1, \ldots, q_N, I_1, \ldots, I_N)}{\partial I_i}. \tag{62}$$

The constancy of the $\widetilde{Q}_i$'s gives

$$\widetilde{Q}_i = \text{constant} = \frac{\partial \widetilde{S}}{\partial I_i} = \frac{\partial \widetilde{W} - at}{\partial I_i} = \theta_i - \Big[\frac{\partial a(I_1, \ldots I_N)}{\partial I_i}\Big]t = \widetilde{\beta} \tag{63}$$

Note that

$$E = H = a(I_1, \ldots, I_N), \tag{64}$$

because H satisfies (59). From this is seen that

$$\begin{aligned}\theta_i &= \nu_i t + \widetilde{\beta}, \\ \nu_i &\equiv \frac{\partial a(I_1, \ldots I_N)}{\partial I_i} = \frac{\partial}{\partial I_i} H.\end{aligned} \tag{65}$$

Now we just need to get an intuition about ν_i. A nice way to get this is by computing the change $\Delta\theta_i$ in the angle θ_i over one period using (62) and (49):

$$\Delta\theta_i \equiv \oint d\theta_i = \oint \frac{\partial \theta_i}{\partial q_i} dq_i = \oint \frac{\partial^2 \widetilde{W}}{\partial q_i \partial I_i} dq_i = \frac{\partial}{\partial I_i} \oint \frac{\partial \widetilde{W}}{\partial q_i} dq_i = \frac{\partial}{\partial I_i} \oint p_i dq_i = 2\pi \frac{\partial}{\partial I_i} I_i = 2\pi.$$

Let the period be τ. Then we now have

$$\begin{aligned} 2\pi &= \nu_i \tau \\ \Rightarrow \quad \nu_i &= \frac{2\pi}{\tau}, \end{aligned} \tag{66}$$

so that ν_i is the frequency of the system and we actually have that Hamilton's equations are satisfied by the action-angle coordinates θ and I:

$$\dot{\theta}_i = \nu_i = \frac{\partial H}{\partial I_i} \tag{67}$$

$$\dot{I}_i = 0 = \frac{\partial H}{\partial \theta_i}, \tag{68}$$

since from (64) H depends only on the action coordinates.

[5]Note that $\{\theta_i, I_i\}$ are not canonically conjugate with respect to the constant Hamiltonian produced by the generating function $\widetilde{S}$ since $\theta_i \neq \dfrac{\partial \widetilde{S}}{\partial I_i}$.

Even though I won't use it in this paper there is a very practical formula I will derive now for completion. First note that from Hamilton's equations we have

$$\dot{q} = \frac{\partial H}{\partial p} \quad \Leftrightarrow \quad dt = \left(\frac{\partial H}{\partial p}\right)^{-1} dq$$

This gives us a neat formula for calculating the freqency ω of a system executing periodic motion in one dimension since the period T is

$$T = \int_0^T dt = \oint \left(\frac{\partial H}{\partial p}\right)^{-1} dq.$$

Suppose that $\left(\frac{\partial H}{\partial p}\right)^{-1} = \frac{\partial p}{\partial H}$, which holds for instance when H depends only on p. In that case,

$$T = 2\pi \frac{\partial}{\partial H}\left(\frac{1}{2\pi}\oint p dq\right) = 2\pi \frac{\partial I}{\partial H},$$

which gives the relation

$$\frac{\partial I}{\partial H} = \omega^{-1}, \tag{69}$$

valid in the above mentioned cases.

3.3 Adiabatic invariance

Now we are ready to define adiabatic invariance and prove the classical adiabatic theorem. A classic treatment of the subject is given by Landau and Lifshitz [7] but the derivation is cumbersome and the result is not as strong as it could be. Recently Wells and Siklos [8], using a formulation of adiabatic invariance similar to Arnold [11], have provided a more transparent proof which also gives the slightly stronger result. This section follows in their footsteps.
We wish to understand what happens to a system executing a periodic motion in a single dimension when the Hamiltonian is being slowly changed[6]. The change is introduced by the dependency of the Hamiltonian on a single parameter λ. Generalizing this to an arbitrary number of parameters doesn't change the idea in the proof - it only makes it more tedious.
More formally let $\epsilon > 0$ be arbitrary and consider changes in the parameter of the form $\lambda(t) = f(\epsilon t)$ for som smooth function $f : \mathbb{R} \to \mathbb{R}$ with $\dot{\lambda}(t) = \mathrm{O}(\epsilon)$ and $\ddot{\lambda}(t) = \mathrm{O}(\epsilon^2)$, then a quantity $B(t)$ is said to be an adiabatic invariant if

$$\forall t \in \left[0, \frac{1}{\epsilon}\right] : |B(t) - B(0)| = \mathrm{O}(\epsilon).$$

In the previous section we defined action-angle coordinates for Hamiltonians independent on time. Now we have a time-dependency and slightly more care must be given in the definitions of the coordinates. First we have the action coordinates

$$I \equiv \frac{1}{2\pi}\oint_C p dq, \tag{70}$$

where the integration is taken on a curve C on which the Hamiltonian H is constant and λ is kept constant. For one such value of H and λ we can follow the discussion of the preceding section, giving us a function $\widetilde{W}(q, I, \lambda)$ and an angle coordinate θ defined by

$$\theta \equiv \frac{\partial \widetilde{W}}{\partial I}. \tag{71}$$

[6]The discussion above assumed that the Hamiltonian was time independent. This means that now sets of action-angle coordinates are well-defined in short time intervals at given times, and for such a set the Hamiltonian is assumed constant in the interval.

Then $\{\theta, I\}$ are a canonically conjugate pair of variables satisfying Hamilton's equation with respect to the Hamiltonian given by the canonnical transformation

$$K = H + \frac{\partial \widetilde{W}}{\partial t} = H + \dot{\lambda}\frac{\partial \widetilde{W}}{\partial \lambda}. \tag{72}$$

Using (72) we can make a Taylor expansion of $I(q, H, \lambda)$ yielding

$$I(H,\lambda) = I(K - \frac{\widetilde{W}}{\partial I}, \lambda) = I(K,\lambda) - \dot{\lambda}\frac{\partial \widetilde{W}}{\partial \lambda}\frac{\partial I}{\partial H}\bigg|_{H=K} + \mathrm{O}(\dot{\lambda^2}). \tag{73}$$

The next step is crucial in the proof. We define

$$J(K,\lambda) \equiv \frac{1}{2\pi}\int_0^{2\pi} I d\theta, \tag{74}$$

where the integration is along the curve given by a constant K.
If we plug (73) into (73) we get, remembering that K is kept constant in the integral,

$$\begin{aligned} J &= I(K,\lambda) - \frac{\dot{\lambda}}{2\pi}\frac{\partial I}{\partial H}\bigg|_{H=K}\int_0^{2\pi}\frac{\partial \widetilde{W}}{\partial \lambda} d\theta + \mathrm{O}(\dot{\lambda^2}) \\ &= I(H,\lambda) + \dot{\lambda}\frac{\partial \widetilde{W}}{\partial \lambda}\frac{\partial I}{\partial H} - \frac{\dot{\lambda}}{2\pi}\frac{\partial I}{\partial H}\bigg|_{H=K}\int_0^{2\pi}\frac{\partial \widetilde{W}}{\partial \lambda} d\theta + \mathrm{O}(\dot{\lambda^2}). \end{aligned} \tag{75}$$

This means that to first order $|J - I| \propto \dot{\lambda}$ since I and $\frac{\partial \widetilde{W}}{\partial \lambda}$ don't depend on $\dot{\lambda}$ implicitly so that

$$|J - I| = \mathrm{O}(\epsilon). \tag{76}$$

The hope is to show now that J is an adiabatic invariant since equation (76) then gives that I is an adiabatic invariant. The strategy is to compute $\frac{dJ}{dt}$ and integrate this from 0 to $\frac{1}{\epsilon}$. Remembering that J is only a function of K and λ we have

$$\frac{dJ}{dt} = \left(\frac{\partial J}{\partial \lambda}\right)_{t,K}\frac{d\lambda}{dt} + \left(\frac{\partial J}{\partial K}\right)_{t,\lambda}\left(\frac{\partial K}{\partial t}\right)_{\theta,I} = \left(\frac{\partial J}{\partial t}\right)_K + \left(\frac{\partial J}{\partial t}\right)_{\lambda,\theta,I} = \left(\frac{\partial J}{\partial t}\right)_{\theta,I}. \tag{77}$$

By using equation (73) and (75) we thus obtain

$$\frac{dJ}{dt} = \left(\frac{\partial}{\partial t}\right)_{\theta,I}\left[I(H,\lambda) + \dot{\lambda}\frac{\partial \widetilde{W}}{\partial \lambda}\frac{\partial I}{\partial H} - \frac{\dot{\lambda}}{2\pi}\frac{\partial I}{\partial H}\int_0^{2\pi}\frac{\partial \widetilde{W}}{\partial \lambda} d\theta + \mathrm{O}(\dot{\lambda^2})\right]. \tag{78}$$

The first term drops out because the derivative is taken with constant I. Again because I and $\widetilde{W}$ don't implicitly depend on $\dot{\lambda}$ and because $\dot{\lambda} = \mathrm{O}(\epsilon)$ and $\ddot{\lambda} = \mathrm{O}(\epsilon^2)$ we have

$$\frac{dJ}{dt} = \mathrm{O}(\epsilon^2). \tag{79}$$

Integrating this from $t = 0$ to $t = \frac{1}{\epsilon}$ gives that $|J(t) - J(0)| = \mathrm{O}(\epsilon)$. From this and equation (76), we then have the desired, $|I(t) - I(0)| = \mathrm{O}(\epsilon)$, so I is an adiabatic invariant.

In the classic textbook on classical mechanics by Landau and Lifshitz [7] it is stated that the average of $\frac{dI}{dt}$ is an adiabatic invariant but as has been shown here, the stronger statement that I is an adiabatic invariant is true. Applying this result is a lot easier than having to first deal with the averages. Dealing with the averaging was done instead by the function J introduced in this proof, but as has been seen I itself is adiabatically preserved In the problem in appendix A it is demonstrated how I is adiabatically conserved for a harmonic oscillator with time dependent frequency.

3.4 Example of a geometric phase using adiabatic invariance

As an example of utilizing the adiabatic invariance of the action coordinate, consider a bead sliding without fricion on a circular hoop with radius R as shown on figure 4.

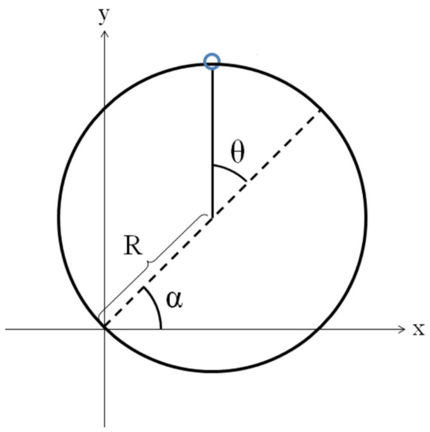

Figure 4: Sketch of the the problem.

The hoop is then slowly rotated with time dependent frequency $\frac{d\alpha}{dt} \equiv \Omega$. (x, y) denote the position in cartesian coordinates of the bead and θ and α are as shown in the figure. From the geometry on the figure it is seen that

$$x = R\cos(\alpha) + R\cos(\alpha + \theta),$$
$$y = R\sin(\alpha) + R\sin(\alpha + \theta). \tag{80}$$

By differentiating with respect to time

$$\dot{x} = -\Omega R\sin(\alpha) - (\Omega + \dot{\theta})R\sin(\alpha + \theta),$$
$$\dot{y} = \Omega R\cos(\alpha) + (\Omega + \dot{\theta})R\cos(\alpha + \theta), \tag{81}$$

and with this the Lagrangian L is obtained:

$$L = T - V = T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) = \frac{1}{2}mR^2\Big(\Omega^2 + (\Omega + \dot{\theta})^2 + 2\Omega(\Omega + \dot{\theta})\cos\theta\Big).$$

The generalized momentum p_θ conjugate to θ therefore is

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \frac{1}{2}mR^2\Big(2(\Omega + \dot{\theta}) + 2\Omega\cos\theta\Big) = mR^2\Big(\dot{\theta} + \Omega(1 + \cos\theta)\Big),$$

and the action associated with this momentum is

$$I = \frac{1}{2\pi}\int_0^{2\pi} p_\theta d\theta = \frac{mR^2}{2\pi}\int_0^{2\pi}\Big(\dot{\theta} + \Omega(1 + \cos\theta)\Big)d\theta.$$

Now imagine that Ω depends very slowly on time, going from 0 up to a certain point which is much lower than the rate the bead (that is, for all t we demand that $\Omega \ll \dot{\theta}$), before going back to 0. Then we can take Ω out of the integral since it is constant during the course of one period for the bead. This gives

$$I = mR^2\Big(\frac{1}{2\pi}\int_0^{2\pi}\dot{\theta}d\theta + \Omega\Big).$$

In this limit we know from the earlier discussion that I is an adiabatic invariant, so $\frac{1}{2\pi}\int_0^{2\pi}\dot{\theta}d\theta+\Omega\equiv C$ is preserved in time during the variation of Ω. More important the constant C is independent on the type of variation of Ω and in particular $\frac{1}{2\pi}\int_0^{2\pi}\dot{\theta}d\theta\equiv\omega_{\mathrm{av}}$ is equal to C when Ω is constantly 0. Thus if the time it takes to complete the variation of Ω from 0 up to some small value and back to 0 is τ then $\int_0^{\tau}\omega_{\mathrm{av}}dt$ is equal to the number of revolutions the small bead makes in the same time. If Ω had been 0 the whole time then $C\tau=\int_0^{\tau}Cdt=\int_0^{\tau}\omega_{\mathrm{av}}$, so $C\tau$ is the number of revolutions the the small bead makes in the time τ if $\Omega(t)=0$ for all $t\in[0,\tau]$. Letting Ω vary as described above and choosing to let it go around exactly once in the time τ, we thus have

$$C\tau=\int_0^{\tau}\omega_{\mathrm{av}}dt+\int_0^{\tau}\Omega dt=\int_0^{\tau}\omega_{\mathrm{av}}dt+2\pi,$$

and this means that the geometric phase of the bead, independent on τ, is $\gamma=-2\pi$ which makes sense especially in the case where the bead is not moving initially in the rest frame of the hoop. In this case the bead is simply moving around the hoop once.

4 The Hannay angle

In the last section I will outline how the geometric phase in classical mechanics is often explained. The reason that this way is chosen is probably that the obtained formula for computing the geometric phase looks a lot lik the formula for computing Berry's phase in quantum mechanics. Consider Hamiltonian H which is dependent on the parameters $\boldsymbol{R}=(R_1,R_2,\ldots)$ permitting a representation in action-angle coordinates. If all the parameters were constant then we would have $\frac{d\theta}{dt}=\frac{\partial\widetilde{W}}{\partial I}$. This mean that when the parameters are time-dependent we therefore have

$$\frac{d\theta}{dt}=\left(\frac{\partial\theta}{\partial t}\right)_{\boldsymbol{R}}+\frac{\partial\theta}{\partial\boldsymbol{R}}\cdot\dot{\boldsymbol{R}}=\frac{\partial\widetilde{W}}{\partial I}+\frac{\partial\theta}{\partial\boldsymbol{R}}\cdot\dot{\boldsymbol{R}}$$

The last term gives the geometric contribution to the phase, so when we change the parameters the change in angle is

$$\Delta\theta=\int d\theta=\int\dot{\theta}dt=\int\frac{\partial H}{\partial I}dt+\int\frac{\partial\theta}{\partial\boldsymbol{R}}\cdot\dot{\boldsymbol{R}}dt=\int\frac{\partial H}{\partial I}dt+\int\frac{\partial\theta}{\partial\boldsymbol{R}}\cdot d\boldsymbol{R},\tag{82}$$

whereby the geometric contribution is found in the last integral. The system before and after the excursion of the parameters is best compared if the parameters end with the same values as they begin. Notice from equation (82) that the path $\boldsymbol{R}$ changes along must enclose an area for it to give a contribution. Now consider cyclical changes in the parameters as described and denote the change from the last term in (82) by γ. Then we have

$$\gamma=\oint\frac{\partial\theta}{\partial\boldsymbol{R}}\cdot d\boldsymbol{R}\tag{83}$$

If the system is only quasi-periodic[7], one may have to remember to keep I constant during the integration since the system has to be periodic, but this is automatically achieved if the parameters are varied slowly enough thanks to the adiabatic theorem.

Equation (83) is reminiscent of the one for Berry's phase in quantum mechanics as described very well in Griffiths [3].

[7]That is, the trajectories in phase space only almost close

The analogy can be stretched quite far when the parameter space is 3-dimensional since we can also define the "connection" as $\boldsymbol{A} \equiv \frac{\partial \theta}{\partial \boldsymbol{R}}$ and "curvature"[8] as $\boldsymbol{B} \equiv \frac{\partial}{\partial \boldsymbol{R}} \times \boldsymbol{A}$. In that case, using Stokes theorem we have

$$\gamma = \int_{\partial \mathcal{S}} \boldsymbol{A} \cdot d\boldsymbol{R} = \int_{\mathcal{S}} \nabla \times \boldsymbol{A} \cdot d\boldsymbol{a} = \int_{\mathcal{S}} \boldsymbol{B} \cdot d\boldsymbol{a}, \tag{84}$$

where $\mathcal{S}$ is the area enclosed in parameter space during the excursion of the parameters and where $d\boldsymbol{a}$ is an infinitesimal area. So γ is the flux through the enclosed area by the curve in parameter space. In the example with the hoop the parameters were x and y and we changed these slowly around from one point and back again[9]. The geometric phase therefore is equal to the flux of $\boldsymbol{B}$ through the hoop, so finding $\boldsymbol{B}$ amounts to solving the problem. When there is references to Hannay's 2-form, what is meant is the "curvature" $\boldsymbol{B}$.

5 Conclusion

We have seen two different strategies to finding the geometric phase picked up by a system depending on parameters being changed slowly. The first strategy is to transform to an accelerated coordinate system and use Newtonian mechanics — with the inclusion of the appropriate fictitious forces — and possibly root out dynamic contributions by making some approximations. This approach was illustrated by solving the Foucault pendulum. The second strategy is to identify an appropriate action-angle pair, let the parameters be changed slowly and envoke the adiabatic theorem, using that the action then is conserved. This approach was useful for dealing with a bead on a rotating hoop. In a given problem one strategy may be more convenient than the other but they may both be viable. Khein and Nelson have solved the Foucault pendulum using action-angle variables [10] and Berry has solved the problem with the bead on the hoop using the Euler force appearing from the time dependent angular velocity of the hoop [12].
Regarding the first strategy, many textbooks on classical mechanics treat the Foucault pendulum in a similar fasion as we initially did and sometimes the word "parallel transport" is used to describe what happens without going into details. My contribution to this was to go through the computation in details and show that the geometric phase can be explained just by parallel transportation, hereby clearly demonstrating the nature of the acquired geometric phase. Regarding the second strategy, the supplied proof of the adiabatic invariance of I, and not just the average of I, makes practical applications of the theorem more transparent. I see many promising avenues for future research in geometric phases, particularly in the intersection between quantum mechanics and classical mechanics. It is my conjecture that insight into quantum mechanical behaviour might be gained by applying the strategies developed in this paper in semi-classical approximations[10]. I hope that this paper will inspire interest and help research in this area.

[8] These are the words commonly used although I'm not sure if they explicitly are related to the connections I have described in section 2.2. Perhaps one could think of it as imposing a non-trivial metric on the parameter space, but I have not been able to find a good rigorous way to do so and compare these with $\boldsymbol{A}$ and $\boldsymbol{B}$.

[9] The example looked deceptively like it was dependent on only one parameter, Ω, but the reason that x and y could be described with a single parameter α and through this Ω was due to the constraint the the bead was confined to move on the hoop.

[10] Of different ideas I persued, I particularly had hoped to find a semi-classical analogue to the anomalous velocity of particles moving in a space with a non-vanishing Berry "curvature" but it didn't work out in time for this project.

References

[1] M.V. Berry, *Quantal phase factors accompanying adiabatic change.* Proceedings of the royal society A: Mathematical and general vol. 392, March 1984.

[2] John H. Hannay, *Angle variable holonomy in adiabatic excursion of an integrable Hamiltonian.* Journal of physics A: Mathematical and general vol. 18 no. 2, Febuary 1985.

[3] David J. Griffiths, *Introduction to quantum mechanics.* Pearson Education, Inc., 2nd edition, international edition, 2004.

[4] A.L. Fetter and J.D. Walecka, *Theoretical mechanics of particles and continua.* Dover Publications, Inc., 1st edition, 2003.

[5] Sean M. Carroll, *An introduction to general relativity: Spacetime and geometry.* Pearson Education, Inc., 1st edition, 2004.

[6] Henrik Schlichtkrull, *Differentiable Manifolds, Lecture notes for geometry 2.* Department of Mathematics, University of Copenhagen, 2013.

[7] L.D. Landau and E.M. Lifshitz, *Course of Theoretical Physics Volume 1: Mechanics.* Butterworth-Heinemann, 3rd edition, 1976.

[8] C.G. Wells and S.T.C. Siklos, *The adiabatic invariance of the action variable in classical dynamics.* European journal of physics A vol. 28 no. 1, January 2007.

[9] H. Godstein, C.P. Poole, jr and J.L. Safko, *Classical mechanics.* Addison Wesley, 3rd edition, international edition, 2002.

[10] Alexander Khein and D.F. Nelson, *Hannay angle study of the Foucault pendulum in action-angle variables.* American journal of physics vol. 61, Febuary 1993.

[11] Vladimir I. Arnold, *Mathematical methods of classical mechanics.* Springer, 2nd edition, 1989.

[12] M.V. Berry, *Classical adiabatic angles and quantal adiabatic phase.* Journal of physics A: Mathematical and general vol. 18 no. 1, January 1985.

A Harmonic oscillator with time dependent frequency

In this section we will study as an example the harmonic oscillator with a time dependent frequency. The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2(t)x^2 \equiv E(t)$$

The goal is to find a new Hamiltonian K so that $\dot{\theta} = \frac{\partial K}{\partial I}$ and $\dot{I} = -\frac{\partial K}{\partial \theta}$. Since

$$K = E + \frac{\partial S}{\partial t}, \tag{85}$$

this goal boils down to finding S. For a regular Harmonic oscillator the solution is

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin\theta,$$
$$p = \sqrt{2mE}\cos\theta.$$

In the course of one period the trajectory in phase space is almost elliptical with half of the major axes $\sqrt{\frac{2E}{m\omega^2}}$ and $\sqrt{2mE}$ respectively. The area of this ellipsis thus is $\pi\sqrt{\frac{2E}{m\omega^2}} \cdot \sqrt{2mE} = 2\pi\frac{E}{\omega}$. I is defined as $I \equiv \frac{1}{2\pi}\oint pdq$, so it is per definition equal to that area divided by 2π. This means that

$$I = \frac{E}{\omega}.$$

From (49) we get

$$p = \sqrt{2mE}\cos\theta = \frac{\partial S(q, I, \omega)}{\partial q} = \frac{\partial S(q, I, \omega)}{\partial \theta},$$

but

$$q = \sqrt{\frac{2E}{m\omega^2}}\sin\theta \quad \Leftrightarrow \quad \theta = \arcsin\Big(\sqrt{\frac{m\omega^2}{2E}}q\Big),$$

$$\Rightarrow \quad \frac{\partial\theta}{\partial q} = \sqrt{\frac{m\omega^2}{2E}}\frac{1}{\sqrt{1 - \frac{m\omega^2}{2E}q^2}} = \sqrt{\frac{m\omega^2}{2E}}\frac{1}{\sqrt{1-\sin^2\theta}} = \sqrt{\frac{m\omega^2}{2E}}\frac{1}{\cos\theta},$$

so

$$\frac{\partial S}{\partial \theta} = \sqrt{2mE}\sqrt{\frac{2E}{m\omega^2}}\cos^2\theta$$
$$\Leftrightarrow \quad S = 2I\int\cos^2\theta d\theta.$$

Thus we can find $\dot{\theta}$ by using (85):

$$\dot{\theta} = \frac{\partial}{\partial I}\Big(E + \frac{\partial S}{\partial t}\Big) = \frac{\partial E}{\partial I} + \frac{\partial}{\partial I}\Big(\frac{\partial S}{\partial \omega}\frac{\partial \omega}{\partial t}\Big) = \omega + \dot{\omega}\frac{\partial}{\partial I}\Big[\frac{\partial S}{\partial \theta}\frac{\partial \theta}{\partial \omega}\Big] = \omega + \dot{\omega}\frac{\partial}{\partial I}\Big[2I\cos^2(\theta)\frac{\partial\theta}{\partial\omega}\Big].$$

To determine $\frac{\partial\theta}{\partial\omega}$ note that

$$q^2 = \frac{2I}{m\omega}\sin^2\theta \quad \Leftrightarrow \quad \sin^2\theta = \frac{m\omega q^2}{2I},$$

so the total differential is

$$2\sin\theta\cos\theta d\theta = \frac{mq^2}{2I}d\omega + \frac{m\omega q}{I}dq - \frac{m\omega q^2}{2I^2}dI,$$

$$\Rightarrow \quad \frac{\partial\theta}{\omega} = \frac{mq^2}{4I\sin\theta\cos\theta} = \frac{m}{4I}\frac{2I}{m\omega}\sin^2\theta\frac{1}{\sin\theta\cos\theta} = \frac{\tan\theta}{2\omega}.$$

Hereby the result is obtained

$$\dot\theta = \omega + \dot\omega\frac{\partial}{\partial I}\left[2I\cos^2\theta\cdot\frac{\tan\theta}{2\omega}\right] = \omega + \dot\omega\cdot 2\cdot\frac{1}{2\omega}\cos\theta\sin\theta,$$

$$\Leftrightarrow \quad \dot\theta = \omega + \frac{\dot\omega}{2\omega}\sin 2\theta,$$

One could likewise calculate the change in action:

$$\dot I = -\frac{\partial}{\partial\theta}\left[E + \frac{\partial S}{\partial t}\right] = 0 - \dot\omega\frac{\partial}{\partial\theta}\left[2I\cos^2\theta\frac{\tan\theta}{2\omega}\right],$$

$$\Leftrightarrow \quad \dot I = -I\frac{\dot\omega}{\omega}\cos 2\theta.$$

This also shows that if the change in ω is comparatively much smaller than ω then I is conserved.