Part II: Classical Dynamics - Revision

Lectures by Maciej Dunajski, notes by James Moore

1 Lagrangian mechanics

1.1 Definitions

Definition: Consider N particles with positions $\mathbf{r}_i$. Let x^A be their coordinates in the space $\mathbb{R}^{3N}$. In this context, we refer to $C = \mathbb{R}^{3N}$ as the *configuration space* of the system.

Definition: The *Lagrangian* of a system of particles (with masses m_A) is a function $\mathcal{L} : C \times C \to \mathbb{R}$ on configuration space C of the form

$$\mathcal{L}(x^A, \dot{x}^A) = \underbrace{\frac{1}{2}\sum_A m_A \dot{x}^A \dot{x}^A}_{\text{kinetic}} - \underbrace{V(x^A, \dot{x}^A)}_{\text{potential}}.$$

That is, the Lagrangian is the *kinetic minus the potential energy*.

Definition: Let γ^A be a curve in configuration space, parametrised by time t, with coordinates $(x^A(t), \dot{x}^A(t))$. We define the *action* of the curve as the functional:

$$S[\gamma] = \int_{t_{\text{initial}}}^{t_{\text{final}}} \mathcal{L}(x^A(t), \dot{x}^A(t))\, dt.$$

1.2 The Principle of Least Action

Theorem: The curve taken in configuration space by a system is an extremum of the action functional S.

Proof: Replace $x^A(t)$ by $x^A(t) + \delta x^A(t)$ in the action functional, and assume end-points are fixed, i.e. $\delta x^A(t_{\text{init}}) = 0$ and $\delta x^A(t_{\text{fin}}) = 0$. Then $S \mapsto S + \delta S$ where:

$$\begin{aligned}\delta S &= \int_{t_{\text{init}}}^{t_{\text{fin}}} \left(\frac{\partial \mathcal{L}}{\partial x^A}\delta x^A + \frac{\partial \mathcal{L}}{\partial \dot{x}^A}\delta \dot{x}^A\right) dt \\ &= \int_{t_{\text{init}}}^{t_{\text{fin}}} \left(\delta x^A\left(\frac{\delta \mathcal{L}}{\delta x^A} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^A}\right)\right)\right) dt + \left[\frac{\partial \mathcal{L}}{\partial \dot{x}^A}\delta x^A\right]_{t_{\text{init}}}^{t_{\text{fin}}},\end{aligned}$$

by parts. Assuming the end-points are fixed, boundary term vanishes. Demanding the action is an extremum, we have $\delta S = 0$ for all δx^A. Thus we must have:

$$\frac{\partial \mathcal{L}}{\partial x^A} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^A}\right) = 0.$$

It remains to check the equivalence of these equations to Newton's equations.

We have

$$\frac{\partial \mathcal{L}}{\partial x^A} = -\frac{\partial V}{\partial x^A}, \qquad \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^A}\right) = \frac{d}{dt}(m_A \dot{x}_A),$$

and so the above equation holds if and only if Newton's equations hold. □

Definition: We call the equations

$$\frac{\partial \mathcal{L}}{\partial x^A} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^A}\right) = 0$$

the *Euler-Lagrange equations*, or just *Lagrange's equations*.

Theorem: The Euler-Lagrange equations hold in any coordinate system.

Proof: Let $q_a(x^1, x^2, ..., x^{3N}, t)$ be some coordinate transformation. Assuming these relationships are invertible, we can write $x^A = x^A(q_1, q_2, ..., q_{3N})$. Then by the chain rule:

$$\dot{x}^A = \frac{\partial x^A}{\partial q_a}\dot{q}_a + \frac{\partial x^A}{\partial t}.$$

Now consider the derivatives of $\mathcal{L}$. We have:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial q_a} &= \frac{\partial \mathcal{L}}{\partial x^k}\frac{\partial x^k}{\partial q_a} + \frac{\partial \mathcal{L}}{\partial \dot{x}^k}\frac{\partial \dot{x}^k}{\partial q_a} \\ &= \frac{\partial \mathcal{L}}{\partial x^k}\frac{\partial x^k}{\partial q_a} + \frac{\partial \mathcal{L}}{\partial \dot{x}^k}\underbrace{\left(\frac{\partial^2 x^k}{\partial q_a \partial q_b}\dot{q}^b + \frac{\partial^2 x^k}{\partial q_a \partial t}\right)}_{\text{using } \dot{x}^A \text{ expression}}.\end{aligned}$$

$$\frac{\partial \mathcal{L}}{\partial \dot{q}_a} = \frac{\partial \mathcal{L}}{\partial \dot{x}^k}\underbrace{\frac{\partial \dot{x}^k}{\partial \dot{q}_a}}_{\substack{\text{can remove}\\ \text{dots by chain}\\ \text{rule}}} = \frac{\partial \mathcal{L}}{\partial \dot{x}^k}\frac{\partial x^k}{\partial q_a}.$$

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_a}\right) = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^k}\right)\frac{\partial x^k}{\partial q_a} + \frac{\partial \mathcal{L}}{\partial \dot{x}^k}\underbrace{\left(\frac{\partial^2 x^k}{\partial q_a \partial q_b}\dot{q}_b + \frac{\partial^2 x^k}{\partial q_a \partial t}\right)}_{\text{using } \dot{x}^A \text{ expression}}.$$

Thus, putting everything together, we have:

$$\frac{\partial \mathcal{L}}{\partial q_a} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_a}\right) = \frac{\partial x^k}{\partial q_a}\left(\frac{\partial \mathcal{L}}{\partial x^k} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^k}\right)\right) = 0. \ \square$$

1.3 Constraints & generalised coordinates

Definition: A set of *holonomic constraints* is a set of relations of the form $\{f_\alpha(x^A, t) = 0\}$, $A = 1, 2, ..., 3N$, where $\alpha = 1, 2, ..., 3N - n$ for some $0 \leq n \leq 3N$.

Definition: If a set of holonomic constraints can be solved in principle by parametrising $x^A = x^A(q_1, q_2, ..., q_n)$ (note n degrees of freedom available), then the q_i are called *generalised coordinates*.

We can incorporate holonomic constraints into the Lagrangian formalism in two ways:

1. Lagrange multipliers: Define a new Lagrangian

$$\mathcal{L}'(x^A, \dot{x}^A, \lambda_\alpha, t) = \mathcal{L}(x^A, \dot{x}^A, t) + \sum_{\alpha=1}^{3N-n} \lambda_\alpha f_\alpha(x^A, t),$$

so that the Euler-Lagrange equations for $\mathcal{L}'$ are given by:

$$\frac{\partial \mathcal{L}'}{\partial \lambda_\alpha} = 0 \quad \Rightarrow \quad f_\alpha(x^A, t) = 0,$$

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^A}\right) - \frac{\partial \mathcal{L}}{\partial x^A} = \underbrace{\sum_{\alpha=1}^{3N-n} \lambda_\alpha \frac{\partial f_\alpha}{\partial x^A}}_{\text{constraint forces}}.$$

We can now solve these equations as usual.

Example: Consider the pendulum shown.

The kinetic energy is $\frac{1}{2}m(\dot{x}^2 + \dot{y}^2)$, the potential energy is $-mgy$ and the holonomic constraint is $x^2 + y^2 - l^2 = 0$. Hence by the above method of Lagrange multipliers, we obtain the three equations:

$$\begin{aligned} x^2 + y^2 &= l^2 \\ m\ddot{x} &= \lambda x \\ m\ddot{y} &= mg + \lambda y. \end{aligned}$$

The Lagrange multiplier is proportional to the tension, $\lambda = -T/l$, in the Newtonian formalism, so we indeed see that the Lagrange multipliers correspond to constraint forces.

2. Generalised coordinates: This method is useful when we don't care about the constraint forces' values.

Theorem: For constrained systems, the equations of motion can be derived directly from the Lagrangian

$$L(q_i, \dot{q}_i, t) := \mathcal{L}(x^A(q_i, t), \dot{x}^A(q_i, \dot{q}_i, t)),$$

where the q_i are the generalised coordinates.

Proof: Define $\mathcal{L}' = \mathcal{L} + \lambda_\alpha f_\alpha = L + \lambda_\alpha f_\alpha$ (summation convention applies) and use the Lagrange multiplier method. We can first change coordinates to

$$y^A := \begin{cases} q_i \text{ for } i = 1, ..., n, \quad A = 1, ..., n \\ f_\alpha \text{ for } \alpha = 1, 2, ..., 3N - n, \quad A = n+1, n+2, ..., 3N. \end{cases}$$

By coordinate invariance of the Euler-Lagrange equations, we know that the equations of motion are, for the first n coordinates:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = \sum_\alpha \lambda_\alpha \frac{\partial f_\alpha}{\partial q_i} = 0,$$

since f_α independent of q_i, by definition. So we have equations for the dynamics entirely in q_i, which hold only on the surface of constraints. □

Example: Consider the same pendulum as before. The constraint is solved by introducing the generalised coordinate θ so that $x = l\sin(\theta)$ and $y = l\cos(\theta)$. Then the Lagrangian becomes $\mathcal{L} = \frac{1}{2}m\dot{\theta}^2 l^2 + lmg\cos(\theta)$, which gives the Euler-Lagrange equation:

$$m\ddot{\theta} = -\frac{mg}{l}\sin(\theta).$$

Example: Consider a bead on a wire:

This has Lagrangian

$$\mathcal{L} = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + gy.$$

We can parametrise the wire as $x = q$, $y = f(q)$ say, giving $\dot{y} = \frac{df}{dq}\dot{q}$. Then the Lagrangian can be written as:

$$\mathcal{L} = \frac{1}{2}\left(\dot{q}^2 + \left(\frac{df}{dq}\dot{q}\right)^2\right) + gf(q),$$

and the equation of motion is then easy to derive.

1.4 Noether's theorem and symmetries

Definition: A function $F(q_a, \dot{q}_a, t)$ is called a *first integral* or *constant of the motion* if $\frac{dF}{dt} = 0$ holds whenever the Euler-Lagrange equations hold.

Definition: If $\frac{\partial \mathcal{L}}{\partial q_b} = 0$, then the momentum $p_b = \frac{\partial \mathcal{L}}{\partial \dot{q}_b}$ is clearly a first integral. We call q_b *ignorable* or *cyclic*.

Definition: Suppose that there exists a one-parameter family of transformations $q_a(t) \mapsto Q_a(s,t)$ such that $Q_a(0,t) = q_a(t)$, where $s \in \mathbb{R}$ is a parameter for the family of transformations. We call this transformation a *continuous symmetry* of the Lagrangian $\mathcal{L}$ if

$$\frac{d}{ds}(\mathcal{L}(Q_a(s,t), \dot{Q}_a(s,t), t)) = 0.$$

Theorem (Noether's theorem): If $Q_a(s,t)$ is a continuous symmetry of $\mathcal{L}$, then

$$\sum_a \frac{\partial \mathcal{L}}{\partial \dot{q}_a} \frac{\partial Q_a}{\partial s}\bigg|_{s=0}$$

is a first integral.

Proof: By the definition of a continuous symmetry, we have

$$0 = \frac{d\mathcal{L}}{ds}\bigg|_{s=0} = \frac{\partial \mathcal{L}}{\partial q_a}\frac{\partial Q_a}{\partial s}\bigg|_{s=0} + \frac{\partial \mathcal{L}}{\partial \dot{q}_a}\frac{\partial \dot{Q}_a}{\partial s}\bigg|_{s=0},$$

where we can replace the derivatives $\frac{\partial \mathcal{L}}{\partial Q_a}, \frac{\partial \mathcal{L}}{\partial \dot{Q}_a}$ with $\frac{\partial \mathcal{L}}{\partial q_a}, \frac{\partial \mathcal{L}}{\partial \dot{q}_a}$ because $q_a = Q_a$ at $s = 0$ (also note summation convention applies). Now use the Euler Lagrange equations:

$$0 = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_a}\right)\frac{\partial Q_a}{\partial s}\bigg|_{s=0} + \frac{\partial \mathcal{L}}{\partial \dot{q}_a}\frac{d}{dt}\left(\frac{\partial \dot{Q}_a}{\partial s}\right)\bigg|_{s=0}$$

$$= \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_a}\frac{\partial Q_a}{\partial s}\right)\bigg|_{s=0}. \quad \square$$

1.5 Applications of Noether's theorem

Theorem: Spatially homogeneous systems obey conservation of momentum.

Proof: Consider a spatially homogeneous Lagrangian:

$$\mathcal{L} = \frac{1}{2}\sum m_i |\dot{\mathbf{r}}_i|^2 - \sum_{i,j} V(|\mathbf{r}_i - \mathbf{r}_j|).$$

A continuous symmetry of the system is $\mathbf{R}_i(s,t) = \mathbf{r}_i + s\mathbf{n}$, where $\mathbf{n}$ is an arbitrary constant vector. So by Noether's Theorem,

$$\sum_i \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} \cdot \frac{\partial \mathbf{R}_i}{\partial s}\bigg|_{s=0} = \sum_i \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} \cdot \mathbf{n} = \sum_i \mathbf{p}_i \cdot \mathbf{n}$$

is a conserved quantity. Since $\mathbf{n}$ was arbitrary, total momentum is conserved. $\square$

Theorem: Spatially isotropic systems obey conservation of angular momentum.

Proof: This time, a continuous symmetry is $\mathbf{R}_i(s,t) = A(s)\mathbf{r}_i$ where $A(s)$ is an orthogonal matrix, and $A(0) = I$. Taylor expanding, we find:

$$\mathbf{R}_i = \mathbf{r}_i + s\hat{\mathbf{n}}(s) \times \mathbf{r}_i + O(s^2),$$

where $\hat{\mathbf{n}}(s)$ is the axis of rotation. The proof now proceeds just as before. $\square$

Theorem: Homogeneity in time implies the Hamiltonian

$$H = \sum_n \dot{q}_n \frac{\partial \mathcal{L}}{\partial \dot{q}_n} - \mathcal{L}$$

is a conserved quantity.

Proof: Homogeneity in time requires

$$\frac{\partial \mathcal{L}}{\partial t} = 0.$$

Thus (with the summation convention applying):

$$\frac{dH}{dt} = \ddot{q}_n \frac{\partial \mathcal{L}}{\partial \dot{q}_n} + \dot{q}_n \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_n}\right) - \frac{\partial \mathcal{L}}{\partial t} - \frac{\partial \mathcal{L}}{\partial q_n}\dot{q}_n - \frac{\partial \mathcal{L}}{\partial \dot{q}_n}\ddot{q}_n.$$

Now simply use the Euler-Lagrange equations, and we're done. (Note: this didn't use Noether's theorem.) $\square$

1.6 Applications of Lagrangian mechanics

Example: Geodesics
We call Lagrangians of the form $\mathcal{L} = \frac{1}{2}g_{ab}(\mathbf{q})\dot{q}^a\dot{q}^b$, where g_{ab} is symmetric, *kinetic Lagrangians*. The Euler-Lagrange equations are given by:

$$\frac{1}{2}\dot{q}^a\dot{q}^b\frac{\partial g_{ab}}{\partial q^c} - g_{cb}\ddot{q}^b - \dot{q}^b\dot{q}^a\frac{\partial g_{cb}}{\partial q^a} = 0$$

$$\Rightarrow \quad \frac{1}{2}\dot{q}^a\dot{q}^b\frac{\partial g_{ab}}{\partial q^c} - g_{cb}\ddot{q}^b - \frac{1}{2}\dot{q}^a\dot{q}^b\frac{\partial g_{cb}}{\partial q^a} - \frac{1}{2}\dot{q}^b\dot{q}^a\frac{\partial g_{bc}}{\partial q^a} = 0,$$

splitting the term up using $g_{cb} = g_{bc}$. Now relabel $a \mapsto b$, $b \mapsto a$ in the third term.

We get:

$$\frac{1}{2}\dot{q}^a\dot{q}^b\left(\frac{\partial g_{ab}}{\partial q^c}-\frac{\partial g_{cb}}{\partial q^a}-\frac{\partial g_{ac}}{\partial q^b}\right)-g_{cb}\ddot{q}^b=0.$$

Multiplying through by $-g^{kc}$, we obtain the equation

$$\ddot{q}^k+\dot{q}^a\dot{q}^b\Gamma^k_{ab}=0,$$

where Γ^k_{ab} are the *Christoffel symbols*:

$$\Gamma^a_{bc}=\frac{1}{2}g^{ad}\left(\frac{\partial g_{bd}}{\partial q^c}+\frac{\partial g_{cd}}{\partial q^b}-\frac{\partial g_{bc}}{\partial q^d}\right).$$

Example: Electromagnetism
The Lagrangian for particles in an electromagnetic field is

$$\mathcal{L}=\frac{1}{2}m|\dot{\mathbf{r}}|^2-\phi+\dot{\mathbf{r}}\cdot\mathbf{A},$$

where ϕ is the scalar potential, and $\mathbf{A}$ is the vector potential. This recovers the Lorentz force law via the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{r}}}\right)=\frac{\partial\mathcal{L}}{\partial\mathbf{r}}\quad\Rightarrow\quad\frac{d}{dt}\left(m\dot{\mathbf{r}}+\mathbf{A}\right)=\frac{\partial\mathcal{L}}{\partial\mathbf{r}}.$$

Writing out in suffix notation:

$$m\ddot{r}_a+\left(\frac{\partial A_a}{\partial r_b}-\frac{\partial A_b}{\partial r_a}\right)\dot{r}_b+\frac{\partial A_a}{\partial t}+\frac{\partial\phi}{\partial r_a}=0.$$

Now recall

$$E_a=-\frac{\partial\phi}{\partial r_a}-\frac{\partial A_a}{\partial t},\qquad B_c=\epsilon_{cde}\frac{\partial A_e}{\partial r_d}.$$

This gives the Lorentz force law as expected.

---

2 Stability analysis

2.1 Eigenvalue method

Suppose that the Euler-Lagrange equations for a system reduce to the form

$$\ddot{q}_i=f_i(q_1,q_2,...q_n),$$

for generalised coordinates q_i.

Definition: We say $\mathbf{q}=\mathbf{q}^0$ is an *equilibrium solution* of this system if $f_i(\mathbf{q}^0)=0$ for all i.

Theorem: A small perturbation $\boldsymbol{\eta}(t)$ to an equilibrium solution at $\mathbf{q}^0$ evolves as

$$\ddot{\boldsymbol{\eta}}=\mathbf{F}\boldsymbol{\eta},$$

where $\mathbf{F}$ is the matrix with elements

$$F_{ij}=\left.\frac{\partial f_i}{\partial q_j}\right|_{\mathbf{q}=\mathbf{q}^0}.$$

Proof: Let $\mathbf{q}=\mathbf{q}^0+\boldsymbol{\eta}(t)$. Then the equations $\ddot{\mathbf{q}}=\mathbf{f}(\mathbf{q})$ become:

$$\ddot{\boldsymbol{\eta}}=\mathbf{f}(\boldsymbol{\eta}+\mathbf{q}^0)=\underbrace{f(\mathbf{q}^0)}_{0}+(\boldsymbol{\eta}\cdot\nabla)\mathbf{f}(\mathbf{q}^0)+...$$

hence to first order, $\ddot{\boldsymbol{\eta}}=\mathbf{F}\boldsymbol{\eta}$ for the given matrix $\mathbf{F}$, as required. $\square$

We analyse stability by looking at the eigenvalues of $\mathbf{F}$. First, we should check that the eigenvalues are real (since $\mathbf{F}$ is not necessarily symmetric).

Theorem: For the Lagrangian

$$\mathcal{L}=\frac{1}{2}T_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j-V(\mathbf{q}),$$

with T_{ij} symmetric, positive-definite and invertible at $\mathbf{q}^0$, the eigenvalues of the matrix $\mathbf{F}$ (governing stability at $\mathbf{q}^0$) are real.

Proof: The Euler-Lagrange equations for $\mathcal{L}$ are:

$$\frac{1}{2}\frac{\partial T_{ij}}{\partial q_k}\dot{q}_i\dot{q}_j-\frac{\partial V}{\partial q_k}-\frac{d}{dt}\left(T_{kj}\dot{q}_j\right)=0.$$

Expanding with $\mathbf{q}=\mathbf{q}^0+\boldsymbol{\eta}(t)$, we have to first order in $\boldsymbol{\eta}$,

$$-\left.\frac{\partial^2V}{\partial q_i\partial q_j}\right|_{\mathbf{q}=\mathbf{q}^0}\eta_j=T_{ij}(\mathbf{q}^0)\ddot{\eta}_j=T_{ij}F_{jk}\eta_k,$$

where in the last equality we used the equation of motion for η_k from the above Theorem. Redefining:

$$T_{ij}=T_{ij}(\mathbf{q}^0),\qquad V_{ij}=\left.\frac{\partial^2V}{\partial q_i\partial q_j}\right|_{\mathbf{q}=\mathbf{q}^0},$$

we can rewrite our equation as the matrix equation:

$$-\mathbf{V}=\mathbf{TF}\qquad\Rightarrow\qquad-\mathbf{T}^{-1}\mathbf{V}=\mathbf{F},$$

since $\mathbf{T}$ is invertible at $\mathbf{q}^0$. Now suppose that $\mathbf{F}\boldsymbol{\mu}=\lambda^2\boldsymbol{\mu}$. This occurs if and only if $\mathbf{V}\boldsymbol{\mu}=-\lambda^2\mathbf{T}\boldsymbol{\mu}$. Take the inner product with $\boldsymbol{\mu}$ to obtain:

$$\overline{\boldsymbol{\mu}}^T\mathbf{V}\boldsymbol{\mu}=-\lambda^2\overline{\boldsymbol{\mu}}^T\mathbf{T}\boldsymbol{\mu}.$$

Now since $\mathbf{V}$ and $\mathbf{T}$ are symmetric, both $\overline{\boldsymbol{\mu}}^T\mathbf{V}\boldsymbol{\mu}$ and $\overline{\boldsymbol{\mu}}^T\mathbf{T}\boldsymbol{\mu}$ are real. Since $\mathbf{T}$ is positive definite, $\overline{\boldsymbol{\mu}}^T\mathbf{T}\boldsymbol{\mu}\neq 0$, and hence it follows that $-\lambda^2$ must be real (if $\overline{\boldsymbol{\mu}}^T\mathbf{T}\boldsymbol{\mu}$ was zero, $-\lambda^2$ could be anything we liked!). $\square$

This Theorem allows us to conduct the following analysis. Suppose $\mathbf{F}\boldsymbol{\mu}_a=\lambda_a^2\boldsymbol{\mu}_a$, $a=1,2...m$ (where $m\leq n$).

Definition: The vectors $\boldsymbol{\mu}_a$ satisfying $\mathbf{F}\boldsymbol{\mu}_a=\lambda_a^2\boldsymbol{\mu}_a$ are called the *normal modes*.

The normal modes are important, since the most general solution of $\ddot{\boldsymbol{\eta}} = \mathbf{F}\boldsymbol{\eta}$ can be written as:

$$\boldsymbol{\eta}(t) = \sum_a \boldsymbol{\mu}_a \left(A_a e^{\lambda_a t} + B_a e^{-\lambda_a t}\right).$$

The A_a and B_a are unimportant constants of integration. The λ_a^2 determine the behaviour of the perturbation. There are two cases:

Case 1 - $\lambda_a^2 < 0$ for some a: In this case, $\lambda_a = \pm\omega_a i$, for ω_a real. Hence the perturbation is linearly stable in the direction $\boldsymbol{\mu}_a$.

Case 2 - $\lambda_a^2 > 0$ for some a: In this case, $\lambda_a = \pm\omega_a$, for ω_a real. Hence the perturbation is linearly unstable in the direction $\boldsymbol{\mu}_a$.

Definition: An equilibrium point $\mathbf{q}^0$ is called *linearly stable* if $\lambda_a^2 < 0$ for all a; otherwise, it is called *linearly unstable*.

2.2 Example: linear triatomic molecules

Consider a triatomic molecule with atoms of masses m, M and m respectively (see diagram).

The Lagrangian is

$$\mathcal{L} = \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}M\dot{x}_2^2 + \frac{1}{2}m\dot{x}_3^2 - V(x_1 - x_2) - V(x_2 - x_3),$$

where the V are chemical potentials (generally complicated). Let $x_i = x_i^0$ be an equilibrium. By symmetry, we expect $|x_1^0 - x_2^0| = |x_2^0 - x_3^0| = r_0$ at equilibrium. Consider small perturbations $x_i(t) = x_i^0 + \eta_i(t)$. Then:

$$V(r) = V(r_0) + \underbrace{\left.\frac{\partial V}{\partial r}\right|_{r=r_0}}_{0}(r - r_0) + \underbrace{\left.\frac{\partial^2 V}{\partial r^2}\right|_{r=r_0}}_{=:\,k}\frac{(r-r_0)^2}{2} + \ldots$$

Close the equilibrium point, the Lagrangian then becomes:

$$\mathcal{L} \approx \frac{1}{2}m\dot{\eta}_1^2 + \frac{1}{2}M\dot{\eta}_2^2 + \frac{1}{2}m\dot{\eta}_3^2 - \frac{1}{2}k\left((\eta_1 - \eta_2)^2 + (\eta_2 - \eta_3)^2\right),$$

which gives an equation of the form $\ddot{\boldsymbol{\eta}} = \mathbf{F}\boldsymbol{\eta}$ when we find the Euler-Lagrange equations. We find that there are three normal modes, one corresponding to translation, and the other to vibrations.

3 Rigid body dynamics

3.1 Kinematics

Definition: A *rigid body* is a collection of N points constrained such that the distances between the points are fixed, $|\mathbf{r}_i - \mathbf{r}_j| = \text{constant},\ i, j = 1, 2, ..., N.$

Continuous configurations are also possible, where we have infinitely many points; throughout, to change from a discrete to a continuous distribution we replace masses by *mass density* and sums by integrals.

Definition: The position of a rigid body fixed at a point P (i.e. non-translating, just rotating) can be specified by its position relative to either a fixed orthonormal frame $\{\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \tilde{\mathbf{e}}_3\}$, called the *space frame*, or a moving orthonormal frame $\{\mathbf{e}_1(t), \mathbf{e}_2(t), \mathbf{e}_3(t)\}$ that moves with the body, called the *body frame*.

In particular, we have $\mathbf{e}_a(t) \cdot \mathbf{e}_b(t) = \delta_{ab} = \tilde{\mathbf{e}}_a \cdot \tilde{\mathbf{e}}_b$. We also choose the vectors to be *right-handed*, in the sense that $\mathbf{e}_a(t) \times \mathbf{e}_b(t) = \epsilon_{abc}\mathbf{e}_c(t)$ (and similarly for the space frame).

The first question we should ask is how to go between the two frames.

Theorem: The 3×3 matrix $R(t)$ defined by $R_{ab} = \mathbf{e}_a(t)\cdot\tilde{\mathbf{e}}_b$ is the *unique* orthogonal transformation such that $\mathbf{e}_a(t) = R_{ab}(t)\tilde{\mathbf{e}}_b$.

Proof: Uniqueness is immediate, since

$$\mathbf{e}_a(t) = R_{ab}(t)\tilde{\mathbf{e}}_b \qquad \Rightarrow \qquad \mathbf{e}_a(t)\cdot\tilde{\mathbf{e}}_c = R_{ab}(t)\tilde{\mathbf{e}}_b\cdot\tilde{\mathbf{e}}_c = R_{ac},$$

by orthonormality of the frames. Orthogonality follows from:

$$[R^T R]_{ab} = R_{ac}R_{bc} = R_{ac}R_{bd}\delta_{cd} = R_{ac}R_{bd}\tilde{\mathbf{e}}_c\cdot\tilde{\mathbf{e}}_d = \mathbf{e}_a\cdot\mathbf{e}_b = \delta_{ab}. \ \square$$

Using this Theorem, we can count the number of degrees of freedom of the system. Notice that R_{ab} has 9 elements, but the equation $R^T R = I$ for orthogonality gives 6 constraints, hence there are 3 degrees of freedom. We will characterise these by *Euler angles* later on.

3.2 Angular velocity

Definition: We notice that

$$\frac{d\mathbf{e}_a(t)}{dt} = \frac{d}{dt}(R_{ab}\tilde{\mathbf{e}}_b) = \dot{R}_{ab}\tilde{\mathbf{e}}_b = \dot{R}_{ab}R^{-1}_{bc}\mathbf{e}_c(t).$$

Define $\omega_{ac}(t) = \dot{R}_{ab}R^{-1}_{bc} = \dot{R}_{ab}R_{cb}$ (using $\mathbf{R}^{-1} = \mathbf{R}^T$). In dyadic notation, $\boldsymbol{\omega} = \dot{\mathbf{R}}\mathbf{R}^T$.

The matrix $\omega_{ac}(t)$ has some useful properties:

Theorem: We have

(i) $\boldsymbol{\omega}^T = -\boldsymbol{\omega}$;

(ii) for any point $\mathbf{r}(t) = r_a\mathbf{e}_a(t)$ on a rotating rigid body, we have
$$\frac{d\mathbf{r}(t)}{dt} = r_a\omega_{ac}\mathbf{e}_c(t).$$

Proof: (i) follows immediately by differentiating the relationship $\mathbf{R}\mathbf{R}^T = I$, which gives:

$$0 = \dot{\mathbf{R}}\mathbf{R}^T + \mathbf{R}\dot{\mathbf{R}}^T = \boldsymbol{\omega} + (\dot{\mathbf{R}}\mathbf{R}^T)^T = \boldsymbol{\omega} + \boldsymbol{\omega}^T.$$

For (ii), we have:

$$\frac{d\mathbf{r}(t)}{dt} = r_a\frac{d\mathbf{e}_a(t)}{dt} = r_a\omega_{ac}\mathbf{e}_c(t). \quad \square$$

From the matrix $\boldsymbol{\omega}$, we can construct the *angular velocity* of a rotating rigid body.

Definition: The *angular velocity components* of a rotating rigid body are defined by

$$\omega_a = \frac{1}{2}\epsilon_{abc}\omega_{bc}.$$

Conversely, $\omega_{ab} = \epsilon_{abc}\omega_c$. Written out in full, we have:

$$\omega_{ab} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}.$$

The *angular velocity vector* is defined by $\boldsymbol{\omega} = \omega_a\mathbf{e}_a(t)$, i.e. the components ω_a are components with respect to the body frame.

Theorem: We have, for any vector $\mathbf{r}(t)$ in the body frame:

$$\frac{d\mathbf{r}(t)}{dt} = \boldsymbol{\omega}\times\mathbf{r}(t).$$

Proof: We have

$$\frac{d\mathbf{e}_a}{dt} = \omega_{ac}\mathbf{e}_c(t) = -\epsilon_{abc}\omega_b\mathbf{e}_c(t) = -\omega_b\mathbf{e}_a(t)\times\mathbf{e}_b(t) = \boldsymbol{\omega}\times\mathbf{e}_a(t),$$

using right-handedness of the frame. Now use $\mathbf{r}(t) = r_a\mathbf{e}_a(t)$ and linearity. $\square$

3.3 The inertia tensor

Definition: The *inertia tensor* is defined by

$$I_{ab} = \sum_i m_i((\mathbf{r}_i\cdot\mathbf{r}_i)\delta_{ab} - (\mathbf{r}_i)_a(\mathbf{r}_i)_b).$$

Notice that this tensor is manifestly symmetric.

Theorem: The kinetic energy of a rotating body is given by $\frac{1}{2}\boldsymbol{\omega}^T\mathbf{I}\,\boldsymbol{\omega}$, where $\mathbf{I}$ is the inertia tensor, and $\boldsymbol{\omega}$ is the angular momentum vector. In components, this is $\frac{1}{2}\omega_aI_{ab}\omega_b$.

Proof: The kinetic energy of the body is given by:

$$\begin{aligned} T &= \frac{1}{2}\sum_i m_i|\dot{\mathbf{r}}_i|^2 \\ &= \frac{1}{2}\sum_i m_i|\boldsymbol{\omega}\times\mathbf{r}_i|^2 && \text{(by above)} \\ &= \frac{1}{2}\sum_i m_i(\epsilon_{abc}\omega_b(\mathbf{r}_i)_c\cdot\epsilon_{ade}\omega_d(\mathbf{r}_i)_e) \\ &= \frac{1}{2}\sum_i m_i\omega_a\omega_b((\mathbf{r}_i\cdot\mathbf{r}_i)\delta_{ab} - (\mathbf{r}_i)_a(\mathbf{r}_i)_b). && \square \end{aligned}$$

In practice, we usually work with the continuous analogue of the inertia tensor:

$$\mathbf{I} = \iiint \rho(\mathbf{r})\begin{pmatrix} y^2+z^2 & -xy & -xz \\ -xy & x^2+z^2 & -yz \\ -xz & -yz & x^2+y^2 \end{pmatrix} d^3\mathbf{r},$$

which can easily be obtained from the discrete formula. Since $\mathbf{I}$ is symmetric and real-valued, it can be diagonalised by an orthogonal transformation and has real eigenvalues.

Definition: The axes with respect to which $\mathbf{I}$ is diagonal are called the *principal axes*. The diagonal elements of $\mathbf{I}$, when expressed in the principal axis basis, are called the *principal moments of inertia*.

Example: We compute the inertia tensor of a disc of mass M, radius r, fixed at its centre.

3.4 The parallel axis theorem

Theorem: Let I_{ab} be the inertia tensor of a body fixed about its centre of mass. Then if P' is displaced $\mathbf{c}$ from the body's centre of mass, the inertia tensor of the body about P' is given by:

$$(I_{\mathbf{c}})_{ab} = I_{ab} + M(|\mathbf{c}|^2\delta_{ab} - c_a c_b).$$

Proof: We have

$$\begin{aligned}(I_{\mathbf{c}})_{ab} &= \sum_i m_i \left(|\mathbf{r}_i - \mathbf{c}|^2\delta_{ab} - (\mathbf{r}_i - \mathbf{c})_a(\mathbf{r}_i - \mathbf{c})_{\mathbf{b}}\right) \\ &= \sum_i m_i(|\mathbf{r}_i|^2\delta_{ab} - (\mathbf{r}_i)_a(\mathbf{r}_i)_b) + \sum_i m_i(|\mathbf{c}|^2\delta_{ab} - c_a c_b) \\ &\quad + \underbrace{\sum_i m_i\mathbf{r}_i \cdot (\text{stuff only dependent on } \mathbf{c})}_{\text{vanishes in COM frame}}. \quad \square\end{aligned}$$

3.5 Euler's equations

Theorem: The angular momentum of a rotating rigid body is given by $\mathbf{L} = \mathbf{I}\,\boldsymbol{\omega}$.

Proof: We have:

$$\mathbf{L} = \sum_{n=1}^{N} m_i\mathbf{r}_i\times\dot{\mathbf{r}}_i = \sum_{i=1}^{N} m_i\mathbf{r}_i\times(\boldsymbol{\omega}\times\mathbf{r}_i) = \sum_{i=1}^{N} m_i(|\mathbf{r}_i|^2\boldsymbol{\omega}-(\mathbf{r}_i\cdot\boldsymbol{\omega})\mathbf{r}_i),$$

which gives $\mathbf{L} = \mathbf{I}\,\boldsymbol{\omega}$. $\square$

Theorem: The motion of a free rigid body under the action of no external forces, with no translation, is determined by *Euler's equations*:

$$\begin{aligned}I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) &= 0 \\ I_2\dot{\omega}_2 + \omega_3\omega_1(I_1 - I_3) &= 0 \\ I_3\dot{\omega}_3 + \omega_1\omega_2(I_2 - I_1) &= 0,\end{aligned}$$

where I_i are the principal moments of inertia of the body, and ω_i are the components of the angular momentum in the principal axis basis.

Proof: For a free rigid body, we have conservation of total angular momentum by Noether's Theorem. Hence writing $\mathbf{L} = L_a(t)\mathbf{e}_a(t)$.

$$\begin{aligned}\mathbf{0} = \frac{d\mathbf{L}}{dt} &= \frac{dL_a(t)}{dt}\mathbf{e}_a(t) + L_a(t)\frac{d\mathbf{e}_a(t)}{dt} \\ &= \dot{L}_a\mathbf{e}_a(t) + L_a(t)\boldsymbol{\omega}\times\mathbf{e}_a(t) \\ &= \dot{L}_a\mathbf{e}_a(t) + L_a(t)(\omega_b\epsilon_{bac}\mathbf{e}_c(t)).\end{aligned}$$

Thus we have $\dot{L}_c + \epsilon_{cba}\omega_b L_a = 0$. In the principal axes, $L_1 = I_1\omega_1$, $L_2 = I_2\omega_2$ and $L_3 = I_3\omega_3$. Substituting this in, we have Euler's equations. $\square$

3.6 First integrals

There are two useful first integrals of Euler's equations:

Theorem: The kinetic energy $\frac{1}{2}(I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2)$, and the square norm of the angular momentum, $|\mathbf{L}|^2 = I_1^2\omega_1^2 + I_2^2\omega_2^2 + I_3^2\omega_3^2$, are both constants of the motion.

Proof: Multiply by $I_1\dot{\omega}_1 + \omega_2\omega_3(I_3 - I_2) = 0$ by ω_1, and similarly for others, and then add all of Euler's equations together. Then integrate directly. Same for angular momentum but multiply through by $I_i\omega_i$. $\square$

3.7 Light spinning tops

We now see some example applications of Euler's equations to spinning tops.

Example 1: Sphere
For a sphere, $I_1 = I_2 = I_3 = I$. Hence Euler's equations imply $\dot{\omega}_a = 0$ for all a. Thus the angular velocity is constant.

Example 2: Symmetric top
For a symmetric top, we have $I_1 = I_2$, $I_1 \neq I_3$ and $I_2 \neq I_3$. Since $I_1 = I_2$, by Euler's equations we have $\dot{\omega}_3 = 0$. The remaining equations reduce to:

$$\dot{\omega}_1 = \left(\frac{I_1 - I_3}{I_1}\right)\omega_2\omega_3, \quad \dot{\omega}_2 = -\left(\frac{I_1 - I_3}{I_1}\right)\omega_1\omega_3.$$

Set $\Omega = \omega_3(I_1 - I_3)/I_1$, which is a constant. Then we can write the equations as $\dot{\omega}_1 = \Omega\omega_2$, $\dot{\omega}_2 = -\Omega\omega_3$. Decoupling, we have $\ddot{\omega}_1 = \Omega\dot{\omega}_2 = -\Omega^2\omega_1$.

In general, we have:

$$(\omega_1, \omega_2, \omega_3) = (A\sin(\Omega t), B\cos(\Omega t), \omega_3(0)).$$

Hence there is spin precession around the $\hat{\mathbf{z}}$ axis.

Example 3: Asymmetric top
For an asymmetric top, $I_1 \neq I_2$, $I_2 \neq I_3$, $I_1 \neq I_3$. Consider the special case $\omega_1 = \Omega$, $\omega_2 = \omega_3 = 0$; this is an equilibrium position. We ask if it is stable.

Let $\omega_1(t) = \Omega + \eta_1(t)$, $\omega_2(t) = \eta_2(t)$ and $\omega_3(t) = \eta_3(t)$. Euler's equations become:

$$\begin{aligned}I_1\dot{\eta}_1 &= O(\eta^2) \\ I_2\dot{\eta}_2 &= \Omega\eta_3(t)(I_3 - I_1) + O(\eta^2) \\ I_3\dot{\eta}_3 &= \Omega\eta_2(t)(I_1 - I_2) + O(\eta^2).\end{aligned}$$

Differentiating the second equation and inserting into the third, we have:

$$I_2\ddot{\eta}_2 = \Omega^2\frac{(I_1 - I_2)(I_3 - I_1)}{I_3}\eta_2 + O(\eta^2).$$

Hence there is an instability if $I_2 < I_1 < I_3$ or $I_3 < I_1 < I_2$.

3.8 Euler angles

We now try and deal with heavy tops, rather than just free tops. This requires *even further machinery*.

Theorem (Euler's Theorem): A general rotation in $\mathbb{R}^3$ may be expressed as a product of three successive rotations about 3 (in general) different axes.

Proof: We want to find the matrix $R_{ab}(t)$ such that $\mathbf{e}_a(t) = R_{ab}(t)\tilde{\mathbf{e}}_b$, the matrix we showed existed and was unique earlier on. We proceed in three steps:

Step 1: Rotation of ϕ about the $\tilde{\mathbf{e}}_3$ axis:

Step 2: Rotation of θ about the $\tilde{\mathbf{e}}_1'$ axis:

Step 3: Rotation of ψ about the $\tilde{\mathbf{e}}_3''$ axis:

Composing all the matrices of these rotations gives an overall matrix:

$$\mathbf{R}(\phi,\theta,\psi) = \mathbf{R}(\tilde{\mathbf{e}}_3'',\psi)\mathbf{R}(\tilde{\mathbf{e}}_1',\theta)\mathbf{R}(\tilde{\mathbf{e}}_3,\phi),$$

and so indeed we can represent the rotation matrix as a product of three successive rotations, with varying angles ϕ, θ and ψ. $\square$

Definition: We call the angles ϕ, θ and ψ the *Euler angles*.

3.9 Angular velocity revisited

Theorem: The angular momentum in the body frame may be expressed in terms of the Euler angles as:

$$\boldsymbol{\omega} = (\dot{\phi}\sin(\theta)\sin(\psi) + \dot{\theta}\cos(\theta))\mathbf{e}_1(t)$$
$$+(\dot{\phi}\sin(\theta)\cos(\psi) - \dot{\theta}\sin(\psi))\mathbf{e}_2(t) + (\dot{\psi} + \dot{\phi}\cos(\theta))\mathbf{e}_3(t).$$

Proof: We prove the result through two steps:

Step 1: We show that if the rotation matrix is $\mathbf{R}(\hat{\mathbf{n}},\phi)$, then angular velocity is $\boldsymbol{\omega} = \dot{\phi}\hat{\mathbf{n}}$. WLOG, take $\hat{\mathbf{n}} = (1,0,0)$, and so

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix}.$$

Compute $\dot{\mathbf{R}}\mathbf{R}^T$ to get:

$$\dot{\mathbf{R}}\mathbf{R}^T = \dot{\phi}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Compare to expected form of ω_{ab}, and then use

$$\omega_a = \frac{1}{2}\epsilon_{abc}\omega_{bc} = \dot{\phi}n_a.$$

Step 2: Show $\boldsymbol{\omega}$ is additive under composition of rotations. We have:

$$\frac{d}{dt}(R_1R_2)(R_1R_2)^T = (\dot{R}_1R_2 + R_1\dot{R}_2)R_2^TR_1^T = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2.$$

where ω_1 and ω_2 are the angular velocity matrices. Hence the angular velocity vectors are also additive.

Step 3: From Steps 1 and 2, we can immediately conclude that the rotation $\mathbf{R}(\tilde{\mathbf{e}}_3'',\psi)\mathbf{R}(\tilde{\mathbf{e}}_1',\theta)\mathbf{R}(\tilde{\mathbf{e}}_3,\phi)$ gives angular velocity vector

$$\boldsymbol{\omega} = \dot{\psi}\tilde{\mathbf{e}}_3'' + \dot{\theta}\tilde{\mathbf{e}}_1' + \dot{\phi}\tilde{\mathbf{e}}_3.$$

Now we carefully construct $\tilde{\mathbf{e}}_3''$, $\tilde{\mathbf{e}}_1'$ and $\tilde{\mathbf{e}}_3$ in the body frame using the Euler angle diagrams. We have:

$$\tilde{\mathbf{e}}_3'' = \mathbf{e}_3(t), \quad \tilde{\mathbf{e}}_1' = \cos(\psi)\mathbf{e}_1(t) - \sin(\psi)\mathbf{e}_2(t),$$

$$\tilde{\mathbf{e}}_3 = \cos(\theta)\tilde{\mathbf{e}}_3'' + \sin(\theta)\tilde{\mathbf{e}}_2''$$
$$= \cos(\theta)\mathbf{e}_2(t) + \sin(\theta)(\cos(\psi)\mathbf{e}_2(t) + \sin(\psi)\mathbf{e}_1(t)).$$

Combining all this gives the result. $\square$

3.10 The heavy symmetric top

We finally have enough machinery to deal with the heavy symmetric top. Consider the heavy symmetric top as shown:

We assume that the top has principal moments of inertia $I_1 = I_2$ and $I_1 \neq I_3$, and the top has mass M.

The Lagrangian is:

$$\mathcal{L} = \frac{1}{2}I_1(\omega_1^2 + \omega_2^2) + \frac{1}{2}I_3\omega_3^2 - Mgl\cos(\theta).$$

Using our Euler-angle expressions for ω_1 and ω_2, we get:

$$\omega_1^2 + \omega_2^2 = \dot{\phi}^2\sin^2(\theta) + \dot{\theta}^2.$$

Substituting this, and the expression for ω_3 in terms of Euler angles, into the the Lagrangian, we have:

$$\mathcal{L} = \frac{1}{2}I_1(\dot{\theta}^2 + \sin^2(\theta)\dot{\phi}^2) + \frac{1}{2}I_3(\dot{\psi}^2 + \dot{\phi}^2\cos^2(\theta) + 2\dot{\psi}\dot{\phi}\cos(\theta)) \\ -Mgl\cos(\theta).$$

Idea: Our aim is to isolate the dynamics in terms of θ using first integrals.

Theorem: The dynamics can be reduced to 1D motion of the form:

$$I_1\ddot{\theta} = -\frac{dV_{\text{eff}}}{d\theta}$$

where

$$V_{\text{eff}}(\theta) = \frac{1}{2}I_1\sin^2(\theta)\left(\frac{b - a\cos(\theta)}{\sin^2(\theta)}\right)^2 + Mgl\cos(\theta)$$

for constants a and b.

Proof: There are three first integrals: the momenta p_ϕ, p_ψ and the total energy E. Constructing the first integrals p_ϕ and p_ψ, we have:

$$\begin{aligned} I_1 b = p_\phi &= I_1\sin^2(\theta)\dot{\phi} + I_3\cos^2(\theta)\dot{\phi} + I_3\dot{\psi}\cos(\theta), \\ I_1 a = p_\psi &= I_3\dot{\psi} + I_3\dot{\phi}\cos(\theta), \end{aligned}$$

for constants a and b. Solving for $\dot{\phi}$ and $\dot{\psi}$ in terms of θ, we obtain:

$$\dot{\phi} = \frac{b - a\cos(\theta)}{\sin^2(\theta)}, \quad \dot{\psi} = \frac{I_1}{I_3}a - \left(\frac{b\cos(\theta) - a\cos^2(\theta)}{\sin^2(\theta)}\right).$$

Trick: we notice that $\omega_3 = \dot{\phi}\cos(\theta) + \dot{\psi} = I_1 a/I_3$, so is a constant (called the *spin*). Now use the total energy as the final first integral:

$$\begin{aligned} E &= \dot{\theta}\frac{\partial\mathcal{L}}{\partial\dot{\theta}} + \dot{\phi}\frac{\partial\mathcal{L}}{\partial\dot{\phi}} + \dot{\psi}\frac{\partial\mathcal{L}}{\partial\dot{\psi}} - \mathcal{L} \\ &= \frac{1}{2}I_1(\dot{\theta}^2 + \sin^2(\theta)\dot{\phi}^2) + \frac{1}{2}I_3\omega_3^2 + Mgl\cos(\theta). \end{aligned}$$

Define $E' = E - \frac{1}{2}I_3\omega_3^2$, which is also necessarily constant, since the spin ω_3 is constant. It then follows that:

$$E' = \frac{1}{2}I_1\dot{\theta}^2 + V_{\text{eff}}(\theta),$$

where V_{eff} is of the required form. Differentiate, and we're done. □

Theorem: Defining $u = \cos(\theta)$, the equations of motion for the top can be written in the form

$$\begin{aligned} \dot{u}^2 &= f(u), \\ \dot{\phi} &= \frac{b - au}{1 - u^2}, \\ \dot{\psi} &= \frac{I_1 a}{I_3} - \frac{u(b - au)}{1 - u^2}, \end{aligned}$$

where $f(u)$ is a cubic polynomial.

Proof: Use the energy equation in the form

$$E' = \frac{1}{2}I_1\dot{\theta}^2 + V_{\text{eff}}(\theta).$$

Redefining $u = \cos(\theta)$, we have

$$\dot{u} = -\dot{\theta}\sin(\theta) \quad \Rightarrow \quad \dot{\theta}^2 = \frac{\dot{u}^2}{1 - u^2}.$$

Also redefining $\alpha = 2E'/I_1$, $\beta = 2Mgl/I_1$, we find that we can put the equations in the required form, with

$$f(u) = (1 - u^2)(\alpha - \beta u) - (b - au)^2. \ \square$$

To plot the cubic described above, we note that $f(u) \to \infty$ as $u \to \infty$, and $f(u) \to -\infty$ as $u \to -\infty$. We also note that $f(\pm 1) < 0$. Physically, we require $-1 \leq u \leq 1$ and $\dot{u}^2 = f(u) > 0$. Hence $f(u)$ looks generally like:

3.11 Types of heavy top motion

The types of motion the top undergoes depend on the sign of $\dot{\phi}$ at the roots of $f(u)$ in $-1 \leq u \leq 1$, say u_1 and u_2, corresponding to maximum and minimum θ, say θ_1 and θ_2. There are three cases:

- $\dot{\phi} > 0$ at u_1 and u_2;
- $\dot{\phi} > 0$ at u_1 and $\dot{\phi} < 0$ at u_2;
- $\dot{\phi} > 0$ at u_1 and $\dot{\phi} = 0$ at u_2.

Note we don't include $\dot{\phi} < 0$ at u_1, u_2 etc, as these are the same motions but in different directions. Sketching gives the diagrams:

We have names for the different types of motion:

Definition: Motion in the ϕ direction is called *precession*. Motion in the θ direction is called *nutation*.

We now examine *uniform precession*, i.e. for all time we have $\dot{\phi} = \text{constant}$ and $\dot{\theta} = 0$.

Theorem: The requirement for uniform precession is

$$\omega_3^2 \geq \frac{4MglI_1\cos(\theta_0)}{I_3^2}.$$

Proof: For uniform precession to occur, $f(u)$ must have a double root u_0. So $f(u_0) = f'(u_0) = 0$. From $f(u_0) = 0$, we have

$$0 = (1-u_0^2)(\alpha-\beta u_0)-(b-au_0)^2 = (1-u_0^2)(\alpha-\beta u_0)-(1-u_0^2)^2\dot{\phi}^2$$

Hence

$$\dot{\phi} = \frac{\alpha - \beta u_0}{1-u_0^2}.$$

Substituting into the condition $f'(u_0) = 0$ gives a quadratic in $\dot{\phi}$, and the condition on its discriminant gives the uniform precession condition. □

Finally we look at the *sleeping top*, i.e. we initially have $\dot{\theta} = \theta = 0$.

Theorem: The sleeping top is stable if

$$\omega_3^2 > \frac{4I_1Mgl}{I_3^2},$$

and unstable otherwise.

Proof: We know that $f(u)$ needs a root at $\theta = 0$, i.e. $u = +1$. This implies $a = b$. From the definition of α and β, we have $\alpha = \beta$ in this case. In particular, these properties show that $f(u)$ has a double root at $+1$. Thus there are two cases:

Let the other root be $u_2 = a^2/\alpha - 1$. The graphs show that if $u_2 > 1$, the motion is stable, since we are forced to stay in $-1 \leq u \leq 1$ and $f(u) > 0$; we can also see that if $u_2 < 1$, the motion is unstable. □

---

4 Hamiltonian mechanics

4.1 Phase space

Definition: Let $\mathcal{L}(q_i, \dot{q}_i, t)$ be a Lagrangian for N particles. We call

$$\mathbf{p}_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$$

the *conjugate momentum* associated to q_i. We call the $6N$-dimensional space $\mathbb{R}^{6N}$ with coordinates (q_i, p_i) *phase space*.

4.2 The Legendre transform

Definition: Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f = f(x, y)$. Define

$$u = \frac{\partial f}{\partial x},$$

and use u and y as coordinates. Define

$$g(u, y) = ux(u, y) - f(x(u, y), y),$$

where to obtain $x(u, y)$ we solved the relation $u = \frac{\partial f}{\partial x}$. We call g the *Legendre transform* of f.

Theorem: The Legendre transform is involutive, i.e. the transform of the transform gives back the original function.

Proof: We take the Legendre transform of the above g. We have:

$$x = \frac{\partial g}{\partial u},$$

The Legendre transform is then:

$$\begin{aligned} h(x,y) &= xu(x,y) - g(u(x,y),y) \\ &= xu(x,y) - ux(u,y) + f(x(u,y),y) = f(x,y). \quad \square \end{aligned}$$

4.3 Hamilton's equations

Definition: The *Hamiltonian* is the Legendre transform of the Lagrangian, given by

$$H(\mathbf{q},\mathbf{p},t) = \sum_i p_i \dot{q}_i - \mathcal{L}(q_i,\dot{q}_i,t),$$

where $\dot{q}_i = \dot{q}_i(\mathbf{q},\mathbf{p},t)$ are implicitly given by

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}.$$

Theorem: *Hamilton's equations* hold:

$$\dot{q}_i = \frac{\partial H}{\partial \dot{p}_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial \dot{q}_i}, \qquad \frac{\partial H}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}.$$

Proof: We have

$$\begin{aligned} dH &= p_i d\dot{q}_i + \dot{q}_i dp_i - \frac{\partial \mathcal{L}}{\partial q_i} dq_i - \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \cancel{\frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i} + \dot{q}_i dp_i - \underbrace{\frac{\partial \mathcal{L}}{\partial q_i}}_{\text{use Euler-Lagrange}} dq_i - \cancel{\frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i} - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \dot{q}_i dp_i - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) dq_i - \frac{\partial \mathcal{L}}{\partial t} dt \\ &= \dot{q}_i dp_i - \dot{p}_i dq_i - \frac{\partial \mathcal{L}}{\partial t} dt. \end{aligned}$$

Compare with

$$dH = \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt. \quad \square$$

Example: Electromagnetism
Consider the Lagrangian for a particle in a magnetic field:

$$\mathcal{L} = \frac{1}{2} m |\dot{\mathbf{r}}|^2 - e(\phi(\mathbf{r}) - \dot{\mathbf{r}} \cdot \mathbf{A}).$$

We construct the Hamiltonian. The conjugate momentum is:

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}} = m\dot{\mathbf{r}} + e\mathbf{A}.$$

Thus the Hamiltonian is

$$\begin{aligned} H &= \mathbf{p} \cdot \dot{\mathbf{r}} - \mathcal{L} \\ &= \mathbf{p} \cdot \left(\frac{\mathbf{p} - e\mathbf{A}}{m}\right) - \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\left(\phi - \frac{(\mathbf{p} - e\mathbf{A}) \cdot \mathbf{A}}{m}\right) \\ &= \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + e\phi. \end{aligned}$$

Hamilton's equations recover the Lorentz force law.

4.4 The principle of least action

The principle of least action can be recovered from the Hamiltonian formalism.

Theorem: Particles take the path in phase space that extremises the action

$$S[\mathbf{q},\mathbf{p}] = \int_{t_1}^{t_2} (p_i \dot{q}_i - H(\mathbf{q},\mathbf{p})\, dt.$$

Note: The integrand is just the Lagrangian, i.e. the Legendre transform of the Hamiltonian.

Proof: We vary p_i and q_i separately to get:

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \left(\dot{q}_i \delta p_i + p_i \delta \dot{q}_i - \frac{\partial H}{\partial p_i} \delta p_i - \frac{\partial H}{\partial q_i} \delta q_i\right) dt \\ &= \int_{t_1}^{t_2} \left(\left(\dot{q}_i - \frac{\partial H}{\partial p_i}\right) \delta p_i + \left(-\dot{p}_i - \frac{\partial H}{\partial q_i}\right) \delta q_i\right) dt + \left[p_i \delta q_i\right]_{t_1}^{t_2}. \end{aligned}$$

Assuming the variation has fixed end-points, the boundary terms vanish. Thus Hamilton's equations are recovered, which are equivalent to Newton's equations. $\square$

4.5 Liouville's theorem

Theorem: Consider a point in phase space $(\mathbf{q}_0, \mathbf{p}_0)$. Let the system evolve via Hamilton's equations, so that at time t, the particle is at $(\mathbf{q}(t), \mathbf{p}(t))$. This is called *Hamiltonian flow*.

Liouville's theorem states that Hamiltonian preserves volumes in phase space: $\mathrm{Vol}(D(0)) = \mathrm{Vol}(D(t))$, where D is a region in phase space that changes with time according to Hamiltonian flow.

Proof: Let

$$\mathrm{Vol}(D(t)) = \int_{D(t)} dq_1 dq_2 ... dq_n\, dp_1 dp_2 ... dp_n = \int_{D(t)} dV(t).$$

We have $dV(0) = dq_1...dq_n\ dp_1...dp_n\big|_{t=0}$. As we evolve according to Hamiltonian flow, we have

$$q_i(t) = q_i(0) + t\frac{\partial H}{\partial p_i}(0) + O(t^2) =: \tilde{q}_i,$$
$$p_i(t) = p_i(0) - t\frac{\partial H}{\partial q_i}(0) + O(t^2) =: \tilde{p}_i.$$

We wish to compute $\det(J)$, where $d\tilde{q}_1...d\tilde{q}_n\ d\tilde{p}_1...d\tilde{p}_n = dV(t) = \det(J)dV(0)$, i.e.

$$\det(J) = \begin{pmatrix} \frac{\partial \tilde{q}_i}{\partial q_j} & \frac{\partial \tilde{q}_i}{\partial p_j} \\ \frac{\partial \tilde{p}_i}{\partial q_j} & \frac{\partial \tilde{p}_i}{\partial p_j} \end{pmatrix} = \det \begin{pmatrix} \delta_{ij} + \frac{\partial^2 H}{\partial p_i q_j}t & \frac{\partial^2 H}{\partial p_i \partial p_j}t \\ \frac{\partial^2 H}{\partial q_i \partial q_j}t & \delta_{ij} - \frac{\partial^2 H}{\partial q_i \partial p_j}t \end{pmatrix}$$

Now use $\det(1 + tM) = 1 + \epsilon \mathrm{tr}(M) + O(t^2)$. It then follows that

$$\det(J) = 1 + t\sum_i \left(\frac{\partial^2 H}{\partial p_i \partial q_j} - \frac{\partial^2 H}{\partial q_i \partial p_j}\right) + O(t^2) = 1 + O(t^2). \quad \square$$

4.6 The Poincaré recurrence theorem

Theorem: Let P be a point in a finite volume phase space. For any neighbourhood D_0 of P, there exists $P' \in D_0$ that will return to D_0 under a Hamiltonian flow in finite time.

Proof: Evolve D_0 to D_1 for a time T, then D_1 to D_2 for a time T, etc, until be get to D_k. Liouville's Theorem implies all of these regions have the same volume.

Suppose that $D_k \cap D_{k'} = \emptyset$ for all distinct integers k and k'. Then

$$\mathrm{Vol}\left(\bigcup_{k=0}^{\infty} D_k\right) = \sum_{k=0}^{\infty} \mathrm{Vol}(D_k) = \infty,$$

which contradicts the finite volume assumption. So there exist distinct k and k' with $D_k \cap D_{k'} \neq \emptyset$. Suppose $k' > k$ without loss of generality, and let $\Omega_{kk'} = D_k \cap D_{k'}$.

Since Hamiltonian flow is invertible, by uniqueness of solution to ODEs, we have $\Omega_{0,k'-k} = D_0 \cap D_{k'-k} \neq \emptyset$. Hence there exists $P' \in D_0$ such that P' returns to D_0 after $k' - k$ steps of time T. $\square$

4.7 Poisson brackets

Definition: The *Poisson bracket* $\{f, g\}$ of f and g, two functions on phase space, is defined by

$$\{f, g\} = \sum_{j=1}^{n} \left(\frac{\partial f}{\partial q_j}\frac{\partial g}{\partial p_j} - \frac{\partial g}{\partial q_j}\frac{\partial f}{\partial p_j}\right).$$

Theorem: The Poisson bracket obeys:

(i) antisymmetry, $\{f, g\} = -\{g, f\}$;

(ii) linearity, $\{\alpha f + \beta g, h\} = \alpha\{f, h\} + \beta\{g, h\}$;

(iii) the Leibniz rule, $\{f, gh\} = \{f, g\}h + \{f, h\}g$;

(iv) the Jacobi identity, $\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\}$.

Proof: Trivial from the definition. $\square$

Theorem: If Hamilton's equations hold, then

$$\frac{df}{dt} = \{f, H\} + \frac{\partial f}{\partial t}.$$

Proof: Just use chain rule and Hamilton's equations. $\square$

Definition: A function f is a first integral of the Hamiltonian H if f is constant under Hamiltonian flow. If $\frac{\partial f}{\partial t} = 0$, then by the above Theorem, f is a first integral if and only if $\{f, H\} = 0$.

In particular, if $\frac{\partial H}{\partial t} = 0$, then H is a first integral, which we call the *energy*.

More generally, if two functions f and g satisfy $\{f, g\}$ they are said to *Poisson commute* or to be *in involution*.

4.8 Canonical transformations

Notice if we define $\mathbf{x} = (\mathbf{q}, \mathbf{p})$, Hamilton's equations can be written more compactly as

$$\dot{\mathbf{x}} = J\frac{\partial H}{\partial \mathbf{x}},$$

where

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

Theorem: The coordinate transform $\mathbf{y} = \mathbf{y}(\mathbf{x})$ preserves the form of Hamilton's equations if and only if $D\mathbf{y}$ is a symplectic matrix, i.e. $D\mathbf{y}JD\mathbf{y}^T = J$.

Proof: We have

$$\dot{y}_a = \frac{\partial y_a}{\partial x_b}\dot{x}_b = \frac{\partial y_a}{\partial x_b}J_{bc}\frac{\partial H}{\partial x_c} = \frac{\partial y_a}{\partial x_b}J_{bc}\frac{\partial y_d}{\partial x_c}\frac{\partial H}{\partial y_d}.$$

Hence

$$\dot{\mathbf{y}} = (D\mathbf{y}JD\mathbf{y}^T)\frac{\partial H}{\partial \mathbf{y}}. \quad \square$$

Theorem: A transformation is canonical if and only if the Poisson bracket structure is conserved.

Proof: Write out the previous Theorem in matrix form. $\square$

4.9 Infinitesimal canonical transformations

Consider the transformation, for $\epsilon \ll 1$:

$$q_i \mapsto Q_i = q_i + \epsilon F_i(\mathbf{q}, \mathbf{p}) + O(\epsilon^2),$$
$$p_i \mapsto P_i = p_i + \epsilon E_i(\mathbf{q}, \mathbf{p}) + O(\epsilon^2).$$

Theorem: If this transformation is canonical, then there exists a function $G(\mathbf{q}, \mathbf{p})$ such that

$$F_i = \frac{\partial G}{\partial p_i}, \quad E_i = -\frac{\partial G}{\partial q_i}.$$

Proof: Use that the Poisson bracket structure is conserved. We have:

$$0 = \{Q_i, Q_j\} = \frac{\partial Q_i}{\partial q_k}\frac{\partial Q_j}{\partial p_k} - \frac{\partial Q_i}{\partial p_k}\frac{\partial Q_i}{\partial q_k}$$

$$= \epsilon \delta_{ik}\frac{\partial F_j}{\partial p_k} - \epsilon \frac{\partial F_i}{\partial p_l}\delta_{jk} + O(\epsilon^2).$$

Similarly,

$$\delta_{ij} = \{Q_i, P_j\} = \delta_{ij} + \epsilon\left(\frac{\partial E_j}{\partial p_i} + \frac{\partial F_i}{\partial q_j}\right) + O(\epsilon^2).$$

Hence we get the conditions:

$$\frac{\partial F_j}{\partial p_i} = \frac{\partial F_i}{\partial p_j}, \qquad \frac{\partial F_i}{\partial q_j} = -\frac{\partial E_j}{\partial p_i}.$$

These consistency conditions imply the existence of the required G. If such a G exists, the condition $0 = \{P_i, P_j\}$ is also fulfilled. $\square$

Idea: Infinitesimal canonical transformations are generated by a Hamiltonian flow, with Hamiltonian G. A special case is time evolution, where $G = H$.

4.10 Noether's theorem

Definition: A *symmetry* of the Hamiltonian is an infinitesimal canonical transformation generated by some G that gives $\delta H = 0$ (to order ϵ).

Theorem: If G generates a symmetry, $\{G, H\} = 0$.

Proof: We have

$$\delta H = \frac{\partial H}{\partial q_i}\delta q_i + \frac{\partial H}{\partial p_i}\delta p_i$$
$$= \epsilon\left(\frac{\partial H}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial G}{\partial q_i}\right) + O(\epsilon^2) = \epsilon\{H, G\} + O(\epsilon^2). \quad \square$$

Theorem (Noether's Theorem): If G generates a symmetry of the Hamiltonian, and is time-independent, then G is a first integral.

Proof: Follows immediately from the above, since $\{G, H\} = 0$. $\square$

4.11 Generating functions

This is a general method of constructing canonical transformations.

Theorem: Let $F(\mathbf{q}, \mathbf{Q})$ be an arbitrary function. Then if we define

$$\mathbf{p} = \frac{\partial F}{\partial \mathbf{q}}, \quad -\mathbf{P} = \frac{\partial F}{\partial \mathbf{Q}},$$

and solve for $\mathbf{Q}$ and $\mathbf{P}$, the result is a canonical transformation.

Proof: Write the actions for the two sets of variables as

$$S = \int_{t_1}^{t_2} (p_i \dot{q}_i - H(\mathbf{q}, \mathbf{p}))\, dt = \int_{t_1}^{t_2} (\mathbf{p}\cdot d\mathbf{q} - H(\mathbf{q}, \mathbf{p})dt),$$

and

$$S' = \int_{t_1}^{t_2} (P_i \dot{Q}_i - H'(\mathbf{Q}, \mathbf{P}))\, dt = \int_{t_1}^{t_2} (\mathbf{P}\cdot d\mathbf{Q} - H'(\mathbf{Q}, \mathbf{P})dt).$$

Setting $\delta S = 0$ and $\delta S' = 0$ gives Hamilton's equations for the two sets of variables. But then $\delta(S - S') = 0$, which implies that the integrands differ by any total derivative,

$$dF = \mathbf{p}\cdot d\mathbf{q} - \mathbf{P}\cdot d\mathbf{Q} - (H - H')dt.$$

Comparing with:

$$dF = \frac{\partial F}{\partial \mathbf{q}}\cdot d\mathbf{q} + \frac{\partial F}{\partial \mathbf{Q}}\cdot d\mathbf{Q},$$

gives the result. In particular, we get $H = H'$. $\square$

By taking the Legendre transform, we can obtain the generating function $F(\mathbf{q}, \mathbf{P})$ with

$$\mathbf{p} = \frac{\partial F}{\partial \mathbf{q}}, \quad \mathbf{Q} = \frac{\partial F}{\partial \mathbf{P}},$$

which we know from Integrable Systems.

4.12 Integrability and action-angle variables

Definition: Suppose there exists a canonical transformation $(\mathbf{q},\mathbf{p}) \mapsto (\boldsymbol{\theta},\mathbf{I})$ with $H \equiv H(\mathbf{I})$. Then $(\boldsymbol{\theta},\mathbf{I})$ are called *action-angle variables*.

Trivially, Hamilton's equations in action-angle variables can be integrated.

Definition: An *integrable system* if a $2n$-dimensional phase space M together with n first integrals $f_i : M \to \mathbb{R}$, $i = 1, 2, ..., n$ such that:

- ∇f_i are linearly independent vectors at all points in M;
- the functions f_i are all in involution.

This definition makes sense because of...

4.13 The Arnold-Liouville theorem

Theorem: Let $(M, f_1, f_2, ..., f_n)$ be an integrable system with $H = f_1$, say. Define

$$M_{\mathbf{c}} = \{(\mathbf{q},\mathbf{p}) \in M : f_i(\mathbf{q},\mathbf{p}) = c_i\}.$$

If $M_{\mathbf{c}}$ is compact, then

(i) $M_{\mathbf{c}}$ is diffeomorphic to the torus T^n;

(ii) there exists a local canonical transformation $(\mathbf{q},\mathbf{p}) \mapsto (\boldsymbol{\theta},\mathbf{I})$ where $\boldsymbol{\theta})$ are coordinates on $M_{\mathbf{c}}$, $\mathbf{I}$ are first integrals, and $H = H(\mathbf{I})$ (i.e. these are action-angle coordinates, so we can easily integrate Hamilton's equations).

Proof: See Integrable Systems. Here, we only construct the $(\boldsymbol{\theta},\mathbf{I})$. Define

$$I_k = \frac{1}{2\pi}\oint_{\Gamma_k} \mathbf{p}\cdot d\mathbf{q},$$

where Γ_k is the kth cycle on the torus. This definition does not depend on the choice of cycle (see Integrable Systems - use Green's Theorem). Since the actions depend only on $\mathbf{c}$, they are first integrals.

Use the generating function

$$F(\mathbf{q},\mathbf{I}) = \int_{\mathbf{q}_0}^{\mathbf{q}} \mathbf{p}\cdot d\mathbf{q}$$

to construct the angle coordinates via

$$\theta_k = \frac{\partial F}{\partial I_k}.$$

Since we used a generating function, the transformation is canonical. □

4.14 Adiabatic invariants

Definition: A function $I(\mathbf{q},\mathbf{p},\lambda)$, with $\dot{\lambda} = O(\epsilon)$, $\epsilon \ll 1$, is an *adiabatic invariant* of a system with Hamiltonian $H(\mathbf{q},\mathbf{p},\lambda)$ if $|I(t) - I(0)| = O(\epsilon)$ for all $0 < t < T/\epsilon$.

Idea: I doesn't change very much. We can treat it as almost constant.

Theorem: The action variable $\mathbf{I}$ is an adiabatic invariant.

Proof: Use action variables $H = H(\mathbf{I},\lambda)$. Introduce the generating function $F(\mathbf{q},\mathbf{I},\lambda)$. If $\lambda = \lambda(t)$, then

$$\frac{\partial F}{\partial t} \neq 0,$$

so in our earlier construction of the generating function, we need instead

$$H' = H + \frac{\partial F}{\partial t} = H + \epsilon\dot{\lambda}\frac{\partial F}{\partial \lambda}.$$

Hence by Hamilton's equations:

$$\dot{\mathbf{I}} = -\frac{\partial H'}{\partial \theta} = -\epsilon\dot{\lambda}\frac{\partial^2 F}{\partial\lambda\partial\theta} = O(\epsilon),$$

assuming $\dot{\lambda} = O(\epsilon)$. Hence $\mathbf{I}$ is slowly-varying. □

Example: Let $H = \frac{1}{2}p^2 + \frac{1}{2}\lambda(t)q^4$. We can easily construct the action:

$$I = \frac{1}{2\pi}\oint \mathbf{p}\cdot\mathbf{q} = \frac{2}{2\pi}\int_{q_1}^{q_2}\sqrt{2E - \lambda(t)q^4}\,dq.$$

where q_1, q_2 are the real roots of $2E - \lambda(t)q^4 = 0$. Changing variables, we can put this in the form:

$$\mathbf{I} = \frac{\sqrt{2E}}{\pi}\left(\frac{2E}{\lambda}\right)^{1/4}\int_{-1}^{1}\sqrt{1 - x^4}\,dx.$$

Since $\mathbf{I}$ is almost constant, it follows that $E \sim \lambda^{1/3}$.

4.15 Poisson structures

Definition: A *Poisson structure* is a pair (M, J) such that M is a phase space of dimension m, $J = J^{ab}$ is a skew-symmetric matrix with components that are functions of M, and the bracket

$$\{f, g\} = \sum_{a,b=1}^{m} J^{ab}\frac{\partial f}{\partial x^a}\frac{\partial g}{\partial x^b},$$

obeys the Jacobi identity. For the Poisson structure, *Hamilton's equations* are $\dot{\mathbf{x}} = \{\mathbf{x}, H\}$.

Most properties of Poisson structures are proved using $\{x^i, x^j\} = J^{ij}$, where x^i and x^j are the coordinate functions of the phase space.