Chapter 3

Variational theory of motion

Here we introduce some more general Lagrangians depending on vector valued functions and functions of several variables.

3.1 Hamilton's action integral

In Newtonian mechanics one considers an idealized mass m concentrated at a point and moving in three-dimensional Euclidean space with position $\mathbf{x} = \mathbf{x}(t)$ for $t \in (a, b)$. The **total kinetic energy** ssociated with a particular motion is defined by

$$T = \frac{1}{2} m \int_a^b |\mathbf{x}'(t)|^2 \, dt.$$

Notice, first of all, that the kinetic energy is a scalar function of time, and T is something derived from the kinetic energy associated with a total motion; it is not a function of time. A second thing to notice is that T is essentially the Dirichlet energy associated with the map $\mathbf{x}$.

Let us also assume there is a *spatially dependent potential energy function* $\Phi = \Phi(\mathbf{x})$ defined on the space in which the particle moves and such that the **total potential energy** associated with a particular motion is

$$V = \int_a^b \Phi(\mathbf{x}(t)) \, dt.$$

For example, if the mass moves in the gravitational field of another mass M at the origin, then

$$\Phi(\mathbf{x}) = -\frac{mMG}{|\mathbf{x}|}$$

where G is a (universal) gravitational constant. This quantity (like any potential energy) is only defined up to an arbitrary additive constant, and any particular sphere centered at the origin may be normalized to have zero potential.

Hamiltons action integral is defined to be

$$A = T - V = \int_a^b \left[\frac{1}{2}m|\mathbf{x}'(t)|^2 - \Phi(\mathbf{x}(t))\right] dt.$$

defined on an admissible class of motions with fixed starting and ending points

$$\mathcal{A} = \{\mathbf{x} \in C^1([a,b] \to \mathbb{R}^3) : \mathbf{x}(a) = \mathbf{x}_a,\, \mathbf{x}(b) = \mathbf{x}_b\}.$$

The first variation is defined by

$$\delta A_{\mathbf{x}}[\vec{\phi}] = \frac{d}{d\epsilon} A[\mathbf{x} + \epsilon\vec{\phi}]\Big|_{\epsilon=0}$$

for vector values variations

$$\vec{\phi} = (\phi_1, \phi_2, \phi_3) \in C_c^\infty(a,b) \times C_c^\infty(a,b) \times C_c^\infty(a,b).$$

A calculation (see Exercise 25) shows that C^2 extremals for A satisfy

$$\delta A_{\mathbf{x}}[\vec{\phi}] = -\int_a^b [m\mathbf{x}'' + D\Phi(\mathbf{x})] \cdot \vec{\phi}\, dt = 0 \quad \text{for all } \vec{\phi}.$$

Applying the fundamental lemma in each component, we obtain Newton's second law of motion:

$$m\mathbf{x}'' = -D\Phi(\mathbf{x}).$$

The vector field $-D\Phi$ is called the potential field in which the motion takes place.

The idea here is that among all possible motions from $\mathbf{x}_a$ to $\mathbf{x}_b$, the action integral finds or chooses the motion which actually occurs within the field $-D\Phi$.

This is Example 3 on page 14 of BGH and makes another appearance in Example 5 on page 18 of BGH.

3.2 Continuum motion

Hamilton's principle can be used to great advantage in a variety of curcumstances. One of the most notable is to determine the equation of motion for an extended body which deforms subject to an elastic energy associated with deformation. Here we present a very simple example. Let $\mathcal{A}$ represent a class of planar curves of fixed length parameterized by arclength. For example, we may consider

$$\mathcal{A} = \left\{ \mathbf{x} \in C^2([0,\ell] \times [0,T] \to \mathbb{R}^2) : \left| \frac{\partial \mathbf{x}}{\partial s} \right| = 1 \right\}$$

where s represents arclength along the continuum. We have assumed some extra regularity here because the elastic energy is going to be made dependedent on the magnitude of the curvature:

$$k = \left| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right|.$$

Various endpoint conditions or other boundary conditions could be imposed. We have left them out for the moment.

Assuming a linear density ρ along the continuum, we may form a total kinetic energy expression

$$T = \frac{1}{2} \int_0^T \int_0^\ell \rho \left| \frac{\partial \mathbf{x}}{\partial t} \right|^2 ds\, dt.$$

A simple elastic potential energy is

$$V = \int_0^T \int_0^\ell \left| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right| ds\, dt.$$

The action integral is then

$$A = T - V = \int_0^T \int_0^\ell \left[\frac{\rho}{2} \left| \frac{\partial \mathbf{x}}{\partial t} \right|^2 - \left| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right| \right] ds\, dt.$$

The first variation is defined as usual by

$$\delta A_{\mathbf{x}}[\vec{\phi}] = \frac{d}{d\epsilon} A[\mathbf{x} + \epsilon \vec{\phi}] \Big|_{\epsilon=0}.$$

If the action is stationary, then

$$\int_0^T \int_0^\ell \left[\rho \frac{\partial \mathbf{x}}{\partial t} \cdot \frac{\partial \vec{\phi}}{\partial t} - \frac{\partial^2 \mathbf{x}}{\partial s^2} \cdot \frac{\partial^2 \vec{\phi}}{\partial s^2} \Big/ \left| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right| \right] ds\, dt = 0$$

for all admissible variations $vec\phi$. In this way, we obtain a nonlinear equation of motion for the elastic rod:

$$\frac{\partial}{\partial t} \left(\rho \frac{\partial \mathbf{x}}{\partial t} \right) = \frac{\partial^2}{\partial s^2} \left(\frac{\partial^2 \mathbf{x}}{\partial s^2} \Big/ \left| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right| \right). \tag{3.1}$$

This is called a nonlinear evolution equation. Notice that it is a fourth order partial differential equation (PDE). See Exercise 26.

Exercises

Exercise 25. *Carry out the details of computing the first variation of the action integral, and derive Newton's second law of motion. What do you get in the case of a centrally symmetric gravitational field. (One can go on with this equation to derive Kepler's laws of planetary motion.)*

Exercise 26. *Verify that (3.1) is the Euler-Lagrange equation associated to the action integral of the elastic rod. Can you model some simple solution of this PDE numerically?*

Exercise 27. *If the elastic energy for the elastic rod is*

$$V = \int_0^T \int_0^\ell \left| \frac{\partial^2 \mathbf{x}}{\partial s^2} \right|^2 ds\, dt.$$

what is the resulting evolution equation?

3.3 Lagrangian constraints

The question we address here is of more applicability than to Hamiltonian mechanics. In particular, the main result can be used to determine the differential equations for geodesics on a surface.

Our basic objective is to give a version of Theorem 9 in which the integral constraint is replaced with a traditional analytic constraint $\Phi(\mathbf{x}) = c$. It is natural to consider this problem for a vector valued function $\mathbf{x} \in C^1([a,b] \to \mathbb{R}^n)$ as considered in Hamiltonian mechanics. We will not strive for minimal regularity.

Theorem 12. *Let*

$$\mathcal{A} = \left\{ \mathbf{x} \in C^1([a,b] \to \mathbb{R}^n) : \mathbf{x}(a) = \mathbf{p},\ \mathbf{x}(b) = \mathbf{q},\ \textit{and } \Phi(\mathbf{x}) = c \right\}$$

where

$$\Phi \in C^2(\mathbb{R}^n)$$

is given. Assume $\mathbf{x}_0 \in \mathcal{A}$ satisfies

$$\mathcal{F}[\mathbf{x}_0] \le \mathcal{F}[\mathbf{x}] \qquad \textit{for all } \mathbf{x} \in \mathcal{A}$$

where

$$\mathcal{F}[u] = \int_a^b F(t, \mathbf{x}, \mathbf{x}')\, dt \qquad \textit{with } F \in C^2([a,b] \times \mathbb{R}^n \times \mathbb{R}^n)$$

and

$$D\Phi(\mathbf{x}_0) \neq 0 \qquad \textit{for all } t \in [a,b]. \tag{3.2}$$

Then there is some $\lambda \in C^0[a,b]$ such that

$$\frac{d}{dt}\left[F_{p_j}(t, \mathbf{x}_0, \mathbf{x}_0') \right] = F_{x_j}(t, \mathbf{x}_0, \mathbf{x}_0') + \lambda \Phi_{x_j}(\mathbf{x}_0)$$

for all $x \in [a,b]$ and $j = 1, 2, \ldots, n$. Equivalently, $\mathbf{x}_0$ is an extremal for the functional $\mathcal{F} + \mathcal{G}$ where

$$\mathcal{G}[\mathbf{x}] = \int_a^b \lambda \Phi(\mathbf{x})\, dt.$$

Here we are considering $F = F(t, \mathbf{x}, \mathbf{p})$ with $\mathbf{x}, \mathbf{p} \in \mathbb{R}^n$, and we have used various notations for derivatives. In particular, the Euler-Lagrange equations may also be written as

$$\left(\frac{\partial F}{\partial p_j}(t, \mathbf{x}_0, \mathbf{x}_0') \right)' = \frac{\partial F}{\partial x_j}(t, \mathbf{x}_0, \mathbf{x}_0') + \lambda(t) \frac{\partial \Phi}{\partial x_j} \mathbf{x}_0$$

for $j = 1, 2, \ldots, n$.

Proof: The condition (3.2) ensures that for each $t_0 \in (a,b)$, there is at least one index j for which $\Phi_{x_j}(\mathbf{x}_0(t)) \neq 0$. This means

$$I = I(t) = \{j : \Phi_{x_j}(\mathbf{x}_0(t)) \neq 0\} \neq \phi.$$

We will establish the following result:

Lemma 6. *If $j, k \in I(t)$, then*

$$\frac{1}{\Phi_{x_k}(\mathbf{x}_0(t))}\left[\frac{d}{dt}[F_{p_k}(t,\mathbf{x}_0,\mathbf{x}_0')] - F_{x_k}(t,\mathbf{x}_0,\mathbf{x}_0')\right]$$
$$\equiv \frac{1}{\Phi_{x_j}(\mathbf{x}_0(t))}\left[\frac{d}{dt}[F_{p_j}(t,\mathbf{x}_0,\mathbf{x}_0')] - F_{x_j}(t,\mathbf{x}_0,\mathbf{x}_0')\right].$$

As a consequence, the expression

$$\lambda(t) = \frac{1}{|I|}\sum_{j\in I}\left(\frac{1}{\Phi_{x_j}(\mathbf{x}_0(t))}\left[\frac{d}{dt}[F_{p_j}(t,\mathbf{x}_0,\mathbf{x}_0')] - F_{x_j}(t,\mathbf{x}_0,\mathbf{x}_0')\right]\right),$$

i.e., the average of the equal values, defines a single continuous function given locally by any one of the functions

$$\lambda(t) = \frac{1}{\Phi_{x_j}(\mathbf{x}_0(t))}\left[\frac{d}{dt}[F_{p_j}(t,\mathbf{x}_0,\mathbf{x}_0')] - F_{x_j}(t,\mathbf{x}_0,\mathbf{x}_0')\right]$$

with $j \in I$. In the course of the proof we will also use (and establish) the Euler-Lagrange equations as asserted in the theorem. To this end, let $t_0 \in (a,b)$ be fixed. We know $I(t_0)$ is nonempty. For ease of presentation, let us assume $n \in I(t_0)$. The argument for other indices is essentially the same. This is also a reasonable time to introduce some simplifying notation:

$\mathbf{x}_0 = \mathbf{x}_0(t) = (x_1^0, \dots, x_n^0)$ where each of the functions x_j^0 is a function of t.

The projection of a point $\mathbf{x} = (x_1, \dots, x_n)$ into $\mathbb{R}^{n-1}$ given by $\mathbf{x} \mapsto (x_1, \dots, x_{n-1})$ will be denoted by $\mathrm{proj}(\mathbf{x})$. We will also freely append coordinates, so that $\mathbf{x} = (\mathrm{proj}(\mathbf{x}), x_n)$.

By continuity there is in interval (α, β) with $a < \alpha < t_0 < \beta < b$ such that $n \in I(t)$ for $\alpha < t < \beta$. Furthermore, by the implicit function theorem, since

$$\Phi_{x_n}(t_0) \neq 0 \qquad \text{and} \qquad \Phi(\mathbf{x}_0) = c,$$

there is some $\delta > 0$ and a unique function $\psi \in C^2(R)$ satisfying

$$\Phi(x_1, \ldots, x_{n-1}, \psi(x_1, \ldots, x_{n-1}; \tau)) = \tau$$

where $Q = \prod_{j=1}^{n-1}(x_j(t_0) - \delta, x_j(t_0) + \delta) \times (c - \delta, c + \delta)$. Using the notation introduced above, ψ satisfies

$$\Phi(\mathrm{proj}(\mathbf{x}), \psi(\mathrm{proj}(\mathbf{x}); \tau)) = \tau.$$

For $\tau = c$ fixed, ψ may be considered on cube/square $Q_0 = \mathrm{proj}(Q) = \prod_{j=1}^{n-1}(x_j(t_0) - \delta, x_j(t_0) + \delta) \subset \mathbb{R}^{n-1}$ as indicated in Figure 3.1, and the level set

$$\{\mathbf{x} \in \mathbb{R}^n : \Phi(\mathbf{x}) = c\}$$

is given locally by the graph of ψ over Q_0. We may also take δ_1 so that

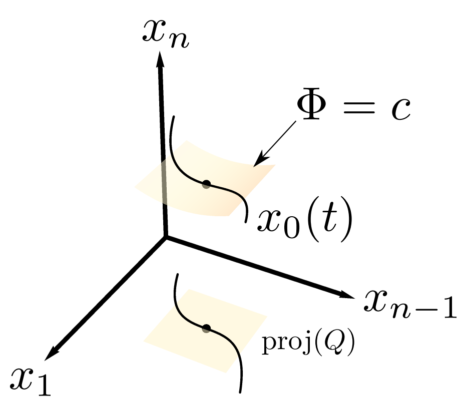

Figure 3.1: construction of a local admissible variation near $\mathbf{x}(t_0)$

$(t_0 - \delta_1, t_0 + \delta_1) \subset (\alpha, \beta)$ and

$$\mathbf{x}_0(t) \subset \Psi(Q) \qquad \text{for } t \in (t_0 - \delta_1, t_0 + \delta_1)$$

where $\Psi(Q)$ is the open image

$$\Psi(Q) = \{(\mathrm{proj}(\mathbf{x}), \psi(\mathrm{proj}(\mathbf{x}); \tau)) : \mathrm{proj}(\mathbf{x}) \in \mathrm{proj}(Q) \text{ and } \tau \in (c - \delta, c + \delta)\}.$$

Letting $\xi \in C_c^\infty(t_0 - \delta_1, t_0 + \delta_1)$, we form for each $\ell = 1, 2, \ldots, n-1$ the variation

$$\mathbf{x}(t) = \mathbf{x}(t; \epsilon) = (\mathrm{proj}(\mathbf{x}_0) + \epsilon\xi\mathbf{e}_\ell, \psi(\mathrm{proj}(\mathbf{x}_0) + \epsilon\xi\mathbf{e}_\ell))$$

where $\psi(\text{proj}(\mathbf{x}_0)+\epsilon\xi\mathbf{e}_\ell)=\psi(\text{proj}(\mathbf{x}_0)+\epsilon\xi\mathbf{e}_\ell;c)$ so that

$$\Phi(\mathbf{x})\equiv c,$$

and $\mathbf{x}$ is well-defined and admissible for ϵ small. Also, there is no variation outside $(t_0-\delta_1,t_0+\delta_1)$. In particular, we have

$$0=\frac{d}{d\epsilon}\int_\alpha^\beta F(t,\mathbf{x},\mathbf{x}')\,dt\Big|_{\epsilon=0}. \tag{3.3}$$

We have restricted to (α,β) for notational convenience, but we could also restrict the integration to the smaller interval $(t_0-\delta_1,t_0+\delta_1)$. Notice that if we write $\mathbf{x}=(x_1,\ldots,\mathbf{x}_n)$, then the coordinates $\mathbf{x}_j$ for $j=1,\ldots,n$ are being used in two different ways. They denote coordinates in $\mathbb{R}^n$, e.g., $F=F(t,\mathbf{x},\mathbf{p})$, as well as the functions $x_j=x_j(t)$ satisfying

$$x_j(t)=\begin{cases} x_j^0(t), & j=1,\ldots,n-1,\ j\neq\ell\\ x_\ell^0(t)+\epsilon\xi(t), & j=\ell\\ \psi(\text{proj}(\mathbf{x}_0)+\epsilon\xi\mathbf{e}_\ell), & j=n.\end{cases}$$

In particular, the quantity on the right in (3.3) is

$$\begin{aligned}\int_\alpha^\beta\Bigg[&F_{x_\ell}(t,\mathbf{x}_0,\mathbf{x}'\big|_{\epsilon=0})\xi+F_{x_n}(t,\mathbf{x}_0,\mathbf{x}'\big|_{\epsilon=0})\frac{d}{d\epsilon}x_n\Big|_{\epsilon=0}\\ &+F_{p_\ell}(t,\mathbf{x}_0,\mathbf{x}'\big|_{\epsilon=0})\xi'+F_{p_n}(t,\mathbf{x}_0,\mathbf{x}'\big|_{\epsilon=0})\frac{d}{d\epsilon}x_n'\Big|_{\epsilon=0}\Bigg]\end{aligned}$$

where $x_n=\psi(\text{proj}(\mathbf{x}_0)+\epsilon\xi\mathbf{e}_\ell)$ and

$$x_n'=\frac{d}{dt}\psi(\text{proj}(\mathbf{x}_0)+\epsilon\xi\mathbf{e}_\ell).$$

We know, from the construction/definition of ψ that setting $\epsilon=0$ gives

$$\psi(\text{proj}(\mathbf{x}_0))=x_n^0.$$

Thus,

$$\mathbf{x}\Big|_{\epsilon=0}=\mathbf{x}_0\qquad\text{and}\qquad\frac{d}{dt}\mathbf{x}\Big|_{\epsilon=0}=\mathbf{x}'\Big|_{\epsilon=0}=\mathbf{x}_0'.$$

In particular,

$$\frac{d}{dt}x_n^0=(x_n^0)'=D\psi\cdot\text{proj}(\mathbf{x}_0').$$

In view of this calculation the quantity on the right in (3.3) simplifies to

$$\int_{\alpha}^{\beta} \Big[F_{x_\ell}(t, \mathbf{x}_0, \mathbf{x}_0')\xi + F_{x_n}(t, \mathbf{x}_0, \mathbf{x}_0') \frac{d}{d\epsilon} x_n\big|_{\epsilon=0} \\ + F_{p_\ell}(t, \mathbf{x}_0, \mathbf{x}_0')\xi' + F_{p_n}(t, \mathbf{x}_0, \mathbf{x}_0') \frac{d}{d\epsilon} x_n'\big|_{\epsilon=0} \Big].$$

Roughly speaking, the two quantities still appearing with an evaluation at $\epsilon = 0$ require some calculation involving the defining equation for ψ, namely

$$\Phi(\mathrm{proj}(\mathbf{x}), \psi(\mathrm{proj}(\mathbf{x}))) = c. \tag{3.4}$$

Differentiating this expression with respect to ϵ, we get

$$\Phi_{x_\ell}\xi + \Phi_{x_n} \frac{d}{d\epsilon} x_n = \Phi_{x_\ell}\xi + \Phi_{x_n}\psi_{x_\ell}\xi = 0.$$

Evaluating at $\epsilon = 0$, we have

$$\Phi_{x_\ell}(\mathbf{x}_0)\xi + \Phi_{x_n}(\mathbf{x}_0) \frac{d}{d\epsilon} x_n\big|_{\epsilon=0} = \Phi_{x_\ell}(\mathbf{x}_0)\xi + \Phi_{x_n}(\mathbf{x}_0)\psi_{x_\ell}(\mathrm{proj}(\mathbf{x}_0))\xi = 0.$$

In particular, on the interval of interest we can write

$$\frac{d}{d\epsilon} x_n\big|_{\epsilon=0} = \psi_{x_\ell}(\mathrm{proj}(\mathbf{x}_0))\xi = -\frac{\Phi_{x_\ell}(\mathbf{x}_0)}{\Phi_{x_n}(\mathbf{x}_0)}\,\xi. \tag{3.5}$$

Differentiating this with respect to t we obtain

$$\frac{d}{d\epsilon} x_n'\big|_{\epsilon=0} = \frac{d}{dt}\left[\psi_{x_\ell}(\mathrm{proj}(\mathbf{x}_0))\xi\right] = -\frac{d}{dt}\left[\frac{\Phi_{x_\ell}(\mathbf{x}_0)}{\Phi_{x_n}(\mathbf{x}_0)}\right]\xi - \frac{\Phi_{x_\ell}(\mathbf{x}_0)}{\Phi_{x_n}(\mathbf{x}_0)}\,\xi'. \tag{3.6}$$

Returning to our expression for the right side (3.3) and substituting from (3.5) and (3.6), the Euler-Lagrange equation determined locally near $t = t_0$ by the variation $\mathbf{x}$ in the x_ℓ coordinate is

$$\frac{d}{dt}\left[F_{p_\ell}(t, \mathbf{x}_0, \mathbf{x}_0')\right] = F_{x_\ell}(t, \mathbf{x}_0, \mathbf{x}_0') - F_{x_n}(t, \mathbf{x}_0, \mathbf{x}_0')\,\frac{\Phi_{x_\ell}(\mathbf{x}_0)}{\Phi_{x_n}(\mathbf{x}_0)} \\ - F_{p_n}(t, \mathbf{x}_0, \mathbf{x}_0')\,\frac{d}{dt}\left[\frac{\Phi_{x_\ell}(\mathbf{x}_0)}{\Phi_{x_n}(\mathbf{x}_0)}\right] + \frac{d}{dt}\left[F_{p_n}(t, \mathbf{x}_0, \mathbf{x}_0')\,\frac{\Phi_{x_\ell}(\mathbf{x}_0)}{\Phi_{x_n}(\mathbf{x}_0)}\right].$$

This immediately simplifies to

$$\frac{d}{dt}\left[F_{p_\ell}(t,\mathbf{x}_0,\mathbf{x}_0')\right] = F_{x_\ell}(t,\mathbf{x}_0,\mathbf{x}_0') + \frac{1}{\Phi_{x_n}(\mathbf{x}_0)}\left[\frac{d}{dt}F_{p_n}(t,\mathbf{x}_0,\mathbf{x}_0') - F_{x_n}(t,\mathbf{x}_0,\mathbf{x}_0')\right]\Phi_{x_\ell}.$$

This equation holds for each $\ell = 1, 2, \ldots, n-1$ and for $t_0 - \delta_1 < t < t_0 + \delta_1$. Thus, setting

$$\lambda(t) = \frac{1}{\Phi_{x_n}(\mathbf{x}_0)}\left[\frac{d}{dt}F_{p_n}(t,\mathbf{x}_0,\mathbf{x}_0') - F_{x_n}(t,\mathbf{x}_0,\mathbf{x}_0')\right],$$

which is independent of ℓ and ξ, we have

$$\frac{d}{dt}\left[F_{p_j}(t,\mathbf{x}_0,\mathbf{x}_0')\right] = F_{x_j}(t,\mathbf{x}_0,\mathbf{x}_0') + \lambda(t)\Phi_{x_j} \quad \text{for } j = 1, 2, \ldots, n$$

with the last equation following directly from the definition of λ. If $j \in I(t_0)$, the set of indices for which $\Phi_{x_j}(t_0) \neq 0$, then there is some open interval around t_0 on which

$$\lambda(t) = \frac{1}{\Phi_{x_j}(\mathbf{x}_0)}\left[\frac{d}{dt}F_{p_j}(t,\mathbf{x}_0,\mathbf{x}_0') - F_{x_j}(t,\mathbf{x}_0,\mathbf{x}_0')\right].$$

As noted above, this means the definition of $\lambda(t)$ given in terms of the average over indices in $I(t)$ is global and unambiguous. $\square$

Exercise 28. *State and prove a version of Theorem 12 which applies to applies to constrained minimization with respect to a constraint of the form* $\Phi(\mathbf{x},\mathbf{x}') = c$.