Dirac-Bergmann and Hamiltonian Field Theory

Niels Slotboom
slotboom.n@gmail.com

Contents

1 Hamiltonian Mechanics

1.1 Legendre Transform

1.1.1 One-Dimensional Case

Say we have a function $\mathcal{L}(q, \dot{q})$. Our goal is to encode the information it contains in terms of slopes with respect to $\dot{q}$ in place of this function depending on $\dot{q}$. This is done using a *Legendre-transformation*. In the following, we suppress the dependence on q since the Legendre transformation acts only on the $\dot{q}$-dependence. The q-dependence is treated parametrically. The Legendre transformation does the following: Take a slope p, and let $H(p)$ be such that

$$p\dot{q} - \mathcal{H}(p) \leq \mathcal{L}(\dot{q}) \tag{1.1.1}$$

everywhere, with equality in exactly one point $\dot{q}$ (This gives implicit assumptions on $\mathcal{L}$, e.g. that it is strictly convex). This is equivalent to

$$p\dot{q} - \mathcal{L}(\dot{q}) \leq \mathcal{H}(p). \tag{1.1.2}$$

Since we have equality in exactly one point $\dot{q}$ and $<$ everywhere else, the point of equality is a maximum (in the case of differentiable functions and that point not being at infinity), so we must have

$$\mathcal{H}(p) = \max_{\dot{q}}[p\dot{q} - \mathcal{L}(\dot{q})]. \tag{1.1.3}$$

The maximality condition hence demands that

$$\frac{\partial \mathcal{L}}{\partial \dot{q}}(\dot{q}(p)) = p, \tag{1.1.4}$$

where $\dot{q}(p)$ is the point at which the maximum (1.1.3) is attained, i.e.

$$\dot{q}(p) = \arg\max_{\dot{q}}[p\dot{q} - \mathcal{L}(\dot{q})]. \tag{1.1.5}$$

For a function $\mathcal{L}$ which is convex in $\dot{q}$, this is an invertible expression, which implicitly defines $\dot{q} = \dot{q}(p)$ and hence

$$\mathcal{H}(p) = p\dot{q}(p) - \mathcal{L}(\dot{q}(p)). \tag{1.1.6}$$

Now, one might wonder why $\mathcal{H}(p)$, the Legendre transform of $\mathcal{L}(\dot{q})$, is of any interest. The answer to this question is: It contains the same information as $\mathcal{L}(\dot{q})$, just phrased differently: since $p = p(\dot{q})$ is invertible, we may just as well write eq. (1.1.6) as

$$\mathcal{L}(\dot{q}) = p(\dot{q})\dot{q} - \mathcal{H}(p(\dot{q})), \tag{1.1.7}$$

which recovers the original $\mathcal{L}(\dot{q})$. This implies that no information was lost in transitioning to $\mathcal{H}(p)$, since we can recover the full $\mathcal{L}(\dot{q})$ from it.

1.1.2 Higher, Finite-Dimensional Case

In higher dimensions, say with $q = q^\mu$, $\dot{q} = \dot{q}^\mu$, the condition we start with is

$$p_\mu \dot{q}^\mu - \mathcal{H}(p) \leq \mathcal{L}(\dot{q}). \tag{1.1.8}$$

with the assumption that we have equality in exactly one point. Here, the $p_\mu \dot{q}^\mu$ generates a hyperplane in place of just a line. The condition that we have equality in exactly one point $\dot{q}$ again translates to

$$\mathcal{H}(p) = \max_{\dot{q}}[p\dot{q} - \mathcal{L}(\dot{q})]. \tag{1.1.9}$$

This time, the maximum is attained where

$$\frac{\partial \mathcal{L}}{\partial \dot{q}^\mu}(\dot{q}(p)) = p_\mu. \tag{1.1.10}$$

Under the condition that this relationship is invertible, which is equivalent to the condition

$$\det\left(\frac{\partial^2 \mathcal{L}}{\partial \dot{q}^\mu \partial \dot{q}^\nu}\right) \neq 0 \tag{1.1.11}$$

on the Jacobian of relation (1.1.10) (or equivalently, the Hessian of $\mathcal{L}$ with respect to $\dot{q}$), we can express

$$\dot{q}^\mu = \dot{q}^\mu(p). \tag{1.1.12}$$

Consequently, we have

$$\mathcal{H}(p) = p_\mu \dot{q}^\mu(p) - \mathcal{L}(\dot{q}(p)). \tag{1.1.13}$$

1.2 Equations of Motion from Lagrangian Mechanics

In Lagrangian mechanics, we work with a Lagrangian $\mathcal{L}(q, \dot{q})$, where $q = q^\mu$ are the coordinates of a particle. The Lagrangian defines an action functional via

$$S[q(\tau)] = \int \mathcal{L}(q, \dot{q})\, d\tau, \tag{1.2.1}$$

where $q(\tau)$ is a parametrisation of the path. The equations of motion are obtained by varying $q \to q + \delta q$ and requiring $S \to S$, i.e. that $\delta S = 0$. This leads to the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial q^\mu} = \frac{d}{d\tau}\frac{\partial \mathcal{L}}{\partial \dot{q}^\mu}. \tag{1.2.2}$$

These are typically second-order differential equations, since $\frac{\partial \mathcal{L}}{\partial \dot{q}^\mu}$ usually depends on $\dot{q}$.

We can now use these equations to derive the canonical equations of motion for the variables q^μ and p_μ associated with a Hamiltonian $\mathcal{H}$. Define the canonical momentum as

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu}. \tag{1.2.3}$$

Inserting this into Equation (1.2.2) yields

$$\dot{p}_\mu = \frac{\partial \mathcal{L}}{\partial q^\mu} = \frac{\partial}{\partial q^\mu}[p_\nu \dot{q}^\nu - \mathcal{H}(q, p)] = -\frac{\partial \mathcal{H}}{\partial q^\mu}, \tag{1.2.4}$$

which is the equation of motion for p_μ.

To derive the equation for q^μ, we begin with the definition of the Hamiltonian as a Legendre transform,

$$\mathcal{H}(q, p) = p_\mu \dot{q}^\mu - \mathcal{L}(q, \dot{q}). \tag{1.2.5}$$

Taking the derivative with respect to p_μ, we obtain

$$\frac{\partial \mathcal{H}}{\partial p_\mu} = \frac{\partial}{\partial p_\mu}[p_\nu \dot{q}^\nu - \mathcal{L}(q, \dot{q})] = \dot{q}^\mu + \underbrace{p_\nu \frac{\partial \dot{q}^\nu}{\partial p_\mu} - \frac{\partial \mathcal{L}}{\partial \dot{q}^\nu}\frac{\partial \dot{q}^\nu}{\partial p_\mu}}_{=0} = \dot{q}^\mu, \tag{1.2.6}$$

where the last equality follows from the identity $p_\nu = \frac{\partial \mathcal{L}}{\partial \dot{q}^\nu}$.

In summary, we have derived the canonical equations

$$\dot{q}^\mu = \frac{\partial \mathcal{H}}{\partial p_\mu}, \quad \dot{p}_\mu = -\frac{\partial \mathcal{H}}{\partial q^\mu}, \tag{1.2.7}$$

which are a set of first-order differential equations, in contrast to the second-order Euler-Lagrange equations (1.2.2). This makes them typically more convenient for numerical integration.

1.3 Poisson Brackets

An important concept in the Hamiltonian formalism is that of Poisson brackets. For any two observables A, B, functions of q, p, i.e.

$$A = A(q,p), \quad B = B(q,p), \tag{1.3.1}$$

their *Poisson bracket* is defined as

$$\{A,B\} = \frac{\partial A}{\partial q^\mu}\frac{\partial B}{\partial p_\mu} - \frac{\partial B}{\partial q^\mu}\frac{\partial A}{\partial p_\mu}. \tag{1.3.2}$$

Before we establish why this is useful, let us compute the Poisson brackets of the fundamental observables q^μ and p_μ. This is rather straightforward:

$$\begin{aligned}
\{q^\mu, q^\nu\} &= \frac{\partial q^\mu}{\partial q^\lambda}\underbrace{\frac{\partial q^\nu}{\partial p_\lambda}}_{=0} - \frac{\partial q^\nu}{\partial q^\lambda}\underbrace{\frac{\partial q^\mu}{\partial p_\lambda}}_{=0} = 0,\\
\{p_\mu, p_\nu\} &= \underbrace{\frac{\partial p_\mu}{\partial q^\lambda}}_{=0}\frac{\partial p_\nu}{\partial p_\lambda} - \underbrace{\frac{\partial p_\mu}{\partial q^\lambda}}_{=0}\frac{\partial p_\nu}{\partial p_\lambda} = 0,\\
\{q^\mu, p_\nu\} &= \underbrace{\frac{\partial q^\mu}{\partial q^\lambda}}_{=\delta^\mu_\lambda}\underbrace{\frac{\partial p_\nu}{\partial p_\lambda}}_{=\delta^\lambda_\nu} - \underbrace{\frac{\partial p_\nu}{\partial q^\lambda}\frac{\partial q^\mu}{\partial p_\lambda}}_{=0} = \delta^\mu_\nu.
\end{aligned} \tag{1.3.3}$$

To be able to reference this more compactly later on, we repeat this result:

$$\{q^\mu, q^\nu\} = \{p_\mu, p_\nu\} = 0, \quad \{q^\mu, p_\nu\} = \delta^\mu_\nu. \tag{1.3.4}$$

We are now ready to see why the notion of the Poisson bracket is useful. To this end, we consider a particular observable—the Hamiltonian $\mathcal{H}(q,p)$, and see how its Poisson bracket with another observable $A(q,p)$ behaves. To this end we recall the canonical equations (1.2.7) and compute

$$\{A,\mathcal{H}\} = \frac{\partial A}{\partial q^\lambda}\underbrace{\frac{\partial \mathcal{H}}{\partial p_\lambda}}_{=\dot{q}^\lambda} - \underbrace{\frac{\partial \mathcal{H}}{\partial q^\lambda}}_{=-\dot{p}_\lambda}\frac{\partial A}{\partial p_\lambda} = \frac{\partial A}{\partial q^\lambda}\frac{dq^\lambda}{d\tau} + \frac{\partial A}{\partial p_\lambda}\frac{dp_\lambda}{d\tau} = \frac{dA}{d\tau}. \tag{1.3.5}$$

This is a profound result, put more succinctly as

$$\frac{dA}{d\tau} = \{A,\mathcal{H}\}. \tag{1.3.6}$$

Put into words, this means the following: the evolution of any observable $A(q(\tau),p(\tau))$ along the particle's path is given by its Poisson bracket with the Hamiltonian (as long as the particle obeys the equations of motion). In particular, for the observables q^μ and p_μ we confirm

$$\{q^\mu,\mathcal{H}\} = \delta^\mu_\lambda\frac{\partial \mathcal{H}}{\partial p_\lambda} = \dot{q}^\mu, \quad \{p_\mu,\mathcal{H}\} = -\frac{\partial \mathcal{H}}{\partial q^\lambda}\delta^\lambda_\mu = \dot{p}_\mu. \tag{1.3.7}$$

We have hence established the Poisson bracket $\{\cdot,\mathcal{H}\}$ to be a valuable tool in obtaining the first-order evolution equation for any observable. One particular result that follows from the anti-symmetry of the Poisson bracket is that

$$\dot{\mathcal{H}} = \{\mathcal{H},\mathcal{H}\} = 0, \tag{1.3.8}$$

implying that along any physical path, the Hamiltonian is conserved.

1.4 Free Non-Relativistic Particle

We now have the necessary theory to begin looking at some examples. In this section, we consider a free particle—in the next, we will put it into a quadratic potential, turning the problem into a harmonic oscillator. A non-relativistic free particle has the Lagrangian

$$\mathcal{L} = \frac{1}{2} m \dot{q}^\mu \dot{q}_\mu, \tag{1.4.1}$$

assuming a flat Euclidean space (for simplicity, let us not try to do things with Minkowski-space for now. We will revisit that in Hamiltonian Field Theory). Here, m denotes the mass of the particle. We want to first derive the Hamiltonian, and then the concrete form of the equations of motion for the free particle. For this, we need the conjugate momenta, which we compute as

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} = m \dot{q}_\mu \quad \Leftrightarrow \quad \dot{q}_\mu = \frac{1}{m} p_\mu \tag{1.4.2}$$

This is the expected formula from classical mechanics: momentum equals mass times velocity. Using this momentum, we can compute the Hamiltonian,

$$\mathcal{H}(q,p) = p_\mu \dot{q}^\mu - \mathcal{L}(q,\dot{q}) = \frac{1}{m} p_\mu p^\mu - \frac{1}{2m} p_\mu p^\mu = \frac{1}{2m} p_\mu p^\mu. \tag{1.4.3}$$

This is also the anticipated result—the Hamiltonian is the total energy of the system, which we know from classical mechanics to be $E = \frac{1}{2} m v^2 = p^2/2m$. Let us now move to the equations of motion. The canonical equations can straightforwardly be derived as

$$\dot{q}^\mu = \frac{\partial \mathcal{H}}{\partial p_\mu} = \frac{1}{m} p^\mu, \qquad \dot{p}_\mu = -\frac{\partial \mathcal{H}}{\partial q^\mu} = 0. \tag{1.4.4}$$

The latter equation has the simple interpretation that the momentum does not change. This makes sense for a free particle, since changes in momentum are caused by forces, of which there act none. The former equation is simply restating the relationship between velocity and momentum we had already found before.

1.5 Harmonic Oscillator

To make our previous example a bit more exciting, we can add a potential term. In general, the Lagrangian of a particle in a potential V has the form

$$\mathcal{L} = \frac{1}{2} m \dot{q}^\mu \dot{q}_\mu - V(q). \tag{1.5.1}$$

A harmonic oscillator has a particularly simple potential, with the full Lagrangian reading

$$\mathcal{L} = \frac{1}{2} m \dot{q}^\mu \dot{q}_\mu - \frac{1}{2} m \omega^2 q^\mu q_\mu. \tag{1.5.2}$$

The conjugate momenta evaluate to the same as in the free case,

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} = m \dot{q}_\mu, \tag{1.5.3}$$

since no additional dependence on $\dot{q}^\mu$ is introduced into the Lagrangian. We can again compute the Hamiltonian, finding

$$\begin{aligned}\mathcal{H} = p_\mu \dot{q}^\mu - \mathcal{L} &= \frac{1}{m} p_\mu p^\mu - \frac{1}{2m} p_\mu p^\mu + \frac{1}{2} m\omega^2 q_\mu q^\mu \\ &= \frac{1}{2m} p_\mu p^\mu + \frac{1}{2} m\omega^2 q_\mu q^\mu\end{aligned} \tag{1.5.4}$$

This should look familiar, from e.g. the Hamiltonian operator of a harmonic oscillator in quantum mechanics, where

$$\widehat{\mathcal{H}} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{q}^2 \tag{1.5.5}$$

Returning to Hamiltonian mechanics, let us now compute the canonical equations. This is also rather straightforward, leading to

$$\dot{q}^\mu = \frac{\partial \mathcal{H}}{\partial p_\mu} = \frac{1}{m} p^\mu, \qquad \dot{p}_\mu = -\frac{\partial \mathcal{H}}{\partial q^\mu} = m\omega^2 q_\mu. \tag{1.5.6}$$

These equations have by oscillatory solutions,

$$\begin{aligned} q^\mu(\tau) &= A^\mu \cos(\omega\tau) + B^\mu \sin(\omega\tau), \\ p^\mu(\tau) &= m\omega[-A^\mu \sin(\omega\tau) + B^\mu \cos(\omega\tau)], \end{aligned} \tag{1.5.7}$$

where $A^\mu, B^\mu \in \mathbb{R}^n$ are arbitrary constant vectors determined by initial conditions.

2 Hamiltonian Field Theory: Basics

2.1 Review: Lagrangian Field Theory

Before constructing the Hamiltonian formalism for classical field theories, we first review the Lagrangian approach. As in mechanics, the Legendre transform will then allow us to transition to the Hamiltonian picture. This process introduces some complications, which we will address in subsequent sections.

Let us begin by recalling the essential elements of Lagrangian mechanics, and then draw analogies with Lagrangian field theory.

In Lagrangian mechanics, we consider a one-dimensional parameter τ living in a connected subset of $\mathbb{R}$. The dynamical variables are the coordinate embeddings $x^\mu(\tau)$, and the Lagrangian $\mathcal{L}$ is a function of these variables and their derivatives,

$$\mathcal{L} = \mathcal{L}(x, \dot{x}, ...), \tag{2.1.1}$$

typically involving only first-order derivatives. From this, we define the action functional,

$$S[x] = \int d\tau\, \mathcal{L}(x(\tau), \dot{x}(\tau)). \tag{2.1.2}$$

Infinitesimally varying the path as $x^\mu \to x^\mu + \delta x^\mu$ results in a variation of the action,

$$\delta S = \int d\tau\, \delta\mathcal{L}(x, \dot{x}), \tag{2.1.3}$$

where the variation of $\mathcal{L}$ with respect to $x^\mu(\tau)$ is given by

$$\delta\mathcal{L} = \left[\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{d}{d\tau}\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu}\right]\delta x^\mu. \tag{2.1.4}$$

Requiring the action to remain invariant under all variations of the path leads to the Euler-Lagrange equations of motion

$$\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{d}{d\tau}\frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = 0. \tag{2.1.5}$$

We now turn to field theory. A field is, loosely speaking, a function defined on a spacetime manifold M. Whether the spacetime is flat or curved will not concern us here, as our focus is the structure of the formalism rather than its geometric details. Locally, the manifold provides coordinate charts, and fields become coordinate-dependent functions.

In contrast to mechanics, where the parameter τ was one-dimensional, field theory takes its parameters from the manifold itself: the space of independent variables is N-dimensional (for an N -dimensional manifold). The dynamical objects are the fields we denote as $\Phi^A(x)$, where A labels field components (or species).

This setup leads to several important differences:

1. **Multiple Derivatives**: Since the fields depend on spacetime coordinates x^μ, their derivatives $\partial_\mu \Phi^A(x)$ are now gradients rather than time derivatives. Depending on the theory, higher-order derivatives may also appear.
2. **Infinite Degrees of Freedom**: In field theory, each field $\Phi^A(x)$ is defined over all of spacetime, which means specifying its value at every point $x \in M$. This leads to an uncountably infinite number of degrees of freedom (This was also the case in classical mechanics when formulated for continuous systems, though over a lower-dimensional base space.).

The Lagrangian $\mathcal{L}$ must now be a scalar function of the fields and their derivatives,

$$\mathcal{L} = \mathcal{L}\big(\Phi^A(x), \partial_\mu \Phi^{A(x)}, ...\big). \tag{2.1.6}$$

Accordingly, the action becomes an integral over the full manifold,

$$S\big[\Phi^A\big] = \int_M d^N x\, \mathcal{L}\big(\Phi^A, \partial_\mu \Phi^A, ...\big). \tag{2.1.7}$$

As before, we determine the equations of motion by requiring stationarity of the action under infinitesimal variations $\Phi^A \to \Phi^A + \delta\Phi^A$. For simplicity, let us assume dependence of $\mathcal{L}$ on only Φ^A and its gradient $\partial_\mu \Phi^A$. We thus vary

$$\begin{aligned} \Phi^A &\to \Phi^A + \delta\Phi^A, \\ \partial_\mu \Phi^A &\to \partial_\mu \Phi^A + \partial_\mu \delta\Phi^A. \end{aligned} \tag{2.1.8}$$

Under such a variation, the action changes as $S \to S + \delta S$, where

$$\delta S = \int_M d^N x \left[\frac{\partial \mathcal{L}}{\partial \Phi^A}\delta\Phi^A + \frac{\partial \mathcal{L}}{\partial\big(\partial_\mu \Phi^A\big)}\partial_\mu \delta\Phi^A\right]. \tag{2.1.9}$$

Using integration by parts on the second term and assuming that $\delta\Phi^A(x)$ vanishes on the boundary (or that the manifold is compact without boundary), we obtain

$$\delta S = \int_M d^N x \left[\frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_\mu\left(\frac{\partial \mathcal{L}}{\partial\big(\partial_\mu \Phi^A\big)}\right)\right]\delta\Phi^A. \tag{2.1.10}$$

Since $\delta\Phi^A(x)$ is arbitrary (within suitable regularity conditions), stationarity $\delta S = 0$ implies the Euler-Lagrange equations for fields,

$$\frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^A)} \right) = 0. \tag{2.1.11}$$

This equation is often written in terms of the *functional derivative* of the action[1],

$$\frac{\delta S}{\delta \Phi^A(x)} := \left[\frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^A)} \right) \right] \big(\Phi^A(x), \partial_\mu \Phi^A(x)\big). \tag{2.1.12}$$

We will consider the functional derivative in more detail in the next section.

2.2 Functional Derivatives

Let $\Phi^A(x)$ be a field configuration over a spacetime manifold M, and let $S\big[\Phi^A\big]$ be a functional —that is, a map from functions to real numbers, typically given as an integral over M of a local function of Φ^A and its derivatives,

$$S\big[\Phi^A\big] = \int_M d^N x \mathcal{L}\big(\Phi^A(x), \partial_\mu \Phi^A(x), ...\big). \tag{2.2.1}$$

The *functional derivative* of S with respect to $\Phi^A(x)$, denoted by $\frac{\delta S}{\delta \Phi^A(x)}$, is defined by the condition that for any test variation $\delta\Phi^A(x)$ with compact support,

$$\delta S = \int_M d^N x \frac{\delta S}{\delta \Phi^A(x)} \delta \Phi^A(x). \tag{2.2.2}$$

In practice, when $\mathcal{L}$ depends on Φ^A and its first derivatives $\partial_\mu \Phi^A$, the functional derivative takes the familiar Euler-Lagrange form,

$$\frac{\delta S}{\delta \Phi^A(x)} = \left[\frac{\partial \mathcal{L}}{\partial \Phi^A(x)} - \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial_\mu \Phi^A} \right) \right](x), \tag{2.2.3}$$

evaluated at the point x. This is the same expression that appears in the Euler-Lagrange equation for fields.

Conceptually, the functional derivative measures the sensitivity of the action $S\big[\Phi^A\big]$ to infinitesimal local changes in the field. It plays the role of a gradient in infinite-dimensional function space, and it vanishes precisely on field configurations that extremise the action.

The definition (2.2.2) tells us that in order to compute the functional derivative of an action, we may compute its variation δS by the standard rules of variational calculus, collect terms and integrate by parts such that we isolate a $\delta\Phi^A(x)$ factor, and subsequently read it off.

Alternatively, we can derive some computation rules for the functional derivative. Firstly, note that $\Phi^A(x)$ can be seen as an integral functional when expressing it as

$$\Phi^A(x) = \int_M d^N y \, \delta(x-y) \delta^{AB} \Phi^B(y). \tag{2.2.4}$$

Using this, the variation becomes

$$\delta\Phi^A(x) = \int_M d^N y \, \delta^{(N-1)}(x-y) \delta^{AB} \delta\Phi^B(y), \tag{2.2.5}$$

from which we can read off that

[1] More generally, this contains derivatives of the integrand $\mathcal{L}$ with respect to all derivatives of Φ^A, not just Φ^A and $\partial_\mu \Phi^A$.

$$\frac{\delta\Phi^A(x)}{\delta\Phi^B(y)} = \delta^{(N-1)}(x-y)\delta^{AB} \tag{2.2.6}$$

This makes intuitive sense: the field value $\Phi^A(x)$ depends on $\Phi^B(y)$ if and only if $A = B$ and $x = y$—otherwise, field values are independent of each other.

Using

$$\partial_\mu \Phi^A(x) = \int_M d^N y\, \partial_\mu \delta^{(N-1)}(x-y)\delta^{AB}\Phi^B(y), \tag{2.2.7}$$

we can further conclude

$$\delta\big(\partial_\mu \Phi^A(x)\big) = \int_M d^N y\, \partial_\mu \delta^{(N-1)}(x-y)\delta^{AB}\delta\Phi^B(y). \tag{2.2.8}$$

This allows us to read off

$$\frac{\delta}{\delta\Phi^B(y)}\partial_\mu\Phi^A(x) = \partial_\mu\delta(x-y)\delta^{AB} = \partial_\mu\frac{\delta}{\delta\Phi^B(y)}\Phi^A(x). \tag{2.2.9}$$

We have hence shown that the partial and functional derivatives commute.

Now that we know how the functional derivative acts on the fields and their derivatives, we can also consider how it acts on the Lagrangian itself. For this, we need to further impose a product and chain rule on it, which leads us to the following:

$$\begin{aligned}
\frac{\delta}{\delta\Phi^B(y)}\mathcal{L}\big(\Phi^A(x), \partial_\mu\Phi^A(x)\big) &= \frac{\partial\mathcal{L}}{\partial\Phi^A}\frac{\delta\Phi^A(x)}{\delta\Phi^B(y)} + \frac{\partial\mathcal{L}}{\partial\big(\partial_\mu\Phi^A\big)}\frac{\delta\big(\partial_\mu\Phi^A\big)(x)}{\delta\Phi^B(y)} \\
&= \frac{\partial\mathcal{L}}{\partial\Phi^A}\delta^{AB}\delta^{(N-1)}(x-y) + \frac{\partial\mathcal{L}}{\partial\big(\partial_\mu\Phi^A\big)}\delta^{AB}\partial_\mu\delta^{(N-1)}(x-y) \\
&= \left[\frac{\partial\mathcal{L}}{\partial\Phi^A} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\big(\partial_\mu\Phi^A\big)}\right](x)\,\delta^{AB}\delta^{(N-1)}(x-y).
\end{aligned} \tag{2.2.10}$$

This is consistent with

$$\begin{aligned}
\frac{\delta S}{\delta\Phi^B(y)} &= \int_M d^N x \frac{\delta}{\delta\Phi^B(y)}\mathcal{L}\big(\Phi^A(x), \partial_\mu\Phi^A(x)\big) \\
&= \int_M d^N x \left[\frac{\partial\mathcal{L}}{\partial\Phi^A} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\big(\partial_\mu\Phi^A\big)}\right](x)\,\delta^{AB}\delta^{(N-1)}(x-y) \\
&= \left[\frac{\partial\mathcal{L}}{\partial\Phi^B} - \partial_\mu\frac{\partial\mathcal{L}}{\partial\big(\partial_\mu\Phi^B\big)}\right](y).
\end{aligned} \tag{2.2.11}$$

So, we now know how the functional derivatives acts on functionals as well as (local) functions of the fields, which can be viewed as functionals as well (in appropriate distribution spaces).

All of this has been very informal, and I'd need to take a proper course on calculus of variations to provide a better insight. But this was more of an interlude so it doesn't really matter.

2.3 Legendre Transformation and the Hamiltonian (Density)

The Legendre transformation of a field theory Lagrangian proceeds very similarly to how it works in mechanics. We start from a Lagrangian density,

$$\mathcal{L} = \mathcal{L}\big(\Phi^A(x), \partial_\mu \Phi^A(x)\big), \tag{2.3.1}$$

and define conjugate momenta as

$$\Pi_A(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi^A)}\big(\Phi^A(x), \partial_\mu \Phi^A(x)\big). \tag{2.3.2}$$

Here, $\partial_0 = \frac{\partial}{\partial x^0}$ denotes the derivative with respect to one chosen coordinate x^0 (not necessarily timelike), such that the full set of coordinates is given by $x^\mu = (x^0, x^i)$. In mechanics, we never had to choose a parameter/direction for the Legendre transform, since the parameter space is one-dimensional, just a single parameter τ, so there's no freedom in choosing a Legendre direction. As we will see later on, the Hamiltonian that comes out of the Legendre transform allows one to evolve arbitrary observables along x^0. Because of this, typically, one chooses x^0 as a timelike coordinate (e.g. the parameter t from a chosen foliation of a spacetime Σ_t, or a time coordinate in cartesian Minkowski-space coordinates), which then yields time-evolution equations. One could, however—just as well—evolve along a spacelike or even null direction. Time-evolution is particularly useful in practice, as it directly encodes causality and allows for the imposition of initial conditions. Before moving on, let us briefly remark that this choice of coordinate to take the Legendre transformation with respect to explicitly breaks Lorentz- or diffeomorphism invariance (depending on context). This is a tradeoff one has to make to get access to evolution equations.

The equation Equation (2.3.2) give a relationship between the field momenta $\Pi_A(x)$ and the velocities $\partial_0 \Phi^A(x)$. To make the Legendre transform invertible, this relationship must be invertible too, requiring that

$$\det\left(\frac{\partial^2 \mathcal{L}}{\partial(\partial_0 \Phi^A)\partial(\partial_0 \Phi^B)}\right) \neq 0 \quad \forall x \in M. \tag{2.3.3}$$

This is equivalent to the variable transformation (2.3.2) having nonsingular Jacobian everywhere, which in turn is equivalent to $\Pi_A(x)$ and $\partial_0 \Phi^A(x)$ having a bijective relationship. The Legendre transform of the Lagrangian (read: the Hamiltonian (density)) takes the form

$$\mathcal{H}\big(\Phi^A(x), \partial_i \Phi^A(x), \Pi_A(x)\big) = \Pi_A(x)\partial_0 \Phi^A(x) - \mathcal{L}\big(\Phi^A(x), \partial_\mu \Phi^A(x)\big), \tag{2.3.4}$$

or more abridgedly,

$$\mathcal{H}\big(\Phi^A, \partial_i \Phi^A, \Pi_A\big) = \Pi_A \partial_0 \Phi^A - \mathcal{L}\big(\Phi^A, \partial_\mu \Phi^A\big). \tag{2.3.5}$$

Notably, the derivatives $\partial_i \Phi^A$ still appear as arguments and are not replaced by momenta. This reflects the fact that this canonical transformation is only partial, i.e. only in the x^0-direction, and not in all spacetime directions. If one decides to Legendre-transform with respect to *all* directions, then one arrives at the *de Donder-Weyl* covariant Legendre transform.

2.4 Equations of Motion from Lagrangian Field Theory

Now that we have a Hamiltonian density, let us derive the equations of motion for it. This proceeds in a manner similar to the calculation in mechanics. We begin with the Euler-Lagrange equations and substitute the expression for the momentum as well as the Lagrangian. More concretely,

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\big(\partial_\mu \Phi^A\big)} = \frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_0 \Pi_A(x) - \partial_i \frac{\partial \mathcal{L}}{\partial(\partial_i \Phi^A(x))} \\ &= -\frac{\partial \mathcal{H}}{\partial \Phi^A} - \partial_0 \Pi_A(x) + \partial_i \frac{\partial \mathcal{H}}{\partial(\partial_i \Phi^A)} \end{aligned} \tag{2.4.1}$$

which implies that

$$\partial_0 \Pi_A(x) = -\left[\frac{\partial \mathcal{H}}{\partial \Phi^A} - \partial_i \frac{\partial \mathcal{H}}{\partial(\partial_i \Phi^A)}\right]. \tag{2.4.2}$$

To find the equations for Φ^A, we simply take the derivative of Equation (2.3.5) with respect to Π_A, which yields

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial \Pi_A} &= \partial_0 \Phi^A + \underbrace{\Pi_B \frac{\partial(\partial_0 \Phi^B)}{\partial \Pi_A} - \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi^B)} \frac{\partial(\partial_0 \Phi^B)}{\partial \Pi_A}}_{=0} \\ &= \partial_0 \Phi^A. \end{aligned} \tag{2.4.3}$$

In summary, we have found the canonical equations for the fields and their momenta,

$$\partial_0 \Phi^A = \frac{\partial \mathcal{H}}{\partial \Pi_A}, \qquad \partial_0 \Pi_A = -\left[\frac{\partial \mathcal{H}}{\partial \Phi^A} - \partial_i \frac{\partial \mathcal{H}}{\partial(\partial_i \Phi^A)}\right]. \tag{2.4.4}$$

This is a good time to introduce the Hamiltonian (not density), which is defined as

$$H(x^0) = \int_{\Sigma_{x^0}} d^{N-1}x \, \mathcal{H}\big(\Phi^A(x), \partial_i \Phi^A(x), \Pi_A(x)\big). \tag{2.4.5}$$

Here, Σ_{x^0} denotes the family of spacelike hypersurfaces with ∂_{x^0} as a normal vector (a foliation, probably somewhat wrong the way I put it but we will get to that in Cambridge I suppose). In terms of this, we can write the equations above as

$$\partial_0 \Phi^A(x) = \frac{\delta H}{\delta \Pi_A(x)}, \quad \partial_0 \Pi_A = -\frac{\delta H}{\delta \Phi^A}, \tag{2.4.6}$$

which in this form are much more analogous to those we found in mechanics (cf. Equation (1.2.7))

2.5 Hamiltonian Mechanics as a 1D Field Theory

Before moving on to Poisson brackets in field theories, we should review the connection between Hamiltonian mechanics and Hamiltonian field theory. This is because there is an important observation to make: Hamiltonian mechanics is a special case of Hamiltonian field theory, where the parameter space is one-dimensional—with τ as the only parameter.

This has a couple of consequences, one of which we already touched upon briefly when computing the Legendre transform for the Lagrangian of a field theory. If there is only one parameter, there is only one choice of what direction to perform the transform with respect to. Another important observation is that of initial conditions. The phase space is essentially the space of solutions. It is parameterised by initial conditions (under sufficient assumptions that they define unique solutions to the equations of motion), meaning that the dimension of the phase space is the dimension of the space of allowed initial conditions. Initial conditions are specified on a surface of codimension 1, i.e. of one dimension less than the parameter space. In the case of mechanics, with one parameter τ, this means that initial conditions are prescribed on dimension-0 surfaces, also known as *points*. This makes the space of initial conditions and thus phase space *finite-dimensional*. As soon as we have a proper field theory, with ≥ 2 parameters or coordinates, the dimension of the initial condition surface is ≥ 1. The space of, say, L^2 functions on such a surface is infinite-dimensional, making the phase space infinite-dimensional as well.

With this perspective in mind, we are ready to extend familiar concepts like Poisson brackets to the field-theoretic setting, where they retain their role as generators of time evolution—albeit now in infinite-dimensional phase space.

2.6 Poisson Brackets

Recall that in mechanics, for two observables $f(q,p)$ and $g(q,p)$, the Poisson bracket is defined as

$$\{f,g\} = \frac{\partial f}{\partial q^\mu}\frac{\partial g}{\partial p_\mu} - \frac{\partial g}{\partial q^\mu}\frac{\partial f}{\partial p_\mu}. \tag{2.6.1}$$

To generalise this to field theory, we should first identify which structural properties are essential and must be preserved:

- **Antisymmetry and bilinearity**: The Poisson bracket is an antisymmetric bilinear operation on the space of observables. This algebraic structure should carry over to the field-theoretic setting.
- **Dependence on canonical variables**: The bracket encodes how observables vary with respect to the canonical variables q^μ and p_μ. In field theory, the canonical variables are the fields $\Phi^A(x)$ and their conjugate momenta $\Pi_A(x)$, so the bracket must similarly reflect how observables depend on these.
- **Sensitivity to local structure**: Observables in field theory are typically functionals that may depend on the fields, their momenta, and possibly spatial derivatives such as $\partial_i\Phi^A$. To capture this dependence, the role of partial derivatives in mechanics should be played by functional derivatives in the generalisation.

Putting all of this together, and sprinkling in an integral for good measure (haha get it, measure), we get to the field-theory definition of the Poisson bracket for two observable functionals $F\big[\Phi^A, \partial_i\Phi^A, \Pi_A\big]$ and $G\big[\Phi^A, \partial_i\Phi^A, \Pi_A\big]$ given by

$$\begin{aligned} F(x^0) &= \int_{\Sigma_{x^0}} d^{N-1}x\, \mathcal{F}\big(\Phi^A(x), \partial_i\Phi^A(x), \Pi_A(x)\big), \\ G(x^0) &= \int_{\Sigma_{x^0}} d^{N-1}x\, \mathcal{G}\big(\Phi^A(x), \partial_i\Phi^A(x), \Pi_A(x)\big), \end{aligned} \tag{2.6.2}$$

then the Poisson bracket reads

$$\{F,G\} = \int_{\Sigma_{x^0}} d^{N-1}x \left[\frac{\delta F}{\delta\Phi^A(x)}\frac{\delta G}{\delta\Pi_A(x)} - \frac{\delta G}{\delta\Phi^A(x)}\frac{\delta F}{\delta\Pi_A(x)}\right]. \tag{2.6.3}$$

The Poisson bracket $\{F,G\}$ is again a functional of the canonical variables with a residual dependence on x^0. The functional derivative here is to be interpreted as

$$\frac{\delta F}{\delta\Phi^A(x)} = \left[\frac{\partial\mathcal{F}}{\partial\Phi^A} - \partial_i\frac{\partial\mathcal{F}}{\partial(\partial_i\Phi^A)}\right](x) \tag{2.6.4}$$

etc. Let us now compute the fundamental Poisson brackets of the canonical variables, with the standard reinterpretation in the functional sense:

$$\begin{aligned} \Phi^A(x^0,x^i) &= \int_{\Sigma_{x^0}} d^{N-1}y\, \Phi^A(x^0,y^i)\delta^{(N-1)}(x^i-y^i), \\ \Pi_A(x^0,x^i) &= \int_{\Sigma_{x^0}} d^{N-1}y\, \Pi_A(x^0,y^i)\delta^{(N-1)}(x^i-y^i). \end{aligned} \tag{2.6.5}$$

With these, we find

$$\big\{\Phi^A, \Phi^B\big\} = \{\Pi_A, \Pi_B\} = 0 \tag{2.6.6}$$

which follows from the Π^A-independence of Φ^B and vice versa. The interesting thing happens with the remaining Poisson bracket, where we can make use of

$$\frac{\delta\Phi^A(x^0,x^i)}{\delta\Phi^B(x^0,y^i)} = \delta^A_B\delta^{(N-1)}(x^i-y^i), \quad \frac{\delta\Pi_A(x^0,x^i)}{\delta\Pi_B(x^0,y^i)} = \delta^B_A\delta^{(N-1)}(x^i-y^i), \tag{2.6.7}$$

and zero for $\delta\Phi/\delta\Pi$ and $\delta\Pi/\delta\Phi$ (this follows from eq. (2.6.4) and eq. (2.6.5)). This allows us to compute

$$\begin{aligned}\big\{\Phi^A(x^0,x^i),\Pi_B(x^0,y^i)\big\} &= \int_{\Sigma_{x^0}} d^{N-1}z \left[\frac{\delta\Phi^A(x^0,x^i)}{\delta\Phi^C(x^0,z^i)}\frac{\delta\Pi_B(x^0,y^i)}{\delta\Pi_C(x^0,z^i)} - \frac{\delta\Pi_B}{\delta\Phi^C}\frac{\delta\Phi^A}{\delta\Pi_C}\right] \\ &= \int_{\Sigma_{x^0}} d^{N-1}z\, \delta^A_C\delta^{(N-1)}(x^i-z^i)\delta^C_B\delta^{(N-1)}(y^i-z^i) \\ &= \delta^A_B\delta^{(N-1)}(x^i-y^i).\end{aligned} \tag{2.6.8}$$

This is analogous to the Poisson bracket in mechanics.

More generally, we can recover classical mechanics by considering the special case $N=1$, where $x^0=\tau$ and there are no spatial coordinates x^i. In this case, $N-1=0$, so the integrals over the spatial slice and delta functions disappear, and the Poisson bracket becomes purely algebraic. The functional derivative with respect to Φ^A and Π_A reduce to ordinary partial derivatives, since the terms involving spatial derivatives (such as the second term in equation (2.6.4)) vanish. Identifying $q^A(\tau)=\Phi^A(x^0)$ and $p_A(\tau)=\Pi_A(\tau)$, we recover the usual Poisson bracket structure of classical mechanics.

So, we would also suspect the Poisson bracket of an observable with the Hamiltonian to correspond to a time evolution. Let us investigate this now. To this end, we simply compute

$$\begin{aligned}\{F,H\} &= \int_{\Sigma_{x^0}} d^{N-1}x \left[\frac{\delta F}{\delta\Phi^A(x)}\underbrace{\frac{\delta H}{\delta\Pi_A(x)}}_{=\partial_0\Phi^A(x)} - \underbrace{\frac{\delta H}{\delta\Phi^A(x)}}_{=-\partial_0\Pi_A(x)}\frac{\delta F}{\delta\Pi_A(x)}\right] \\ &= \int_{\Sigma_{x^0}} d^{N-1}x\left[\frac{\delta F}{\delta\Phi^A(x)}\partial_0\Phi^A(x) + \frac{\delta F}{\delta\Pi_A(x)}\partial_0\Pi_A(x)\right] \\ &= \frac{dF}{dx^0}.\end{aligned} \tag{2.6.9}$$

The last step might seem a little bit abrupt, so let us go through it more carefully. Evaluating the total derivative with respect to x^0, making use of the symmetry of partial derivatives, we get

$$\begin{aligned}\frac{dF}{dx^0} &= \int_{\Sigma_{x^0}} d^{N-1}x\frac{d}{dx^0}\mathcal{F}\big(\Phi^A(x),\partial_i\Phi^A(x),\Pi_A(x),\partial_i\Pi_A(x)\big) \\ &= \int_{\Sigma_{x^0}} d^{N-1}x\left[\frac{\partial\mathcal{F}}{\partial\Phi^A}\partial_0\Phi^A + \frac{\partial\mathcal{F}}{\partial(\partial_i\Phi^A)}\partial_0\partial_i\Phi^A + \frac{\partial\mathcal{F}}{\partial\Phi^A}\partial_0\Phi^A + \frac{\partial\mathcal{F}}{\partial(\partial_i\Pi_A)}\partial_0\partial_i\Pi_A\right] \\ &= \int_{\Sigma_{x^0}} d^{N-1}x\left[\left[\frac{\partial\mathcal{F}}{\partial\Phi^A} - \partial_i\frac{\partial\mathcal{F}}{\partial(\partial_i\Phi^A)}\right]\partial_0\Phi^A + \left[\frac{\partial\mathcal{F}}{\partial\Pi_A} - \partial_i\frac{\partial\mathcal{F}}{\partial(\partial_i\Pi_A)}\right]\partial_0\Pi_A\right] \\ &= \int_{\Sigma_{x^0}} d^{N-1}x\left[\frac{\delta F}{\delta\Phi^A}\partial_0\Phi^A + \frac{\delta F}{\delta\Pi_A}\partial_0\Pi_A\right].\end{aligned} \tag{2.6.10}$$

This completes the analogy to mechanics.

3 Hamiltonian Field Theory: Examples and Constraints

In this section, we apply the formalism developed previously to specific examples of field theories. We begin with a real scalar field theory and then move on to Maxwell theory. For each case, we start from the Lagrangian, derive the canonical variables, and compute the corresponding equations of motion.

The scalar field example proceeds smoothly and serves to illustrate the mechanics of the formalism in a straightforward setting. In contrast, Maxwell theory reveals a subtlety: the Legendre transformation becomes singular due to the presence of gauge freedom. This breakdown motivates the introduction of constraints, which will extend the canonical formalism to accommodate gauge-invariant systems.

3.1 Naive Example: Real Scalar Field Theory

We begin with one of the simplest field theories: a massive real scalar field in flat Minkowski spacetime. The Lagrangian density is given by

$$\mathcal{L} = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2, \tag{3.1.1}$$

where we adopt the metric signature used by adults—mostly plus, of course ($\eta_{\mu\nu} = \text{diag}\{-1, +1, ..., +1\}$). The associated action reads

$$S = \int d^N x\, \mathcal{L}, \tag{3.1.2}$$

and its variation under $\phi \to \phi + \delta\phi$ is

$$\delta S = \int d^N x \big[\underbrace{\partial_\mu\partial^\mu}_{=\Box}\phi - m^2\phi\big]\delta\phi \tag{3.1.3}$$

From this, we extract the Euler-Lagrange equation of motion,

$$(\Box - m^2)\phi = 0. \tag{3.1.4}$$

While this covariant equation is elegant, it is second-order in time and does not make the dynamics or initial value structure immediately transparent. Moreover, it is ill-suited for direct numerical implementation. To address this, we now transition to the Hamiltonian formalism, which, though non-covariant, naturally casts the dynamics in a first-order form better adapted to time evolution.

To derive the Hamiltonian, we need to perform a Legendre transform. To this end, we derive the canonical momenta,

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial(\partial_0\phi)}(x) = -\partial^0\phi(x) = \partial_0\phi(x). \tag{3.1.5}$$

It is pretty clear already that this is an invertible relationship, but to be sure we can compute the Hessian (in this case, a 1×1 matrix) as well. This yields

$$\frac{\partial^2\mathcal{L}}{\partial(\partial_0\phi)^2} = 1 \tag{3.1.6}$$

which is clearly non-singular. Thus, the Legendre-transform of $\mathcal{L}$ (aka the Hamiltonian density) is well-defined and given by

$$\begin{aligned}\mathcal{H} = \pi\partial_0\phi - \mathcal{L} &= \pi^2 + \frac{1}{2}\eta^{\mu\nu}\partial_\mu\phi\partial_\nu\phi + \frac{1}{2}m^2\phi^2 \\ &= \frac{1}{2}\pi^2 + \frac{1}{2}\delta^{ij}\partial_i\phi\partial_j\phi + \frac{1}{2}m^2\phi^2.\end{aligned} \tag{3.1.7}$$

The Hamiltonian (non-density) is simply the integral over a spatial slice,

$$H(t) = \int_{\mathbb{R}^3} d^3x\, \mathcal{H}(x). \tag{3.1.8}$$

From it, we can derive the equations of motion for the canonical variables ϕ and π. These read

$$\begin{aligned}\partial_0\phi &= \frac{\delta H}{\delta\pi} = \frac{\partial\mathcal{H}}{\partial\pi} - \partial_i\frac{\partial\mathcal{H}}{\partial(\partial_i\pi)} = \pi, \\ \partial_0\pi &= -\frac{\delta H}{\delta\phi} = \delta^{ij}\partial_i\partial_j\phi - m^2\phi.\end{aligned} \tag{3.1.9}$$

More compactly, we can write

$$\partial_0\phi = \pi, \quad \partial_0\pi = \Delta\phi - m^2\phi \tag{3.1.10}$$

Clearly, these equations can be combined to find

$$\partial_0^2\phi = \partial_0\pi = \Delta\phi - m^2\phi, \tag{3.1.11}$$

which can be rearranged to recover the Euler-Lagrange equation

$$(\Box - m^2)\phi = 0. \tag{3.1.12}$$

This system is unconstrained and features a simple phase space: one real degree of freedom per point in space x^i, evolving in the direction x^0, described by the field $\phi(x^0, x^i)$ and its conjugate momentum $\pi(x)$. The Hamiltonian equations of motion form a well-posed Cauchy problem, making the theory suitable for both analytical and numerical treatment—initial data specified on a spatial slice determines the evolution uniquely. Moreover, this canonical structure forms the basis for canonical quantisation, where one promotes ϕ and π to operators and imposes equal-time commutation relations,

$$\left[\hat{\phi}(x^0, x^i), \hat{\pi}(x^0, y^i)\right] = i\delta^{(N-1)}(x^i - y^i), \tag{3.1.13}$$

leading to a quantised scalar field theory in the Schrödinger or Fock picture. In contrast, our next example—Maxwell theory—features gauge freedom, which renders the Legendre transform singular and demands a more careful analysis involving constraints.

3.2 First Attempt at Maxwell Theory

We now attempt to repeat the procedure from the previous section to Maxwell theory in Minkowski space. The fields involved in Maxwell theory are the components of a 1-form,

$$A = A_\mu\, dx^\mu, \tag{3.2.1}$$

which define a field strength via the exterior derivative,

$$F = dA = \frac{1}{2}F_{\mu\nu}\, dx^\mu \wedge dx^\nu, \tag{3.2.2}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{3.2.3}$$

Though it is by no means necessary to use the language of differential forms here, it gives us the advantage of making gauge invariance obliviously obvious. A gauge transformation is a transformation of A as

$$A \to A + d\Psi, \tag{3.2.4}$$

where Ψ is a 0-form (aka a scalar). On the level of the components, this reads

$$A_\mu \to A_\mu + \partial_\mu \Psi. \tag{3.2.5}$$

The reason this makes gauge invariance of F obvious is because of the exactness of the exterior derivative, $d^2 = 0$. Making use of that, we immediately deduce the field strength's gauge transformation behaviour,

$$F = dA \to d(A + d\Psi) = dA = F. \tag{3.2.6}$$

Hence, F, as well as any quantity constructed from it, is gauge invariant. In particular, we can define an action as

$$S[A] = \int F \wedge \star F = \int d^N x \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right] \tag{3.2.7}$$

where $\star F$ is the Hodge dual of F. We won't get any further into differential form language here, so we won't concern ourselves with the Hodge dual anymore. This was just a neat way to briefly introduce some very basic differential form theory from what I learned the past half year or so. Either way, from now on, we just continue with the Lagrangian density

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \tag{3.2.8}$$

To derive the Euler-Lagrange equations of motion, it is most convenient to derive the variation of the action under $A_\mu \to A_\mu + \delta A_\mu$. This yields

$$\begin{aligned} \delta S &= -\int d^N x \, \frac{1}{2} F^{\mu\nu} \delta F_{\mu\nu} = -\int d^N x \, F^{\mu\nu} \partial_{[\mu} \delta A_{\nu]} \\ &= -\int d^N x \, F^{\mu\nu} \partial_\mu \delta A_\nu = \int d^N x \left[\partial_\mu F^{\mu\nu} \right] \delta A_\nu. \end{aligned} \tag{3.2.9}$$

The requirement of stationarity, $\delta S = 0$, thus gives rise to the Euler-Lagrange equations of motion,

$$\partial_\mu F^{\mu\nu} = 0. \tag{3.2.10}$$

In the following, it will be useful to decompose the action in terms of the time-space and space-space components of the field strength (there are no time-time components due to anti-symmetry). More concretely, we expand

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = -\frac{1}{2} F_{0i} F^{0i} - \frac{1}{4} F_{ik} F^{ik}. \tag{3.2.11}$$

Notably, time-derivatives of the field A_μ only appear in the first term. This allows us to compute the conjugate momenta. For the spatial components, we find

$$\pi^i = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} = -F^{0i} = F_{0i} = \partial_0 A_i - \partial_i A_0 = E^i, \tag{3.2.12}$$

where E^i denotes the i-th component of the electric field. For the time component, π^0, we encounter something unexpected (or expected, if you've given it some thought before). Namely, following the definition, we find that

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_0)} = 0 \tag{3.2.13}$$

identically since $F_{00} = 0$ is the only component that could contain any dependence on $\partial_0 A_0$, but is zero due to antisymmetry. This means the π^0 momentum is constrained and clearly leads to a non-invertible relationship between the velocities $\partial_0 A_\mu$ and the canonical momenta π^μ. This is also clear from

$$\frac{\partial^2 \mathcal{L}}{\partial(\partial_0 A_\mu)\partial(\partial_0 A_\nu)} = \text{diag}\{0, 1, ..., 1\} = \begin{cases} 0, & \mu, \nu = 0, \\ 1, & \mu = \nu \neq 0. \end{cases} \tag{3.2.14}$$

This is a rank $N-1$ matrix which has zero determinant. This reflects the fact that the velocity $\partial_0 A_0$ does not appear in the Lagrangian, and hence its conjugate momentum vanishes identically. In conclusion, we are unable to perform a non-singular Legendre transform.

Let us now reinterpret this algebraic obstruction from a more physical point of view. From the expression $\pi^i = F_{0i} = \partial_0 A_i - \partial_i A_0$, we see that the relationship between the momenta π^i and the velocities $\partial_0 A_i$ is invertible, provided A_0 is treated as a given background field. In contrast, the time component yields $\pi^0 = 0$, since no term in the Lagrangian depends on $\partial_0 A_0$. This means that A_0 does not appear with any time derivatives in the action and hence does not appear with any time derivatives in the equations of motion either. In other words, it has no dynamics of its own.

Nonetheless, A_0 can still influence the evolution fo the spatial components A_i through its spatial derivatives $\partial_i A_0$, which enter the equations of motion via the expression for π^i. Thus, while A_0 cannot evolve independently, it acts as a source for the dynamics of the spatial sector.

This non-dynamical character allows us to remove A_0 entirely by a gauge choice. For example, if we take

$$\Psi = \int dx^0 A_0, \tag{3.2.15}$$

then under the gauge transformation $A \to A - d\Psi$, we obtain

$$A_0 \to A_0 - \partial_0 \Psi = 0. \tag{3.2.16}$$

This is often referred to as the *temporal gauge*. Of course, this is not the only way to fix the gauge redundancy; other choices such as the Coulomb gauge, $\partial_i A^i = 0$, or Lorenz gauge, $\partial_\mu A^\mu = 0$, are equally valid. In the end, they simply relate the non-dynamical component A_0 to the dynamical A_i.

From the variation of the action in (3.2.9), we can now distinguish two qualitatively different types of equation. A variation with $\delta A_0 = 0$, $\delta A_i \neq 0$, yields

$$0 = \partial_\mu F^{\mu i} = \partial_0 F^{0i} + \partial_k F^{ki} = -\partial_0 \pi^i + \partial_k F^{ki}, \tag{3.2.17}$$

which is a genuine evolution equation, since it involves a time derivative of the momentum. In contrast, a variation with $\delta A_0 \neq 0$, $\delta A_i = 0$, gives

$$0 = \partial_\mu F^{\mu 0} = \partial_k F^{k0} = \partial_k \pi^k. \tag{3.2.18}$$

This equation contains no time derivatives at all—it is a constraint rather than a dynamical law. Recognising $\pi^k = E^k$, we identify this constraint as Gauss' law.

In summary, we attempted a Legendre transformation of the Maxwell Lagrangian to obtain a Hamiltonian formulation. This revealed that the momentum π^0 is constrained to vanish identically, indicating that A_0 is a non-dynamical field. This is reflected both in the structure of the Lagrangian and in the equations of motion. In addition to the primary constraint $\pi^0 = 0$, we also uncovered a

secondary constraint $\partial_i \pi^i = 0$, corresponding to Gauss' law. Both constraints arise naturally from the underlying gauge redundancy, and their proper treatment requires a more systematic theory of constrained Hamiltonian systems, to which we now turn.

3.3 Introduction to Constraints in Mechanics

In this subsection, we introduce the concept of constraints in classical systems. We begin with mechanical systems, where they are easier to motivate and analyse, and then generalise the formalism to field theories with an N-dimensional parameter space. This general framework will then allow us to revisit the situation we encountered earlier with Maxwell theory from a broader perspective and develop the Dirac-Bergmann algorithm.

3.3.1 Constraints in Lagrangian Mechanics

Consider a mechanical system with a configuration described by coordinates $q^\mu(\tau)$, governed by a Lagrangian of the form

$$\mathcal{L} = \mathcal{L}(q, \dot{q}). \tag{3.3.1}$$

Now suppose that the system is subject to an additional condition—such as being confined to a lower-dimensional submanifold of the ambient configuration space, or having fixed components of its velocity. Such conditions are formalised as constraints of the form

$$F(q, \dot{q}) = 0, \tag{3.3.2}$$

where F is a function of the coordinates and velocities.

To make this idea concrete, consider the case of a free particle in $\mathbb{R}^2$, with coordinates $q^\mu = (x, y)$ and a free Lagrangian

$$\mathcal{L}(x, y, \dot{x}, \dot{y}) = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2. \tag{3.3.3}$$

A possible constraint could be that the motion is restricted to the unit circle, i.e.,

$$F(x, y, \dot{x}, \dot{y}) = x^2 + y^2 - 1 = 0. \tag{3.3.4}$$

This condition restricts the admissible trajectories in configuration space. A function like F, which defines such a restriction, is called a *constraint function*, and the condition itself is called a *constraint*.

While defining a constraint is straightforward—one simply specifies a function $F(q, \dot{q})$ and demands it to vanish—the more subtle question is how such a condition affects the equations of motion. This leads to the formalism of *Lagrange multipliers*, which we now develop.

Let us first try to get an intuitive picture. You may have noticed something already: above, we explicitly took a *free* Lagrangian—only to then tie our particle to a leash and drag it around the unit circle. The only truly "free" solution, in the sense of satisfying the unconstrained Euler-Lagrange equations

$$\ddot{x} = \ddot{y} = 0, \tag{3.3.5}$$

would be a stationary particle. But motion along a circle is never inertial—it is always accelerating. So the free dynamics and the constrained dynamics are clearly not the same.

The takeaway is this: *the equations of motion get modified.* Something must supply the force that keeps the particle on the constraint surface $F = 0$. These additional terms act exactly so as to cancel the parts of the dynamics that would otherwise pull the system away from $F = 0$, while leaving all motion within the constraint surface untouched.

On a superficial level, this is straightforward. If we want new terms in the equations of motion, we add new terms to the Lagrangian. That much is clear. But what is not yet clear is which terms to add. We want just enough to enforce the constraint, and nothing more—no overcorrections, no spurious effects. More formally, we look for paths $q(\tau)$ that minimise the action functional

$$S[q] = \int d\tau\, \mathcal{L}(q, \dot{q}), \tag{3.3.6}$$

while simultaneously being subject to the constraint. Informally, we can write this as

$$q(\tau) = \underset{f(q,\dot{q})=0}{\arg\min}\, S[q]. \tag{3.3.7}$$

We aren't minimising the functional S on the whole configuration space anymore, only on a restricted subset.

At this point, our intuition has taken us as far as it can. Time to be precise—and introduce the method of *Lagrange multipliers*. For this, consider the following preposterous idea: what if we could turn the constraint equation

$$F(q, \dot{q}) = 0 \tag{3.3.8}$$

into an equation of motion itself? Then it would automatically be satisfied whenever the equations of motion hold. That sounds promising—but clearly cannot work *as is*. So let us reflect on what it actually means to be an equation of motion, or just a plain ol' equation. The answer is kind of obvious from the previous section, but still worth reiterating: equations of motion arise from varying the degrees of freedom in the action and requiring

$$\delta S = 0. \tag{3.3.9}$$

This gives a collection of conditions which we call the equations of motion. Now that we reminded ourselves of what that means, we can continue on our "fantasy journey" of turning the constraint (3.3.8) into one.

This means we must somehow vary a degree of freedom in such a way that the constraint (3.3.8) pops out. Varying with respect to q, that *might* be possible in some cases—but there is a much simpler trick. Just introduce a new degree of freedom—let us call it $\lambda(\tau)$. Then consider an action of the form

$$S_C[q; \lambda] = \int d\tau\, \lambda(\tau) F(q(\tau), \dot{q}(\tau)). \tag{3.3.10}$$

Varying this degree of freedom as $\lambda \to \lambda + \delta\lambda$ gives

$$\delta S_C = \int d\tau\, F(q, \dot{q}) \delta\lambda, \tag{3.3.11}$$

and requiring this to vanish for arbitrary $\delta\lambda(\tau)$ gives back the constraint $F = 0$. Voilà—We have turned it into an equation of motion for this new action S_C.

Naturally, S_C carries none of the physics of the original system. But since S does not depend on λ, varying the total action

$$S_T = S - S_C = \int d\tau [\mathcal{L}(q, \dot{q}) - \lambda F(q, \dot{q})] \equiv \int d\tau\, \mathcal{L}_T(q, \dot{q}; \lambda). \tag{3.3.12}$$

with respect to λ gives

$$\delta S_T = 0 \iff \delta S_C = 0 \iff F(q, \dot{q}) = 0. \tag{3.3.13}$$

So the constraint is now *built into* the variational principle of S_T—just as we wanted. The minus sign here is pure convention; it will become more natural later. As a side note: the symbol $\mathcal{L}_\tau(q, \dot{q}; \lambda)$ we introduced in passing above will from now on be referred to as the *total Lagrangian*.

Clearly, modifying the action in this way affects the equations of motion for q as well. After all, we are introducing a new term that depends on q and $\dot{q}$. A straightforward calculation shows that under $q \to q + \delta q$, the variation of the total action becomes

$$\delta S_T = \int d\tau \left[\frac{\delta \mathcal{L}}{\delta q^\mu} - \lambda \frac{\delta F}{\delta q^\mu} \right] \delta q^\mu. \tag{3.3.14}$$

Writing out the full equations of motion, we find

$$\left[\frac{\partial \mathcal{L}}{\partial q^\mu} - \frac{d}{d\tau} \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} \right](\tau) = \lambda(\tau) \left[\frac{\partial F}{\partial q^\mu} - \frac{d}{d\tau} \frac{\partial F}{\partial \dot{q}^\mu} \right](\tau). \tag{3.3.15}$$

This is just the regular Euler-Lagrange expression on the left-hand side, with a correction term on the right that enforces the constraint. That correction represents the *constraint force* needed to keep the system on the surface defined by $F = 0$. Also, the conventional minus sign becomes of use here: it allows us to take the constraint force to the right-hand side without introducing additional signs.

The new degree of freedom $\lambda(\tau)$ we have introduced is not a physical coordinate like $q(\tau)$, but rather a *Lagrange multiplier*. It enforces a constraint by adjusting itself to ensure $F(q, \dot{q}) = 0$ throughout the motion. We should note here that it is not dynamical—the full action does not incorporate any of its derivatives, only λ itself.

Whether one must explicitly solve for $\lambda(\tau)$ depends on the problem. Sometimes, λ can be eliminated by combining the equations of motion with the constraints and its derivatives, effectively reducing the system to unconstrained variables. In other cases, λ carries meaningful information and must be kept as an auxiliary field to correctly describe the dynamics.

Either way, the introduction of λ provides a powerful and systematic way to incorporate constraints directly into the variational principle, allowing us to handle complex constrained systems with the familiar tools of Lagrangian and Hamiltonian mechanics.

3.3.2 Constraints in Hamiltonian Mechanics

The first sections of these notes aim to develop Hamiltonian field theory from the ground up. So far, we have laid out the foundations and encountered our first major complication: a singular relationship between velocities and momenta in Maxwell theory. This led us into this detour on constrained dynamics. However, up to this point, our discussion of constraints has remained entirely within the framework of *Lagrangian* mechanics.

In this section, we begin the translation of that framework into the Hamiltonian picture. When the relationship between velocities $\dot{q}^\mu$ and canonical momenta

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} \tag{3.3.16}$$

is non-singular, the Legendre transform can be performed directly, and the treatment of constraints carries over in a rather straightforward way. That will be the focus here.

The singular case is more delicate. There, the Legendre transform fails to be invertible, and constraints arise intrinsically from the degeneracy of the Lagrangian's Hessian with respect to the velocities. We will return to that scenario later, in the context of field theories, when we introduce the Dirac-Bergmann algorithm.

Recall from the previous section that starting from a Lagrangian $\mathcal{L}(q, \dot{q})$ and a constraint

$$F(q, \dot{q}) = 0, \tag{3.3.17}$$

we may define a constrained *total Lagrangian* by

$$\mathcal{L}_T(q, \dot{q}; \lambda) = \mathcal{L}(q, \dot{q}) - \lambda F(q, \dot{q}), \tag{3.3.18}$$

where $\lambda = \lambda(\tau)$ is a newly introduced degree of freedom enforcing the constraint. This total Lagrangian describes the dynamics of the constrained system, and may itself admit a Legendre transform, yielding the constrained Hamiltonian $\mathcal{H}_T(p, q; \lambda)$. This construction, as always, presupposes that the relation

$$p_\mu = \frac{\partial \mathcal{L}_T}{\partial \dot{q}^\mu} = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} - \lambda \frac{\partial F}{\partial \dot{q}^\mu} \tag{3.3.19}$$

between the momenta and the velocities is invertible. Since λ appears without derivatives, it does not contribute a momentum and is non-dynamical; we choose not to Legendre-transform with respect to its velocity.

Applying the standard Legendre transform yields

$$\begin{aligned} \mathcal{H}_T &= p_\mu \dot{q}^\mu - \mathcal{L}_T = \left[\frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} \dot{q}^\mu - \lambda \frac{\partial F}{\partial \dot{q}^\mu} \dot{q}^\mu \right] - [\mathcal{L} - \lambda F] \\ &= \mathcal{H} - \lambda \mathcal{F}, \end{aligned} \tag{3.3.20}$$

where $\mathcal{H}$ is the Legendre transform of the unconstrained Lagrangian, i.e. the unconstrained Hamiltonian

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu} \dot{q^\mu} - \mathcal{L}, \tag{3.3.21}$$

and $\mathcal{F}$ is the Legendre transform of the constraint function F,

$$\mathcal{F} = \frac{\partial F}{\partial \dot{q}^\mu} \dot{q}^\mu - F \tag{3.3.22}$$

As a side note (hopefully clarifying rather than confusing), we now have three momentum-like quantities in play:

- the *total* canonical momentum from $\mathcal{L}_T$,

$$p_\mu = \frac{\partial \mathcal{L}_T}{\partial \dot{q}^\mu}, \tag{3.3.23}$$

- the *unconstrained* momentum, arising from $\mathcal{L}$ alone,

$$p_\mu^{\mathrm{UC}} = \frac{\partial \mathcal{L}}{\partial \dot{q}^\mu}, \tag{3.3.24}$$

- and the constraint momentum, defined by

$$p_\mu^C = \frac{\partial F}{\partial \dot{q}^\mu}. \tag{3.3.25}$$

From equation (3.3.19), we find that they are related via

$$p_\mu = p_\mu^{\mathrm{UC}} - \lambda p_\mu^C. \tag{3.3.26}$$

This matters because the Legendre transforms $\mathcal{L} \to \mathcal{H}$ and $F \to \mathcal{F}$ are constructed with respect to p^{UC} and p^C, respectively. By the above relationship between the momenta, this causes the full Legendre transform $\mathcal{L}_T \to \mathcal{H}_T$ to be with respect to the total momentum p_μ, as required.

3.3.3 Example: Constrained Dynamics in $\mathbb{R}^2$

In this section, we apply our theory of Lagrange multipliers to the free action

$$S[q] = \int d\tau \left[\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2\right] \tag{3.3.27}$$

with $q^\mu = (x, y) \in \mathbb{R}^2$, subject to the constraint

$$0 = F(x, y, \dot{x}, \dot{y}) = x^2 + y^2 - 1. \tag{3.3.28}$$

We will first write down the total Lagrangian and its associated equations of motion, then compute the total Hamiltonian and derive its canonical equations. After that, we will analyse these results in detail—the constraint force will turn out to have a particularly intuitive interpretation. Finally, we will solve the equations explicitly, obtaining both the particle's trajectory $q(\tau)$ and the corresponding Lagrange multiplier $\lambda(\tau)$.

We begin with the total action, whose integrand is the total Lagrangian. From the general formalism developed earlier, we know that it is given by

$$\begin{aligned} S_T[q;\lambda] &= \int d\tau [\mathcal{L}(q,\dot{q}) - \lambda F(q,\dot{q})] \\ &= \int d\tau \left[\frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 - \lambda(x^2+y^2-1)\right]. \end{aligned} \tag{3.3.29}$$

The total Lagrangian is therefore

$$\mathcal{L}_T(q,\dot{q};\lambda) = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 - \lambda(x^2+y^2-1). \tag{3.3.30}$$

Under a variation $x \to x + \delta x$, the action S_T changes as

$$\delta S_T = \int d\tau [\dot{x}\delta\dot{x} - 2\lambda x\delta x] = -\int d\tau [\ddot{x} + 2\lambda x]\delta x, \tag{3.3.31}$$

where we have integrated by parts and discarded boundary terms. By symmetry under $x \leftrightarrow y$, the full Euler-Lagrange equations of motion are

$$\ddot{x} = -2\lambda x, \qquad \ddot{y} = -2\lambda y. \tag{3.3.32}$$

Before analysing these equations, let us also compute the Hamiltonian and the associated canonical equations. The canonical momenta, derived from $\mathcal{L}_T$, are

$$p_x = \frac{\partial \mathcal{L}_T}{\partial \dot{x}} = \dot{x}, \qquad p_y = \frac{\partial \mathcal{L}_T}{\partial \dot{y}} = \dot{y} \tag{3.3.33}$$

which is clearly an invertible relationship. Consequently, the total Hamiltonian is

$$\begin{aligned} \mathcal{H}_T(p,q;\lambda) &= p_\mu \dot{q}^\mu - \mathcal{L}_T(q,\dot{q}(p);\lambda) \\ &= p_x^2 + p_y^2 - \left[\frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 - \lambda(x^2+y^2-1)\right] \\ &= \frac{1}{2}p_x^2 + \frac{1}{2}p_y^2 + \lambda(x^2+y^2-1). \end{aligned} \tag{3.3.34}$$

The associated canonical equations are

$$\begin{aligned} \dot{x} &= \frac{\partial \mathcal{H}}{\partial p_x} = p_x, & \dot{y} &= \frac{\partial \mathcal{H}}{\partial p_y} = p_y, \\ \dot{p}_x &= -\frac{\partial \mathcal{H}}{\partial x} = -2\lambda x, & \dot{p}_y &= -\frac{\partial \mathcal{H}}{\partial y} = -2\lambda y. \end{aligned} \tag{3.3.35}$$

These are clearly consistent with the Euler-Lagrange equations (3.3.32) derived from the Lagrangian formulation.

Let us now examine the physical content of the equations of motion, particularly the nature of the constraint force. Using vector notation, we may rewrite the Euler-Lagrange equations compactly as

$$\ddot{q}^\mu = -2\lambda q^\mu. \tag{3.3.36}$$

This shows that the acceleration is proportional to the position vector, with proportionality factor -2λ. Thus, the force is directed radially—towards the origin if $\lambda > 0$, or away from it if $\lambda < 0$ (though only the former is consistent with circular motion on the constraint surface).

This behaviour aligns with the geometry of the problem: the constraint surface is the unit circle in configuration space, or more precisely, a submanifold of $\mathbb{R}^2 \times \mathbb{R}^2$, parameterised by $(q, \dot{q})$, defined by the condition $x^2 + y^2 = 1$. In order to remain on that circle, the particle must experience a force that points normal to the surface at all times. The normal vector to the circle at any point is proportional to the position vector itself, which is exactly the direction of the constraint force appearing on the right-hand side of equation (3.3.36).

The magnitude of this foce is determined by the Lagrange multiplier λ. Physically, λ must take a value that ensures the centripetal acceleration precisely matches the kinematics of the motion. If λ is too small, the particle spirals outward; if too large, it collapses inward. It is thus intuitively clear that only the correct value of $\lambda(\tau)$ maintains the particle on a stable circular path.

We now turn to solving the equations explicitly and determining the form of $\lambda(\tau)$. We consider three different aspects: the purely theoretical existence-of-solutions, a geometric derivation of λ which will allow the easy solution of the full equations of motion, and lastly a shortcut in which we explicitly parameterise the constraint surface to reduce the system to a simpler one where the constraints are implemented automatically.

1. Structurally, equation (3.3.36) resembles a 2D time-independent Schrödinger equation with potential $V(\tau) = 2\lambda(\tau)$. As a differential equation, solutions always exist for arbitrary $\lambda(\tau)$, but these trajectories may not satisfy the constraint $F = 0$. Only if $\lambda(\tau)$ is determined as the correct Lagrange multiplier does a given solution remain confined to the circle. So while equation (3.3.36) admits solutions for any function λ, only for the physically meaningful one the system is actually constained.

2. As a second point of view, we can derive solutions of equation (3.3.36) that actually satisfy the constraint by first computing $\lambda(\tau)$. The Lagrange multiplier turns out to be very simple for this system: a non-negative constant. Intuitively this makes sense—after all, we are restricting a free particle to a circle. This means that there should be no change in velocity *along* the circle, meaning that the circular motion has constant angular velocity and hence the centripetal force must have a constant magnitude as well. The Lagrange multiplier λ constitutes that magnitude, and hence, it should be constant. Moreover, it cannot be negative, as otherwise the particle would be accelerated away from the origin, which cannot lead to bound circular motion.

   Though these are compelling arguments, let us make this mathematically rigorous as well. First of all, we note that we can write the constraint $F = 0$ as

$$q \cdot q = |q|^2 = 1. \tag{3.3.37}$$

Differentiating once with respect to τ gives us the orthogonality relation

$$\dot{q} \cdot q = 0, \tag{3.3.38}$$

and a second differentiation leads to

$$\ddot{q} \cdot q + |\dot{q}|^2 = 0. \tag{3.3.39}$$

Inserting the equation of motion (3.3.36) into this further leads to

$$-2\lambda \underbrace{|q|^2}_{=1} + |\dot{q}|^2 \quad \Longleftrightarrow \quad \lambda = \frac{1}{2}|\dot{q}|^2. \tag{3.3.40}$$

This is already an explicit expression for λ. Without much more effort, we can even show that this is constant—let us do this as well. We simply differentiate λ to find

$$\dot{\lambda} = \frac{d}{d\tau}\frac{1}{2}|\dot{q}|^2 = \ddot{q} \cdot \dot{q} = -2\lambda \underbrace{q \cdot \dot{q}}_{=0} = 0. \tag{3.3.41}$$

This tells us that $\lambda = \text{const.}$, drastically simplifying the remaining equation (3.3.36) into one solvable in closed form. It is also easy to see now that $\lambda < 0$ leads to exponential/hyperbolic solutions, which cannot be constrained—we must hence have $\lambda \geq 0$. Note here that, however, λ can still not be chosen freely. It depends on the initial velocity of the particle, the magnitude of which is preserved in the circular motion, as we will see momentarily. Nonetheless, for $\lambda > 0$ constant we can write down the solutions as

$$q(\tau) = q_0 \cos\sqrt{\lambda}\tau + \frac{1}{\sqrt{\lambda}}\dot{q}_0 \sin\sqrt{\lambda}\tau, \tag{3.3.42}$$

where q_0 and $\dot{q}_0$ are the values of $q(\tau)$ and $\dot{q}(\tau)$ at $\tau = 0$, respectively. By our derivations above, they must satisfy $\dot{q}_0 \cdot q_0 = 0$. Moreover, to satisfy the constraint, it needs to hold that

$$1 = |q(\tau)|^2 = |q_0|^2 \cos^2\sqrt{\lambda}\tau + \frac{|\dot{q}_0|^2}{\lambda}\sin^2\sqrt{\lambda}\tau, \tag{3.3.43}$$

which can only be valid if $|q_0|^2 = 1$ and $\lambda = |\dot{q}_0|^2$. We have thus completely solved the constrained dynamics, arriving at the solution

$$q(\tau) = q_0 \cos(|\dot{q}_0|\tau) + \frac{\dot{q}_0}{|\dot{q}_0|}\sin(|\dot{q}_0|\tau). \tag{3.3.44}$$

Lastly, for $\lambda = 0$, we find the constant solutions $q(\tau) = q_0 \in S^1 \subset \mathbb{R}^2$.

3. Another way to approach the problem is to move to new dynamical variables that automatically encode the constraint. This is not always as straightforward, particularly if F involves velocities, but nonetheless is a powerful trick to keep in mind. In our case, where the constraint is

$$x^2 + y^2 = 1, \tag{3.3.45}$$

we can reduce the system to a single dynamical variable $\varphi(\tau)$, related to $x(\tau)$ and $y(\tau)$ via

$$x(\tau) = \cos\varphi(\tau), \quad y(\tau) = \sin(\tau). \tag{3.3.46}$$

By the trigonometric identity $\sin^2 + \cos^2 = 1$, this reparametrisation always satisfies the constraint, and the system formulated in terms of φ becomes unconstrained. In particular, we have

$$\dot{x} = -\dot{\varphi}\sin\varphi, \quad \dot{y} = \dot{\varphi}\cos\varphi, \tag{3.3.47}$$

implying that the unconstrained free Lagrangian becomes

$$\mathcal{L} = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 = \frac{1}{2}\dot{\varphi}^2 \sin^2\varphi + \frac{1}{2}\dot{\varphi}^2 \cos^2\varphi = \frac{1}{2}\dot{\varphi}^2. \tag{3.3.48}$$

This is the simplest non-trivial system one can encounter: A free particle in one dimension, with equation of motion

$$\ddot{\varphi} = 0 \tag{3.3.49}$$

and solutions

$$\varphi(\tau) = \dot{\varphi}_0 \tau + \varphi_0. \tag{3.3.50}$$

In terms of the variables $x(\tau)$ and $y(\tau)$, the solutions read

$$x(\tau) = \cos(\dot{\varphi}_0 \tau + \varphi_0), \qquad y(\tau) = \sin(\dot{\varphi}_0 \tau + \varphi_0), \tag{3.3.51}$$

which are precisely the solutions found earlier in equation (3.3.44).

3.4 Constraints in Field Theories

3.4.1 Generalising Mechanical Constraints to Field Theories

In the previous section, we constructed the mathematical tool of Lagrange multipliers to deal with constraints in mechanical systems. In this section, we generalise these results to field theories, again starting from the Lagrangian picture and transitioning to the Hamiltonian formalism for non-singular Lagrangians. The case of singular Lagrangians will be treated separately in the section after the next.

Recall that in field theory, the spacetime Manifold M serves as the parameter space for the dynamical variables. It is locally coordinatised by x^μ, and the fields (i.e. dynamical variables) are functions

$$\Phi^A(x) \in C^\infty(M). \tag{3.4.1}$$

The Lagrangian density $\mathcal{L}$ is a local function of the fields and their derivatives,

$$\mathcal{L} = \mathcal{L}\big(\Phi^A(x), \partial_\mu \Phi^A(x)\big). \tag{3.4.2}$$

The corresponding unconstrained action is given by

$$S\big[\Phi^A\big] = \int_M d^N x\, \mathcal{L}\big(\Phi^A, \partial_\mu \Phi^A\big). \tag{3.4.3}$$

Constraints generalise naturally in this setting: as in mechanics, they take the form of equations that must be satisfied on shell, but now depend on both the fields and their derivatives,

$$F\big(\Phi^A(x), \partial_\mu \Phi^A(x)\big) = 0, \tag{3.4.4}$$

for F a function of the fields and their derivatives.

To enforce this constraint, we follow the same procedure as in mechanics: introduce a new, non-dynamical degree of freedom $\lambda(x)$ acting as a Lagrange multiplier. Define the total action as

$$S_T\big[\Phi^A\big] = \int_M d^N x \left[\mathcal{L}\big(\Phi^A, \partial_\mu \Phi^A\big) - \lambda F\big(\Phi^A, \partial_\mu \Phi^A\big)\right]. \tag{3.4.5}$$

This promotes the constraint (3.4.4) to an equation of motion—varying $\lambda \to \lambda + \delta\lambda$ arbitrarily, while keeping the fields Φ^A fixed, leads to

$$0 \stackrel{!}{=} \delta S_T = \int_M d^N x \, F \delta\lambda, \tag{3.4.6}$$

which requires

$$F = 0 \tag{3.4.7}$$

must hold identically.

As in the mechanical case, the equations of motion for the fields Φ^A are also modified. Under a variation with respect to the fields, $\Phi^A \to \Phi^A + \delta\Phi^A$, the total action varies as

$$0 \stackrel{!}{=} \delta S_T = \int_M d^N x \left[\frac{\delta \mathcal{L}}{\delta \Phi^A} - \lambda \frac{\delta F}{\delta \Phi^A} \right]. \tag{3.4.8}$$

We can read off the modified equations of motion to be

$$\left[\frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\big(\partial_\mu \Phi^A\big)} \right](x) = \lambda(x) \left[\frac{\partial F}{\partial \Phi^A} - \partial_\mu \frac{\partial F}{\partial\big(\partial_\mu \Phi^A\big)} \right](x). \tag{3.4.9}$$

As before, the left-hand side gives the standard field equations for the unconstrained system, while the right-hand side encodes the "normal direction" to the constraint surface[2], scaled pointwise by the multiplier $\lambda(x)$ to ensure the appropriate magnitude of the constraint force.

Of course, the Lagrangian picture can also be translated to the Hamiltonian setting. We pursue this for non-singular Lagrangians here, and will return to the singular case—which introduces additional constraints—in a later section. This translation casts the constrained dynamics in terms of phase space variables and prepares the ground for a systematic treatment of constraints, where their classification and algebraic structure will become central in the Dirac-Bergmann formalism.

Suppose that the total Lagrangian has a non-singular Hessian with respect to its velocities,

$$\det\left(\frac{\partial^2 \mathcal{L}_T}{\partial(\partial_0 \Phi^A)\partial(\partial_0 \Phi^B)} \right) \neq 0. \tag{3.4.10}$$

This gives an inveritble relationship between the velocities and the momenta,

$$\Pi_A = \frac{\partial \mathcal{L}_T}{\partial(\partial_0 \Phi^A)} = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi^A)} - \lambda \frac{\partial F}{\partial(\partial_0 \Phi^A)}. \tag{3.4.11}$$

The Hamiltonian density associated with $\mathcal{L}_T$ is hence

$$\begin{aligned} \mathcal{H}_T &= \Pi_A \partial_0 \Phi^A - \mathcal{L}_T \\ &= \left[\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi^A)} \partial_0 \Phi^A - \mathcal{L} \right] - \lambda \left[\frac{\partial F}{\partial(\partial_0 \Phi^A)} \partial_0 \Phi^A - F \right] \\ &= \mathcal{H} - \lambda \mathcal{F}, \end{aligned} \tag{3.4.12}$$

where $\mathcal{H}\big(\Phi^A, \Pi_A\big)$ is the unconstrained Hamiltonian associated with $\mathcal{L}$ and $\mathcal{F}\big(\Phi^A, \Pi_A\big)$ is the Legendre transform of the constraint function F.

As a last step, we can compute the canonical equations of motion of $\mathcal{H}_T$ as

[2]More precisely, this is variation orthogonal to the surface of admissible configurations

$$\begin{aligned}\partial_0\Phi^A &= \frac{\delta\mathcal{H}_T}{\delta\Pi_A} = \frac{\delta\mathcal{H}}{\delta\Pi_A} - \lambda\frac{\delta\mathcal{F}}{\delta\Pi_A},\\ \partial_0\Pi_A &= -\frac{\delta\mathcal{H}_T}{\delta\Phi^A} = -\left[\frac{\delta\mathcal{H}}{\delta\Phi^A} - \lambda\frac{\delta\mathcal{F}}{\delta\Phi^A}\right].\end{aligned} \tag{3.4.13}$$

These equations extend the unconstrained canonical equations (2.4.6) by terms involving the transformed constraint function $\mathcal{F}$.

3.4.2 Example: Constrained Biscalar Field Theory

3.4.2.1 Lagrangian Picture

Now that we have developed the general formalism for constraints in field theory, let us consider a concrete example. The natural field-theoretic analogue of the mechanical system studied in Section 3.3.3 is a theory of two free real scalar fields, φ_1 and φ_2, in Minkowski spacetime, subject to a constraint. The unconstrained free Lagrangian density is

$$\mathcal{L}(\varphi_i, \partial_\mu\varphi_i) = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi_1\partial_\nu\varphi_1 - \frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi_2\partial_\nu\varphi_2, \tag{3.4.14}$$

where we omit the inclusion of mass terms for simplicity. As a constraint, we impose the field-theoretic analogue of motion restricted to a circle,

$$0 = F(\varphi_1, \varphi_2) = \varphi_1^2 + \varphi_2^2 - 1. \tag{3.4.15}$$

This is not meant to be physically deep (though one could view it as a complex scalar field constrained to take values on the unit circle), but it serves as a useful toy model to explore constrained dynamics in a field theory.

To enforce the constraint, we introduce a Lagrange multiplier field $\lambda(x)$ to promote $F = 0$ to an equation of motion. The total Lagrangian becomes

$$\mathcal{L}_T(\varphi_i, \partial_\mu\varphi_i; \lambda) = -\frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi_1\partial_\nu\varphi_1 - \frac{1}{2}\eta^{\mu\nu}\partial_\mu\varphi_2\partial_\nu\varphi_2 - \lambda(\varphi_1^2 + \varphi_2^2 - 1). \tag{3.4.16}$$

Varying the total action with respect to φ_1 and φ_2, we obtain the modified Euler-Lagrange equations,

$$\Box\varphi_i = 2\lambda\varphi_i. \tag{3.4.17}$$

On the lef-hand side, we recover the standard wave operator; the right-hand side encodes the constraint force. In this case, it acts as a restoring force, normal to the constraint surface $\varphi_1^2 + \varphi_2^2 = 1$ at each spacetime point.

This constraint surface is an infinite-dimensional generalisation of a circle: a submanifold of the field configuration space, formally something like

$$\{(\varphi_1, \varphi_2) \in L^2(\mathbb{R}^{N-1}) \times L^2(\mathbb{R}^{N-1}) \mid \varphi_1(x)^2 + \varphi_2(x)^2 = 1 \text{ for all } x \in \mathbb{R}^{N-1}\}, \tag{3.4.18}$$

or a more regularised version admitting weak derivatives.

There are two natural perspectives on what the equations of motion (3.4.17) tell us:

1. **Geometric Constraint Force**: The term $-2\lambda\varphi_i$ is the analogue of a centripetal force: it keeps the field values $(\varphi_1(x), \varphi_2(x))$ on the unit circle at each spacetime point. At any x, it points in the "normal direction" to the constraint surface and adjusts its magnitude via $\lambda(x)$ to maintain the constraint dynamically. This is the direct, though abstract, geometric interpretation, mirroring the mechanical case.

2. **Effective mass interpretation**: Alternatively, we can rearrange the equation as

$$(\Box - m^2(x))\varphi_i = 0, \quad \text{with} \quad m^2(x) = 2\lambda(x). \tag{3.4.19}$$

This resembles a Klein-Gordon equation with a spacetime-dependent mass. While a bit abstract to visualise, this perspective highlights how the constraint enforces field behaviour by adjusting the local "mass" to keep the fields bounded to the constraint surface. Here, $\lambda(x)$ plays the role of a dynamically adjusting mass profile.

3.4.2.2 Hamiltonian Picture

Let us now switch to the Hamiltonian picture, as it illustrates a couple of things nicely which we will see again in the Dirac-Bergmann algorithm.

To begin, we compute the conjugate momenta

$$\pi_i = \frac{\partial \mathcal{L}_T}{\partial(\partial_0 \varphi_i)} = -\eta^{0\mu}\partial_\mu \varphi_i = \partial_0 \varphi_i. \tag{3.4.20}$$

Using these, the total Hamiltonian density can be determined as

$$\begin{aligned} \mathcal{H}_T(\varphi_i, \pi_i, \partial_k \varphi_i) &= \pi_i \partial_0 \varphi_i - \mathcal{L}_T\big(\varphi_i, \partial_\mu \varphi_i\big) \\ &= \sum_{i=1,2} \left[\frac{1}{2}(\pi_i)^2 + \frac{1}{2}(\nabla \varphi_i)^2\right] + \lambda(\varphi_1^2 + \varphi_2^2 - 1) \end{aligned} \tag{3.4.21}$$

With the Hamiltonian H_T defined as

$$H_T = \int d^3x\, \mathcal{H}_T(\varphi_i(x), \pi_i(x), \nabla\varphi_i(x)) \tag{3.4.22}$$

this leads to the canonical equations of motion

$$\partial_0 \varphi_i = \frac{\delta H_T}{\delta \pi_i} = \pi_i, \quad \partial_0 \pi_i = -\frac{\delta H_T}{\delta \varphi_i} = \Delta\varphi_i - 2\lambda\varphi_i \tag{3.4.23}$$

By inserting the former equation into the latter, it is easily verified that the second-order Euler-Lagrange equation (3.4.17) is recovered. These equations describe the evolution of a field configuration through phase space, i.e. the space of functions (φ_i, π_i) defined over a spatial hypersurface at fixed x^0. This space may be modelled as the product $L^2(\mathbb{R}^{N-1}) \times L^2(\mathbb{R}^{N-1})$, provided we choose appropriate boundary conditions. Put differently, the canonical equations are evolution equations in the vector space $L^2(\mathbb{R}^{N-1}) \times L^2(\mathbb{R}^{N-1})$.

Similarly, after transforming to the Hamiltonian picture, the constraints now also live in phase space—in the previous sections, where we developed the general theory, we saw that the constraint function F also undergoes a Legendre transformation,

$$F\big(\varphi_i, \partial_\mu \varphi_i\big) \to \mathcal{F}(\varphi_i, \pi_i, \partial_k \varphi_i) = \frac{\partial F}{\partial(\partial_0 \varphi_i)}\partial_0 \varphi_i - F. \tag{3.4.24}$$

The general form reflects that constraints, in the Hamiltonian formalism, are functions on phase space involving φ, π and possibly spatial derivatives thereof. In our case, the constraint is purely positional and so the Legendre transform is trivial, $\mathcal{F} = -F$.

With the upcoming Dirac-Bergmann algorithm in mind, let us investigate second-class constraints. Suppose we know that the primary constraint $\mathcal{F} = 0$ is satisfied at some initial time x_0^0. This is nothing outlandish—if we seek solutoins that respect the constraint, the initial data must lie on the constraint surface. To remain on that surface, the constraint must be preserved in time. That is, we must demand

$$\partial_0 \mathcal{F} = 0 \tag{3.4.25}$$

for all configurations such that $\mathcal{F} = 0$. In general, this condition corresponds to the requirement that the Poisson bracket $\{\mathcal{F}, \mathcal{H}_T\}$ vanishes, but here we can compute the time derivative directly as

$$0 \overset{!}{=} \partial_0 \mathcal{F} = 2\varphi_i \partial_0 \varphi_i = 2\varphi_i \pi_i. \tag{3.4.26}$$

This leads to a secondary constraint,

$$0 = \mathcal{G}(\varphi_i, \pi_i) = \varphi_i \pi_i. \tag{3.4.27}$$

To ensure this constraint also holds for all times, we now demand $\partial_0 \mathcal{G} = 0$, or more concretely

$$0 = \partial_0 \mathcal{G} = (\partial_0 \varphi_i)\pi_i + \varphi_i(\partial_0 \pi_i). \tag{3.4.28}$$

Substituting the canonical equations of motion (cf. (3.4.23)), we find

$$0 = \pi_i \pi_i + \varphi_i[\Delta\varphi_i - 2\lambda\varphi_i] = [\pi_i \pi_i + \varphi_i \Delta\varphi_i] - 2\lambda\varphi_i\varphi_i. \tag{3.4.29}$$

Using the primary constraint $\mathcal{F} = 0 \Rightarrow \varphi_1^2 + \varphi_2^2 = 1$, we can solve for the Lagrange multiplier λ, finding

$$\lambda = \frac{1}{2}[\pi_1^2 + \varphi_1 \Delta\varphi_1 + \pi_2^2 + \varphi_2 \Delta\varphi_2]. \tag{3.4.30}$$

Thus, the consistency condition for the secondary constraint does not yield a new equation to impose, but instead determines the Lagrange multiplier λ dynamically. This is characteristic of second-class constraints and will appear again in more general form when we turn to the Dirac-Bergmann algorithm.

3.5 The Dirac-Bergmann Algorithm

As we have seen in the case of Maxwell theory, field theories with singular Lagrangians are not esoteric mathematical curiosities—they occur in physically relevant settings. In such cases, the relationship between momenta and velocities becomes non-invertible, making a straightforward Legendre transformation impossible. This, however, is not the end of the Hamiltonian story. A Hamiltonian formulation of Maxwell theory does exist—it simply requires a more intricate construction, via the so-called Dirac–Bergmann algorithm.

While one might feel this effort is unnecessary—after all, Maxwell's equations themselves provide a consistent first-order evolution system—certain theories, like General Relativity, force the issue. In those contexts, transitioning to the Hamiltonian picture becomes essential for deriving evolution equations. This motivates us to generalise the Hamiltonian formalism to accommodate singular Lagrangians. The Dirac–Bergmann algorithm provides the necessary framework to do just that.

We will first cover the algorithm in mechanics, where the phase space is finite-dimensional and easier to visualise (though one might argue that singular Lagrangians are more esoteric there). Then we generalise to field theories, and finally treat Maxwell theory in the way it deserves—by extracting a clean Hamiltonian field theory from it.

3.5.1 Construction of the Algorithm in Mechanics

The setting of the Dirac-Bergmann algorithm is as follows. Suppose we have coordinates $q^\mu(\tau)$, where τ parameterises a one-dimensional base manifold (e.g. time), and a Lagrangian

$$\mathcal{L} = \mathcal{L}(q, \dot{q}). \tag{3.5.1}$$

Suppose further that $\mathcal{L}$ is singular, i.e. that the Hessian determinant of the Lagrangian with respect to velocities

$$\det\left(\frac{\partial^2\mathcal{L}}{\partial\dot{q}^\mu\partial\dot{q}^\nu}\right) = 0 \tag{3.5.2}$$

vanishes at some configurations and for certain values of τ. In other words, the relationship between velocities and momenta, given by

$$p_\mu = \frac{\partial\mathcal{L}}{\partial\dot{q}^\mu}, \tag{3.5.3}$$

is not invertible.

This non-invertibility implies that certain combinations of the momenta are not independent but instead satisfy *constraints*, which we may—without loss of generality—express as

$$\phi_a(q,p) = 0, \tag{3.5.4}$$

where the index a labels the constraints, running from 1 to the number m of such relations. This number equals the number of configuration variables minus the rank of the Hessian above. Note that the constraint functions ϕ_a may depend on both the coordinates q and momenta p, but not on the velocities $\dot{q}$.

The geometric picture of this setup is as follows: With singular Lagrangians, we are no longer free to explore the entire phase space $\Gamma = \{(q,p)\}$. Each constraint restricts us to a submanifold $\tilde{\Gamma}$ of codimension m—we lose one dimension per independent constraint.

Let us consider how such constraints can be enforced. We can mimic what happens when constraints are introduced at the Lagrangian level by adding non-dynamical degrees of freedom to the system, acting as Lagrange multipliers. Specifically, we add a term $\lambda^a(\tau)\phi_a(q,p)$ to the Hamiltonian. This introduces the necessary constraint forces into the system, though with undetermined multipliers. So, how do we determine them?

A natural starting point is to require the initial condition to satisfy all the constraints ϕ_a, i.e. to choose it on the constrained submanifold $\tilde{\Gamma}$. Beyond that, it is only reasonable to demand that these constraints remain valid under time evolution: we should impose $\frac{d}{d\tau}\phi_a = 0$ initially. The same goes for the second derivative, the third, and so on.

This might sound alarming—infinitely many constraints?! But not to worry: at some point, usually sooner rather than later, these conditions begin to close among themselves, turning into combinations of constraints already imposed, or leading to equations that determine the λ^a. In effect, one always finds a finite, closed set of constraints. The procedure for systematically identiying them—laid out here, so far, in slightly informal terms—is known as the Dirac-Bergmann algorithm. We shall now embark on its formal introduction.

The algorithm is best broken down into steps:

1. **Canonical Momenta:** We begin by computing the Legendre relations between the momenta and velocities,

$$p_\mu = \frac{\partial\mathcal{L}}{\partial\dot{q}^\mu}. \tag{3.5.5}$$

Let N be the amount of coordinates, i.e. $\mu = 1, ..., N$. If the Hessian matrix

$$W_{\mu\nu} := \frac{\partial^2\mathcal{L}}{\partial\dot{q}^\mu\partial\dot{q}^\nu} \tag{3.5.6}$$

is not invertible, i.e. if

$$\operatorname{rank} W \equiv r < N, \tag{3.5.7}$$

then the map $\dot{q}^\mu \mapsto p_\mu$ is not globally invertible.

Accordingly, we split the momenta into two subsets:

- **Regular Sector:** momenta p_α for which the relation $\dot{q}^\alpha \leftrightarrow p_\alpha$ can be inverted. This may require a reparametrisation of the coordinates q^μ to separate regular from singular directions cleanly.
- **Singular Sector:** momenta p_a that cannot be solved for in terms of velocities. This gives rise to *primary constraints*,

$$\phi_a^{(0)}(q,p) = 0. \tag{3.5.8}$$

This split is often straightforward in practice. In many physical systems, the constraints manifest as simple conditions such as a momentum vanishing identically.

The set of primary constraints $\phi_a^{(0)}$ defines the *primary constraint surface* in phase space, specifically

$$\Gamma^{(0)} \equiv \Big\{(q,p) \in \Gamma \,\Big|\, \phi_a^{(0)}(q,p) = 0\Big\}. \tag{3.5.9}$$

2. **Total Hamiltonian:** On the regular sector, where the Legendre transform is valid, we define the *canonical Hamiltonian*

$$\mathcal{H}_C = p_\alpha \dot{q}^\alpha - \mathcal{L}. \tag{3.5.10}$$

We then promote the primary constraints $\phi_a^{(0)} = 0$ to equations of motion via Lagrange multipliers $\lambda^a(\tau)$. This yields the *total Hamiltonian*

$$\mathcal{H}_T \equiv \mathcal{H}_C + \lambda^a \phi_a^{(0)}. \tag{3.5.11}$$

The time evolution of the canonical variables is governed by

$$\dot{q}^\mu = \{q^\mu, \mathcal{H}_T\} = \frac{\partial \mathcal{H}_T}{\partial p_\mu}, \qquad \dot{p}_\mu = \big\{p_\mu, \mathcal{H}_T\big\} = -\frac{\partial \mathcal{H}_T}{\partial q^\mu}. \tag{3.5.12}$$

This defines off-shell evolution—i.e. before all constraints have been enforced.

3. **Constraint Consistency and Secondary Constraints:** To ensure that the primary constraints are preserved under time evolution, we impose

$$\frac{d}{d\tau}\phi_a^{(0)}(q,p) \approx 0. \tag{3.5.13}$$

Here, $\approx$ denotes *weak equality*, i.e. equality only on the primary constraint surface $\Gamma^{(0)}$. The consequences fall into three categories:

- **Fixing a Lagrange multiplier:** the consistency condition determines a λ^a. In that case, we simply solve for it and proceed.
- **Closure of a constraint chain:** the time derivative of a constraint reduces to a combination of existing constraints,

$$\frac{d}{d\tau}\phi_a^{(0)} = f\big(\varphi^{(0)}\big), \tag{3.5.14}$$

with $f(0) = 0$. Then no new constraint arises.

- **New (secondary) constraint:** if neither of the above applies, the consistency condition yields a genuinely new constraint:

$$\phi_a^{(1)}(q,p) \equiv \frac{d}{d\tau}\phi_a^{(0)}(q,p) \approx 0. \tag{3.5.15}$$

In case secondary constraints do arise, we further define the secondary constraint surface as

$$\begin{aligned}\Gamma^{(1)} &= \Big\{(q,p) \in \Gamma^{(0)} \,\Big|\, \phi_a^{(1)} \approx 0\Big\} \\ &= \Big\{(q,p) \in \Gamma \,\Big|\, \phi_a^{(0)} = 0, \phi_a^{(1)} \approx 0\Big\}.\end{aligned} \tag{3.5.16}$$

4. **Higher-Level Constraints:** We iterate the same procedure: demand that secondary constraints be preserved under time evolution, potentially generating *tertiary constraints* $\phi_a^{(2)}$, and so on. Each step defines a corresponding surface,

$$\Gamma^{(i+1)} \equiv \Big\{(q,p) \in \Gamma^{(i)} \,\Big|\, \phi_a^{(i+1)}(q,p) \approx 0\Big\}. \tag{3.5.17}$$

The process terminates when no new constraints arise.

At the end of this procedure, we obtain:

- A full set of constraints:

$$\mathcal{C} = \Big\{\phi_a^{(0)}, \phi_a^{(1)}, \phi_a^{(2)}, ...\Big\}, \tag{3.5.18}$$

- A nested sequence of constraint surfaces…

$$\Gamma \supset \Gamma^{(0)} \supset \Gamma^{(1)} \supset \Gamma^{(2)} \supset ..., \tag{3.5.19}$$

- …whose intersection defines the *final constraint surface*:

$$\Gamma^{\infty} \equiv \bigcap_i \Gamma^{(i)} = \Big\{(q,p) \in \Gamma \,\Big|\, \phi_a^{(i)}(q,p) = 0, \forall \phi_a^{(i)} \in \mathcal{C}\Big\}. \tag{3.5.20}$$

By construction, all Lagrange multipliers have now either been fixed or remain completely undetermined, in which case they generate gauge freedom. These unfixed multipliers do not affect evolution on Γ^{∞}.

And now the key result: If the initial data satisfy all constraints $\phi_a^{(i)} = 0$, then the evolution under $\mathcal{H}_T$ will preserve them. That is, the trajectory remains entirely within Γ^{∞}, and all constraints are automatically maintained over time.

It is rather simple to see that this is the case: the way we constructed Γ^{∞} directly implies that on this surface, any derivative of the primary constraints with respect to τ, regardless of order (including the "zeroth derivative", the constraint itself), is zero. This can only be the case if $\phi_a^{(0)} = 0$ for all τ, and hence the (primary) constraints are always satisfied if they are satisfied initially.

There are two more questions we should at least mention here. Firstly, why do we only include constraint terms of the primary constraints in the Lagrangian, but not of higher order constraints? This can be answered as follows. The primary constraints are those that arise from the singularity of the Lagrangian. Not all constraints are independent—rather, some combinations of them must vanish. This tells us the following: Our evolution cannot happen in the whole (naive) phase space $\Gamma = \{(q,p)\}$, but only a subset thereof. Not all choices of momentum values make sense, as not all of them can be inverted for their velocity configurations. In other words, not all choices of momenta even have a solution of the equations of motion $q(\tau)$ that could produce those momenta. This means, parts of the unconstrained phase space Γ are *unphysical*—they don't really exist. So, to have something meaningful, we need to promote the primary constraints.

Secondary and higher order constraints are of a fundamentally different nature, though. Rather than being a strict requirement from the singular Legendre map, they simply arise from wanting the primary constraints to *remain* fulfilled if they are fulfilled initially. If there are such constraints, that tells us that even if a phase space position (q, p) satisfies the primary constraints, i.e. lies on $\Gamma^{(0)}$, it *might* still drift off that hypersurface through evolution—hence isn't a physically meaningful configuration either. A particle's path cannot be physically meaningful in one instant and become unreason in the second. We thus don't only want solutions that are physically meaningful initially, but that remain physically meaningful—or in more precise terms, we only care about phase space trajectories that satisfy the primary constraints at all times. This simply reduces the relevant portions of phase space to Γ^{∞}. This submanifold constitutes all phase space configurations for which the evolution preserves the primary constraints for all τ.

At this point, notice the fundamental difference between primary and higher order constraints: primary constraints need to be satisfied for a point (q, p) to be physically meaningful at all—thus must be part of the evolution and hence the Hamiltonian. Secondary constraints, however, only further "weed out" the trajectories that still fail to "stay on course", so to speak.

Then, the second question: We constructed a mathematically meaningful way of extracting Hamiltonian evolution from a singular Lagrangian in a rather canonical way. But, how can we trust that we did not change the physical content of the theory while doing it? Why does the new formulation describe the same physics? *This is something I would like to maybe get into at one point, not now—I'm not sure if I could properly prove this as of right now. But, I feel like this question deserves to at least be asked here, since it is rather fundamental. An intuitive answer was partially given by answering the first question, but a rigorous answer remains to be provided.*

It is rather simple to see that this is the case: the way we constructed Γ^{∞} directly implies that on this surface, any derivative of the primary constraints with respect to τ, regardless of order (including the constraints themselves), is zero. This implies that $\phi_a^{(0)} = 0$ for all τ, i.e. the primary constraints vanish identically along the evolution, and hence are always satisfied if they are satisfied initially.

There are two more questions we should at least mention here. Firstly, why do we only include constraint terms of the primary constraints in the Lagrangian, but not of higher order constraints? This can be answered as follows. The primary constraints are those that arise from the degeneracy of the Lagrangian. Not all momenta are independent—rather, some combinations of them must vanish. This tells us the following: our evolution cannot happen in the whole (naive) phase space $\Gamma = \{(q, p)\}$, but only a subset thereof. Not all choices of momentum make values make sense, as not all of them can be inverted for their velocity configurations (notice that the converse situation always makes sense—any velocities always lead to valid momenta). In other words, not all choices of momenta even correspond to consistent solutions $q(\tau)$ of the equations of motion that could produce those momenta. This means, parts of the unconstrained phase space Γ are unphysical—they do not really exist. So, to have something meaningful, we need to incorporate the primary constraints directly into the Hamiltonian through Lagrange multipliers.

Secondary and higher order constraints are of a fundamentally different nature. Rather than being a strict requirement from the Legendre map, they simply arise from wantin the primary constraints to *remain* fulfilled if they are fulfilled initially. If there are such constraints, that tells us that even if a phase space position (q, p) satisfies the primary constraints, i.e. lies on $\Gamma^{(0)}$, it *might* still drift off that hypersurface through evolution—hence it isn't a physically meaningful trajectory either. A particle's path cannot be physically meaningful in one instant and beocome unphysical in the next. We thus don't only want solutions that are physically meaningful initially, but that remain physically meaningful—more precisely, we only care about phase space trajectories that satisfy the primary constraints at all times. These are precisely those trajectories that satisfy all higher order constraints as well, which restricts physically admissible trajectories to lie within Γ^{∞}.

This submanifold constitutes all phase space configurations for which the evolution preserves the primary constraints for all τ.

At this point, notice the fundamental difference between primary and higher order constraints: primary constraints need to be satisfied for a point (q, p) to be physically meaningful at all—thus must be part of the evolution and hence included in the Hamiltonian. Secondary constraints, however, only further "weed out" the trajectories that still fail to remain within the constraint surface under time evolution, even when the primary constraints are promoted.

Then, the second question: We constructed a mathematically meaningful way of extracting Hamiltonian evolution from a singular Lagrangian in a rather canonical way. But, how can we trust that we did not change the physical content of the theory while doing it? Why does the new formulation describe the same physics? *This is something I would like to get into at some point in the future, not now—I'm not sure if I could properly prove this as of right now. But, I feel like this question deserves to be at least asked here, since it is rather fundamental. An intuitive answer was partially given by answering the first question, but a rigorous answer remains to be provided.*

3.5.2 Example: Mechanical System with a Singular Lagrangian

To elucidate the somewhat abstractly introduced Dirac-Bergmann algorithm, let us explore a simple mechanical system with a singular Lagrangian. We consider a two-dimensional system with only one dynamical degree of freedom, defined by

$$\mathcal{L} = \frac{1}{2}\dot{x}^2 - V(x, y). \tag{3.5.21}$$

Technically speaking, $y(\tau)$ is an externally defined control function. We will not concern ourselves with that here and still treat it as a coordinate $y(\tau)$.

Let us begin by reviewing the Euler-Lagrange equations to get a sense of the theory. The equation of motion for $x(\tau)$ reads

$$\ddot{x} = -\partial_x V(x, y). \tag{3.5.22}$$

Treating $y(\tau)$ the same way gives another equation,

$$0 = -\partial_y V(x, y) \tag{3.5.23}$$

This tells us two things: on one hand, $y(\tau)$ is completely arbitrary—on the other, the potential V must be independent of it. In other words, consistency of the Euler-Lagrange equations forces $\partial_y V = 0$, so our Lagrangian does not involve y at all. We are thus describing a one-dimensional particle in a potential, with an additional nondynamical function $y(\tau)$ that has no effect on the dynamics of x.

Regardless of the questionable physical relevance of such a system, it is clear that the Lagrangian is singular. To see this, compute the Hessian determinant with respect to the velocities,

$$\det\left[\begin{pmatrix}\partial_{\dot{x}}\partial_{\dot{x}} & \partial_{\dot{x}}\partial_{\dot{y}} \\ \partial_{\dot{y}}\partial_{\dot{x}} & \partial_{\dot{y}}\partial_{\dot{y}}\end{pmatrix}\mathcal{L}\right] = \det\begin{pmatrix}1 & 0 \\ 0 & 0\end{pmatrix} = 0. \tag{3.5.24}$$

This means that the Legendre transformation cannot be invertible.

We now now compute the canonical momenta and identify the primary constraints. For x, we have

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = \dot{x}. \tag{3.5.25}$$

For y, we get

$$p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}} = 0. \tag{3.5.26}$$

This tells us that $p_y = 0$ is a primary constraint. The regular sector consists of the pair (x, p_x), and the singular sector of $\big(y, p_y\big)$ with the constraint

$$\phi_1^{(0)}(q,p) = p_y = 0. \tag{3.5.27}$$

This defines the primary constraint surface

$$\Gamma^{(0)} = \big\{\big(x, y, p_x, p_y\big) \mid p_y = 0\big\}. \tag{3.5.28}$$

Next, we construct the canonical and total Hamiltonians. The former is simply given by

$$\mathcal{H}_C = p_x\dot{x} - \mathcal{L} = \frac{1}{2}p_x^2 + V(x). \tag{3.5.29}$$

To obtain the total Hamiltonian, we add a Lagrange multiplier $\lambda(\tau)$ enforcing the constraint,

$$\mathcal{H}_T = \mathcal{H}_C + \lambda^a\phi_a^{(0)} = \frac{1}{2}p_x^2 + V(x) + \lambda p_y \tag{3.5.30}$$

The canonical equations of motion are follow as

$$\begin{aligned} \dot{x} &= \frac{\partial \mathcal{H}_T}{\partial p_x} = p_x, \qquad & \dot{p}_x &= -\frac{\partial \mathcal{H}_T}{\partial x} = -\partial_x V(x), \\ \dot{y} &= \frac{\partial \mathcal{H}_T}{\partial p_y} = \lambda, \qquad & \dot{p}_y &= -\frac{\partial \mathcal{H}_T}{\partial y} = 0. \end{aligned} \tag{3.5.31}$$

In particular, the multiplier λ governs the evolution of $y(\tau)$ via $\dot{y} = \lambda$.

We then move to secondary constraints. Demanding consistency of $\phi_1^{(0)}$ yields

$$0 \overset{!}{=} \frac{d}{d\tau}\phi_1^{(0)}(q,p) = \dot{p}_y = 0, \tag{3.5.32}$$

which is trivially satisfied by the equations of motion. Hence, we do not get any secondary or higher order constraints, and the multiplier λ remains arbitrary. By the relation $\dot{y}(\tau) = \lambda(\tau)$ we get from the equations of motion, this means that $y(\tau)$ is arbitrary as well, reflecting our findings from the Lagrangian formulation.

So, while this may not be the most interesting system physically, it illustrates some key features of the Dirac-Bergmann algorithm in a simple mechanical context: the appearance of primary constraints, arbitrainess of certain degrees of freedom, and the absence of further consistency conditions.

3.5.3 Example II: Mechanical System with a Gauge Potential

Things become more interesting in field theory, where derivatives with respect to parameters beyond time can give rise to richer constraint structures. To prepare for that, we will generalise the Dirac-Bergmann algorithm to field-theoretic systems in the next section. Before doing so, however, we examine a second mechanical example—slightly more intricate than the previous—to further develop our intuition and illustrate the algorithm in action once more.

The Lagrangian of this system is given by

$$\mathcal{L} = \frac{1}{2}(\dot{x} - y)^2. \tag{3.5.33}$$

It describes one dynamical variable, $x(\tau)$, and one non-dynamical variable, $y(\tau)$. Structurally, this resembles a kinetic term modified by a gauge potential.

Let us begin with the Lagrangian formulation. The Euler-Lagrange equations read

$$\begin{aligned} 0 &= \underbrace{\frac{\partial \mathcal{L}}{\partial x}}_{=0} - \frac{d}{d\tau}\frac{\partial \mathcal{L}}{\partial \dot{x}} = \ddot{x} - \dot{y} \qquad \Longrightarrow \qquad \ddot{x} = \dot{y} \\ 0 &= \frac{\partial \mathcal{L}}{\partial y} - \frac{d}{d\tau}\underbrace{\frac{\partial \mathcal{L}}{\partial \dot{y}}}_{=0} = -(\dot{x} - y) \qquad \Longrightarrow \qquad \dot{x} = y \end{aligned} \tag{3.5.34}$$

Since the second equation implies the first, the equations of motion reduce to $\dot{x} = y$, with $y(\tau)$ arbitrary. The solution is simply

$$x(\tau) = c + \int d\tau\, y(\tau) \tag{3.5.35}$$

where c is an integration constant determined by initial data.

Now we turn to the Hamiltonian picture. The canonical momenta are

$$\begin{aligned} p_x &= \frac{\partial \mathcal{L}}{\partial \dot{x}} = \dot{x} - y \quad \Longrightarrow \quad \dot{x} = p_x + y \\ p_y &= \frac{\partial \mathcal{L}}{\partial \dot{y}} = 0. \end{aligned} \tag{3.5.36}$$

We obtain a primary constraint, $\phi_1^{(0)} = p_y = 0$, and the Legendre transform with respect to $\dot{x}$ yields the canonical Hamiltonian

$$\mathcal{H}_C = p_x \dot{x} - \mathcal{L} = \frac{1}{2}p_x^2 + y p_x \tag{3.5.37}$$

Enforcing the constraint $\phi_1^{(0)} = 1$ via a Lagrange multiplier $\lambda(\tau)$ leads to the total Hamiltonian

$$\mathcal{H}_T = \frac{1}{2}p_x^2 + y p_x + \lambda p_y. \tag{3.5.38}$$

This Hamiltonian governs the evolution through phase space. The corresponding canonical equations are

$$\begin{aligned} \dot{x} &= \frac{\partial \mathcal{H}}{\partial p_x} = p_x + y, \qquad & \dot{p}_x &= -\frac{\partial \mathcal{H}}{\partial x} = 0, \\ \dot{y} &= \frac{\partial \mathcal{H}}{\partial p_y} = \lambda, & \dot{p}_y &= -\frac{\partial \mathcal{H}}{\partial y} = -p_x. \end{aligned} \tag{3.5.39}$$

These will guide our analysis of consistency conditions.

Demanding preservation of the primary constraint $\phi_1^{(1)} = 0$ under time evolution yields

$$0 \overset{!}{=} \frac{d}{d\tau}\phi_1^{(0)} = \dot{p}_y = -p_x \tag{3.5.40}$$

by making use of the canonical equations derived above. This is a genuinely new constraint—it is neither a combination of known constraints nor does it fix λ. Thus, we have to introduce the secondary constraint

$$\phi_1^{(1)} = p_x = 0. \tag{3.5.41}$$

At this point, one might start to feel a bit suspicious. These two conditions limit us to the secondary constraint surface

$$\Gamma^{(1)} = \big\{(x, y, p_x, p_y) \,\big|\, p_x = p_y = 0\big\}. \tag{3.5.42}$$

While this may seem overly restrictive at first glance, it is in fact consistent with the Lagrangian solution: $p_x = 0$ implies $\dot{x} = y$, in perfect agreement with the Euler-Lagrange equation.

There are no tertiary constraints—the condition

$$0 \overset{!}{=} \frac{d}{d\tau}\phi_1^{(1)} = \dot{p}_x = 0 \tag{3.5.43}$$

is automatically satisfied according to the canonical equations. Notice that the Lagrange multiplier λ remained undetermined.

In summary, we have now constructed the constraint surface $\Gamma^\infty = \Gamma^{(1)}$, on which the evolution remains constrained. Hence, such solutions always satisfy $p_x = p_y = 0$, reducing the canonical equations to

$$\dot{x} = y, \quad \dot{y} = \lambda, \tag{3.5.44}$$

with the other two equations being fulfilled trivially. Since λ was found to be arbitrary, this reproduces the Lagrangian picture exactly. Compared to the previous example, this system features a primary *and* a secondary constraint, giving a hint of more intricate behaviour—but still falling short of the complexity encountered in field theory.

3.5.4 Dirac-Bergmann in Field Theory

Without motivation, here we simply state the procedure behind the Dirac-Bergmann algorithm for field theories. As a starting point, suppose we have dynamical variables (fields) $\Phi^A(x)$, which take values on an N-dimensional spacetime manifold with local coordinates x^μ. We further consider a Lagrangian density

$$\mathcal{L} = \mathcal{L}\big(\Phi^A, \partial_\mu \Phi^A\big) \tag{3.5.45}$$

which defines an action given by

$$S[\Phi] = \int d^N x \, \mathcal{L}\big(\Phi^A(x), \partial_\mu \Phi^A(x)\big). \tag{3.5.46}$$

We work under the assumption that the Lagrangian is singular, i.e. that the velocity Hessian

$$W_{AB} = \frac{\partial^2 \mathcal{L}}{\partial(\partial_0 \Phi^A)\partial(\partial_0 \Phi^B)} \tag{3.5.47}$$

does not have full rank and hence zero determinant. The procedure then goes as follows:

1. **Canonical Momenta:** As a first step, the Legendre relations between the momenta and velocities,

$$\Pi_A = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi^A)} \tag{3.5.48}$$

are to be computed. Using $r \equiv \operatorname{rank} W_{AB} < N$, we can split the momenta into two subsets:

- **Regular Sector:** Momenta $\Pi_{\mathcal{A}}$ for which the relation $\partial_0 \Phi^{\mathcal{A}} \leftrightarrow \Pi_{\mathcal{A}}$ can be inverted. This may require a reparametrisation of the fields Φ^A to separate regular from singular directions cleanly.

- **Singular Sector:** Momenta Π_a that cannot be solved for in terms of the velocities $\partial_0\Phi^A$. This gives rise to *primary constraints*,

$$\phi_a^{(0)}\big(\Phi^A, \Pi_A\big) = 0 \tag{3.5.49}$$

  In many cases, these manifest directly as identities such as $\Pi_a = 0$.

Just as in mechanics, the set of primary constraints $\phi_a^{(0)}$ defines the *primary constraint surface* in phase space, given by

$$\Gamma^{(0)} \equiv \Big\{\big(\Phi^A, \Pi_A\big) \in \Gamma \,\Big|\, \phi_a^{(0)}\big(\Phi^A, \Pi_A\big) = 0\Big\}. \tag{3.5.50}$$

Notably, in contrary to mechanics, this surface (as well as the space it Γ it lives in), is infinite-dimensional.

2. **Canonical and Total Hamiltonian:** With respect to the regular sector, the Legendre transformation can be performed, yielding the *canonical Hamiltonian density* $\mathcal{H}_C$, given by

$$\mathcal{H}_C = \Pi_{\mathcal{A}}\partial_0\Phi^{\mathcal{A}} - \mathcal{L}. \tag{3.5.51}$$

To enforce the primary constraints, we introduce Lagrange multipliers $\lambda^a(x)$ and use them to promote the constraints to equations of motion by defining the *total Hamiltonian density*

$$\mathcal{H}_T = \mathcal{H}_C + \lambda^a\phi_a^{(0)}. \tag{3.5.52}$$

From this, we further define the total Hamiltonian as

$$H_T = \int d^{N-1}x\, \mathcal{H}\big(\Phi^A, \Pi_A\big) \tag{3.5.53}$$

This Hamiltonian is what we use to define off-shell evolution; the equations of motion read

$$\partial_0\Phi^A = \big\{\Phi^A, H_T\big\} = \frac{\delta\mathcal{H}_T}{\delta\Pi_A}, \qquad \partial_0\Pi_A = \{\Pi_A, H_T\} = -\frac{\delta\mathcal{H}_T}{\delta\Phi^A}. \tag{3.5.54}$$

3. **Constraint Consistency and Secondary Constraints:** We demand constraints to remain satisfied if they are satisfied initially. More concretely, we require

$$\partial_0\phi_a^{(0)}\big(\Phi^A, \Pi_A\big) \overset{!}{\approx} 0. \tag{3.5.55}$$

where $\approx$ indicates weak equality, i.e. equality only on $\Gamma^{(0)}$. There are three possible outcomes for this:

   1. **Determination of a Multiplier:** Such a condition can lead to a relationship between a multiplier λ^a and the canonical variables. In this case, we solve for it and may ignore the condition.

   2. **Closure:** When the expression $\partial_0\phi_a^{(0)}$ can be written as a function of known constraints,

$$\partial_0\phi_a^{(0)} = f\big(\phi_a^{(0)}\big) \tag{3.5.56}$$

   —with $f(0) = 0$—the condition is automatically satisfied. We may thus ignore it.

   3. **Emergence of a Secondary Constraint** If neither of the first two outcomes apply, then the condition (3.5.55) constitutes a new *secondary constraint* $\phi_a^{(1)} \approx 0$, where

$$\phi_a^{(1)}\big(\Phi^A, \Pi_A\big) = \partial_0\phi^{(0)}\big(\Phi^A, \Pi_A\big). \tag{3.5.57}$$

If there are secondary constraints, they lead to the definition of the *secondary constraint surface*

$$\Gamma^{(1)} = \Big\{ \big(\Phi^A, \Pi_A\big) \in \Gamma^{(0)} \,\Big|\, \phi_a^{(1)}\big(\Phi^A, \Pi_A\big) \approx 0 \Big\}. \tag{3.5.58}$$

4. **Higher-Order Constraints:** The previous step is iterated: demand that secondary constraints be preserved under time evolution, potentially generating *tertiary constraints* $\phi_a^{(2)}$, and so on. At each step, define the corresponding surface,

$$\Gamma^{(i+1)} = \Big\{ \big(\Phi^A, \Pi_A\big) \in \Gamma^{(i)} \,\Big|\, \phi_a^{(i+1)}\big(\Phi^A, \Pi_A\big) \approx 0 \Big\} \tag{3.5.59}$$

   The procedure is stopped when an iteration does not introduce new constraints.

Just as in mechanics, completion of this procedure leads to:

- A full set of constraints:

$$\mathcal{C} = \Big\{ \phi_a^{(0)}, \phi_a^{(1)}, \phi_a^{(2)}, ... \Big\} \tag{3.5.60}$$

- A nested sequence of (infinite dimensional) constraint surfaces,

$$\Gamma \supset \Gamma^{(0)} \supset \Gamma^{(1)} \supset ... \supset \Gamma^{\infty}, \tag{3.5.61}$$

- where Γ^∞ is the *final constraint surface*, defined as

$$\Gamma^{\infty} = \bigcap_i \Gamma^{(i)} = \Big\{ \big(\Phi^A, \Pi_A\big) \in \Gamma \mid \phi_a^{(i)}\big(\Phi^A, \Pi_A\big) = 0, \forall \phi_a^{(i)} \in \mathcal{C} \Big\}. \tag{3.5.62}$$

The consequences are the same as in mechanics: By construction, all Lagrange multipliers have now either been fixed or remain completely undetermined, in which case they generate gauge freedom. These unfixed multipliers do not affect evolution on Γ^∞. Further, starting from an initial configuration $\big(\Phi^A, \Pi_A\big) \in \Gamma^\infty$ and evolving along the flow field generated by $\mathcal{H}_T$, the constraints are preserved and we are in a formulation that is physically equivalent to the Lagrangian picture.

Put more concisely,

$$\big(\Phi^A(0, x^i), \Pi_A(0, x^i)\big) \in \Gamma^{\infty}, \qquad \partial_0 \Phi^A = \frac{\delta \mathcal{H}_T}{\delta \Pi_A}, \qquad \partial_0 \Pi_A = -\frac{\delta \mathcal{H}_T}{\delta \Phi^A} \tag{3.5.63}$$

implies that

$$\frac{\partial \mathcal{L}}{\partial \Phi^A} - \partial_\mu \frac{\partial \mathcal{L}}{\partial\big(\partial_\mu \Phi^A\big)} = 0. \tag{3.5.64}$$

This construction is what will finally allow us to derive the Hamiltonian formulation of Maxwell theory in the next section.

3.5.5 Example: Maxwell Theory, Revisited Properly

We have endeavoured plenty to get to this point, but we can now do it: Treat Maxwell theory in the Hamiltonian picture, with the Dirac-Bergmann algorithm. For completeness, let us start from the beginning. In the mostly plus signature of the Minkowskian metric, the Maxwell Lagrangian density reads

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \tag{3.5.65}$$

where $F_{\mu\nu}$ is the field strength associated with the 1-form $A = A_\mu \, dx^\mu$,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{3.5.66}$$

To make the subsequent analysis simpler, we provide a number of alternative but equivalent expressions for this Lagrangian:

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2}F_{0i}F^{0i} - \frac{1}{4}F_{ik}F^{ik} \\ &= \frac{1}{2}E_iE_i - \frac{1}{4}\varepsilon_{ikl}B_l\varepsilon_{ikp}B_p \\ &= \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2)\end{aligned} \tag{3.5.67}$$

Here we made use of the expressions for the electric field components E_i and magnetic field components B_i, which are given by

$$E_i = \partial_0 A_i - \partial_i A_0 = F_{0i}, \qquad \varepsilon_{ijk}B_k = \partial_i A_j - \partial_j A_i = F_{ij}, \tag{3.5.68}$$

as well as the identity

$$\varepsilon_{ijk}\varepsilon_{ijl} = 2\delta_{kl}. \tag{3.5.69}$$

For inclusiveness, we restate the Euler-Lagrange equations for Maxwell theory, which read

$$\partial_\mu F^{\mu\nu} = 0. \tag{3.5.70}$$

Dissecting this a bit further, we find two types of equations:

$$0 = \partial_i F^{i0} = \nabla \cdot \mathbf{E}, \tag{3.5.71}$$

which is simply Gauss' law, and

$$0 = \partial_0 F^{0k} + \partial_i F^{ik} = \partial_0 E_k + \varepsilon_{ikj}\partial_i B_j = \left(\partial_0 \mathbf{E} - \nabla \times \mathbf{B}\right)_k. \tag{3.5.72}$$

In the following, we will rediscover these equations in the pursuit of the Dirac-Bergmann algorithm.

1. **Canonical Momenta:** Without further ado, let us proceed with its procedure. We begin by evaluating the Legendre map to derive the momenta. For the spatial momenta, we get

$$\Pi_i = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_i)} = -\frac{\partial F_{0i}}{\partial(\partial_0 A_i)}F^{0i} = -F^{0i} = E_i. \tag{3.5.73}$$

   In words: the momenta corresponding to the spatial velocities $\partial_0 A_i$ are the electric field components E_i. We may rearrange this equation to find the velocity in terms of the momentum,

$$E_i = F_{0i} = \partial_0 A_i - \partial_i A_0 \quad \Leftrightarrow \quad \partial_0 A_i = E_i + \partial_i A_0 \tag{3.5.74}$$

   For the temporal momentum, we find

$$\Pi_0 = \frac{\partial \mathcal{L}}{\partial(\partial_0 A_0)} = 0. \tag{3.5.75}$$

   To summarise, we have found three regular/invertible momenta, E_i, and a singular momentum Π_0. The latter gives us a primary constraint,

$$\phi_1^{(0)} = \Pi_0 = 0. \tag{3.5.76}$$

2. **Canonical and Total Hamiltonian Densities:** Performing the partial Legendre transform with respect to the regular sector yields the canonical Hamiltonian density

$$\begin{aligned}\mathcal{H}_C &= E_i\partial_0 A_i - \mathcal{L} = E_i(E_i + \partial_i A_0) - \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2) \\ &= \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + E_i\partial_i A_0\end{aligned} \tag{3.5.77}$$

We are now ready to define the total Hamiltonian density by introducing a Lagrange multiplier $\lambda(x)$ to enforce the constraint $\phi_a^{(0)} = \Pi_0 = 0$. It reads

$$\mathcal{H}_T = \mathcal{H}_C + \lambda\Pi_0 = \frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + E_i\partial_i A_0 + \lambda\Pi_0. \tag{3.5.78}$$

The associated Hamiltonian is given by

$$H_T = \int d^{N-1}x\, \mathcal{H}_T = \int d^{N-1}x \Big[\frac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) - A_0(\nabla\cdot\mathbf{E}) + \lambda\Pi_0\Big], \tag{3.5.79}$$

where we integrated by parts, discarding vanishing boundary terms. The reason for doing so will become evident later.

3. **Secondary Constraints:** Demanding consistency of $\Pi_0 = 0$ leads to

$$0 \overset{!}{=} \partial_0\Pi_0 = -\frac{\delta H_T}{\delta A_0} = \nabla\cdot\mathbf{E}, \tag{3.5.80}$$

which is simply Gauss' law. This constitutes a secondary constraint,

$$\phi_1^{(1)} = \nabla\cdot\mathbf{E} = 0. \tag{3.5.81}$$

Observe here that H_T contains the term $A_0(\nabla\cdot\mathbf{E})$, which is of the form of a constraint term for $\phi_1^{(1)}$—with A_0 acting as the Lagrange multiplier enforcing it.

4. **Higher Order Constraints:** Imposing that $\phi_1^{(1)}$ be consistent as well, we find

$$\begin{aligned} 0 &\overset{!}{=} \partial_0\partial_i E_i = \partial_i(\partial_0 E_i) = -\partial_i\frac{\delta H_T}{\delta A_i} \\ &= -\partial_i\Bigg[\underbrace{\frac{\partial\mathcal{H}_T}{\partial A_i}}_{=0} - \partial_k\frac{\partial\mathcal{H}_T}{\partial(\partial_k A_i)}\Bigg] = \partial_i\partial_k F^{ik} = 0 \end{aligned} \tag{3.5.82}$$

This means that the constraint chain closes, and we do not get any tertiary constraints.

Notice that we have not found any conditions on the two Lagrangian multipliers A_0 and λ. They hence remain free, and we may choose them to be whatever we please to make calculations easier.

Let us now compute the canonical equations of motion. For the spatial components these read

$$\partial_0 A_i = \frac{\delta H_T}{\delta E_i} = E_i + \partial_i A_0, \qquad \partial_0 E_i = -\frac{\delta H_T}{\delta A_i} = -\Bigg[\underbrace{\frac{\partial\mathcal{H}_T}{\partial A_i}}_{=0} - \partial_k\frac{\partial\mathcal{H}_T}{\partial(\partial_k A_i)}\Bigg] = \partial_k F^{ki} \tag{3.5.83}$$

In index-free notation, these are simply the equations

$$\partial_0\mathbf{A} = \mathbf{E} + \nabla A_0 \quad\Leftrightarrow\quad \mathbf{E} = \partial_0\mathbf{A} - \nabla A_0, \tag{3.5.84}$$

$$\partial_0\mathbf{E} = \nabla\times\mathbf{B}, \tag{3.5.85}$$

which we already know and love from Maxwell theory in its original formulation.

The canonical equations for the temporal components read

$$\partial_0 A_0 = \frac{\delta\mathcal{H}_T}{\delta\Pi_0} = \lambda, \qquad \partial_0\Pi_0 = -\frac{\delta\mathcal{H}_T}{\delta A_0} = \partial_i E_i \tag{3.5.86}$$

For configurations on Γ^∞, the right-hand side of the latter equation vanishes and hence $\Pi_0 = 0$ is preserved automatically. The former show that λ governs the (otherwise unconstrained) time evolution of A_0.

3.5.6 Revisiting Poisson Brackets

Let us revisit the definition of Poisson brackets in field theory from first principles and establish a set of practical computation rules.

Consider a Hamiltonian system with canonical fields $\Phi^A(x)$ and their conjugate momenta $\Pi_A(x)$, defined on a spatial hypersurface of dimension $N-1$. For two differentiable local functionals $\mathcal{A}\big[\Phi^A, \Pi_A\big]$ and $\mathcal{B}\big[\Phi^A, \Pi_A\big]$, the Poisson bracket is defined as

$$\{\mathcal{A}, \mathcal{B}\} = \int d^{N-1}x \left[\frac{\delta \mathcal{A}}{\delta \Phi^A(x)} \frac{\delta \mathcal{B}}{\delta \Pi_A(x)} - \frac{\delta \mathcal{B}}{\delta \Phi^A(x)} \frac{\delta \mathcal{A}}{\delta \Pi_A(x)}\right], \tag{3.5.87}$$

This generalises the finite-dimensional Poisson bracket to infinite dimensional field theory, replacing partial with functional derivatives and summations with spatial integrals.

From this definition, the equal-time fundamental Poisson brackets follow as[3]

$$\begin{aligned} \big\{\Phi^A(x), \Pi_B(y)\big\} &= \delta^A_B \delta^{N-1}(x-y), \\ \big\{\Phi^A(x), \Phi^B(y)\big\} &= \{\Pi_A(x), \Pi_B(y)\} = 0. \end{aligned} \tag{3.5.88}$$

There are a few computation rules that one can further derive to make calculations with Poisson brackets easier. Let us now go through these.

- **Leibniz Rule** The Poisson bracket satisfies a Leibniz rule,

$$\{\mathcal{A}\mathcal{B}, \mathcal{C}\} = \mathcal{A}\{\mathcal{B}, \mathcal{C}\} + \{A, \mathcal{C}\}\mathcal{B}. \tag{3.5.89}$$

  This is rather simple to see, as by the Leibniz rule for the functional derivative we find

$$\begin{aligned} \{\mathcal{A}\mathcal{B}, \mathcal{C}\} &= \int d^{N-1}x \left[\frac{\delta(\mathcal{A}\mathcal{B})}{\delta \Phi^A(x)} \frac{\delta \mathcal{C}}{\delta \Pi_A(x)} - \frac{\delta \mathcal{C}}{\delta \Phi^A(x)} \frac{\delta(\mathcal{A}\mathcal{B})}{\delta \Pi_A(x)}\right] \\ &= \int d^{N-1}x \left[\left(\frac{\delta \mathcal{A}}{\delta \Phi^A(x)}\mathcal{B} + \mathcal{A}\frac{\delta \mathcal{B}}{\delta \Phi^A(x)}\right)\frac{\delta \mathcal{C}}{\delta \Pi_A(x)}\right. \\ &\qquad \left. - \frac{\delta \mathcal{C}}{\delta \Phi^A(x)}\left(\frac{\delta \mathcal{A}}{\delta \Pi_A(x)}\mathcal{B} + \mathcal{A}\frac{\delta \mathcal{B}}{\delta \Pi_A(x)}\right)\right] \\ &= \int d^{N-1}x \left[\frac{\delta \mathcal{A}}{\delta \Phi^A(x)} \frac{\delta \mathcal{C}}{\delta \Pi_A(x)} - \frac{\delta \mathcal{C}}{\delta \Phi^A(x)} \frac{\delta \mathcal{A}}{\delta \Pi_A(x)}\right]\mathcal{B} \\ &\quad + \int d^{N-1}x\, \mathcal{A}\left[\frac{\delta \mathcal{B}}{\delta \Phi^A(x)} \frac{\delta \mathcal{C}}{\delta \Pi_A(x)} - \frac{\delta \mathcal{C}}{\delta \Phi^A(x)} \frac{\delta \mathcal{B}}{\delta \Pi_A(x)}\right] \\ &= \{\mathcal{A}, \mathcal{C}\}\mathcal{B} + \mathcal{A}\{\mathcal{B}, \mathcal{C}\}, \end{aligned} \tag{3.5.90}$$

  where we made use of the fact that $\mathcal{A}$ and $\mathcal{B}$ do not depend on the integration variable x^i and can hence be pulled out of the integrals. Naturally, by the anti-symmetry of the Poisson bracket, this identity generalises to the second argument,

$$\{\mathcal{A}, \mathcal{B}\mathcal{C}\} = \mathcal{B}\{\mathcal{A}, \mathcal{C}\} + \{\mathcal{A}, \mathcal{B}\}\mathcal{C}. \tag{3.5.91}$$

  Notably, this is the same structure as one finds for commutators—as is a necessity for canonical quantisation.

[3] Here we consider Φ^A and Π_B as functions of x^i only, at fixed x^0. We have the implicit convention that all participating functions are evaluated at the same value of x^0.

- **Power Rule** From the Leibniz rule, the power rule

$$\{\mathcal{A}^n, \mathcal{B}\} = n\mathcal{A}^{n-1}\{\mathcal{A}, \mathcal{B}\}, \quad n \geq 2. \tag{3.5.92}$$

follows. This can be shown by induction, starting with the base case $n = 2$:

$$\{\mathcal{A}^2, \mathcal{B}\} = \{\mathcal{A}\mathcal{A}, \mathcal{B}\} = \mathcal{A}\{\mathcal{A}, \mathcal{B}\} + \{\mathcal{A}, \mathcal{B}\}\mathcal{A} = 2\mathcal{A}\{\mathcal{A}, \mathcal{B}\}. \tag{3.5.93}$$

We can then perform the induction step as

$$\begin{aligned}\{\mathcal{A}^{n+1}, \mathcal{B}\} &= \{\mathcal{A}\mathcal{A}^n, \mathcal{B}\} = \mathcal{A}\{\mathcal{A}^n, \mathcal{B}\} + \{\mathcal{A}, \mathcal{B}\}\mathcal{A}^n \\ &= n\mathcal{A}\mathcal{A}^{n-1}\{\mathcal{A}, \mathcal{B}\} + \{\mathcal{A}, \mathcal{B}\}\mathcal{A}^n = (n+1)\mathcal{A}^n\{\mathcal{A}, \mathcal{B}\}.\end{aligned} \tag{3.5.94}$$

- **Interaction with Derivatives** Suppose we have a y^i-dependent local functional $\mathcal{A}(y)$. We may still compute Poisson brackets such as

$$\{\mathcal{A}(y), \mathcal{B}\} = \int d^{N-1}x \left[\frac{\delta\mathcal{A}(y)}{\delta\Phi^A(x)}\frac{\delta\mathcal{B}}{\delta\Pi_A(x)} - \frac{\delta\mathcal{B}}{\delta\Phi^A(x)}\frac{\delta\mathcal{A}(y)}{\delta\Pi_A(x)}\right] \tag{3.5.95}$$

(and in fact, we already did this when computing identities (3.5.88)). This bracket $\{\mathcal{A}(y), \mathcal{B}\}$ is, in particular, a function of y^i. Thus, we can take derivatives with respect to it—let us see how they interact with the bracket. Making use of the fact that partial and functional derivatives commute, we derive

$$\begin{aligned}\partial_{y^i}\{\mathcal{A}(y), \mathcal{B}\} &= \int d^{N-1}x\, \partial_{y^i}\left[\frac{\delta\mathcal{A}(y)}{\delta\Phi^A(x)}\frac{\delta\mathcal{B}}{\delta\Pi_A(x)} - \frac{\delta\mathcal{B}}{\delta\Phi^A(x)}\frac{\delta\mathcal{A}(y)}{\delta\Pi_A(x)}\right] \\ &= \int d^{N-1}x \left[\frac{\delta\big(\partial_{y^i}\mathcal{A}(y)\big)}{\delta\Phi^A(x)}\frac{\delta\mathcal{B}}{\delta\Pi_A(x)} - \frac{\delta\mathcal{B}}{\delta\Phi^A(x)}\frac{\delta\big(\partial_{y^i}\mathcal{A}(y)\big)}{\delta\Pi_A(x)}\right] \\ &= \{\partial_i\mathcal{A}(y), \mathcal{B}\}.\end{aligned} \tag{3.5.96}$$

- **Interaction with Integrals:** By the same reasoning as for derivatives—because integration and functional derivation can be exchanged—it can be shown that

$$\int d^{N-1}y \{\mathcal{A}(y), \mathcal{B}\} = \left\{\int d^{N-1}y\, \mathcal{A}(y), \mathcal{B}\right\}. \tag{3.5.97}$$

This is particularly useful for brackets with the Hamiltonian, as one has

$$\{\mathcal{A}, H\} = \left\{\mathcal{A}, \int d^{N-1}y\, \mathcal{H}(y)\right\} = \int d^{N-1}y\, \{\mathcal{A}, \mathcal{H}(y)\} \tag{3.5.98}$$

which turns calculations with the Hamiltonian into calculations with the Hamiltonian density which are typically simpler to handle.

For convenience, we summarise the above results in a cheat sheet:

$$\{\mathcal{A}\mathcal{B}, \mathcal{C}\} = \mathcal{A}\{\mathcal{B}, \mathcal{C}\} + \{\mathcal{A}, \mathcal{C}\}\mathcal{B}, \tag{3.5.99}$$

$$\{\mathcal{A}^n, \mathcal{B}\} = n\mathcal{A}^{n-1}\{\mathcal{A}, \mathcal{B}\}, \tag{3.5.100}$$

$$\{\partial_i\mathcal{A}(y), \mathcal{B}\} = \partial_{y^i}\{\mathcal{A}(y), \mathcal{B}\}, \tag{3.5.101}$$

$$\left\{\int d^{N-1}y\, \mathcal{A}(y), \mathcal{B}\right\} = \int d^{N-1}y\, \{\mathcal{A}(y), \mathcal{B}\} \tag{3.5.102}$$

To get more familiar with these calculation rules, let us reconsider the free scalar field theory

$$\mathcal{L} = -\eta^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - \frac{1}{2}m^2\varphi^2, \tag{3.5.103}$$

for which we have previously derived the canonical momentum $\pi(x) = \partial_0\varphi(x)$ as well as the Hamiltonian

$$\mathcal{H}(x) = \frac{1}{2}\pi^2 + \frac{1}{2}\delta^{ij}\partial_i\varphi\partial_j\varphi + \frac{1}{2}m^2\varphi^2. \tag{3.5.104}$$

We now recompute the canonical equations of motion using Poisson brackets, purely algebraically —using the rules above. This leads to

$$\begin{aligned}
\partial_0\varphi(x) &= \{\varphi(x), H\} = \int d^{N-1}y\{\varphi(x), \mathcal{H}(y)\} \\
&= \int d^{N-1}y\left[\frac{1}{2}\{\varphi(x), \pi(y)^2\} + \frac{1}{2}\delta^{ij}\underbrace{\{\varphi(x), \partial_i\varphi(y)\partial_j\varphi(y)\}}_{=0} + \frac{1}{2}m^2\underbrace{\{\varphi(x), \varphi(y)^2\}}_{=0}\right] \\
&= \int d^{N-1}y\,\pi(y)\{\varphi(x), \pi(y)\} = \int d^{N-1}y\,\pi(y)\delta^{(N-1)}(x-y) = \pi(x),
\end{aligned} \tag{3.5.105}$$

which is the correct equation of motion for φ. Let us repeat the same for π:

$$\begin{aligned}
\partial_0\pi(x) &= \{\pi(x), H\} = \int d^{N-1}y\{\pi(x), \mathcal{H}(y)\} \\
&= \int d^{N-1}y\left[\frac{1}{2}\underbrace{\{\pi(x), \pi(y)^2\}}_{=0} + \frac{1}{2}\delta^{ij}\{\pi(x), \partial_i\varphi(y)\partial_j\varphi(y)\} + \frac{1}{2}m^2\{\pi(x), \varphi(y)^2\}\right] \\
&= \int d^{N-1}y\left[\delta^{ij}\partial_{y^i}\varphi(y)\underbrace{\{\pi(x), \partial_j\varphi(y)\}}_{=-\partial_{y^j}\delta^{(N-1)}(x-y)} + m^2\varphi(y)\underbrace{\{\pi(x), \varphi(y)\}}_{=-\delta^{(N-1)}(x-y)}\right] \\
&= \int d^{N-1}y\big[-\delta^{ij}\partial_{y^i}\varphi(y)\partial_{y^j}\delta^{(N-1)}(x-y) - m^2\varphi(y)\delta^{(N-1)}(x-y)\big] \\
&= \int d^{N-1}y\big[\delta^{ij}\partial_{y^i}\partial_{y^j}\varphi(y) - m^2\varphi(y)\big]\delta^{(N-1)}(x-y) \\
&= \Delta\varphi(x) - m^2\varphi(x)
\end{aligned} \tag{3.5.106}$$

which also reproduces our result from equation (3.1.10). Notice that in these derivations we never had to make explicit use of the Poisson bracket's definition—it was possible to simply use the established rules of calculation as well as the fundamental brackets of the canonical variables.

In this sense, the Hamiltonian formalism can be viewed as *generated by* the fundamental relation

$$\big\{\Phi^A(x), \Pi_B(y)\big\} = \delta^A_B\delta^{(N-1)}(x-y), \tag{3.5.107}$$

together with the dynamical prescription

$$\partial_0\mathcal{A} = \{\mathcal{A}, H\} \tag{3.5.108}$$

for any observable $\mathcal{A}$, and the algebraic calculation rules derived above.