Appendix A

Constrained Dynamics

In the current Appendix, we will introduce the key concept of constrained dynamics, indispensable for any system with gauge redundancy. A common feature of gauge theories in general is that they are formulated with more field variables than the actual degrees of freedom counting suggests. There exists an elaborate Hamiltonian formulation of such theories with extra degrees of freedom built-in, due to Dirac, at both classical and quantum level. Although this round-about attitude, of reaching the physical degrees freedom starting from a redundant formulation, may look odd, there are much advantage in doing so if we wish to take a full benefit of the symmetry structures.

Dirac's theory of constrained dynamics we will cover in this Appendix is also important and unavoidable for theories formulated with a single time-derivative in the Lagrangian. The latter includes typical fermionic theories, although in this appendix we will fall short of the latter. Much of what happen with fermions can be seen indirectly with the Chern-Simons theories, which is our last example in this Appendix.

A.1 Mechanical Prototypes

Let us first recall the Hamiltonian view on classical mechanics, starting with the multi-particle action,

$$\int L\,dt = \int \left(\frac{1}{2}M_{ab}(q)\dot{q}^a\dot{q}^b - V(q)\right) dt \ , \tag{A.1.1}$$

with its Euler-Lagrange equation

$$\delta \int L\, dt = 0 \quad \rightarrow \quad \frac{d}{dt}(M_{ab}\dot{q}^b) = \frac{1}{2}\partial_a M_{cd}\dot{q}^c\dot{q}^d - \frac{\partial V}{\partial q^a} \ . \tag{A.1.2}$$

The same can be done alternatively in a first-order formulation with canonical variables,

$$\begin{aligned} p_a \ &\equiv \ \frac{\delta}{\delta \dot{q}^a}\int L \ , \\ H(q;p) \equiv (p_a\dot{q}^a - L(q,\dot{q}))\Big|_{\text{extremize w.r.t } \dot{q}^a} &= \frac{1}{2}(M^{-1})^{ab}p_a p_b + V(q) \ . \end{aligned} \tag{A.1.3}$$

It follows also that

$$L(q,\dot{q}) \ = \ (p_a\dot{q}^a - H(q;p))\Big|_{\text{extremize w.r.t } p_a} \ . \tag{A.1.4}$$

One may take a reverse view by taking

$$p_a\dot{q}^a - H(q;p) \tag{A.1.5}$$

as the definition of the Lagrangian, now with q and p considered as fundamental variables, and extremize

$$\begin{aligned} \frac{\delta}{\delta p}\int(p\dot{q} - H(q;p))dt = 0 \ , \\ \frac{\delta}{\delta q}\int(p\dot{q} - H(q;p))dt = 0 \end{aligned} \tag{A.1.6}$$

and obtain the evolution equation for p and q as

$$\begin{aligned} \frac{\delta}{\delta p_a} \quad &\rightarrow \quad \dot{q}^a = \frac{\delta H}{\delta p_a} = (M^{-1})^{ab}p_b \ , \\ \frac{\delta}{\delta q^a} \quad &\rightarrow \quad \dot{p}_a = -\frac{\delta H}{\delta q^a} = -\frac{1}{2}\partial_a(M^{-1})^{cd}p_c p_d - \frac{\partial V}{\partial q^a} \ . \end{aligned} \tag{A.1.7}$$

This choice of variables where the evolution is dictated by H is what we mean by Hamiltonian dynamics. This is clearly equivalent to the original Euler-Lagrange equation; Combining the two and using

$$\partial_a M^{-1} = -M^{-1}(\partial_a M)M^{-1} \ , \tag{A.1.8}$$

we recover the above second-order equation of motion easily.

The above Hamiltonian dynamics has an elegant geometrical realization that starts with the Poisson Bracket defined for functions on the phase space, i.e., the space spanned by q^a's and p_a's. Given a pair of functions f and g on the phase space, the Poisson bracket is,

$$[\, f, g\,]_{\text{P.B}} \equiv \frac{\partial f}{\partial q^a}\frac{\partial g}{\partial p_a} - \frac{\partial f}{\partial p_a}\frac{\partial g}{\partial q^a} \tag{A.1.9}$$

In particular the coordinates and the conjugate momenta thereof obey,

$$[\, q^a, p_b\,]_{\text{P.B.}} = \delta^a_b \tag{A.1.10}$$

which of course elevates to the quantum canonical commutator by multiplying the right hand side by $\mathbb{i}\hbar$.

The above canonical equations of the motion are then recast as

$$\dot{q}^a = [\, q^a, H(q;p)\,]_{\text{P.B.}} \ , \qquad \dot{p}_a = [\, p_a, H(q;p)\,]_{\text{P.B.}} \tag{A.1.11}$$

Since all dynamical variables are functions of q's and p, this means that

$$\dot{f}(q;p) = [\, f(q;p), H(q;p)\,]_{\text{P.B.}} \tag{A.1.12}$$

This phase space dynamics can also be interpreted more geometrically in terms of the symplectic structure,

$$\Omega = \sum_a dq^a \wedge dp_a \ , \qquad d\Omega = 0 \ , \tag{A.1.13}$$

and the so-called Hamiltonian flow thereof. We will not dwell on this geometric interpretation of the Hamiltonian dynamics, as it is available almost in any graduate textbooks on classical mechanics.

A.1.1 A Prototype L_1

The most familiar prototype of constrained dynamics are mechanics equipped with Lagrange multipliers. As we are familiar from classical mechanics, sometimes it is advantageous to keep redundant set of variables only to impose a restriction. For instance, take a free particle motion on a circle,

$$\begin{aligned} L_1 &= \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - \frac{\lambda}{2}(x^2 + y^2 - R^2) \\ &= \frac{m}{2}\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{\lambda}{2}(r^2 - R^2) \end{aligned} \tag{A.1.14}$$

with

$$H_1 = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) + \frac{\lambda}{2}(r^2 - R^2) \ . \tag{A.1.15}$$

The Lagrange multiplier λ imposes

$$r^2 - R^2 = 0 \ . \tag{A.1.16}$$

Although one may immediately reduce this as $x + iy = Re^{\mathrm{i}\theta}$ and keep θ as the only surviving variable,

$$L_1' = \frac{mR^2}{2}\dot{\theta}^2 \ , \qquad H_1' = \frac{1}{2mR^2}p_\theta^2 \ , \tag{A.1.17}$$

we will stick to this redundant form, L_1, as it proves to be quite instructive.

Recasting this into the canonical form

$$L_1 = p_r\dot{r} + p_\theta\dot{\theta} + p_\lambda\dot{\lambda} - H_1(r, \theta, \lambda; p_r, p_\theta, p_\lambda) \ , \tag{A.1.18}$$

we immediately find a constraint,

$$\varphi_1 \equiv p_\lambda \quad \rightarrow \quad \varphi_1 \approx 0 \ , \tag{A.1.19}$$

where we introduced $\approx$ for the so-called weak equality, meaning that the equality holds upon imposing the complete set of constraints. By definition, classical trajectories live on the hypersurface carved out by the constraints.

$\varphi_1 \approx 0$ is called the primary constraint in that its follows immediately when we move over to the canonical formulation. Starting with such a primary constraint we often discover secondary constraints as its consequences. For example, $\varphi_1 \approx 0$ has to mean $\dot{\varphi}_1 \approx 0$ as well. We actually already knew this since

$$\varphi_2 \equiv -\dot{\varphi}_1 = -[\,p_\lambda, H_1\,]_{\text{P.B.}} = \frac{r^2 - R^2}{2} \tag{A.1.20}$$

with the innocent overall sign for later convenience, is the λ equation of motion. So we have

$$\varphi_2 \approx 0 \ , \tag{A.1.21}$$

as a secondary constraint. It does not stop here since we need to make sure $\dot{\varphi}_2 \approx 0$ as well, which gives,

$$\varphi_3 \equiv \dot{\varphi}_2 = \frac{1}{2}[\,r^2 - R^2, H_1\,]_{\text{P.B.}} = \frac{rp_r}{m} \ . \tag{A.1.22}$$

$\varphi_3 \approx 0$ demands the radial momentum to be absent, as to be expected for a circular motion.

We have one more step to go,

$$\dot{\varphi}_3 = \frac{1}{m}[\,rp_r, H_1\,]_{\text{P.B.}} = 2\left(\frac{H_1}{m} - \frac{\lambda}{m}(2\varphi_2 + R^2/2)\right) \ , \tag{A.1.23}$$

so we might as well take

$$\varphi_4 \equiv \frac{H_1}{m} - \frac{\lambda R^2}{2m} \ . \tag{A.1.24}$$

$\varphi_4 \approx 0$ determines the value of the Lagrange multiplier λ, given a trajectory. Since H commute with itself, the generation of secondary constraints stops here. Now we can self-consistently impose

$$\varphi_{1,2,3,4} \approx 0 \ , \tag{A.1.25}$$

bringing us back effectively to a single pair of canonical coordinates θ and its momentum p_θ, as in $x + \mathbb{i}y = Re^{\mathbb{i}\theta}$.

A significant fact to note for the prototype is that the Poisson bracket among these constraints do not vanish even weakly,

$$[\varphi_1, \varphi_4]_{\text{P.B.}} \approx \frac{R^2}{2m}, \qquad [\varphi_2, \varphi_3]_{\text{P.B.}} \approx \frac{R^2}{2m}. \tag{A.1.26}$$

Other combinations vanish weakly. From this one can see that the matrix,

$$[\varphi_i, \varphi_j]_{\text{P.B.}} \tag{A.1.27}$$

is a non-singular matrix, even weakly. Constraints of this type are called the second class. Next, we come to classification of constraints to the first-class and the second class, by Dirac, and how one may deal with such second-class constraints more effectively.

Before proceeding further, we wish to warn the readers that the procedure we used above for L_1 and H_1 is actually incomplete. This has something to do with a necessity to modify the Hamiltonian H to include terms linear in the constraints, which can affect generation of secondary constraints and how we view λ above. Although we will initially work with H_1 and the four accompanying constraint $\varphi_{1,2,3,4}$, a more systematic approach would have treated the dynamics a little differently. We will revisit L_1 once each in next two subsections, to illustrate what we mean by this. Nevertheless, this more systematic alternatives do not represent corrections, but merely alternatives.

A.2 First-Class vs. Second-Class

A.2.1 The Total Hamiltonian and the Dirac Bracket

Of course one naturally wonders why in the world we have gone through this gymnastics where, at least with the prototype here, we may obviate all these by solving λ equation first $x + \mathbb{i}y = Re^{\mathbb{i}\theta}$ and keeping θ only. The answer to this is two-fold. One is that it is not always easy, or even downright impossible to "solve" in many cases. The latter in particular applies to all of fermionic quantum mechanics and fermionic quantum field theories. The other is that there are certain kinds of constraints, to be studied in next subsection, called the first-class constraints connected to gauge redundancies. Here again, 'solving" first to remove the unphysical part of the variables is often cumbersome but, more importantly, not profitable.

These two kinds of situations may be separated via a classification of the constraints themselves. Given the Poisson bracket we started with, suppose we compute the pairwise bracket of these constraints,

$$[\boldsymbol{\varphi}_I, \boldsymbol{\varphi}_J]_{\text{P.B.}} , \tag{A.2.1}$$

and ask whether these vanish weakly, i.e., vanish on the hypersurface $\varphi_J \approx 0$. For some pairs, it may while for others it may not.

The constraints are classified into the first-class and the second-class by collecting a maximal subset whose element commute with all the constraints. Such commuting constraints are called the first class, while the rest are called the second class. Since a first-class constraint must commute with all possible constraints, and we may determine the first-class constraints only after determining the entire set of $\boldsymbol{\varphi}$'s.

From now on, we will denote these two types of constraints by different symbols,

$$\boldsymbol{\varphi}_I \qquad \rightarrow \qquad \boldsymbol{\chi}_A \ , \ \ \varphi_i \ . \tag{A.2.2}$$

In other words,

$$[\boldsymbol{\chi}_A, \boldsymbol{\chi}_B]_{\text{P.B.}} \approx 0 \ , \qquad [\boldsymbol{\chi}_A, \varphi_i]_{\text{P.B.}} \approx 0 \ , \tag{A.2.3}$$

while $[\varphi_i, \varphi_j]_{\text{P.B.}}$ is nonsingular on the constrained surface. Given the antisymmetric property of the Poisson bracket, it should be clear that the number of second-class constraints, φ_i's, are always even.* With our prototype example above, all four $\varphi_{i=1,2,3,4}$ are second-class, although we will later see that there are alternative descriptions where $\varphi_{1,4}$ are traded off in favor of a single first-class constraint.

One puzzling aspect of dynamics with the constraints, at the first sight, is that physical expressions are in principle ambiguous along the directions normal to the constrained hypersurface. On the other hand, we cannot impose the constraints too early, since a Poisson bracket involving a constraint need not always vanish even if we impose all the constraints. In fact the Hamiltonian is itself a function on the phase

*For mechanics and field theories based on Grassmann numbers, necessary for fermions, this is no longer true as the Poisson bracket becomes symmetric.

space, so one must wonder whether the following type of shift

$$H \quad \to \quad \mathbb{H} \equiv H + \sum_I \boldsymbol{\lambda}_I \boldsymbol{\varphi}_I \ , \tag{A.2.4}$$

makes any difference. In fact, the prototype Hamiltonian H_1 came with $\lambda\varphi_2$ built in, so why not do the same with other constraints. The rule of the game can become confusing rather quickly.

For a clarification why such shifts are often necessary, let us consider another simple example,

$$L_2 = y\dot{x} - \frac{1}{2m}y^2 - V(x) \ , \qquad H_2 = \frac{1}{2m}y^2 + V(x) \ , \tag{A.2.5}$$

which is clearly equivalent to, upon using equation of motion for y,

$$L_2' = \frac{m}{2}\dot{x}^2 - V(x) \ , \qquad H_2' = \frac{1}{2m}p_x^2 + V(x) \ , \tag{A.2.6}$$

the usual Newtonian mechanics for a particle of mass m on a line and the potential $V(x)$. As such, we can expect the constraints of L_2 to reduce the canonical degrees of freedom by two. Indeed, L_2 comes with a pair of primary and second-class constraints,

$$\varphi_1 \equiv p_y \approx 0 \ , \qquad \varphi_2 \equiv p_x - y \approx 0 \ , \qquad [\,\varphi_1, \varphi_2\,]_{\text{P.B.}} = 1 \ . \tag{A.2.7}$$

However, an immediate puzzle appears from how $\dot{\varphi}_1$ seems to generate a secondary constraint, since

$$[\,p_y, H_2\,]_{\text{P.B.}} = -\frac{y}{m} \ , \tag{A.2.8}$$

whose weakly vanishing means $p_x = 0$ or $\dot{x} = 0$, contrary to the dynamical content of L_2'. If we had used H_2' in place of H_2, the result looks proper, but this would have been an illegal thing to do since we decided to keep all four canonical degrees of freedom in play when we were asking about the above Poisson bracket with p_y in it.

This oddity tells us that the time-evolution with the Poisson bracket, driven by the Hamiltonian derived in the usual manner, is not the entire story. Suppose that we decided to extend the Hamiltonian as

$$\mathbb{H}_2 = H_2 + u_1\varphi_1 + u_2\varphi_2 \ , \tag{A.2.9}$$

with unknown functions $u_{1,2}$, perform the Poisson bracket with this "total" Hamiltonian for the time evolution as

$$\dot{\varphi}_1 = [\,\varphi_1, \mathbb{H}_2\,]_{\text{P.B.}} = -\frac{y}{m} + u_2 \ , \qquad \dot{\varphi}_2 = [\,\varphi_2, \mathbb{H}_2\,]_{\text{P.B.}} = -u_1 - V'(x) \ . \tag{A.2.10}$$

Note that demanding $\dot{\varphi}_{1,2} = 0$ actually fixes the value of $u_1 = -V'(x)$, $u_2 = y/m$ here, instead of generating secondary constraints.

These u's are also called the Lagrange multipliers, as it multiplies the constraint, just as λ did for H_1. This is in part motivated by how in the very first example, we had the constraint φ_2 already appearing additively in the Hamiltonian, multiplied by λ. The phenomenon of the Lagrange multipliers becoming fixed on shell is familiar from classical mechanics, although u's here are a little different from λ. One major difference is that u_i's added this way are not meant to be a phase space coordinate at all, while λ of L_1 and its conjugate momentum p_λ were treated so. Later, we will dwell on related subtleties after presenting a more complete picture of constrained systems.

The same can be done with the prototype L_1 above, with $\mathbb{H}_1$ with all constraints and accompanying u_i's built-in. For this example, however, we need to keep in mind L_1 came with a Lagrange multiplier of its own λ times the constraint called φ_2 in that example. Because of this, one has two different options for the total Hamiltonian $\mathbb{H}_1$, depending on whether or not we treat λ as part of initial phase space variables or not. One simple option is that we consider λ as u_2 from the beginning in the sense of the "total" Hamiltonian $\mathbb{H}_1$ with its conjugate momentum disregarded.

In the latter viewpoint, λ and $\varphi_1 = p_\lambda$ are not part of the unconstrained phase space, so that only φ_2 and φ_3 need to be counted as the constraints. In the latter viewpoint, our "total" Hamiltonian is

$$\mathbb{H}_1' \ = \ H_1\Big|_{\lambda\to u_2} + u_3\varphi_3 \ . \tag{A.2.11}$$

We may do this if we decide to forget about the original Lagrangian side. Either way, u_2 is fixed by either $\dot{\varphi}_3 = [\,\varphi_3, \mathbb{H}_1\,]_{\text{P.B.}} \approx 0$ or $\dot{\varphi}_3 = [\,\varphi_3, \mathbb{H}_1'\,]_{\text{P.B.}} \approx 0$, resulting in the same dynamics in the end.†

†The other option of treating λ as part of the phase space, now with the total Hamiltonian understood, incurs yet another line of thought to be pursued in next subsection, where we need to deal with cases where some Lagrange multipliers are not fixed but rather become a gauge ambiguity.

The "total" Hamiltonian thus can contain many new u's which do not appear in the original Lagrangian and some of them seem essential while others are mere formality with not much of consequences. Some of them starts out a part of the configuration variable, as in λ of L_1. Others, like u_i's above are just unknown functions on the phase space, only to be determined on-shell by the time-independence of the second -class constraints. One can also choose to elevate the latter type of Lagrange multipliers u_i's in $\mathbb{H}$ to canonical variables, as long as we make sure to treat its conjugate momenta p_{u_i} as constraints, whose equation of motion yields $\varphi_i \approx 0$ to begin with. For each u we treat as if dynamical, this of course enlarges the phase space by two more canonical variables, so the process could be repeated forever, although somewhat meaninglessly as it does not affect the true dynamics on the constrained subspace.

One must wonder if all these formalisms, with potentially never-ending yet meaningless growth of the phase space, are worth the trouble. Although there is no real ambiguity here of dynamical consequence, the formalism with the Poisson bracket and the second-class constraints can become quite heavy, more so if we view each of u's in the "total" Hamiltonian as part of the phase space prior to the constraints. Somehow it would be better if we can find a simpler way where the Lagrange multipliers can be evaded altogether. Much of these complications and confusions often originate from the question of when it is safe to impose the constraint and when it is premature to do so, which in turn is due to how the Poisson bracket of a pair of constraint need not vanish even weakly.

Fortunately, there exists a clever alternative to the Poisson bracket that can bypass much of these procedural headaches associated with the second-class constraints. This modified commutator is called the Dirac bracket,

$$[\,f,g\,]_{\rm Dirac} \;\equiv\; [\,f,g\,]_{\rm P.B.} - [\,f,\varphi_i\,]_{\rm P.B.}(C^{-1})^{ij}[\,\varphi_j,g\,]_{\rm P.B.}\;, \tag{A.2.12}$$

with the non-singular matrix,

$$C_{ij} \;\equiv\; [\,\varphi_i,\varphi_j\,]_{\rm P.B.}\;. \tag{A.2.13}$$

The Dirac bracket is, quite clearly, a projection of the Poisson bracket onto $\varphi \approx 0$,

As this single example L_1 shows, the constrained dynamics often involves several different routes to one and the same on-shell dynamics in the end.

which means,

$$[\, f, \varphi\,]_{\rm Dirac} \;\approx\; 0 \tag{A.2.14}$$

for any second-class constraint φ and arbitrary dynamical quantity f.

In a dynamical system with the second-class constraints, the canonical dynamics may be performed with this Dirac's modification of the Poisson bracket,

$$\dot{f}(q;p) = [\, f(q;p), H(q;p)\,]_{\rm Dirac} \ . \tag{A.2.15}$$

One should see immediately that with this modified bracket, the distinction between H and $\mathbb{H}$ disappears as far as the second class constraints are concerned, since any pieces that contain these constraints are projected out. All these extra baggages and ambiguities due to the second-class constraints can be conveniently removed by the Dirac bracket, whereby we are also free to impose these second-class constraints freely, before or after the Dirac bracket.

As an immediate check, let us revisit L_2 above and ask for the equation of motion for the remaining x and p_x. If we use the Poisson bracket and happen to use H_2 rather than $\mathbb{H}_2$, we end up with

$$\dot{p}_x \approx \dot{y} = [\, y, H_2\,]_{\rm P.B.} = 0 \ . \tag{A.2.16}$$

which is clearly a nonsense in view of L_2'. This can be remedied most economically by using the Dirac bracket instead,

$$\begin{aligned} \dot{p}_x \approx \dot{y} \;\; &= \;\; [\, y, H_2\,]_{\rm Dirac} \;=\; [\, y, H_2\,]_{\rm P.B.} - [\, y, \varphi_1\,]_{\rm P.B.}[\, \varphi_2, \varphi_1\,]^{-1}_{\rm P.B.}[\, \varphi_2, H\,]_{\rm P.B.} \\ &= \;\; 0 - [\, y, p_y\,]_{\rm P.B.} \times (-1) \times [\, p_x - y, H\,]_{\rm P.B.} \;=\; -V'(x) \ , \end{aligned} \tag{A.2.17}$$

which is precisely the expected equation of motion for x from L_2'. We can see the same works for L_1; the off-constraint variable r could have easily gone wrong if we had used the Poisson bracket with H_1.

A word of caution is in order. As is clear from the definition of the Dirac bracket, it may be defined only after the complete set of constraints are collected and classified. The Poisson bracket remains the fundamental symplectic structure that governs the Hamiltonian dynamics. It is also with the Poisson bracket that we need to use for discovering the secondary constraints, although we must continue to update $\mathbb{H}$ along the process. The Dirac bracket is not a new symplectic structure we introduced ad hoc but rather an equivalent but cleaner version of the same symplectic structure which obviates much of the unnecessary computational baggages.

One can also come away with the impression that these examples $L_{1,2}$ are artificial, as, in both, we have a perfectly sensible way to reduce the degrees of freedom to unconstrained ones only. However, these simple examples are also quite prototypical; we will encounter field theory examples with single time-derivative, both bosonic and fermionic, where the Dirac bracket becomes practically indispensable.

A.2.2 First-Class and Gauge Redundancies

Along the way we discover constraints, we have seen that we should take into account the potential ambiguities associated with the off-constraint directions. This led us to consider the "total" Hamiltonian,

$$\mathbb{H} \;=\; H + \sum_I \boldsymbol{\lambda}_I \boldsymbol{\varphi}_I \;=\; H + \sum_i u_i \varphi_i + \sum_A \Lambda_A \boldsymbol{\chi}_A \;, \tag{A.2.18}$$

we see that only u_i's and not Λ_A's appear on the right hand side of $[\,\boldsymbol{\varphi}_I, \mathbb{H}]_{\text{P.B.}}$ since $\boldsymbol{\chi}_A$'s commute with all constraints weakly. This means that the coefficients u's of the second-class constraints get fixed on shell.

Alternatively, we have seen also how the Dirac bracket allows us to simplify the dynamics on the constrained hypersurface, by effectively removing φ_i directions from the dynamics. Unless one is for some reason interested in on-shell values of u_i's, the latter is far easier route to the dynamics; the Dirac bracket allows us to forget about φ_i directions and in particular to drop $\sum_i u_i \varphi_i$ from $\mathbb{H}$.

Either way, we are left with the first-class constraint terms $\sum_A \Lambda_A \boldsymbol{\chi}_A$ in $\mathbb{H}$. Although they commute weakly among themselves, the commutators between them and generic dynamical variables vanish do not vanish weakly, so these terms in $\mathbb{H}$ seemingly generate arbitrary shift of variables proportional to Λ_A's.

$$\begin{aligned} \dot{f} \;&=\; [\,f, H\,]_{\text{Dirac}} + \sum_A \,[\,f, \Lambda_A \boldsymbol{\chi}_A\,]_{\text{Dirac}} \\ &\approx\; [\,f, H\,]_{\text{Dirac}} + \sum_A \;\Lambda_A\,[\,f, \boldsymbol{\chi}_A\,]_{\text{Dirac}} \;. \end{aligned} \tag{A.2.19}$$

On the flip side of the coin, there was no compelling reason why we kept these terms in $\mathbb{H}$, unlike those of φ_i's, so the shift due to such terms should not be physically significant either.

The only way to make sense of this odd situation is to say that the canonical variables we started with is ambiguous physically, and we must consider classical trajectories that differ by such shift due to $\boldsymbol{\chi}_A$'s as physically equivalent solutions. Take for instance a trivial toy model,

$$L_3 \;=\; \frac{1}{2}(x\dot{x}+y\dot{y})^2 \;=\; \frac{1}{2}(r^2\,\dot{r}^2)\;, \qquad H_3 \;=\; \frac{p_r^2}{2r^2}\;. \tag{A.2.20}$$

Since the angle θ of $x+\mathbb{i}y = re^{\mathbb{i}\theta}$ does not appear anywhere, we have

$$\boldsymbol{\chi} \equiv p_\theta \tag{A.2.21}$$

as the primary constraint, which commutes with H_3 and with itself and thus does not generates a secondary constraint. The total Hamiltonian is then,

$$\mathbb{H}_3 = H_3 + \Lambda p_\theta \tag{A.2.22}$$

While we have the option of forgetting about θ and p_θ entirely, suppose that for some reason we must keep x and y as the variables. H_3 dictates the radial motion,

$$\begin{aligned}\dot{r} \;&=\; [\,r, \mathbb{H}_3\,]_{\text{P.B.}} \;=\; [\,r, H_3\,]_{\text{P.B.}} \;=\; \frac{p_r}{r^2}\;,\\ \dot{p}_r \;&=\; [\,p_r, \mathbb{H}_3\,]_{\text{P.B.}} \;=\; [\,p_r, H_3\,]_{\text{P.B.}} \;=\; \frac{p_r^2}{r^3}\;,\end{aligned} \tag{A.2.23}$$

whereas the angular part

$$\dot{x}+\mathbb{i}\dot{y} \;=\; [\,re^{\mathbb{i}\theta}, \mathbb{H}_3\,]_{\text{P.B.}} \;=\; \mathbb{i}\Lambda\, re^{\mathbb{i}\theta} \tag{A.2.24}$$

rotate the phase by the integration of the arbitrary function Λ. Of course there is nothing strange about this, since the Lagrangian did not depend on the phase part of $x+\mathbb{i}y = re^{\mathbb{i}\theta}$ at all. This merely reminds us that we started with redundant set of variables, x and y, instead of the single physical variable r.

The toy model teaches us how the first-class constraints remove degrees of freedom. Note that unlike the previous examples with the second-class constraints only, we find exactly one first-class constraint whereas there are two canonical variables θ and p_θ which are unphysical. The single first-class constraint achieve this task, first but imposing $\varphi \approx 0$, and then performing an arbitrary shift of θ and basically telling us not to worry; Configurations with the same $r(t)$ but mutually different phase angles are all supposed to be declared to be one and the same physical trajectories. As such, the canonical degrees of freedom is reduced by twice the number of the first-class constraints.

Combining this with what we learned of the second-class constraint, we conclude that the canonical degree of freedom is reduced by

$$2 \times \# \text{ of } \boldsymbol{\chi}_A \; + \; \# \text{ of } \varphi_i \; . \tag{A.2.25}$$

In terms of configuration space variables, the reduction is by

$$\# \text{ of } \boldsymbol{\chi}_A \; + \; \frac{1}{2} \times \# \text{ of } \varphi_i \; , \tag{A.2.26}$$

instead, which makes sense as the number of second-class constraints is even. For fermionic systems the latter number can be actually odd due to how the Poisson bracket is symmetric rather than anti-symmetric, but in these cases there is no sensible definition of "configuration space" variables, so we safely resort to the first counting.

Finally, with the Dirac bracket (or with the Poisson bracket if φ_i's are absent), note that the weakly commuting nature of $\boldsymbol{\chi}_A$ means

$$[\boldsymbol{\chi}_A, \boldsymbol{\chi}_B]_{\text{Dirac}} \; = \; \sum_C f_{ABC}\, \boldsymbol{\chi}_C + O(\varphi^2) \; , \tag{A.2.27}$$

so they span an algebra of some kind. As we will see presently, in the prototypical field theory examples, the algebra in question is the infinite-dimensional symmetry algebra associated with position-dependent gauge transformations. Although one often refers to the latter as the gauge symmetry, it should be really called the gauge redundancies in view of what we have seen above.

A.3 Ambiguities and Subtleties

Before turning to field theory examples, we wish to illustrate a point here by revisiting a couple of examples above. As we saw earlier with L_1, how we handle constraints and Lagrange multiplier is open to some ambiguity, and in fact the constraint classification itself and what should be considered the constraints are not entirely rigid. Without affecting the dynamics on constraints, the off-constraint part of the phase space comes with some amount of ambiguity.

In fact, if one wishes, we can continue to add more phase variables and more constraints indefinitely, by deciding to treat u_i's as part of the phase space coordinates on par with λ of L_1 example. The ambiguity in question goes beyond such relatively trivial one, and in fact allows us to swap first-class constraints to second-class one and vice versa. In this last part of the section, we will revisit L_1 and L_3 for illustration.

Revisiting L_1: a Single First-Class instead of a Pair of Second-Classes

The question of whether we treat Lagrange multiplier as phase space coordinate or not is often ambiguous. For illustration of this question, L_1 is a little special among the three examples we displayed in that the Lagrangian include an explicit Lagrange multiplier λ. Note how we use the same name "Lagrange multiplier" for the ones that appear in the Lagrangian, such as this λ of L_1, and those we use for extending the Hamiltonian H to $\mathbb{H}$, say, $\{\boldsymbol{\lambda}_I\} = \{u_i\} \cup \{\Lambda_A\}$. This common name is customary, yet the two are rather different. λ and its conjugate p_λ started out as part of the phase space, while u's and Λ's are functions on the latter.

The question of how to view λ was the key to the two different ways we dealt with L_1 dynamics earlier. The first was to regard the Lagrange multiplier λ and its conjugate p_λ as part of phase space, and impose four second-class constraints, $\varphi_{1,2,3,4} \approx 0$. Alternatively, if we disregard the Lagrangian side, we could have replaced λ by u_2 for φ_2. This way, we start with only r, θ, and their conjugate momenta as phase space variables, and impose two second-class constraints $\varphi_{2,3} \approx 0$. The total Hamiltonian in this picture was $\mathbb{H}_1'$ in (A.2.11). The final dynamics remain the same, regardless of these two choices.

We wish to point out yet another description, which in retrospect is perhaps more in line with the general procedure we learned above. There was a potential problem with the first procedure, because we did the exercise before learning about the necessity of the total Hamiltonian $\mathbb{H}$. If we had updated the Hamiltonian to include the secondary constraints as soon as we discovered them, we would have computed $\dot{\varphi}_3$ by performing the Poisson bracket of φ_3 against

$$H_1 + u_1 p_\lambda + u_2 \frac{r^2 - R^2}{2} + u_3 \frac{r p_r}{m} \tag{A.3.1}$$

and realized that $\dot{\varphi}_3 \approx 0$ fixes u_2 via λ and H_1, instead of generating a new constraint. This is because H_1 itself contains $\lambda\,(r^2 - R^2)/2$, so whatever fix u_2 will actually fix $\lambda + u_2$. If we did this, we ended up with three constraints $\varphi_{1,2,3}$, instead of the four second-class $\varphi_{1,2,3,4}$.

On the other hand, with this smaller set, $\varphi_1 = p_\lambda$ becomes a first-class, as it commutes with $\varphi_{2,3}$. As such, $\varphi_1 = p_\lambda$ should be denoted as $\boldsymbol{\chi}_1$, and the dynamics is dictated by

$$\mathbb{H}_1'' = H_1 + \Lambda_1 \boldsymbol{\chi}_1 + u_2 \varphi_2 + u_3 \varphi_3 \tag{A.3.2}$$

The degrees of counting is the same as before; we have one first-class $\boldsymbol{\chi}_1 = p_\lambda$ and two second-class ones $\varphi_{2,3}$, removing the four canonical degrees of freedom in total again. λ remains undetermined, even after the classical trajectories are fixed, but this is precisely what we should expect since p_λ is a first-class constraint and its shifting action on λ is interpreted as a gauge redundancy. This is not what we are accustomed from the usual Lagrange multiplier already present in the action from the beginning. Nevertheless, there is nothing wrong with this 3rd description since λ does not represent an independent physical quantity, classical or quantum.

More generally, if the Lagrangian comes with a term linear in the Lagrange multiplier

$$L = \cdots - \lambda \varphi + \cdots \tag{A.3.3}$$

with no time-derivative present in φ, it can be treated systematically along the same line as with L_1 and $\mathbb{H}_1''$. $p_\lambda \approx 0$ is a primary constraint while $\varphi \approx 0$ is a secondary that follows from $\dot{p}_\lambda \approx 0$. If φ proves to be a first-class, there is not much beyond this. If φ proves to be a second-class, on the other hand, it means that there is another second-class constraints $\varphi' \approx 0$. The total Hamiltonian $\mathbb{H}$ is then include

$$\mathbb{H} = \cdots + \Lambda p_\lambda + (\lambda + u)\varphi + u'\varphi' + \cdots \tag{A.3.4}$$

where we used the notation Λ for the new Lagrange multiplier of p_λ, anticipating that p_λ will prove to be a first-class.

Without $u\varphi$ in $\mathbb{H}$, $\dot{\varphi}' \approx 0$ would have generated another secondary constraint that contains a term linear in λ. If we did this, this new constraint would be a second-class pair with p_λ. With $u\varphi + u'\phi'$ in $\mathbb{H}$, on the other hand, $\dot{\varphi}' \approx 0$ and $\dot{\varphi} \approx 0$ are now equations for u and u', which determines $(\lambda + u)$ and u' given a classical trajectories on the constrained surface, i.e. on-shell, but cannot fix λ separately. At the same time, the generation of the secondary constraint along this particular branch of the procedure ends here, and p_λ remains a first-class. Instead of fixing λ on-shell, therefore, we end up with an arbitrary gauge function Λ which shifts $\dot{\lambda}$ arbitrarily. The end result is that λ and p_λ again decouple from the dynamics, as they should.

In other words, depending on whether $u\varphi$ is included in the middle procedure, the character of p_λ and exactly how λ is dealt with on-shell differ, although the end results on the physical quantities on-shell remain unaffected. As this example shows, the constrained dynamics often come with multiple descriptions where the initial phase space, the subsequent set of constraints, and their classification can be all different. All these, even though no real changes can be found on the dynamics that matter.

Gauge Fixing L_3: a Pair of Second-Class from a Single First-Class

We saw that, depending how we treat the Lagrange multiplier present in the Lagrangian, there might be an option of trade off a pair of second-class constraints to a single first-class. In the latter option, the Lagrangian multiplier λ became an arbitrary quantity, playing the role of a gauge function. Since the value of λ is not important for the final dynamics on the constrained surface, the two options represents mere alternatives. On the flip side, there is a systematic way to convert a first-class to a pair of second-class, namely the gauge-fixing. As a simplest example, let us revisit L_3.

Recall how L_3 started with x and y as the configuration variables and constrained to radial motions only by a first-class constraint $p_\theta \approx 0$. A perfectly sensible thing to do is to introduced $\theta - \theta_0 \approx 0$ as a new constraint, whose Poisson Bracket with p_θ is not weakly zero but rather a unit. Now calling $\varphi \equiv p_\theta$ and $\tilde{\varphi} \equiv \theta - \theta_0$, we have the new total Hamiltonian,

$$\mathbb{H}_3' = H_3 + u\varphi + \tilde{u}\tilde{\varphi} \tag{A.3.5}$$

The Lagrange multipliers get fixed to $u \approx 0 \approx \tilde{u}$ via the usual time-independence of $\tilde{\varphi}$ and φ, respectively.

This gauge-fixing can be incorporated into the Lagrangian side as well, since all we need to do is to add a term

$$L_3 + \tilde{\lambda}(\theta - \theta_0) \tag{A.3.6}$$

whereby $\tilde{\lambda}$ equation of motion will impose $\theta \approx \theta_0$. The gauge symmetry $\theta \to \theta + \alpha$ is broken by such a term, but of course the whole point of gauge-fixing is to reduce such gauge redundancy. Depending on how to treat $\tilde{\lambda}$, we may end up expanding the phase space further and imposing more constraints, but by now it should be clear that such ambiguities are harmless.

The Dirac bracket which starts from

$$[\tilde{\varphi}, \varphi]_{\text{P.B.}} = 1 \tag{A.3.7}$$

becomes trivial for any expression containing θ or p_θ, so in the end, we have

$$\dot{r} = [r, H_3]_{\text{Dirac}} = \frac{p_r}{r^2}\ , \qquad \dot{p}_r = [p_r, H_3]_{\text{Dirac}} = \frac{p_r^2}{r^3} \tag{A.3.8}$$

as the only remaining dynamical content, the same as when we did not gauge-fix.

As this almost trivial example shows, a set of first-class constraint $\boldsymbol{\chi}_A$'s can be converted into second-class φ_A's when augmented by an equal number of gauge-fixing conditions, $\tilde{\varphi}_B$. The gauge-fixing condition can be added to the total Hamiltonian with their Lagrange multipliers $\tilde{u}_B$'s, or even elevated as gauge-fixing terms with $\tilde{\lambda}_B$ in the Lagrangian. The gauge fixing conditions must be such that

$$[\tilde{\varphi}_B, \varphi_A]_{\text{P.B.}} \tag{A.3.9}$$

is of full rank, so that the matrix C built from this and its transverse as the two off-diagonal blocks must be invertible. When we quantize gauged theories in practice, we always do so in the end. Given a gauged system, the choice of such gauge conditions is hardly unique; depending on what $\tilde{\varphi}_B$'s we choose, the quantization proceeds differently but in a manner that the physical part of the dynamics is not be affected by such choices.

Despite these examples, we must not think that the distinction between the first-class and the second-class are irrelevant. First-class constraints can be converted into second-class by introducing gauge-fixing conditions as if they are also constraints, as above. The reverse processes of trying to interpret non-commuting pairs of second-class constraints and half-many first-classes and their gauge-fixing condition, even if possible, are less than straightforward and more to the point may be possible at the cost of other nice properties such as the general covariance.

Requisite Genericity of the Constraints

Dirac's original exposition is often challenged by various "counter-examples", as one can easily design the Lagrangian with the intension to disrupt his procedures and classifications. Such counterexamples are often equipped with singular and fine-tuned form of Lagrangian and unlikely to be relevant in real applications; nevertheless, these cast some residual and often unnecessary doubt on the validity of Dirac's proposal. There is one particular type of such complaints that can be dealt with easily; before we move on to field theory examples a comment on it is worthwhile.

For an illustration, let us again resort back to L_1 which describes a particle on a circle. Suppose that for some reason we decided to use instead,

$$\tilde{L}_1 = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - \frac{\lambda}{2}(x^2 + y^2 - R^2)^2 \tag{A.3.10}$$

The Lagrange multiplier λ imposes the same constraint, $\varphi_2 = (x^2 + y^2 - R^2)/2 \approx 0$, but the routine of reading off the new constraint from $[p_\lambda, \tilde{H}_1]_{\text{P.B.}} \approx 0$ produces

$$\tilde{\varphi}_2 = 2(\varphi_2)^2 \tag{A.3.11}$$

as the secondary constraint, in place of φ_2.

Taking a time-derivative of $\tilde{\varphi}_2 \approx 0$, we obtain

$$\varphi_2\varphi_3 \approx 0 \tag{A.3.12}$$

in place of $\varphi_3 \approx 0$ that would have arisen from the original L_1. This expression vanishes weakly upon $\tilde{\varphi}_2 \approx 0$ since the latter implies $\varphi_2 \approx 0$ as well. Does this mean that the would-be tertiary constraint $\varphi_3 \sim x\dot{x} + y\dot{y}$ is not generated? This cannot be since there is no real change to the dynamics; $\varphi_3 \approx 0$ says the radial momentum must vanish, which clearly must follow from the fixed r^2 condition. The right thing to do is to take φ_2 again as the secondary constraint even though $(\varphi_2)^2$ is seemingly produced as the secondary constraint if we insist on $\tilde{L}_1$ and follow the usual gymnastics.

This problem is really about whether the chosen constraints are good coordinates near the constrained surfaces in the phase space. As a trivial analogy, consider a plane $z = 0$ in $\mathbb{R}^3$, which can be equally described by $z^3 = 0$ but z^3 is a bad coordinate at $z = 0$. If one tries to introduce variables that are degenerate at the constraint, one would end up with such procedural oddities, which has nothing to do with real dynamics but are merely due to bad choice of the phase space coordinates that parameterize directions normal to the constrained surface. One can see from this example that we would end up with various difficulties when we conjure up constraints with degenerate behavior in the small tubular neighborhood of the constrained surface and try to apply Dirac's routine blindly.

The constrained dynamics is as much about the off-constraint variables as on-constraint ones. Much of the Dirac procedure assumes implicitly that we can make all the constraints, which should be regarded as the normal coordinates away from the constrained surface, as nondegenerate as possible. In practice, this requirement translates to the following schematics of the Poisson brackets near the constrained surface,

$$\begin{aligned} [\varphi, \varphi]_{\text{P.B.}} &\sim 1 \, , \\ [\boldsymbol{\chi}, \boldsymbol{\chi}]_{\text{P.B.}} &\sim \boldsymbol{\chi} + \varphi^2 \, , \\ [\boldsymbol{\chi}, \varphi]_{\text{P.B.}} &\sim \boldsymbol{\chi} + \varphi \, . \end{aligned} \tag{A.3.13}$$

The Hamiltonian is also a first-class, in the sense of the weakly vanishing Poisson brackets against the constraints, so we must have $[\mathbb{H}, \varphi]_{\text{P.B.}} \sim \boldsymbol{\chi} + \varphi$ and $[\mathbb{H}, \boldsymbol{\chi}]_{\text{P.B.}} \sim \boldsymbol{\chi} + \varphi^2$ as well.

A.4 Gauge Theories in Canonical Formulation

Extending this to field theory is straightforward. All that's new is that the dynamical variables are functions of all spacetime coordinates, not just a single time coordinate. Maxwell theory with its first class constraints serves as a prototype for gauge theories in general, so we start with this example, and move on the Chern-Simons theory in $d = 3$ where second-class constraints enter as well.

A.4.1 Maxwell Theory and the Gauss Constraint

Starting the Maxwell in flat spacetime, for simplicity,

$$S_{\text{Maxwell}} \;=\; \int d^dx\; \mathcal{L}_{\text{Maxwell}} \;=\; -\frac{1}{4}\int d^dx\; F^2 \;=\; \int d^dx\; \frac{1}{2}(\vec{E}^2 - \vec{B}^2) \tag{A.4.1}$$

with

$$\vec{E} \;=\; \vec{\partial}A_0 - \partial_0\vec{A}\;, \qquad \vec{B} \;=\; \vec{\partial}\times\vec{A}\;. \tag{A.4.2}$$

Note how the time derivative of A_0 never appears anywhere. As such, the conjugate momenta,

$$\pi^\mu \;\equiv\; \frac{\delta}{\delta\dot{A}_\mu}\int\mathcal{L}\;, \tag{A.4.3}$$

are such that

$$\pi^0 = 0\;, \qquad \pi^i = -E^i\;. \tag{A.4.4}$$

The first of which implies that something different happens. The definition of the Hamiltonian or canonical formulation is to use the conjugate momenta as fundamental variables, yet one of them vanishes identically.

Nevertheless, let us try to rewrite the Lagrangian via the canonical variables,

$$\begin{aligned}\int d^dx\;\mathcal{L}_{\text{Maxwell}} \;&=\; \int d^dx\left(\pi^i\dot{A}_i + \cancel{\pi^0\dot{A}_0} - \mathcal{H}(A,\pi)\right)\Big|_{\pi^i\text{ extremization}}\\ &=\; \int d^dx\Big(\pi^i\underbrace{(\dot{A}_i - \vec{\partial}A_0)}_{\pi^i} - \mathcal{H} + \pi^i\partial_iA_0\Big)\Big|_{\pi^i\text{ extremization}}\end{aligned} \tag{A.4.5}$$

Comparing against the configuration space form of the Lagrangian, we find

$$\mathcal{H} = \frac{1}{2}(\pi^2 + B^2) + \pi^i \partial_i A_0 \ , \tag{A.4.6}$$

or by integration by part,

$$\mathcal{H} = \frac{1}{2}(\pi^2 + B^2) + A_0 \underbrace{(-\partial_i \pi^i)}_{\partial_i E_i} \ . \tag{A.4.7}$$

In other words,

$$\int d^d x \ \mathcal{L}_{\text{Maxwell}} = \int d^d x \left(\pi^i \dot{A}_i - \frac{1}{2}(\pi^2 + B^2) + A_0(\partial_i \pi^i) \right) \Bigg|_{\pi^i \text{ equation of motion}} \tag{A.4.8}$$

This shows that, in canonical variables, A_0 is a Lagrange multiplier that imposes $\partial_i \pi^i = 0$ (or $\partial_i E^i = 0$ for the Maxwell theory). The same can be seen by resorting to the Hamiltonian equation of motion,

$$0 = \partial_t \pi^0 = -\frac{\delta \mathcal{H}}{\delta A_0} = \frac{\delta}{\delta A_0}(A_0 \partial_i \pi^i) = \partial_i \pi^i \ , \tag{A.4.9}$$

where we simply used that time derivative of something that vanishes must vanish as well.

Recognizing $\pi^i = -E^i$, the last is nothing but the usual Gauss law, in the absences of charge density,

$$-\partial_i E^i = 0 \ , \tag{A.4.10}$$

it is not difficult to imagine how this extends when we add charged matter fields. We can do the latter by adding to the Lagrangian, for example,

$$\mathcal{L}_{\text{matter}} = -\eta^{\mu\nu}(\partial_\mu + \mathbb{i} q A_\mu)\phi^\dagger (\partial_\nu - \mathbb{i} q A_\nu)\phi \ , \tag{A.4.11}$$

whose conjugate momenta are

$$\begin{aligned} \Pi &= \frac{\delta}{\delta \partial_t \phi} \int \mathcal{L}_{\text{matter}} = (\partial_0 + \mathbb{i} q A_0)\phi^\dagger \ , \\ \Pi^\dagger &= \frac{\delta}{\delta \partial_t \phi^\dagger} \int \mathcal{L}_{\text{matter}} = (\partial_0 - \mathbb{i} q A_0)\phi \ . \end{aligned} \tag{A.4.12}$$

The total Lagrangian in the canonical form then has the form,

$$\int d^d x\, \mathcal{L} \;=\; \int d^d x \left(\pi^i \dot{A}_i + \Pi\dot{\phi} + \Pi^\dagger \dot{\phi}^\dagger - \mathcal{H}(A_\mu, \pi^i; \phi, \phi^\dagger, \Pi, \Pi^\dagger)\right)\Bigg|_{\pi^i,\Pi \text{ extremization}} \qquad (\text{A.4.13})$$

with

$$\mathcal{H} \;=\; \frac{1}{2}(\pi^2 + B^2) + \Pi\Pi^\dagger + \cdots + A_0\left(-\partial_i \pi^i + \mathbb{i}q(\Pi\phi - \Pi^\dagger\phi^\dagger)\right) \ . \qquad (\text{A.4.14})$$

All that happens to π^0's null equation of motion is that the Gauss constraint acquires a charge density,

$$\partial_i \pi^i + \rho \approx 0 \ , \qquad \rho \equiv -\mathbb{i}q(\Pi\phi - \Pi^\dagger\phi^\dagger) \ . \qquad (\text{A.4.15})$$

The fact that $\pi^0 \approx 0$ and the subsequent $\partial_t \pi^0 \approx 0$ imposes a constraint to the dynamics is unchanged. The latter, secondary constraint is called the Gauss constraint.

Let us first sit back and explore the implications of what we have found so far, A_0 proved to be a Lagrange multiplier in the canonical variable choice, such that we find the following set of constraints,

$$\pi^0 \approx 0 \ , \qquad \partial_i \pi^i - \mathbb{i}q(\Pi\phi - \Pi^\dagger\phi^\dagger) \approx 0 \qquad (\text{A.4.16})$$

It is easy to see that together they form the first-class constraints, since the Poisson bracket of the two is identically zero. As we have seen, the first-class constraints remove two degrees of freedom each. For Maxwell theory, we see this reduction fairly explicitly.

The Lagrange multiplier A_0 is arbitrary while $\pi^0 = 0$ identically, leaving us with A_i and π^i. The Gauss constraint generates the gauge transformation,

$$A_i \to A_i + \partial_i \vartheta \qquad (\text{A.4.17})$$

for an arbitrary gauge function ϑ. We again see that each first-class constraint removes a pair of canonical field variables. Even though we started with $2d$ canonical variables A_μ and π^μ, therefore, we end up with $2(d-2)$ canonical field theory degrees of freedom in the phase space, or equivalently $d-2$ field variables in the configuration space.

The Gauss constraint is the generator of the $U(1)$ gauge symmetry and demands that physical states must be invariant under the gauge redundancy. In a gravitational system, such gauge redundancies manifest as general coordinate transformations.

Maxwell Theory under the Stückelberg Mechanism

A particular case of the above deserves further attention. Start with a single charged field ϕ but suppose some underlying potential so that the modulus of the field is frozen to an expectation value,

$$\phi = \frac{m}{\sqrt{2}} e^{\mathbb{i}\theta} \ , \tag{A.4.18}$$

which is nothing but a minimally distilled Higgs mechanism that give a mass m to the gauge field. With $q = 1$, for the sake of simplicity, we have the action

$$\int d^d x \left(-\frac{1}{4} F^2 - \frac{m^2}{2} (\partial\theta - A)^2 \right) \ . \tag{A.4.19}$$

We still find the same pair of first-class constraints,

$$\pi^0 \approx 0 \ , \qquad -\partial_i \pi^i + \Pi_\theta \approx 0 \ , \tag{A.4.20}$$

with

$$\Pi_\theta = m^2(\partial_0\theta - A_0) \ . \tag{A.4.21}$$

All that happens is that the Gauss constraint not only shift A by the gauge transformation but shift θ additively.

The combined system of $d+1$ field variables, A_μ and θ, is reduced to $d-1$ such; By judicious choice of the gauge, we may remove θ and end up with $d-1$ dynamical field variables, A_i, appropriate for a massive vector. Despite how the second term can be interpreted as the mass of the photon, the gauge redundancy is intact, which is the hallmark of the Higgs mechanism.

A.4.2 Chern-Simons Theory and the Dirac Bracket

A little unusual form of gauge theory exists in $d = 3$ spacetime, with the so-called Chern-Simons action. The Maxwell theory in $d = 3$ has $d-2 = 1$ propagating degrees of freedom after two first-class constraints, $\pi^0 \approx 0$ and $\partial_i \pi^i + \cdots \approx 0$ are taken into account. In contrast, the Chern-Simons theory, without an accompanying Maxwell term, carries zero field theory degrees of freedom. This is a simplest field theory example where both first-class and second-class constraints are found together and in particular the Dirac bracket plays a central role. Although one can perform the computation here with both Maxwell and Chern-Simons present, in which case the single massless degrees of freedom of the former become massive, we will consider the pure Chern-Simons as this limit generates the second-class constraints of interest.

The action is

$$S_{\text{C.S.}} \;=\; \frac{\kappa}{4\pi}\int A\wedge F \;=\; \frac{\kappa}{4\pi}\int d^3x\ \left(A_0F_{12} - A_1F_{02} + A_2F_{01}\right)\ , \quad \text{(A.4.22)}$$

where $\kappa/\hbar$ is called the level. This action is not manifestly gauge-invariant because of two reasons. First, under continuous gauge transformation, the integrand shifts by a total derivative, so we need a compact manifold or an additional boundary condition to make $S_{\text{C.S.}}$ invariant. Second, even if the latter conditions are met, $S_{\text{C.S.}}$ shifts under the so-called large gauge transformations by some quantized amount, depending on the topology of the spacetime. We nevertheless encounter such actions because the latter shift may be ignored if we consider $S_{\text{C.S.}}$ at quantum level, i.e., as an exponent in some path integral. For a reason we will not go into here, this forces κ to be an integer multiple of $\hbar$.

As with the Maxwell case, the conjugate momentum π^0 of A_0 is null, such that

$$\begin{aligned}\int d^2\mathbf{x}\,\mathcal{H} \;&=\; \int d^2\mathbf{x}\,(\pi^i\dot{A}_i - \mathcal{L}_{\text{C.S.}}) \;=\; \int d^2\mathbf{x}\,(\pi^i F_{0i} - \mathcal{L}_{\text{C.S.}} + \pi^i\partial_i A_0)\\ &=\; -\int d^2\mathbf{x}\,A_0\left(\partial_i\pi^i + \frac{\kappa}{4\pi}F_{12}\right)\ . \qquad \text{(A.4.23)}\end{aligned}$$

As before $\dot{\pi}^0 \approx 0$ induces a secondary constraint,

$$-\partial_i\pi^i - \frac{\kappa}{4\pi}F_{12} \approx 0 \qquad \text{(A.4.24)}$$

which is a Gauss constraint where we see that the magnetic flux acts like an electric charge.

The real difference from the Maxwell theory comes about from how the conjugate momenta π^i of A_i are not independent variables. From

$$\pi^i \equiv \frac{\delta S_{\text{C.S.}}}{\delta \dot{A}_i} = \frac{\kappa}{4\pi}\epsilon^{ij}A_j \ , \tag{A.4.25}$$

with $\epsilon^{12} = -\epsilon^{21} = 1$, we find two additional primary constraints,

$$\Pi^i \equiv \pi^i - \frac{\kappa}{4\pi}\epsilon^{ij}A_j \approx 0 \ . \tag{A.4.26}$$

Now it is not difficult to see that Π^i's are second-class while π^0 and the Gauss constraint remain first-class. The Poisson bracket between the Gauss constraint and Π^i's are such that there are two potential commutator terms but these two cancel out each other.

The pair of second-class constraints yield,

$$[\,\Pi^1(\mathbf{x}),\Pi^2(\mathbf{y})\,]_{\text{P.B.}} \ = \ -\frac{\kappa}{2\pi}\delta^{(2)}(\mathbf{x}-\mathbf{y}) \ , \tag{A.4.27}$$

and they generate the Dirac bracket,

$$\begin{aligned}[\,A_1(\mathbf{x}),A_2(\mathbf{y})\,]_{\text{Dirac}} \ &= \ -\int d^2\mathbf{z}[\,A_1(\mathbf{x}),\Pi^1(\mathbf{z})]_{\text{P.B.}}\left(\frac{\kappa}{2\pi}\right)^{-1}[\,\Pi^2(\mathbf{z}),A_2(\mathbf{y})]_{\text{P.B.}} \\ &= \ \frac{2\pi}{\kappa}\delta^{(2)}(\mathbf{x}-\mathbf{y}) \ . \end{aligned} \tag{A.4.28}$$

The precise numerical factor on the right carries an important physical meaning at quantum level. Although the field theory degree of freedoms are all removed by these four constraints, two first-class and two second-class, an enumerable number of quantum mechanical states survive if we put the theory on a finite spatial volume. The number of such quantum states is proportional to κ, which follows from quantization of the above Dirac commutator or more simply from the so-called Bohr-Sommerfeld quantization.

Let us again make a cautionary remark, on the matter of premature use of the second-class constraints. We have earlier emphasized how the second-class constraint should not be imposed too early if one is determined to use the Poisson bracket. If we had computed the Poisson bracket *after imposing the secondary constraints*, $\Pi^i \approx 0$, we would have ended up with

$$[\, A_1(\mathbf{x}), A_2(\mathbf{y})\,]_{\rm P.B.} \;\approx\; \frac{4\pi}{\kappa}\delta^{(2)}(\mathbf{x}-\mathbf{y}) \;, \tag{A.4.29}$$

which is *off by factor 2* relative to the Dirac bracket which must be used for the canonical quantization. This also goes against the fundamental premise of the canonical dynamics that a Poisson bracket between a pair of configuration variables or between a pair of momentum variables is always zero, again telling us that this procedure is illegal. One can ostensibly avoid this factor 2, by replacing $-A_1F_{02} + A_2F_{01}$ by $2A_2F_{01}$, citing an integration by part. One should regard the last approach a quick and dirty way out and at best as a mere reminder of what the right procedure would have given.

Such subtleties are always present when the Lagrangian is in the so-called first order formulation, meaning that the time-derivative occurs only once in the kinetic terms. The latter in particular is always the case for fermionic systems, yet the subtleties are often glossed over as most textbooks sidestep the intricacy of the constrained dynamics in favor of a faster progression to the dynamics. The same kind of the factor 2 issue above is present for Majorana fermions, inevitably for exactly the same reasons as the above Chern-Simons example, while for complex fermions, the canonical commutators do not come from the Poisson bracket, despite how one seemingly gets the right answer by a naive procedure if one chooses to treat the fermion fields and their complex conjugate differently.