Chapter 2

Hamilton's Formalism for Constraint Systems

2.1 Singular Lagrangian systems

The most general mathematical setting for gauge theories is Dirac's constraint formalism [1]. Here I review this formalism, also to prepare the ground for the following chapters, and in particular the one on the Hamiltonian reduction of $WZNW$ theories.

Attempts to handle constrained systems date back more than forty years. In his classical works Dirac set up a formalism to treat such systems self-consistently [18]. Later Bergman et.al. in a series of papers investigated the connection between constraints and invariances [3, 11, 13]. After the introduction of Grassmann variables to describe fermions [9], the formalism has been extended to include fields with half-integer spins [25, 14, 10]. The development culminated with the advent of the elegant and powerful BRST formalism [7]. These and other classical results have been a prerequisite for the quantization of gauge theories both in the path integral formalism [20, 6] and in the functional Schrödinger picture [47, 33].

There are several excellent reviews on the treatment of constrained systems of gauge theories besides Dirac well-known lectures [19]. Some focus more on systems with a finite number of degrees of freedom [45], others on field theories [30] and some on both [46, 27, 31]. For generally covariant theories you may consult [26].

We shall be concerned with systems whose dynamics can be derived from Hamilton's variational principle. I assume that all Lagrangians depend

[1] For an alternative approach see [22]

at most on first derivatives, up to divergence terms [2]. Throughout this work I shall use local coordinates, unless I am forced to address global questions, e.g. the Gribov problem or the role of topologically nontrivial field configurations.

With these assumptions the classical trajectories of a system with N degrees of freedom make the action

$$S = \int_{t_1}^{t_2} L(q^i, \dot{q}^i) dt \quad , \quad i = 1, \dots, N \tag{2.1}$$

stationary under variations $\delta q(t)$ which vanish at the endpoints. The q and $\dot{q}$ are local coordinates on the velocity phase space TQ. The necessary conditions for S to be stationary are the *Euler-Lagrange equations*

$$L_i \equiv -\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}^i}\Big) + \frac{\partial L}{\partial q^i} = 0 \tag{2.2}$$

which can be rewritten as

$$L_i = -\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j}\ddot{q}^j - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j + \frac{\partial L}{\partial q^i} \equiv -W_{ij}(q,\dot{q})\ddot{q}^j + V_i = 0. \tag{2.3}$$

We see that the accelerations at a given time are uniquely determined by $(q, \dot{q})$ at that time only if the Hessian (W_{ij}) can be inverted. Such systems are called *regular*.

For *singular systems* $\det W = 0$, and the accelerations and hence time evolution will not be uniquely fixed by $(q, \dot{q})$. For such systems different time evolutions will stem from the same initial conditions.

The rank R of W, which we assume for simplicity to be constant on TQ, being $R<N$ implies the existence of $M=N-R$ null-eigenvectors Y_m of W:

$$Y^i_m(q,\dot{q})W_{ij}(q,\dot{q}) = 0 \quad , \qquad m = 1, \dots, M. \tag{2.4}$$

Contracting the E-L equations (2.3) with the Y_m we get

$$\phi_m(q,\dot{q}) \equiv Y^i_m V_i = 0 \quad , \qquad m = 1, \dots, M. \tag{2.5}$$

These equations do not contain accelerations. Assume that $M' \leq M$ relations

$$\phi_{m'} = 0 \quad , \quad m' = 1, \dots, M', \tag{2.6}$$

are functionally independent on the others, and the remaining ones are either dependent or identically fulfilled. The independent ones are the so-called *Lagrange constraints.*

[2] For higher derivative theories, and in particular for higher derivative gravity, see [26].

For *field theories* the dynamics is described by functions $\varphi^a(x)$ of spacetime with values in a certain target space. The index a may belong to an internal symmetry, it may be a spacetime index or both internal and spacetime index as in non-non-Abelian gauge theories. When going from point mechanics to field theory one may think of replacing the discrete label i by a continuous one $(a, \vec{x})$:

$$q^i(t) = q(t, i) \longrightarrow q(t, a, \vec{x}) = \varphi^a(t, \vec{x}) = \varphi^a(x).$$

Summations become spatial integrals, e.g.

$$\sum_i \dot{q}^i \dot{q}^i \longrightarrow \sum \int dx\ \dot{\varphi}^a(\vec{x})\dot{\varphi}^a(\vec{x})$$

and functions of $(q, \dot{q})$ become functionals of φ and $\dot{\varphi}$. Also, derivatives with respect to q^i or $\dot{q}^i$ become functional derivatives, e.g.

$$\frac{\partial L}{\partial \dot{q}^i} \longrightarrow \frac{\delta L}{\delta \dot{\varphi}^a(\vec{x})}$$

The velocity phase space TQ is chosen so that the Lagrange-functional L is continuous and sufficiently often differentiable. If the target space is linear one may choose a Banach space (typically a Sobolov space), otherwise one tries to model the theory on a C^k-Banach manifold [16, 37] since the implicit function theorem still applies then. Banach manifolds are modeled over Banach spaces and are straightforward generalizations of finite-dimensional manifolds.

A functional on a Banach space X is called continuous if

$$\lim_{n\to\infty} F[\varphi_n] = F[\varphi] \quad \text{for} \quad X \ni \varphi_n \to \varphi.$$

F is called Frechet-differentiable at φ if there exists a linear functional F'_φ such that

$$|F[\varphi + \delta\varphi] - F[\varphi] - F'_\varphi[\delta\varphi]| = o(\parallel \delta\varphi \parallel) \quad \text{for all} \quad \parallel \delta\varphi \parallel \to 0.$$

For *local theories* the Lagrangian has the form

$$L[\varphi, \dot{\varphi}] = \int dx\ \mathcal{L}(\varphi, \partial_i\varphi, \dot{\varphi}) \tag{2.7}$$

with a Lagrangian density $\mathcal{L}$ depending only on the field and its derivatives at the same point. For such theories the Euler-Lagrange equations are

$$L_a \equiv -\frac{\partial}{\partial t}\frac{\delta L}{\delta \dot{\varphi}^a} + \frac{\delta L}{\delta \varphi^a} = -\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi^a)} + \frac{\partial \mathcal{L}}{\partial \varphi^a} = 0, \tag{2.8}$$

where I adopted the common notation

$$F'_\varphi[\delta\varphi] \equiv \int \frac{\delta F}{\delta\varphi(x)}\delta\varphi(x). \tag{2.9}$$

Rewriting the field equations as

$$\begin{aligned} L_a &= -\frac{\partial^2\mathcal{L}}{\partial(\partial_\mu\varphi^a)\partial(\partial_\nu\varphi^b)}\partial_\mu\partial_\nu\varphi^b - \frac{\partial^2\mathcal{L}}{\partial(\partial_\mu\varphi^a)\partial\varphi^b}\partial_\mu\varphi^b + \frac{\partial\mathcal{L}}{\partial\varphi^a} \\ &\equiv -W^{\mu\nu}_{ab}\partial_\mu\partial_\nu\varphi^b + V_a = 0 \end{aligned} \tag{2.10}$$

we can see that theories with x^0 taken as evolution parameter are regular if W^{00}_{ab} is invertible and singular if it is not. For singular systems there exist (for each $\vec{x}$) $M = N - R$ null-vectors

$$Y^a_m(\varphi,\partial\varphi)W^{00}_{ab}(\varphi,\partial\varphi) = 0 \quad , \quad m = 1,\ldots,M \tag{2.11}$$

which lead to nontrivial and independent Lagrangian constraints

$$\phi_{m'}(\varphi,\partial\varphi) \equiv Y^a_{m'}V_a = 0 \quad , \quad m' = 1,\ldots,M' \le M, \tag{2.12}$$

involving only the fields and their first derivatives.

How one proceeds for singular systems is neatly explained in [43, 46]. There are two points which have to be considered. Firstly the rank of the Hessian may decrease if one takes the independent constraints (2.6,12) into account. This may lead to new independent constraints. Again the rank may decrease leading to further constraints, etc. This process terminates as soon as the rank does not change anymore.

Secondly one needs to check whether the constraints one has found after the above algebraic process has terminated are respected by the *time evolution.* These may lead to new constraints. Again and again differentiate newly emerging constraints until no new ones arise. Add those relations involving accelerations to those already present. Consistency of the old relations with the new ones may lead to further constraints. After all that one needs again to check whether the rank of the Hessian has changed. If this is the case one needs to start from the beginning etc.

Generalized Bianchi identities If a theory possesses a *local* gauge invariance we may map solutions into solutions without affecting the initial conditions. Thus we expect that *gauge theories are singular systems.* Actually this follows from the generalized Bianchi identity [50, 48] which we derive next.

The point transformations

$$
\begin{aligned}
x' &= x'(x) \sim x + \delta x \\
\varphi'(x') &= \varphi'(\varphi(x), x) \sim \varphi(x) + \delta\varphi
\end{aligned} \tag{2.13}
$$

which leave the action invariant

$$\int d^d x' \, \mathcal{L}(\varphi', \partial'\varphi', x') = \int d^d x \, \mathcal{L}(\varphi, \partial\varphi, x) \tag{2.14}$$

form a group which we assume to be continuous. For transformations close to the identity $d^dx' = d^dx(1 + \partial_\mu \delta x^\mu)$, and the invariance (2.14) implies

$$\delta\mathcal{L} + \mathcal{L}\partial_\mu \delta x^\mu = \partial_\mu \lambda^\mu \tag{2.15}$$

with some λ. Using

$$\delta\mathcal{L} = \frac{\partial\mathcal{L}}{\partial\varphi^a}\delta\varphi^a + \frac{\partial\mathcal{L}}{\partial(\partial_\nu\varphi^a)}\delta(\partial_\nu\varphi^a) + \partial_\mu\mathcal{L}\delta x^\mu$$

it follows at once that

$$\delta\mathcal{L} + \mathcal{L}\partial_\mu\delta x^\mu = \partial_\mu(\mathcal{L}\delta x^\mu) + \partial_\mu(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\bar{\delta}\varphi^a) + L_a\bar{\delta}\varphi^a,$$

where the Euler derivatives L_a have been defined in (2.8) and

$$\bar{\delta}\varphi^a = \delta\varphi^a - \partial_\mu\varphi^a\delta x^\mu \sim \varphi^{a\prime}(x) - \varphi^a(x) \tag{2.16}$$

is the infinitesimal difference of the old and the transformed files at the same point. We used that $[\bar{\delta}, \partial_\mu] = 0$. Thus the gauge invariance implies

$$\boxed{\partial_\mu(\mathcal{L}\delta x^\mu + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\bar{\delta}\varphi^a - \lambda^\mu) + L_a\bar{\delta}\varphi^a = 0} \tag{2.17}$$

and these are the *generalized Bianchi identities.* Nowhere did we use the equation of motion and thus (2.17) are *off-shell identities.*

First assume that S is invariant under global transformations forming a n-dimensional Lie-group. Then

$$\lambda^\mu = \epsilon_\alpha\lambda^{\alpha\mu}, \quad \delta_\epsilon x^\mu = \epsilon_\alpha A^{\alpha\mu}, \quad \delta_\epsilon\varphi^a = \epsilon_\alpha B^{\alpha a}, \tag{2.18}$$

where the $\epsilon_\alpha, \alpha = 1, \dots, n$ are the constant parameters of the infinitesimal transformations. Inserting this into (2.17) and going on shell, $L_a = 0$, we conclude

$$\partial j^{\alpha\mu} = 0, \quad \text{where} \quad j^{\alpha\mu} = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}(B^{\alpha a} - A^{\alpha\nu}\partial_\nu\varphi^a) + \mathcal{L}A^{\alpha\mu} - \lambda^{\alpha\mu}, \tag{2.19}$$

which *is Noether's first theorem.* Note that we allowed for general point transformation so that (2.19) applies to both space-time and internal symmetries. However, when deriving (2.19) we imposed the equations of motion so that the currents are *conserved only on-shell.*

Let us now assume that the symmetry transformations are local. In that case the parameters become space-time dependent and (2.18) generalizes to

$$\delta_\epsilon x^\mu = \epsilon_\alpha A^{\alpha\mu} \quad , \quad \delta_\epsilon \varphi^a = \epsilon_\alpha B^{\alpha a} + \partial_\mu \epsilon_\alpha C^{\alpha a\mu}, \tag{2.20}$$

where the $\epsilon_\alpha(x)$ parametrize the infinitesimal local gauge transformations and B and C are the so-called *descriptors* [3], which in general depend on the fields and their derivatives. I assumed that no second or higher derivatives of ϵ enter because this covers most interesting examples. With

$$\bar{\delta}_\epsilon \varphi^a = \epsilon_\alpha (B^{\alpha a} - \partial_\mu \varphi^a A^{\alpha\mu}) + \partial_\mu \epsilon_\alpha C^{\alpha a\mu} \tag{2.21}$$

the integrated form of (2.17), after a partial integration, reads

$$0 = \int \epsilon_\alpha [L_a (B^{\alpha a} - \partial_\mu \varphi^a A^{\alpha\mu}) - \partial_\mu (L_a C^{\alpha a\mu})]. \tag{2.22}$$

Since it must hold for arbitrary functions ϵ_α this implies that the expression between the square brackets must vanish. Inserting L_a from (2.10) we end up with

$$\begin{aligned} 0 &= L_a (B^{\alpha a} - \partial_\mu \varphi^a A^{\alpha\mu} - \partial_\mu C^{\alpha a\mu}) - C^{\alpha a\mu} \partial_\mu V_a \\ &\quad + C^{\alpha a\mu} (\partial_\mu W_{ab}^{\rho\sigma\rho} \partial_\rho \partial_\sigma \varphi^b + W_{ab}^{\rho\sigma} \partial_\mu \partial_\rho \partial_\sigma \varphi^b). \end{aligned} \tag{2.23}$$

Since these are off-shell identities we conclude

$$C^{\alpha a(\mu} W_{ab}^{\rho\sigma)} = 0, \tag{2.24}$$

where the brackets around the indices mean symmetrization. In particular, descriptors $C^{\alpha a0}$ which are not identically zero are null-eigenvectors of the Hessian,

$$C^{\alpha a0} W_{ab}^{00} = 0 \tag{2.25}$$

and render the system singular. If all $C^{\alpha a\mu}$ vanish, then (2.23) reduces to

$$0 = (B^{\alpha a} - \partial_\rho \varphi^a A^{\alpha\rho}) L_a \Longrightarrow (B^{\alpha a} - A^{\alpha\rho} \partial_\rho \varphi^a) W_{ab}^{(\mu\nu)} = 0. \tag{2.26}$$

Thus, if $C \equiv 0$ but the $B^{\alpha a} - A^{\alpha\rho}\partial_\rho\varphi^a$ are not identically zero, we conclude again that the system is singular. So we have the important result that *gauge theories are necessarily singular.* However, the converse is not true. Not all conceivable singular systems are gauge theories.

2.2 Primary and secondary constraints

The departing point for the Hamiltonian formalism is to define the canonical momenta (densities) by

$$p_i = \frac{\partial L}{\partial \dot{q}^i}(q, \dot{q}) \quad \text{resp.} \quad \pi_a(\vec{x}) = \frac{\delta L}{\delta \dot{\varphi}^a(\vec{x})}, \tag{2.27}$$

where we assume that $L \in C^2(TQ)$. Only if

$$W_{ij} = \frac{\partial p_i}{\partial \dot{q}^j} \quad \text{resp.} \quad W^{00}_{ab}(\vec{x}, \vec{y}) = \frac{\delta \pi_a(\vec{x})}{\delta \dot{\varphi}^b(\vec{y})} \tag{2.28}$$

is invertible, that is for regular systems, can this relation be solved for all velocities in terms of the phase space variables [3], $\dot{q} = \dot{q}(q, p)$ resp. $\dot{\varphi} = \dot{\varphi}(\varphi, \pi)$. In the other case not all momenta (2.27) are independent, but there are some relations

$$\phi_m(q, p) = 0 \quad \text{resp.} \quad \phi_m(\varphi, \pi) = 0 \quad , \quad m = 1, \ldots, M \tag{2.29}$$

that follow from the definition (2.27) of the momenta. I shall assume that the constraints (2.29) are independent.

In the following we restrict ourselves to finite dimensional systems and only comment on the related results for field theories. The corresponding field theoretical formulas, if they apply, are obtained if one uses deWitt's condensed notation [17] in which i becomes a composite index.

The conditions (2.29) are the $M = N - R$ *primary constraints*. They define the $2N - M$-dimensional *primary constraint surface*, denoted by Γ_p. The equations of motions have not been used to derive them and they imply no restriction on the $(q, \dot{q})$. (2.27) maps the $2N$-dimensional velocity phase space TQ to the lower-dimensional sub-manifold Γ_p in the momentum phase space Γ. Hence the inverse images of a given point in Γ_p form a manifold of dimension M.

To pass to the Hamiltonian formalism we impose some *regularity conditions* on the primary constraints. They can be alternatively formulated as:

1. the independent functions $\phi_m, m = 1, \ldots, M$ can be locally taken as the first M coordinates of a new, regular, coordinate system in the vicinity of Γ_p.

2. The gradients $d\phi_1, \ldots, d\phi_M$ are locally linearly independent on Γ_p; i.e., $d\phi_1 \wedge \ldots \wedge d\phi_M \neq 0$ on Γ_p.

[3] for field theories we assume TQ to be a Banach manifold so that the inverse function theorem applies

For example, if ϕ is an admissible constraint, ϕ^2 is not, since $d(\phi^2)=2\phi d\phi=0$ on Γ_p. If the constraints are regular the following properties hold.

Theorem 1 *If a smooth function $F(q,p)$ vanishes on Γ_p, then $F = f^m\phi_m$ for some functions f^m.*

Theorem 2 *If $\lambda_i\delta q^i + \mu^i\delta p_i = 0$ for arbitrary variations $\delta q^i, \delta p_i$ tangent to the constraint surface, then*

$$\lambda_i = u^m\frac{\partial\phi_m}{\partial q^i} \quad \textit{and} \quad \mu^i = u^m\frac{\partial\phi_m}{\partial p_i} \qquad \textit{on} \quad \Gamma_p$$

for some u^m.

Before proving these two important theorems it is useful to distinguish between weak and strong equations. A function $F(q,p)$ defined in the neighborhood of Γ_p is called *weakly zero* if

$$F|_{\Gamma_p} = 0 \Longleftrightarrow F \approx 0 \tag{2.30}$$

and *strongly zero* if

$$F|_{\Gamma_p} = 0 \quad \text{and} \quad (\frac{\partial F}{\partial q^i}, \frac{\partial F}{\partial p_i})|_{\Gamma_p = 0} \Longleftrightarrow F \simeq 0. \tag{2.31}$$

These definitions are useful since the equations of motion contain gradients of functions on Γ_p. The primary constraint surface can itself be defined by weak equations. We have

$$\phi_m \approx 0 \quad \text{but} \quad \phi_m \not\simeq 0 \tag{2.32}$$

because of our regularity conditions on the constraints.
Since $\nabla_x(f^m\phi_m) \approx f^m\nabla_x\phi_m$, where $x = (q,p)$ denotes the phase space coordinates, the first theorem implies

Lemma 1 $F \approx 0 \Longrightarrow F - f^m\phi_m \simeq 0$ *for some functions f^m.*

To prove the first theorem we choose the independent constraints ϕ_m as first coordinates of a regular coordinate system $x=(\phi,\tilde{x})$ in the neighborhood of Γ_p. Since $F(0,\tilde{x})=0$ we have

$$F(\phi,\tilde{x}) = \int_0^1 \frac{d}{d\tau}F(\tau\phi,\tilde{x})d\tau = \phi_m\int_0^1 F_{,m}(\tau\phi,\tilde{x})d\tau$$

and thus

$$F = f^m\phi_m \quad \text{with} \quad f^m = \int_0^1 F_{,m}(\tau\phi,\tilde{x})d\tau. \tag{2.33}$$

This proves theorem 1 in the neighborhood U of any point on Γ_p. We cover the neighborhood of Γ_p by open sets U_i, on each of which theorem 1 applies. Together with the open sets V_k on which $\phi_k \neq 0$ the U_i cover the whole phase space. On V_k we can set $F = (F/\phi_k)\phi_k$ and theorem 1 holds there. Finally, to guarantee that the f^m are the same on the overlap of U_i and $U_{i'}$ one uses a finite partition of unity.

Theorem 2 follows immediately from the regularity condition which implies that at a given point x on Γ_p a basis of $T_x\Gamma_p$ (the vectors tangent to Γ_p at x), together with the gradients $\nabla_x\phi_m$ form a basis of $T_x\Gamma$. The assumption in theorem 2 means that (λ, μ) are orthogonal to $T_x\Gamma_p$. Thus it must be a linear combination of the gradients $\nabla_x\phi_m$.

For *field theories* one finds

$$F[\phi, \tilde{x}] \approx 0 \Rightarrow F = \int f^m\phi_m \quad , \quad f^m(\vec{x}) = \int d\tau \frac{\delta F}{\delta\phi_m(\vec{x})}[\tau\phi, \tilde{x}] \tag{2.34}$$

and a weakly vanishing functional is a linear combination of smeared constraints. The test functions should lie in the space dual to the space of the constraints [8].

2.2.1 Legendre transformation

The canonical Hamiltonian

$$H = \dot{q}^i p_i - L \quad \text{resp.} \quad H = \int dx\ \pi_a(\vec{x})\dot{\varphi}^a(\vec{x}) - L = \int dx\ \mathcal{H} \tag{2.35}$$

has the remarkable property that $\dot{q}$ enters H only through the combination $p(q, \dot{q})$. This follows from

$$\delta H = \dot{q}^i\delta p_i + \delta\dot{q}^i p_i - \delta\dot{q}^i\frac{\partial L}{\partial\dot{q}^i} - \delta q^i\frac{\partial L}{\partial q^i} = \dot{q}^i\delta p_i - \delta q^i\frac{\partial L}{\partial q^i} \tag{2.36}$$

which shows that H is a function of p and q only. Here δp is to be regarded as linear combination of δq and $\delta\dot{q}$ so that $\delta q, \delta p$ are tangent to Γ_p. H is only defined on Γ_p since we used the constraints. We would like to extend the formalism to the whole phase space Γ.

The equation (2.36) can be rewritten as

$$\left(\frac{\partial H}{\partial q^i} + \frac{\partial L}{\partial q^i}\right)\delta q^i + \left(\frac{\partial H}{\partial p_i} - \dot{q}^i\right)\delta p_i = 0 \tag{2.37}$$

with variations tangent to Γ_p. H may be the restriction to the hyper-surface Γ_p of a function $\tilde{H}$ defined all over phase space. Then (2.37) holds with H replaced by $\tilde{H}$. Applying theorem 2 we conclude that

$$\dot{q}^i \approx \frac{\partial \tilde{H}}{\partial p_i} + u^m \frac{\partial \phi_m}{\partial p_i} \quad , \quad -\frac{\partial L}{\partial q^i} \approx \frac{\partial \tilde{H}}{\partial q^i} + u^m \frac{\partial \phi_m}{\partial q^i}. \tag{2.38}$$

The first set of relations enables us to recover the velocities from the $(q,p) \in \Gamma_p$ and the parameters u^m. Because of the regularity conditions on the constraints two different u yield different $\dot{q}$ and the first relation permits us to express u as function of q and $\dot{q}$. This way one obtains an *invertible Legendre transformation* from the $2N$-dimensional velocity phase space to the $2N$ dimensional space $\Gamma_p \times \{u^m\}$:

$$p_i = \frac{\partial L}{\partial \dot{q}^i}(q,\dot{q}) \quad \text{and} \quad u^m = u^m(q,\dot{q}) \tag{2.39}$$

with inverse transformation

$$\dot{q}^i = \frac{\partial \tilde{H}}{\partial p_i} + u^m \frac{\partial \phi_m}{\partial p_i} \quad \text{and} \quad \phi_m(q,p) = 0. \tag{2.40}$$

We had to extend the Hamiltonian, which was originally defined only on Γ_p, to a neighborhood of Γ_p. According to theorem 1 two possible extensions differ by a term $c^m \phi_m$. Thus the formalism should be unchanged by the replacement

$$\tilde{H} \longrightarrow \tilde{H} + c^m(q,p)\phi_m. \tag{2.41}$$

Indeed, making this transformation in (2.38) just shifts the u to $u + c$.

Finally, the relations (2.38) allow us to rewrite the equation of motion (2.2) in the equivalent Hamiltonian form

$$\dot{q}^i \approx \frac{\partial H}{\partial p_i} + u^m \frac{\partial \phi_m}{\partial p_i} \quad \text{and} \quad \dot{p}_i \approx -\frac{\partial H}{\partial q^i} - u^m \frac{\partial \phi_m}{\partial q^i}, \tag{2.42}$$

where we dropped the tilde atop H. The Lagrangian equations of motion (2.2) are equivalent to (2.42). The phase space function

$$H_p \equiv H + u^m \phi_m \tag{2.43}$$

is the *primary Hamiltonian.*

Introducing the Poisson bracket of two phase space functions

$$\begin{aligned} \{F,G\} &\equiv \frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q^i} \qquad \text{resp.} \\ \{F,G\} &\equiv \int dx \Big(\frac{\delta F}{\delta \varphi^a(\vec{x})}\frac{\delta G}{\delta \pi_a(\vec{x})} - \frac{\delta F}{\delta \pi_a(\vec{x})}\frac{\delta G}{\delta \varphi^a(\vec{x})} \Big) \end{aligned} \tag{2.44}$$

and using $u^m \nabla_x \phi_m \approx \nabla_x(u^m \phi_m)$, the Hamiltonian equations of motion can be rewritten as

$$
\begin{aligned}
\dot{q}^i &\approx \{q^i, H_p\} \approx \{q^i, H\} + \{q^i, \phi_m\}u^m \\
\dot{p}_i &\approx \{p_i, H_p\} \approx \{p_i, H\} + \{p_i, \phi_m\}u^m.
\end{aligned}
\tag{2.45}
$$

Besides there are still the equations defining Γ_p:

$$
\phi_m(q,p) = 0. \tag{2.46}
$$

For an any phase-space function $F(q,p)$ the *time evolution* follows then from

$$
\boxed{\dot{F} \approx \{F, H_p\} \approx \{F, H\} + u^m\{F, \phi_m\}.} \tag{2.47}
$$

2.2.2 Dirac-Bergman algorithm

The constraints must be consistent with the time evolution, that is if initially (q,p) is on Γ_p it should remain there at later times. This means that the equations of motion should preserve the constraints and this gives rise to the *consistency conditions* [18, 3]

$$
\dot{\phi}_m \approx \{\phi_m, H\} + \{\phi_m, \phi_n\}u^n \equiv h_m + C_{mn}u^n \approx 0. \tag{2.48}
$$

For non-admissible Lagrangians these relations will be inconsistent. As an example take $L = \dot{q} - q$ which leads to $H = q$ and $\phi = p - 1$ so that (2.48) reads $1 \approx 0$. For such inconsistent models the action has no stationary points and we shall exclude them.

To discuss the consistency relations (2.48) we distinguish the two following cases:

- $\det C \not\approx 0$:

  In this case u is uniquely fixed by (2.48) to be $u^n \approx C^{nm}h_m$, where C^{nm} is the inverse of C_{nm}. The time evolution (2.47) of a phase space function becomes

$$
\dot{F} \approx \{F, H\} - \{F, \phi_m\}C^{mn}\{\phi_n, H\}. \tag{2.49}
$$

  No additional conditions appear. For any initial data (q,p) on Γ_p the time evolution stemming from (2.49) is unambiguous and stays on Γ_p.

- $\det C \approx 0$:

  In this case u is not fixed and (2.48) is only solvable if $h_m w_a^m \approx 0$ for all left null-eigenvectors w_a of C. Either these equations are fulfilled or they lead to a certain number K_1 of new constraints

$$
\phi_k \approx 0 \quad , \quad k = M+1, \ldots, M+K_1 \equiv J_1,
$$

called *secondary constraints.* The primary and secondary constraints $\phi_j \approx 0$, $j = 1, \ldots, J_1$ define a hyper-surface $\Gamma_1 \subseteq \Gamma_p$.

Now one has to check consistency for the primary and newly generated secondary constraints on Γ_1,

$$\dot{\phi}_j = \{\phi_j, H\} + \{\phi_j, \phi_n\}u^n \equiv h_j + C_{jn}u^n = 0 \quad \text{on} \quad \Gamma_1$$

with the rectangular $J_1 \times M$ matrix C. The left null-eigenvectors w_a^j of C_{jn} imply further conditions $w_a^j h_j = 0$ on Γ_1 and may lead to further, so-called *tertiary constraints* which, together with the primary and secondary constraints, define a hyper-surface $\Gamma_2 \subseteq \Gamma_1$, etc.

This procedure terminates after a finite number of iterations and the following situation is reached: There is a hyper-surface $\Gamma_c \subset \Gamma$ defined by

$$\phi_j \approx 0 \quad , \quad j = 1, \ldots, M + K \equiv J. \tag{2.50}$$

For every left null-eigenvector w_a^j of the rectangular matrix $C_{jm} = \{\phi_j, \phi_m\}$ the conditions $w_a^j\{\phi_j, H\} \approx 0$ are fulfilled. For the multiplier fields there are the equations

$$\{\phi_j, H\} + \{\phi_j, \phi_m\}u^m \approx 0, \tag{2.51}$$

where $\approx$ now means equality on Γ_c. We note that the primary constraints are merely consequences of the definition of the momenta, whereas we used the equations of motion to arrive at the secondary constraints[4].

We make the same regularity assumptions on the full set of constraints ϕ_j defining Γ_c as we made on the primary constraints ϕ_m defining Γ_p. Also, we assume that the rank of C is constant on Γ_c.

2.3 First and second class constraints

The distinction between primary and secondary constraints will be of minor importance in the final form of the Hamiltonian theory. A different classification of constraints, namely into first and second class [19], will play a central part. Let v_a be a basis of the kernel of C,

$$\{\phi_j, \phi_m\}\, v_a^m \approx 0 \quad , \quad a = 1, \ldots, \dim \operatorname{Ker} C = M - \operatorname{rank} C. \tag{2.52}$$

The general solution for the multipliers u in (2.51) has then the form

$$u = \tilde{u} + \mu^a v_a, \tag{2.53}$$

where $\tilde{u}$ is a particular solution. We have separated the part of u that remains undetermined by the consistency conditions. This part contains $M - \operatorname{rank} C$ free functions μ^a.

[4]in the sequel I call all non-primary constraints secondary

The combinations of primary constraints

$$\phi_a = v_a^m \phi_m \tag{2.54}$$

weakly commute with all other constraints,

$$\{\phi_a, \phi_j\} \approx 0 \quad , \quad j = 1, \ldots, J. \tag{2.55}$$

Moreover, since the v^a form a basis of $\operatorname{Ker} C$, the ϕ_a are a complete set of primary constraints with this property. This leads to the concept of *first class functions* and in particular *first class constraints* (FCC). A function $F(q,p)$ is said to be first class if its Poisson bracket with *all constraints* vanish weakly (on Γ_c),

$$\{F, \phi_j\} \approx 0 \quad , \quad j = 1, \ldots, J. \tag{2.56}$$

The set of first class functions is closed under Poisson bracket [19]. This is proved as follows: if F, G are first class, then according to theorem 1

$$\{F, \phi\} = \phi' \quad , \quad \{G, \phi\} = \phi''$$

for any constraint ϕ, where ϕ', ϕ'' are some linear combinations of the constraints. Using the Jacobi identity we have

$$\{\{F, G\}, \phi\} = \{F, \{G, \phi\}\} - \{G, \{F, \phi\}\} = \{F, \phi''\} - \{G, \phi'\} \approx 0.$$

In particular the constraints ϕ_a are a complete set of first class primary constraints (modulo squares of second class constraints). Also, as a result of the Dirac-Bergman algorithm H_p is first class.

A function that is not first class is called *second class.* I use a notation adapted to this new classification. All primary and secondary FCC are denoted by γ_a. The remaining constraints are called second class constraints (SCC) and I denote them by χ_α.

The first property we need is that the matrix of SCC

$$\boxed{\Delta_{\alpha\beta} = \{\chi_\alpha, \chi_\beta\} \quad \text{is non-singular.}} \tag{2.57}$$

Indeed, if it was singular, then there would exist a null vector $r^\alpha \Delta_{\alpha\beta} \approx \{r^\alpha \chi_\alpha, \chi_\beta\} \approx 0$. Since $r^\alpha \chi_\alpha$ also commutes weakly with the FCC (by their first class property) it would weakly commute with all constraints and would be first class which contradicts our assumption. For counting degrees of freedom it is important to note that the number of SCC must be even. Otherwise the antisymmetric Δ would be singular.

Now consider the consistency conditions (2.51). They are identically fulfilled for the γ_a. For the SCC we have

$$\{\chi_\alpha, H\} + \Delta_{\alpha\beta} u^\beta \approx 0, \tag{2.58}$$

where $u^\beta = 0$ if χ^β is a secondary SCC. Solving for the multipliers we obtain

$$\Delta^{\beta\alpha}\{\chi_\alpha, H\} = \left\{ \begin{array}{lll} u^\beta, & \chi_\beta & \text{primary} \\ 0, & \chi_\beta & \text{secondary} \end{array} \right. , \tag{2.59}$$

where $\Delta^{\alpha\beta}\Delta_{\beta\gamma} = \delta^\alpha_\gamma$. Thus all multipliers belonging to the primary SCC are determined by the consistency conditions and we remain with the un-dermined multipliers μ^a in (2.53). We have the important result that the *number of undetermined multipliers is equal to the number of independent primary FCC.*

Inserting that into the equations of motion (2.47) we end up with

$$\boxed{\dot{F} \approx \{F, H\} + \{F, \phi_a\}\mu^a - \{F, \chi_\alpha\}\Delta^{\alpha\beta}\{\chi_\beta, H\},} \tag{2.60}$$

where the ϕ_a are the primary FCC. One can easily check that all constraints are preserved in time.

2.3.1 Second class constraints and Dirac bracket

For purely SC systems no multipliers remain in the time evolution (2.60) and there is no ambiguity in the dynamics. The term in (2.60) containing the inverse of Δ forces the system to stay on Γ_c. This surface is the *reduced phase space* for SC systems.

Motivated by (2.60) one introduces the *Dirac bracket* [18] for two phase space function as

$$\boxed{\{F, G\}^* \equiv \{F, G\} - \{F, \chi_\alpha\}\Delta^{\alpha\beta}\{\chi_\beta, G\},} \tag{2.61}$$

in terms of which

$$\dot{F} \approx \{F, H\}^* \tag{2.62}$$

for SC systems. This bracket possess the same properties as the Poisson bracket, i.e. they are antisymmetric, bilinear and obey the Jacobi identity and product rule. In addition we have

$$\{F, \chi_\alpha\}^* = 0 \;\; , \;\; \{F, G\}^* \approx \{F, G\} \;\; , \;\; \{F, \{G, K\}^*\}^* \approx \{F, \{G, K\}\} \tag{2.63}$$

for arbitrary F and first class G, K. These properties follow easily from the definition (2.61) and the property that first class functions have vanishing Poisson bracket with all constraints, e.g.

$$\{F, \chi_\alpha\}^* = \{F, \chi_\alpha\} - \{F, \chi_\beta\}\Delta^{\beta\gamma}\{\chi_\gamma, \chi_\alpha\} = 0. \tag{2.64}$$

Let us draw an immediate consequence of (2.64). According to theorem 1 any function can be replaced by its restriction to Γ_c, up to a linear combination of the constraints. Thus when calculating the Dirac bracket (2.61) between two functions we may replace them by their restriction to Γ_c since the other brackets vanish on account of (2.64). It follows that the *SCC can be set equal to zero either before or after evaluating the Dirac bracket.*

To understand the *geometric meaning* of SCC we recall some facts from symplectic geometry [4]:

In most cases the phase space Γ is the cotangental bundle T^*Q over the configuration space Q and hence is equipped with a natural symplectic structure (a non-degenerate closed two-form)

$$\omega = \omega_{\mu\nu}\, dx^\mu \wedge dx^\nu \tag{2.65}$$

which, according to Darboux, can locally be written as $\omega = dq^i \wedge dp_i$. Given ω, we can assign to a functions F its corresponding Hamiltonian vector field X_F by

$$i_{X_F}\omega = dF, \tag{2.66}$$

where i_X and d are the interior and exterior derivatives, respectively. In local coordinates we find

$$i_{X_F}\omega(Y) = \omega(X_F, Y) = \omega_{\mu\nu} X_F^\mu Y^\nu \quad \text{and} \quad dF(Y) = \partial_\nu F\, Y^\nu \tag{2.67}$$

for any vector field Y, so that

$$X_F^\mu = -\omega^{\mu\nu}\partial_\nu F, \quad \text{where} \quad \omega^{\mu\nu}\omega_{\nu\rho} = \delta^\mu_{\ \rho}. \tag{2.68}$$

The Poisson bracket of two functions is

$$\{F, G\} = -\partial_\rho F \omega^{\rho\sigma}\partial_\sigma G = \omega_{\mu\nu}\omega^{\mu\rho}\partial_\rho F \omega^{\nu\sigma}\partial_\sigma G = \omega(X_F, X_G). \tag{2.69}$$

In particular, the change of F under the Hamiltonian flow generated by G can be written as

$$F' \equiv \{F, G\} = \omega(X_F, X_G) = i_{X_F}\omega(X_G) = dF(X_G) = X_G^\mu \partial_\mu F. \tag{2.70}$$

In other words, the flows generated by G are just the motions along the Hamiltonian vector field X_G. For $G = H$ these are the Hamiltonian equations of motion.

Finally there is an important relation between the Poisson and Lie bracket,

$$[X_F, X_G] = -X_{\{F,G\}}, \quad \text{where} \quad [X, Y]^\mu = X^\alpha \partial_\alpha Y^\mu - Y^\alpha \partial_\alpha Y^\mu \tag{2.71}$$

are the Lie bracket. This relation follows from the Jacobi identity.

Let us now return to the SC systems. The inclusion map $j : \Gamma_c \longrightarrow \Gamma$ induces a two-form on Γ_c, namely the pull back of the symplectic form ω on Γ, $\omega_c = j^*\omega$. ω_c is closed since ω has this property, but it may be degenerate. In this case it is called *pre-symplectic.* However, for SCC it is indeed symplectic, as follows from

Theorem 3 *The χ_α are second class if and only if $\omega_c = j^*\omega$ is non-degenerate.*

Actually, we shall see that the Dirac bracket belongs to ω_c. Most properties of the Dirac bracket, and in particular the Jacobi identity follow then at once from the corresponding properties of ω.

To prove this theorem we must show that ω is non-degenerate on the vectors tangent to Γ_c. A vector field Y is tangent to Γ_c if $Y^\mu \partial_\mu \chi_\alpha$ vanishes for all constraints χ_α. With (2.70) this is equivalent to

$$\omega(X_\alpha, Y) \approx 0 \quad \text{for all} \quad X_\alpha \equiv X_{\chi_\alpha}. \tag{2.72}$$

On the other hand, from (2.69) follows that

$$\omega(X_\alpha, X_\beta) = \{\chi_\alpha, \chi_\beta\} \equiv \Delta_{\alpha\beta} \not\approx 0 \tag{2.73}$$

so that the Hamiltonian vector fields of the constraints are not tangent. Let us now determine the vectors X which obey

$$\omega(X, Y) \approx 0 \quad \text{for all tangent} \quad Y. \tag{2.74}$$

Since ω is non-degenerate (2.74) can have dimΓ-dimΓ_c independent solutions X. But because of our regularity conditions on the constraints the dimΓ-dimΓ_c Hamiltonian vector fields X_α, which are not tangent, are independent solutions. Thus any X which obeys (2.74) is a combination of the X_α. Hence there cannot be a tangent X obeying (2.74) and this proves that $j^*\omega$ is non-degenerate.

Note that we used the SC nature of the constraints and in particular that the *flows generated by the SCC lead off the constraint surface.*

Now it is easy to prove that the Dirac bracket furnishes an explicit representation for the induced Poisson bracket. For that consider

$$\{F, G\}^* = \omega(X_F, X_G) - \omega(X_F, X_\alpha)\Delta^{\alpha\beta}\omega(X_\beta, X_G) \equiv \omega^*(X_F, X_G). \tag{2.75}$$

It is easy to see, that $\omega^*(X_F + X_\chi, X_G) = \omega^*(X_F, X_G)$ for any Hamiltonian vector field belonging to the constraints. Thus ω^* depends only on the tangent components of X_F, X_G. But for tangent X_F we have $\omega(X_F, X_\alpha) \approx 0$ (see (2.72)) and ω^* can be replaced by ω without changing the value of (2.75). This proves that ω^* is just the pull-back of ω.

2.3.2 First class constraints and gauge transformations

Purely FC systems are relevant since gauge theories are of this type. Gauge related point should be identified and this leads to the problem of gauge invariant functions and/or the gauge fixing problem. The FCC together with a complete set of gauge fixing conditions form then a SC system. Hence for FC systems the gauge fixing define a subset $\Gamma_r \in \Gamma_c$ and this set is the *reduced phase space.*

For purely FC systems the time evolution is governed by

$$\dot{F} \approx \{F, H\} + \{F, \phi_a\}\mu^a, \tag{2.76}$$

with primary FCC ϕ_a. For the same initial conditions we get different evolutions, depending on the multipliers μ^a. The presence of arbitrary functions μ^a in the primary Hamiltonian tells us that not all $x=(q,p)$ are observable, i.e. there are several x representing a given physical state. Assume that the initial value $x(0)$ is given and represents a certain state. Then the equation of motion should fully determine the physical state at later times. So if $x'(t) \neq x(t)$ stem from the same physical state $x(0)$ then they should be identified.

Consider two infinitesimal time evolutions of $F = F(0)$ given by H_p with different values of the multipliers,

$$F_i(t) = t\{F, H\} + t\{F, \phi_a\}\mu_i^a, \quad \text{i=1,2} \quad . \tag{2.77}$$

The difference $F_2(t) - F_1(t)$ between the values is then

$$\delta_\mu F = \{F, \mu^a \phi_a\}, \quad , \quad \mu = t(\mu^2 - \mu^1). \tag{2.78}$$

Since such a transformation does not alter the the physical state at time t it is a *gauge transformation* [19]. Now we calculate

$$[\delta_\mu, \delta_\nu]F = \{\{\mu^a \phi_a, \nu^b \phi_b\}, F\} \tag{2.79}$$

and conclude that the commutator of any two primary FCC also generate gauge transformations. Also, performing a gauge transformation at $t = 0$ with multipliers ν and then time evolve with multipliers μ should lead to the same state as doing these transformations in the reverse order. We find

$$[\delta_{t,\mu}, \delta_\nu]F = t\{\dot{\nu}^a \phi_a - \{\nu^a \phi_a, H\} + \{\nu^a \phi_a, \mu^b \phi_b\}, F\} \tag{2.80}$$

and conclude that the commutators $\{\phi_a, H\}$ also generate gauge transformations.

We have seen that the first class functions are closed with respect to the Poisson bracket and thus the $\{\phi_a, \phi_b\}$ and $\{\phi_a, H\}$ are linear combinations of the FCC. However, in general there will appear secondary FCC in these combinations. Also, if we compared the higher order terms in the time evolutions (2.77) we would find that time derivatives of $\{\phi_a, H_p\}$ generate gauge transformations. This way secondary FCC show up as gauge transformations in all relevant systems and this lead Dirac to conjecture that *all* FCC

γ_a generate gauge transformations. We shall assume this conjecture to hold in what follows although there are some exotic counterexamples [2, 15].

Note, however, that if the structure constants in

$$\{\gamma_a, \gamma_b\} = t_{ab}^c \gamma_c \tag{2.81}$$

depend on the canonical variables, then $[\delta_\mu, \delta_\nu]F$ is a gauge transformation only on the constraint surface. Also, above we made the hidden assumption that time (or the space-time coordinates in field theory) is not transformed. Else we would have to take $F + \delta_\mu F$ at the transformed time $t + \delta_\mu t$ before calculating the second variation δ_ν. We come back to this point when discussing generally covariant theories.

We conclude that the most general physically permissible motion should allow for an arbitrary gauge transformation to be performed during the time evolution. But H_p contains only the primary FCC. We thus have to add to H_p the secondary FCC multiplied by arbitrary functions. This led Dirac to introduce the *extended Hamiltonian*

$$H_p \longrightarrow H_e = H + \mathcal{N}^a \gamma_a \tag{2.82}$$

which contains *all* FCC [19]. H_e accounts for all the gauge freedom.

Clearly, H_p and H_e should imply the same time evolution for the classical observables. *Observables* are gauge invariant functions on Γ_c, that is phase space functions that weakly commute with the gauge generators,

$$\boxed{F \quad \text{observable} \quad \Longleftrightarrow \{F, \gamma_a\} \approx 0 \quad \text{for all FCC} \quad \gamma_a.} \tag{2.83}$$

Since $H_e - H_p$ is a combination of the secondary FCC, we have

$$\dot{F} \approx \{F, H_p\} \approx \{F, H_e\} \tag{2.84}$$

for any observable F, as required. In the extended formalism one makes no distinction between primary and secondary FCC since they are treated symmetrically. The introduction of H_e is a new feature of the Hamiltonian scheme. It does not follow from the Lagrangian formalism.

At this point the following remark is in order. The number of time dependent functions which enter the Lagrangian off mass-shell gauge transformation is equals to the number of Lagrangian constraints and hence equals to the number of primary constraints. Hence, if there are secondary FCC it cannot be that any constraint $G = \mathcal{N}^a(t)\gamma_a$ generate off mass-shell transformations. Indeed, when discussing the consistency conditions we assumed that the primary constraints must be conserved only for on mass-shell trajectories. Hence for arbitrary off mass-shell variations the secondary constraints do not guarantee that the primary constraints are respected. We come back to this important point in chapter 4.

Let us now investigate the *geometric meaning* of FC systems. As a preparation we show:

The induced 2-form $j^*\omega$ has rank $\geq N-2M$, where M is the number of independent first or second class constraints.

Let us assume that the tangent vectors X_p, $p=1,\dots,P$ form a basis for the null-eigenvectors of $j^*\omega$, i.e. $j^*\omega(X_p,Y)=\omega(X_p,Y)=0$ for all tangent Y. Let us now see how big P can be. For that we consider

$$a^p\omega(X_p,Z_q)=0, \quad \text{where the} \quad Z_q, \quad q=1,\dots,M$$

together with the tangent vectors form a basis of $T\Gamma$ at the point under consideration. These are M equations for P unknown. So, if $P\geq M$ then there would always exist a solution $X=a^pX_p$ with $\omega(X,Z_q)=0$ for all Z_q. Being also a null-eigenvector of $j^*\omega$ we would conclude that $\omega(X,Z)=0$ for all vectors $Z\in T\Gamma$ or that ω is degenerate. This then proves the statement above. Now we have the following

Theorem 4 *For a FC system the induced two-form $j^*\omega$ is maximally degenerate. The kernel is spanned by the Hamiltonian vector fields belonging to the FCC.*

First, if $X^\mu\partial_\mu\gamma_b\approx 0$ for all constraints, then X is tangent. But since $X_a^\mu\partial_\mu\gamma_a\approx\{\gamma_a,\gamma_b\}\approx 0$, all Hamiltonian vector fields belonging to the constraints are tangent. Second, for an arbitrary tangent vector X we have $\omega(X_a,X)=i_{X_a}(X)=\partial_\mu\chi_aX^\mu\approx 0$. Thus the M X_a's are null-eigenvectors of the induced two-form and the rank of $j^*\omega$ equals $2N-2M$, i.e. it is maximally degenerate.

Thus we have the following situation: The M FCC generate *flows which stay on the constraint surface* and which we identified with gauge transformations. The Hamiltonian vector fields belonging to the constraints are the null-directions of the induced pre-symplectic 2-form. That these null-vector fields generate gauge orbits follows from

Theorem 5 *On Γ_c the vectors X_a generate M-dimensional manifolds.*

The proof uses the Frobenius integrability condition, which says that M linearly independent vector fields are integrable (through each point in Γ_c there is a surface, the gauge orbit, to which the X_a are tangent) iff all Lie brackets $[X_a,X_b]$ are linear combinations of $(X_1,\dots,X_M)$. Indeed,

$$[X_a,X_b]=-X_{\{\gamma_a,\gamma_b\}}=-t_{ab}^{\ c}X_c+\gamma_c\omega^{\mu\nu}\partial_\nu t_{ab}^{\ c}\approx -t_{ab}^{\ c}X_c, \tag{2.85}$$

where we used (2.81). Note that for (q,p)-dependent structure constant (as in gravity) the null vector fields are integrable only on the constraint surface.

In a next step one wants to eliminate the gauge degrees of freedom that is identify points on the same gauge orbit. This can in principle be achieved by introducing gauge invariant variables, e.g. the transverse potential or holonomies in electrodynamics, or alternatively by fixing the gauge. A *gauge fixing* must obey two conditions: first it must be attainable and second it should fix the gauge uniquely [5]. We can fix the gauge by imposing the independent conditions

$$F_a(q,p) = 0, \quad \text{a=1,\ldots,M}. \tag{2.86}$$

The surface defined by these conditions should intersect every gauge orbit in exactly one point. A necessary condition is that at least one gauge fixing function F_b should change in the direction of all null-vectors X_a. In other words, there is at least one F_b so that

$$\lambda^a (X_a, \nabla F_b) = \lambda^a \{\gamma_a, F_b\} \neq 0 \tag{2.87}$$

for all λ. This implies that

$$\det\{\gamma_a, F_b\} \equiv \det\ F_{ab} \neq 0. \tag{2.88}$$

In particular, if we could choose gauge fixings canonically conjugated to the constraints, $\{\gamma_a, F_b\} = \delta_{ab}$, then the gauge orbits would intersect the gauge fixing surfaces orthogonal and in this case $\det F_{ab} = 1$. The determinant of F plays an important part in the quantization of gauge systems and is the well-known *Faddeev-Popov determinant* [21].

Because of (2.88) the FCC together with the gauge fixings form a SC system and we can take over the result from the previous subsection. The reduced phase space Γ_r consists of the points fulfilling the constraints and gauge fixings. Collecting the γ_a and F_a into one vector, Ω_p, $p = 1, \ldots, 2M$, we find for the Hamiltonian equation of motion for any phase space function

$$\dot{F} = \{F, H\} - \{F, \Omega_p\} G^{pq} \{\Omega_q, H\}. \tag{2.89}$$

Mixed second and first class constraints Before gauge fixing the evolution is governed by the first class partner of the extended Hamiltonian

$$H_e^* = H + \gamma_a \mathcal{N}^a - \chi_\alpha \Delta^{\alpha\beta} \{\chi_\beta, H\} \tag{2.90}$$

since we have

$$\dot{F} = \{F, H_e^*\} = \{F, H\} + \{F, \gamma_a\} \mathcal{N}^a - \{F, \chi_\alpha\} \Delta^{\alpha\beta} \{\chi_\beta, H\}. \tag{2.91}$$

[5] There may be obstructions to fulfilling these requirements as has been demonstrated by Gribov and Singer [28, 44].

For a discussion of the *starred variables* see [12]. After gauge fixing one can again introduce starred variables with respect to the SC system $\psi_I = (\chi_\alpha, \gamma_a, F_b)$. Denoting the Poisson brackets matrix of all these constraints ψ_I by $\tilde{\Delta}$, we have

$$\dot{F} = \{F, H^*\}, \quad \text{where} \quad H^* = H + \psi_I \tilde{\Delta}^{IJ}\{\psi_J, H\}. \tag{2.92}$$

2.4 First order action principles

The solutions to the primary Hamiltonian equations of motion (2.45, 2.46) extremize the primary (or total) first order action,

$$\delta S_p = \delta \int_{t_1}^{t_2} \Big(\dot{q}^i p_i - H - \sum_{primary} u^m \phi_m\Big) dt = 0 \tag{2.93}$$

with respect to variations $\delta q, \delta p, \delta u$ subject only to the restriction $\delta q(t_1) = \delta q(t_2) = 0$. The variables u^m which have been introduced to make the Legendre transformation invertible appear now as Lagrange multipliers enforcing the *primary constraints.* It is clear that the theory is invariant under $H \to c^m \phi_m$ since such a change can be absorbed into the Lagrange multipliers.

The variational principle (2.93) is equivalent to

$$\delta \int \Big(\dot{q}^i p_i - H\Big) dt = 0 \quad \text{subject to} \quad \phi_m = \delta\phi_m = 0. \tag{2.94}$$

There is yet another variational principle which for gauge invariant observables leads to the same time evolution, namely the *extended action principle.* The equations of motion for the extended formalism follow from

$$\boxed{\delta S_e = \delta \int \Big(\dot{q}^i p_i - H - \sum_{all\, constr.} u^j \phi_j\Big) dt = 0,} \tag{2.95}$$

where the sum extends over *all constraints.*

Take the case of purely SCC and let $y^i \to x^\mu(y^p)$ be the embedding of $\Gamma_r \subset \Gamma$. The Lagrange multiplier method guarantees that the implementation of the constraints χ_α, either directly or via the Lagrange multipliers, are equivalent. Now let us solve the constraint directly in (2.95). Recall that a symplectic 2-form can locally be written as $\omega = d\theta$. The pull-back of the one-form potential θ is

$$j^*(\theta) = j^*(p_i dq^i) \equiv j^*(a_\mu dx^\mu) = a_\mu(x(y)) \frac{\partial x^\mu}{\partial y^p} dy^p. \tag{2.96}$$

Inserting this into the extended action one finds

$$S_e[y] = \int \{\theta - H(x(y))dt\} = \int (a_p \dot{y}^p - h)dt. \tag{2.97}$$

The corresponding variational principle yield then the equation of motion for SC systems:

$$\delta S_e[y] = 0 \Longleftrightarrow \dot{y}^p = \{y^p, h(y)\}^*. \tag{2.98}$$

This can be checked directly by using

$$j^*\omega = j^* d\theta = d\, j^*\theta = d\, a_p(y) dy^p. \tag{2.99}$$

The fact that (2.95) yields (2.97) is of practical use when calculating Dirac bracket. One solves for the constraints inside the action and from the new kinetic term one reads off the induced potential form on Γ_c. From (2.99) one computes the induced symplectic form and thus the Dirac bracket.

For FC systems it is also legitimate to solve the FCC inside the action

$$S_e = \int \left(q^i p_i - H - \mathcal{N}^a \gamma_a\right) dt. \tag{2.100}$$

However, since the induced 2 form is degenerate the equations of motion on Γ_c are not canonical. To get Hamiltonian equations one needs to go to Γ_r by imposing additional gauge conditions. Then one may write down the corresponding action for the SC system as discussed above.

2.5 Abelian Chern-Simons Theory with Sources

To see how the general formalism works in an explicit example I consider the Abelian Chern-Simons model [29, 41, 32, 24]. This is a field theory for a gauge potential A_a in 3 space-time dimensions with coordinates $x = (x^0, x^1, x^2) \equiv (t, \vec{x})$ with first order Lagrangian density

$$\begin{aligned} \mathcal{L} \;&= \tfrac{\kappa}{4} A^a \epsilon_{abc} F^{cb} + A^a J_a, \quad \text{where} \\ F_{ab} \;&= \partial_a A_b - \partial_b A_a \quad , \quad \partial^a J_a = 0. \end{aligned} \tag{2.101}$$

Indices are lowered with the metric $\eta_{ab} = \mathrm{diag}(1, -1, -1)$ and ϵ_{abc} is the Levi-Civita symbol, $\epsilon_{012} = 1$. We enclose the system in a finite box $[0, L] \times [0, L]$. The quantum theory is sensitive to the value of the coupling constant κ. For rational $2\pi\kappa$ and vanishing external current J the Hilbert space becomes finite-dimensional [41].

For arbitrary periodic currents the action is invariant under $U(1)$-gauge transformations

$$A^a \to A^a + \partial^a \lambda \quad , \quad S \to S + \oint n^a \left(\frac{\kappa}{4} \lambda \epsilon_{abc} F^{cb} + \lambda J_a\right) \tag{2.102}$$

provided λ vanishes at the initial and final times and λ, F_{01}, F_{02} are periodic in x^1, x^2 with period L. So we shall *assume* that λ and F_{ab} are both periodic.

2.5.1 Dirac theory for $U(1)$-Chern-Simons theory

Since $\mathcal{L}$ is linear in the first derivatives the Hessian vanishes identically and the model is singular. Thus we expect 3 independent primary constraints (per space point). More explicitly, the canonical momentum densities are

$$\pi_a(\vec{x}) = \frac{\delta L}{\delta \dot{A}^a(\vec{x})} = \frac{\kappa}{2} A^b(\vec{x})\epsilon_{ba0} \tag{2.103}$$

and immediately lead to the *primary constraints*

$$\{\phi_m\} = \{\pi_0, \pi_1 + \frac{\kappa}{2}A^2, \pi_2 - \frac{\kappa}{2}A^1\} \quad , \quad m = 1,2,3. \tag{2.104}$$

The canonical Hamiltonian becomes

$$\begin{aligned} H &= \int d^2x \Big(\pi_a(\vec{x})\dot{A}^a(\vec{x}) - \mathcal{L}\Big) = \int d^2x\, \mathcal{H} \\ &= -\int d^2x \Big(\frac{\kappa}{2} A^a \epsilon_{abi}\partial^i A^b + A^a J_a\Big) \quad , \quad i = 1,2. \end{aligned} \tag{2.105}$$

and the time evolution is determined by (2.47) with primary Hamiltonian

$$H_p = \int d^2x\, \mathcal{H}_p \quad , \quad \mathcal{H}_p = \mathcal{H} + u^m \phi_m, \tag{2.106}$$

and fundamental Poisson bracket

$$\{A^a(\vec{x}), \pi_b(\vec{y})\} = \delta^a_b \delta(\vec{x} - \vec{y}). \tag{2.107}$$

Let us now see whether secondary constraints arise from the consistency conditions $\dot{\phi}_m \approx 0$. One computes

$$\begin{aligned} \dot{\phi}_1(\vec{x}) &= \int d^2y\, \{\pi_0(\vec{x}), \mathcal{H}_p(\vec{y})\} \\ &= -\int d^2y\, \{\pi_0(\vec{x}), \kappa A^0(\vec{y})\epsilon_{0ji}\partial^i A^j(\vec{y}) + J_0(\vec{y})A^0(\vec{y})\} \\ &= \int d^2y\, \Big(\kappa\epsilon_{0ji}\partial^i A^j(\vec{y}) + J_0(\vec{y})\Big)\delta(\vec{x} - \vec{y}) \\ &= \epsilon_{0ji}\kappa\partial^i A^j(\vec{x}) + J_0(\vec{x}) \end{aligned}$$

leading to the *secondary constraint*

$$\phi_4(\vec{x}) = \kappa F_{12} - J_0(\vec{x}). \tag{2.108}$$

There is a quicker way to arrive at this conclusion, since $\dot{\phi}_1 = -\partial\mathcal{H}_p/\partial A^0$.

The time derivative of the other two primary constraints are

$$
\begin{aligned}
\dot{\phi}_2(\vec{x}) &= \kappa(u^3 - \partial^2 A^0(\vec{x})) + J_1(\vec{x}) \\
\dot{\phi}_3(\vec{x}) &= \kappa(-u^2 + \partial^1 A^0(\vec{x})) + J_2(\vec{x})
\end{aligned}
\tag{2.109}
$$

and putting them weakly to zero fixes the multipliers u^2, u^3. Finally, we must have

$$
\dot{\phi}_4(\vec{x}) = \kappa\partial^2 u^2(\vec{x}) - \kappa\partial^1 u^3(\vec{x}) + \partial_0 J_0(\vec{x}) \approx 0. \tag{2.110}
$$

Inserting u^2, u^3 from (2.109) this becomes $\partial_a j^a = 0$ and yields no further condition. Thus the Dirac-Bergman algorithm leads to 3 primary and 1 secondary constraint.

Obviously $\phi_1 \equiv \gamma_1$ is first class. Also the combination

$$
\gamma_2 = \partial_1\phi_2 + \partial_2\phi_3 + \phi_4 = \partial_i\pi_i + \frac{\kappa}{2}F_{12} - J_0 \tag{2.111}
$$

is first class and is the analog of the Gauss constraint in electrodynamics. As SCC we may take

$$
\chi_1 = \phi_2 \quad \text{and} \quad \chi_2 = \phi_3 \Longrightarrow \Delta_{\alpha\beta} = \kappa \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \delta(\vec{x} - \vec{y}). \tag{2.112}
$$

Δ has inverse $\Delta^{-1} = -\Delta/\kappa$ and (2.91) reads

$$
\dot{F} \approx \{F, H\} + \{F, \gamma_a\}\mathcal{N}^a + \frac{1}{\kappa}\int d^2y\, \{F, \chi_i(\vec{y})\}\, \epsilon_{ij}\, \{\chi_j(\vec{y}), H\}. \tag{2.113}
$$

Since the FCC commute with all constraints they generate transformations on Γ_c, i.e. if (A, π) is on Γ_c then

$$
\begin{aligned}
\delta_{\mathcal{N}}\, A(\vec{x}) &= \textstyle\int d^2y\, \{A(\vec{x}), \gamma_a(\vec{y})\}\, \mathcal{N}^a(\vec{y}) \\
\delta_{\mathcal{N}}\, \pi(\vec{x}) &= \textstyle\int d^2y\, \{\pi(\vec{x}), \gamma_a(\vec{y})\}\, \mathcal{N}^a(\vec{y})
\end{aligned}
\tag{2.114}
$$

are variations tangent to it. This follows from $\delta_{\mathcal{N}}\phi_j = \int\{\phi_j, \gamma_a\}\mathcal{N}^a \approx 0$. Also, since these transformations commute with H_p, one expects that they are related to infinitesimal gauge transformations (2.102). Indeed, defining

$$
G = \int d^2y \Big(\partial_0\lambda(\vec{y})\gamma_1(\vec{y}) + \lambda(\vec{y})\gamma_2(\vec{y})\Big) \tag{2.115}
$$

one finds

$$
\delta_\lambda A^a(\vec{x}) = \partial^a\lambda(\vec{x}). \tag{2.116}
$$

Only the particular combination (2.115) of the FCC generate the off mass-shell gauge transformations (2.102). This particular combination has the property that $\dot{G}$ is a combination of the *primary* FCC constraints only. In chapter 4 we shall show more generally that for systems with FCC which are linear in the momenta a combination of FCC generate Lagrangian off mass-shell symmetries if its time-derivative is a combination of primary FCC.

2.5.2 Gauge fixing of Chern-Simons theory

We have seen that the Chern-Simons theory possesses two SCC and two FCC. Now we supplement those by two gauge fixing conditions, namely

$$F_1 = A_0 \quad \text{and} \quad F_2 = \partial_i A^i. \tag{2.117}$$

Altogether the conditions $(\chi_1, \chi_2, \gamma_1, F_1, \gamma_2, F_2)$ form a SC system and define Γ_r. The surface defined by χ_i, γ_1 and F_1 can be parametrized by the spatial components of the gauge potential which can be decomposed as

$$A_i = \epsilon_{ij}\partial_j\varphi + \partial_i\lambda + \frac{1}{L}q_i \tag{2.118}$$

with constant q_i [6]. Imposing further γ_2 and F_2 we see that

$$A_i = -\frac{1}{\kappa}\epsilon_{ij}\partial_j\frac{1}{\triangle}J^0 + \frac{1}{L}q_i \tag{2.119}$$

so that $\Gamma_r = \{q_1, q_2\}$ is finite-dimensional. Furthermore, $\gamma_2 = 0$ and the periodicity of the A_i imply that the total charge $Q = \int d^2x J^0$ must vanish.

The inverse Poisson bracket 'matrix' reads

$$(\tilde{\Delta}^{IJ})(x,y) = \frac{1}{\kappa}\begin{pmatrix} 0 & -1 & 0 & 0 & \frac{1}{\triangle}\partial_2 & 0 \\ 1 & 0 & 0 & 0 & -\frac{1}{\triangle}\partial_1 & 0 \\ 0 & 0 & 0 & \kappa & 0 & 0 \\ 0 & 0 & -\kappa & 0 & 0 & 0 \\ \frac{1}{\triangle}\partial_2 & -\frac{1}{\triangle}\partial_1 & 0 & 0 & 0 & -\kappa\frac{1}{\triangle} \\ 0 & 0 & 0 & 0 & \kappa\frac{1}{\triangle} & 0 \end{pmatrix}\delta(x-y)$$

and one finds the following Dirac bracket for the coordinates on Γ_r

$$\{q_i, q_j\} = -\delta_{ij}. \tag{2.120}$$

When calculating the starred Hamiltonian, one should recall that for a periodic function $\triangle^{-1}\triangle f = f - V^{-1}\int f$. After some algebra one finds

$$H^* = -\int d^2x\Big\{A^0(\kappa F_{12} - J_0) + \frac{1}{\kappa}J_0\frac{1}{\triangle}\epsilon_{ij}\partial_i J_j\Big\} - \frac{1}{2}j_i q^i - \frac{1}{\kappa}\epsilon_{ij}j_i p_j, \tag{2.121}$$

where we have introduced the mean 'fluxes'

$$j_i \equiv \frac{1}{L}\int d^2x\, J_i \quad , \quad q^i \equiv \frac{1}{L}\int d^2x\, A^i \quad \text{and} \quad p_i \equiv \frac{1}{L}\int d^2x\, \pi_i. \tag{2.122}$$

[6]The $U(1)$-bundle over the torus is non-trivial and A must be periodic only up to non-trivial gauge transformations [42]. For simplicity we assume here that A is periodic and hence $\int F_{12} = 0$

After imposing the constraints χ_i, γ_2 and F_1 the non-trivial equations of motion take the simple form

$$\dot{q}_i = -\frac{1}{\kappa}\epsilon_{ij} j_j. \tag{2.123}$$

Of course, the evolution belonging to H^* stays on Γ_r. Since Γ_r is 2-dimensional the (topological) Chern-Simons theory is effectively a simple mechanical system. This was expected from the beginning since there are 6 constraints and gauge fixings for 6 degrees of freedom (per $\vec{x}$).

To see the meaning of this result more clearly, let us see what the *observables* are. As coordinates on Γ_c we may take λ and q_i in (2.118), so that we are considering functionals $F[A^0, \lambda, q_i, J_a]$. Such F commute with the FCC if they are independent of λ and A^0. Hence observables have the form

$$F = F[J_a, q_i] \tag{2.124}$$

and depend only on the zero-modes of the A_i.

Let us finally remark that for a pure CS theory ($J=0$) the Lagrangian density is invariant, up to a total time derivative, under gauge transformations for which only $e^{i\lambda}$ must be periodic. This introduces global gauge transformations with windings around the handles of the torus (the box with opposite points identified). Hence we must identify gauge potentials which are related by such global gauge transformations

$$A_i \sim A_i + \frac{2\pi}{L} n_i \quad \text{or} \quad q_i \sim q_i + 2\pi n_i. \tag{2.125}$$

Gauge invariant functionals must be invariant under such transformations. Thus they depend only on $\exp(i\sum m_i q_i)$. For pure Chern-Simons theories we have

$$e^{i\sum m_i q_i} = W(C, A) = \exp\left\{i \oint_C A\right\} \tag{2.126}$$

on the constraint surface ($F_{12}=0$) if the loop C winds m_i-times around the torus in the direction i. For a contractible loop $W(C,A)$ vanishes. Thus, observables have the form

$$F(A) = F\left(e^{i\oint_C A}\right). \tag{2.127}$$

Let C, D be 2 loops which wind m_i, n_i-times around the torus in the direction i. We parametrize them by $x(\tau), y(s)$. We compute

$$\oint_C \oint_D \{A(x(\tau)), A(y(s))\} = -\oint_C \oint_D \epsilon_{ij}\dot{x}^i(\tau)\dot{y}^j(s) d\tau ds = -(n_1 m_2 - n_2 m_1).$$

Upon deformation of the curves the commutator is invariant and therefore is a *topological invariant*. This can be understood by noting that for $J=0$ the

Chern-Simons model (2.101) is invariant under space-time diffeomorphisms. In particular the *spatial* ones are generated by

$$G_{diff} = \int d^2x \, \epsilon^i A_i \gamma_2 \tag{2.128}$$

and hence observables must be invariant under spatial diffeomorphisms.

Finally note, that for the pure Chern-Simons theory the phase space variables q_i lie in $[0, 2\pi]$, that is Γ_r is compact and as a consequence the Hilbert space becomes finite dimensional.

Actually there is a quicker way to arrive at these results if one inserts the fields on the reduced phase space (2.119) into the first order action. One easily finds

$$S = \int dt (\frac{\kappa}{2} q_i \epsilon_{ij} \dot{q}_j - q_i j_i) + \frac{1}{\kappa} \int dt d^2x \, J^0 \frac{1}{\triangle} \epsilon_{ij} \partial_i J_j \tag{2.129}$$

which of course reproduces the correct equations of motion (2.123).

Bibliography

[1] E.S. Albers and B.W Lee, Phys. Rep. **9C** (1973) 1

[2] T.R. Allcock, Phil. Trans. R. Soc. (Lond.) **279** (1975) 33

[3] J.L. Anderson and P.G. Bergman, Phys. Rev. **83** (1951) 1018

[4] V. Arnold, *Mathematical Methods of Classical Mechanics*, Graduate Texts in Math. 60, Springer, Berlin 1978

[5] O. Babelon and C.M. Viallet, Phys. Lett. **85B** (1979) 246

[6] I.A. Batalin and G.A. Vilkovisky, Phys. Lett. **69B** (1977) 309

[7] C. Becchi, C. Rouet and R. Stora, Ann. Phys. **98** (1976) 287; Phys. Lett. **B52** (1974) 344; Commun. Math. Phys. **42** (1975) 127

[8] R. Benguria P. Cordero and C. Teitelboim, Nucl. Phys., **B122** (1977) 61

[9] F.A. Berezin, *The Method of Second Quantization*, New York, Academic Press 1966

[10] F.A. Berezin and M.S. Marinov, Ann. Phys. **104** (1977) 336

[11] P.G. Bergman, I. Goldberg, A. Janis and E. Newman, Phys. Rev. **103** (1956) 807

[12] P.G. Bergman and A. Komar, Phys. Rev. Lett. **4** (1969) 432

[13] P.G. Bergman, Rev. Mod. Phys. **33** (1961) 510

[14] R. Casalbuoni, Nuovo Cimento, **33A** (1976) 115; Nuovo Cimento, **33A** (1976) 389

[15] R. Cawley, Phys. Rev. Lett. **42** (1979) 413

[16] P. Chernoff and J. Marsden, Lecture Notes Math. **425** (1974)

[17] B. DeWitt, *Dynamical Theory of Groups and Fields*, Gordon and Breach, New York 1965

[18] P.A.M. Dirac, Can. J. Math., **2** (1950) 129; Can. J. Math. **3** (1951) 1; Proc. R. Soc. (Lond.) **A246** (1958) 326

[19] P.A.M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Yeshiva University, New York 1964

[20] L.D. Faddeev, Theor. Math. Phys. **1** (1969) 1

[21] L.D. Faddeev and V.N. Popov, Phys. Lett. **25B** (1967) 29

[22] L.D. Faddeev and R. Jackiw, Phys. Rev. Lett. **60** (1988) 1692

[23] L. Feher, L. O'Raifeartaigh, P. Ruelle, I. Tsutsui and A. Wipf, Phys. Rep. **222** (1992) 1

[24] J. Fröhlich and U. Studer, Rev. Mod. Phys. **65** (1993) 733

[25] J.L. Gervais and B. and Sakita, Nucl. Phys. **B64** (1971) 632

[26] D.M. Gitman and I.V. Tyutin, *Quantization of Fields with Constraints*, Berlin, Springer 1990

[27] J. Govaers, *Hamiltonian Quantisation and Constrained Dynamics*, Leuven University Press 1991

[28] V.N. Gribov, Nucl. Phys. **B139** (1978) 1

[29] C.R. Hagen, Ann. Phys. **157** (1984) 342

[30] A.J. Hanson, T. Regge and C. Teitelboim, *Constrained Hamiltonian Systems*, Accademia Nazionale dei Lincei, 1976

[31] M. Henneaux and C. Teitelboim, *Quantization of Gauge Systems*, Princeton Univ. Press, 1992

[32] R. Iengo and K. Lechner, Phys. Rep. **213** (1992) 179

[33] R. Jackiw, Seminair de Mathematiques Superieurs, Montreal, Quebec 1988

[34] C. Kiefer and A. Wipf, *Functional Schrödinger Equation for Fermions in External Gauge Fields*, preprint ETH-TH/1993-17

[35] T. Kimura, Prog. Theor. Phys. **27** (1962) 7478

[36] M. Lüscher, Nucl. Phys. **254** (1985) 52

[37] J. Marsden, *Applications of Global Analysis in Math. Physics*, Publish or Perish, Boston 1974

[38] P.K. Mitter and C.M. Viallet, Comm. Math. Phys. **79** (1979) 457

[39] V. Mukhanov and A. Wipf, *Unified Treatment of local Symmetries*, preprint ETH-TH/92-49

[40] A.M. Polyakov, Mod. Phys. Lett. **A3** (1988) 3258

[41] A.P. Polychronakos, Ann. Phys. **203** (1990) 231

[42] I. Sachs and A. Wipf Helv. Phys. Acta **65** (1992) 652

[43] S. Shanmugadhasan, Proc. Camb. Phil Soc. **59** (1963) 743

[44] I.M Singer, Comm. Math. Phys. **60** (1978) 78

[45] E.C.G. Sudarshan and N. Mukunda, *Classical Dynamics, A Modern Perspective*, Wiley, New York 1974

[46] K. Sundermeyer, *Constrained Dynamics*, Springer Lecture Notes 169, 1982

[47] K. Symanzik, Nucl. Phys. **190** [FS3] (1981) 1

[48] A. Trautman, *Lectures on General Relativity*, Brandeis Summer Institute in Theor. Phys., Prentice-Hall, N.J., 1964

[49] I.V. Tyutin, Lebedev preprint FIAN, (1975) 39

[50] R. Utiyama, Progr. Theor. Phys. Suppl. **9** (1959) 19

[51] C. Wiesendanger and A. Wipf, *Running Coupling Constants from finite Size Effects*, preprint ETH-TH-93/10

[52] E. Witten, Commun. Math. Phys. **121** (1989) 351

[53] T.T. Wu and C.N. Yang, Phys. Rev. **D12** (1975) 3845

[54] D. Zwanziger, Nucl. Phys. **B209** (1982) 336; Nucl. Phys. **B323** (1989) 513; Nucl. Phys. **B345** (1990) 461

[55] C.A.P. Galvao and C. Teitelboim, J. of Math. Phys. **21** (1980) 1863.