UNIVERSITY OF HELSINKI REPORT SERIES IN PHYSICS

HU-P-D126

GAUSS'S LAW IN YANG–MILLS THEORY

ANTTI SALMELA

Theoretical Physics Division
Department of Physical Sciences
Faculty of Science
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION
To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in Auditorium E207 at Physicum, Gustaf Hällströmin katu 2, on 17th October, 2005, at 12 o'clock.

Helsinki 2005

ISBN 952-10-2107-1
ISSN 0356-0961
ISBN 952-10-2108-X (PDF)
http://ethesis.helsinki.fi
Yliopistopaino
Helsinki 2005

Preface

This work is based on research done during the past few years at the Theoretical Physics Division of the Helsinki University Department of Physical Sciences. In the course of the work I have received financial support from the Academy of Finland under grant No. SA 16832 and also from the Magnus Ehrnrooth foundation and the Finnish Cultural Foundation. This support is gratefully acknowledged.

First of all, I would like to thank my supervisor Christofer Cronström. He has both guided me with his experience and encouraged me to develop ideas on my own, and in addition he has made substantial efforts to arrange resources for this research. I am also indebted to the referees of this thesis, Antti Niemi and Stephen Hwang, for their comments and suggestions. The regulars of the Theory Division coffee hour and my room-mates deserve special thanks for creating a nice and not too serious working environment. I am also grateful to each and everyone who has helped me in some way during the preparation of this thesis.

Helsinki, September 2005

Antti Salmela

A. Salmela: Gauss's law in Yang–Mills theory, University of Helsinki, 2005, 57 p, University of Helsinki, Report Series in Physics, HU-P-D126, ISSN 0356-0961, ISBN 952-10-2107-1 (printed version), 952-10-2108-X (PDF)

INSPEC classification: A1110N, A1110E, A1235C, A0350, A1130
Keywords: Yang–Mills theory, Hamiltonian formalism, Gauss's law, quantum chromodynamics

Abstract

Present knowledge of quantum Yang–Mills theory in the continuum is mainly limited to the perturbative weak-coupling regime. In order to study strong-coupling phenomena such as colour confinement and chiral symmetry breaking a nonperturbative formulation of the theory is required at the quantum level. The Hamiltonian formalism seems to be more appropriate for determining the physical states of the theory than the path-integral approach, but it cannot be derived from the classical Lagrangian as easily. Giving up Gauss's law, it is possible to construct an extended Hamiltonian system where certain degrees of freedom are nonphysical. Gauss's law must then be implemented later, either at the classical or at the quantum level.

This thesis investigates the Gauss law from several viewpoints. Two of the research papers concentrate on solving the law at the Lagrangian level. One of them presents an algebraic solution method where a non-Abelian Hodge decomposition is derived for the colour-electric field. The second paper examines the solutions of Gauss's law in the Lagrangian formalism by functional analytic methods. Theorems regarding the existence and smoothness of solutions are stated and proved. The third paper presents a novel method for implementing Gauss's law in a classical Hamiltonian formulation of Yang–Mills theory. This approach is based on a careful analysis of the Poisson bracket algebra of the Gauss law generators. The review part of the thesis gives an overview of classical and quantum Yang–Mills theory in general, and it discusses some of the most important methods by which a Hamiltonian formalism can be derived.

Contents

Preface i
Abstract ii
List of papers iv

1 Introduction **1**

2 Yang–Mills theory **3**
2.1 Some history 3
2.2 Yang–Mills Lagrangian 4
2.3 Equations of motion 7
2.4 Symmetries and conservation laws 9
2.4.1 Noether's theorem 9
2.4.2 Gauge invariance 10
2.4.3 Poincaré invariance 11
2.4.4 Conformal symmetries 14
2.4.5 Fermionic symmetries 16
2.5 Path-integral quantisation 18

3 Hamiltonian formalism **22**
3.1 Motivation 22
3.2 The Gauss law problem 23
3.3 Lagrangian approach 25
3.4 Constrained quantum theory 28
3.4.1 Dirac quantisation 28
3.4.2 BRST formalism 31
3.5 Unconstrained classical theory 34
3.5.1 Dirac's formalism 34
3.5.2 Abelianisation 37
3.5.3 Function group method 39

4 Conclusions **43**

A Lie groups and algebras **44**

B Functional analysis **47**
B.1 Banach and Hilbert spaces 47
B.2 Sobolev spaces 49
B.3 Continuous embeddings 51

Bibliography **53**

List of papers

This thesis consists of an introductory review part, followed by three research papers:

I: A. Salmela,
An algebraic method for solving the SU(3) *Gauss law*,
J. Math. Phys. **44** (2003) 2521.

II: A. Salmela,
Covariant Poisson equation with compact Lie algebras,
J. Math. Phys. **45** (2004) 2844.

III: A. Salmela,
Function group approach to unconstrained Hamiltonian Yang–Mills theory,
hep-th/0409142, to appear in J. Math. Phys. **46** (2005).

Chapter 1

Introduction

Yang–Mills theory has celebrated its 50th anniversary, but it has not ceased to offer challenges to physicists. Maybe surprisingly, a number of basic questions are still without answers. For instance, despite the simple-looking form of the classical Yang–Mills Lagrangian no-one has managed to solve the equations of motion in the general case. At the quantum level it has proved difficult even to give a universal and mathematically solid definition of quantum Yang–Mills theory [1]. The perturbative sector of the theory has been quantised successfully, but the strong-coupling regime is well beyond the reach of ordinary perturbation theory. In asymptotically free theories such as QCD the coupling grows strong at large spatial distances, where phenomena like colour confinement and spontaneous chiral symmetry breaking are expected to occur. These phenomena are properties of hadronic states, which are well-known experimentally but so far poorly explained by QCD. Despite this state of affairs it is widely believed that QCD is the correct theory of strong interactions, mainly because lattice calculations in the strong-coupling regime yield encouraging results. From a theoretical point of view, however, a lattice with a finite number of points is not a good substitute for the continuum, particularly when the power of contemporary computers is limited to relatively small lattices. Therefore one should keep trying to develop calculation methods for the quantum Yang–Mills theory of the continuum, in particular in the non-perturbative sector of the theory.

Traditionally the quantisation of classical systems has relied on the path-integral formalism and the canonical approach. In Yang–Mills theory the path-integral formalism is at present more firmly established than the canonical alternative, probably because perturbative quantum field theory is involved with transition amplitudes more than with state functionals. In the strong-coupling regime of QCD, on the other hand, knowledge of the states would be more important so that comparisons to experimentally observed hadronic states could be made. Canonical quantisation is in principle better suited for this purpose, because the Schrödinger picture is able to give both the physical states and their energies. Unfortunately the field-theoretical Schrödinger formalism is much more ambiguous than its quantum-mechanical counterpart, mainly because regularisation and renormalisation of ill-defined expressions is required. Due to the non-Lorentz-covariant appearance of the canonical formalism the traditional renormalisation methods also become more difficult to apply than in the path-integral formalism. As a matter of fact, the canonical approach faces problems before quantisation is even considered. This is because the standard procedure for transforming Lagrangian systems into a Hamiltonian form fails for gauge theories. Neglecting the problem at first, one can construct a Hamiltonian which reproduces the dynamical Yang–Mills equations, but Gauss's law is then

absent in such a system. No physical applications are possible before the Gauss law is properly incorporated into the Hamiltonian formalism, be it classical or quantal.

This thesis introduces some new methods for implementing the Gauss law either within the Lagrangian formalism or in the Hamiltonian picture. Paper I presents an algebraic method for SU(3) Yang–Mills theory in the first-order Lagrangian formalism, and it produces a non-Abelian extension of the ordinary Hodge decomposition. Paper II provides a mathematical point of view on the covariant Poisson equation, which is a Lagrangian form of Gauss's law. Using methods of functional analysis, sufficient conditions are derived for the existence and smoothness of solutions to this equation. In Paper III the starting point is Hamiltonian, and a novel approach to the Gauss law problem is presented. The construction of physical variables is performed explicitly for pure SU(2) Yang–Mills theory, and an extension to more general theories is outlined. The review part of this thesis then offers a wider perspective on Yang–Mills theory and its Hamiltonian formulation in general. The foundations of the Lagrangian formulation of Yang–Mills theory are presented in Chapter 2. The emphasis is on QCD-type theories, which contain gauge bosons and Dirac fermions, but not scalars. The continuous symmetries of such theories are given together with the corresponding conserved charges, and attention is paid to classical symmetries which do not survive quantisation. Although these symmetries are not directly related to the Gauss law problem, they become important when attempting to quantise Yang–Mills theory. At the end of Chapter 2 the path-integral formalism is presented and critically compared to the canonical approach. The Hamiltonian formalism is the subject of Chapter 3. The Gauss law problem is formulated and the most important solution methods, both classical and quantal, are presented. Conclusions are drawn in the final chapter. There are also two appendices in which the most important mathematical concepts of the research papers are explained. Appendix A is a brief summary of Lie groups and algebras, and Appendix B presents all the mathematics that is required for understanding Paper II.

Chapter 2

Yang–Mills theory

2.1 Some history

The prototype of all Yang–Mills theories is electrodynamics. The breakthrough of its quantised form took place in the late 40's with the discovery of renormalisation procedures that made it possible to remove the infinities encountered in the perturbation expansion (Ref. [2] contains a collection of articles). In 1954 Yang and Mills introduced their model as an SU(2) generalisation of electrodynamics [3]. Their original motivation was to construct a theory of nuclear forces that would be invariant under SU(2) isospin rotations of the nucleons. The model was rather soon extended to more general Lie groups [4], but its physical significance was not recognised at once, mainly because of the problem with particle masses. The short range of weak and strong interactions required massive gauge bosons to mediate them, while the gauge invariance of Yang–Mills theory precluded the insertion of explicit mass terms. This contradiction was settled in favour of Yang–Mills theory by the idea of spontaneous symmetry breaking through the Higgs mechanism [5,6]. Later on, the success of the electroweak theory [7–9] and the proof that non-Abelian theories are renormalisable [10,11] finally confirmed the physical applicability of Yang–Mills theory.

The idea of describing strong interactions with a Yang–Mills theory was put forward in the early 70's [12]. The discovery of asymptotic freedom [13] then led to the breakthrough of this model, nowadays known as quantum chromodynamics (QCD). Unlike the electroweak theory, gluons cannot be made massive by the Higgs mechanism, because the insertion of scalars would result in too large violations of charge conjugation and flavour symmetries [14]. In addition to that, they might render the β-function positive, thus destroying asymptotic freedom. Gluons are thus massless, but the spectrum of QCD is still expected to consist of massive particles only. This is due to the confinement phenomenon, which prevents coloured objects such as quarks and gluons from appearing as free particles. It has been conjectured that the mechanism of confinement is analogous to the Meissner effect in superconductivity with the exception that the roles of electricity and magnetism are reversed [15,16]. Quarks are thus supposed to be bound together by colour-electric strings. Although no actual proof exists, this idea has been widely accepted since it is strongly supported by lattice calculations [17–19]. Confinement and chiral symmetry breaking as a result of magnetic monopole condensation have also been found to be properties of certain supersymmetric gauge models [20].

At present the applications of Yang–Mills theory extend far beyond the Standard Model. Among them there are grand unified theories [21,22], supersymmetric gauge

theories [23–25] and theories with curved or non-commutative space-time [26–28]. The modern view on Yang–Mills theory is that it most likely is not a fundamental theory of elementary particles but rather an effective low-energy limit of some so far unknown unified theory.

2.2 Yang–Mills Lagrangian

There is a large variety of Yang–Mills theories depending on the interactions under consideration and the particle species involved. The Lagrangian densities of these theories can be written as sums of separately gauge-invariant terms in the form

$$\mathcal{L} = \mathcal{L}_{YM} + \mathcal{L}_\theta + \sum_{f=1}^{N_f} \mathcal{L}_f + \sum_{s=1}^{N_s} \mathcal{L}_s + \mathcal{L}_{\text{Yuk}},$$

where the pure Yang–Mills Lagrangian $\mathcal{L}_{YM}$ and the topological term $\mathcal{L}_\theta$ describe gauge bosons. The densities $\mathcal{L}_f$ and $\mathcal{L}_s$ refer to fermions and scalars respectively. The interactions mediated by the gauge bosons are included in the matter Lagrangians, and the term $\mathcal{L}_{\text{Yuk}}$ then contains Yukawa-type couplings between matter fields.

Let us denote the gauge group by G. Pure Yang–Mills theory could be defined with rather mild assumptions on G, but the existence of matter fields requires the representations of G to consist of unitary or orthogonal matrices or possibly their direct products. Gauge bosons are described by Lie algebra -valued fields, also known as gauge potentials

$$A_\mu(x) = A^a_\mu(x)\, T_a,$$

where the T_a's are generators in the fundamental representation of the gauge group. (See Appendix A for detailed definitions). The components A^a_μ are assumed to be real and the generators T_a to be Hermitean. The gauge field tensor is defined by

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu - ig\,[A_\mu, A_\nu], \tag{2.1}$$

which in component form reads

$$F^a_{\mu\nu} = \partial_\nu A^a_\mu - \partial_\mu A^a_\nu + g\, f_{bc}{}^a A^b_\mu\, A^c_\nu.$$

Under local gauge transformations A_μ transforms as a connection in the adjoint representation:

$$A_\mu(x) \to \omega(x) A_\mu(x)\, \omega^{-1}(x) + \frac{i}{g}\Big(\partial_\mu\, \omega(x)\Big)\omega^{-1}(x), \qquad \omega(x) \in G. \tag{2.2}$$

As a result, $F_{\mu\nu}$ transforms homogeneously according to the rule

$$F_{\mu\nu}(x) \to \omega(x) F_{\mu\nu}(x)\, \omega^{-1}(x). \tag{2.3}$$

The Lagrangian density is

$$\mathcal{L}_{YM} = -\frac{1}{4}\, F^a_{\mu\nu}\, F^{\mu\nu}_a,$$

where both repeated Lie algebra and space-time indices are summed over. The Minkowski metric $g_{\mu\nu}$ is taken to have the signature $(-,+,+,+)$, and the Lie algebra metric h_{ab} is used to define an inner product

$$(X, Y) = X^a\, Y_a = h_{ab}\, X^a\, Y^b.$$

This inner product is assumed to be positive-definite and invariant under the adjoint action of the gauge group, i.e.,

$$(\omega X \omega^{-1},\, \omega\, Y \omega^{-1}) = (X, Y) \qquad \text{for all } \ \omega \in G. \tag{2.4}$$

If the gauge group is semisimple, the Killing form can be used to define the metric as $h_{ab} = -c f_{ae}{}^{d} f_{bd}{}^{e}$, where c is some positive constant. With unitary groups h_{ab} can be normalised to δ_{ab}. Using equations (2.3) and (2.4) it is now easy to see that $\mathcal{L}_{YM}$ is gauge-invariant.

The action can be complemented with a topological term known as the second Chern class

$$\mathcal{L}_\theta = \frac{g^2}{32\pi^2}\, \theta F^a_{\mu\nu} *F_a^{\mu\nu}, \tag{2.5}$$

where θ is a constant angle and $*F_{\mu\nu}$ denotes the Hodge dual of the gauge field tensor,

$$*F^{\mu\nu} = \frac{1}{2}\, \varepsilon^{\mu\nu\rho\sigma} F_{\rho\sigma}. \tag{2.6}$$

Being a total divergence,

$$\begin{aligned} \mathcal{L}_\theta &= \partial_\mu K^\mu, \\ K^\mu &= -\frac{g^2}{16\pi^2}\, \theta \varepsilon^{\mu\nu\rho\sigma} A^a_\nu \left(\partial_\sigma A_{\rho a} + \frac{1}{3}\, g\, f_{bca}\, A^b_\rho A^c_\sigma \right), \end{aligned}$$

the θ-term does not affect the equations of motion, but it is related to the topological properties of the theory. This term also introduces violations of P and CP symmetries in QCD, because it changes sign under the transformations

$$\begin{aligned} A^P_\mu(x^0, -\mathbf{x}) &= \mathbb{P}_\mu{}^\nu A_\nu(x^0, \mathbf{x}), \\ A^{CP}_\mu(x^0, -\mathbf{x}) &= -\mathbb{P}_\mu{}^\nu A^*_\nu(x^0, \mathbf{x}), \\ \mathbb{P} &= \mathrm{diag}(1, -1, -1, -1). \end{aligned}$$

Experimental results on P and CP violations imply that the magnitude of θ must be small. Yet it is not possible to neglect the θ-term altogether — even if one decided to do so at the classical level, the term would reappear at the quantum level as a phase characterising topologically nontrivial gauge transformations [29]. Terms of this form also appear in connection with the chiral anomaly [30,31].

Matter fields live in various representation spaces of the gauge group. Dirac fermions ψ_f consist of components of the form

$$\psi_f(x) = \begin{pmatrix} \psi^1_f(x) \\ \psi^2_f(x) \\ \psi^3_f(x) \\ \psi^4_f(x) \end{pmatrix}, \qquad \psi^\alpha_f(x) = \begin{pmatrix} \psi^{\alpha 1}_f(x) \\ \vdots \\ \psi^{\alpha\, d_f}_f(x) \end{pmatrix},$$

where α is a spinor index and d_f denotes the dimension of the group representation space. The generators $T_{f,a}$ of the representation are $d_f \times d_f$ -matrices satisfying the commutation

relations of the gauge algebra. If group elements of the fundamental representation are written in the form

$$\omega = \exp\left(-ig\,\varepsilon^a\,T_a\right), \tag{2.7}$$

then the corresponding fermionic representation is given by

$$\omega_f = \exp\left(-ig\,\varepsilon^a\,T_{f,a}\right).$$

Several coordinate patches $\{\varepsilon^a\}$ may be required if the exponential mapping (2.7) is not one-to-one in the entire group space. Under local gauge transformations ψ_f then transforms as

$$\psi_f^\alpha(x) \to \omega_f(x)\,\psi_f^\alpha(x). \tag{2.8}$$

A Lagrangian for Dirac spinors is given by

$$\mathcal{L}_f = -\bar{\psi}_f\gamma^\mu D_{f,\mu}(A)\psi_f - m_f\,\bar{\psi}_f\psi_f,$$

where $\bar{\psi}_f = i\,\psi_f^\dagger\gamma^0$ and

$$D_{f,\mu}(A) = \partial_\mu + ig\,A_\mu^a\,T_{f,a}.$$

The γ-matrices satisfy the Clifford algebra

$$\{\gamma^\mu,\gamma^\nu\} = 2\,g^{\mu\nu}\,\mathbb{1},$$

and one possible choice is

$$\gamma^0 = -i\begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^k = -i\begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix}, \tag{2.9}$$

where the σ^k's stand for the Pauli matrices. This Lagrangian is invariant under gauge transformations specified by equations (2.2) and (2.8). In the Standard Model and in neutrino physics fermions are Dirac or Majorana spinors, which in the electroweak theory appear either in the fundamental or in the singlet representation, depending on their chirality. In QCD both chiralities transform in the fundamental representation. Supersymmetric theories also contain Weyl spinors in the adjoint representation.

Scalars, when present, are regarded in the group representation space as d_s-dimensional vectors

$$\phi(x) = \begin{pmatrix} \phi^1(x) \\ \vdots \\ \phi^{d_s}(x) \end{pmatrix}.$$

The group action is given by

$$\omega_s = \exp\left(-ig\,\varepsilon^a\,T_{s,a}\right),$$

where the $T_{s,a}$'s are the generators of this representation. Scalars transform under local gauge transformations as

$$\phi(x) \to \omega_s(x)\,\phi(x),$$

and hence a gauge-invariant Lagrangian can be defined by

$$\mathcal{L}_s = -\big(D_s^\mu(A)\phi\big)^\dagger\big(D_{s,\mu}(A)\phi\big) - V\big(\phi^\dagger\phi\big),$$

where

$$D_{s,\mu}(A) = \partial_\mu + ig\, A^a_\mu\, T_{s,a}.$$

The potential V is usually assumed to have a double-well shape in order to generate spontaneous symmetry breaking. In most cases scalars are taken to transform either in the fundamental or in the adjoint representation. Scalars also appear in the Yukawa interactions $\mathcal{L}_{\text{Yuk}}$, which in general can consist of all possible gauge-invariant combinations of matter fields, taking into account the restrictions originating from other symmetries of the theory and usually also from the requirement of renormalisability. However, fermion masses have been included in the Lagrangian $\mathcal{L}_f$, although they are supposed to be generated through spontaneous symmetry breaking in the electroweak sector of the Standard Model.

2.3 Equations of motion

From now on we shall confine ourselves to Lagrangians describing gauge bosons and Dirac fermions only, i.e.,

$$\mathcal{L} = -\frac{1}{4}\, F^a_{\mu\nu}\, F_a^{\mu\nu} - \sum_{f=1}^{N_f} \big(\bar\psi_f \gamma^\mu D_\mu(A) \psi_f + m_f\, \bar\psi_f \psi_f\big), \tag{2.10}$$

where

$$D_\mu(A) = \partial_\mu + ig\, A^a_\mu\, T_a$$

and the fermions transform in the fundamental representation. At the classical level fermions are regarded as complex-valued spinors which become Grassmann fields only after quantisation. Of course, it would be possible to take Grassmann-valued fermions in the very beginning, but this convention would only change some signs in the classical formulas. The Lagrangian (2.10) is a generalised version of the QCD Lagrangian, which has the same form with $G = \mathrm{SU}(3)$ and $N_f = 6$. Scalars and supersymmetric particles have been discarded, because the ultimate goal of the research in this thesis is to apply the results to non-supersymmetric QCD at low energies.

The classical equations of motion of every Lagrangian system follow from the requirement that the action functional

$$S = \int L\, dt = \int \mathcal{L}\, d^4x$$

be stationary against arbitrary variations of the fields. In our case these requirements read

$$\frac{\delta S}{\delta A^a_\mu(x)} = 0, \quad \frac{\delta S}{\delta \bar\psi_f^{\alpha C}(x)} = 0, \quad \frac{\delta S}{\delta \psi_f^{\alpha C}(x)} = 0,$$

and when applied to the Lagrangian (2.10), they give rise to the equations

$$\nabla_\nu(A) F^{\mu\nu} = j^\mu \tag{2.11}$$

$$\big(\gamma^\mu D_\mu(A) + m_f\big)\psi_f = 0, \tag{2.12}$$

where $\nabla_\nu(A)$ stands for the covariant derivative in the adjoint representation

$$\nabla_\nu(A) = \partial_\nu + ig\,[A_\nu, \cdot\,]$$

and j^μ denotes the fermionic matter current

$$j^\mu = \sum_f ig\left(\bar{\psi}_f \gamma^\mu T_a \psi_f\right) T^a. \tag{2.13}$$

In addition to these equations the mere definitions of the gauge field tensor (2.1) and its dual (2.6) result in the Bianchi identity

$$\nabla_\nu(A) *F^{\mu\nu} = 0. \tag{2.14}$$

Through the identification of colour electric and magnetic fields as

$$E^k \;=\; F^{0k}, \qquad B^k \;=\; *F^{0k},$$

the equations (2.11) and (2.14) can be cast in the form of generalised Maxwell equations

$$\partial_k E^k = J^0 \tag{2.15a}$$

$$\partial_k B^k = \widetilde{J}^0 \tag{2.15b}$$

$$\varepsilon^{kl}{}_m \, \partial_l E^m = -\widetilde{J}^k - \partial_0 B^k \tag{2.15c}$$

$$\varepsilon^{kl}{}_m \, \partial_l B^m = J^k + \partial_0 E^k, \tag{2.15d}$$

where the colour electric and magnetic currents read

$$J^\mu \;=\; -ig\,[A_\nu, F^{\mu\nu}] + j^\mu \tag{2.16}$$

$$\widetilde{J}^\mu \;=\; -ig\,[A_\nu, *F^{\mu\nu}]. \tag{2.17}$$

Equation (2.15a) is the non-Abelian Gauss law and equation (2.15b) its magnetic analogue. Equations (2.15c) and (2.15d) correspond to Faraday's and Ampère's laws respectively. Compared to an Abelian theory, the biggest difference is the property that the self-interactions of the gauge bosons now yield contributions to the colour electric and magnetic currents.

The Yang–Mills equations (2.11) look more innocent than they are — in reality they are extremely difficult to solve. Mathematical studies on the existence, uniqueness and smoothness of solutions to these equations have been carried out (Refs. [32–35] and Paper II of this thesis), but the results give no clue for constructing explicit solutions. However, there is a no-go theorem by Coleman [36] stating that the only possible nonsingular solution to the Yang–Mills equations in the sourcefree case ($j^\mu = 0$) is the vacuum solution, provided that the asymptotic conditions

$$\lim_{|\mathbf{x}|\to\infty} |\mathbf{x}|^{3/2+\epsilon} F^a_{\mu\nu}(x^0, \mathbf{x}) = 0, \qquad 0 < \epsilon < \frac{1}{2}, \qquad x^0 > 0$$

hold uniformly in direction and time. Special solutions are known, and usually they describe topologically nontrivial field configurations such as monopoles [37–39], vortices [40], dyons [41] or instantons [42]. In particular, instantons are known to play a role in resolving the so-called U(1) problem of chiral symmetries [43], because they yield nonzero contributions to the second Chern class in Euclidean space. In Minkowski space such Euclidean configurations can be regarded as representing tunnelling processes between states with different topological quantum numbers. In a quantised theory the role of classical solutions is obscured by quantum fluctuations.

2.4 Symmetries and conservation laws

2.4.1 Noether's theorem

The Lagrangian (2.10) respects various physically important symmetries. These symmetries are related to conserved quantities through the celebrated theorem by Noether [44,45]. The theorem applies to continuous infinitesimal transformations of coordinates and fields

$$\begin{cases} x'^{\mu} = x^{\mu} + \delta x^{\mu} \\ A'^{a}_{\mu}(x') = A^{a}_{\mu}(x) + \delta A^{a}_{\mu}(x) \\ \psi_f'^{\alpha C}(x') = \psi_f^{\alpha C}(x) + \delta\psi_f^{\alpha C}(x) \end{cases} \tag{2.18}$$

such that the equations of motion remain invariant. This turns out to be the case if the Lagrangian (2.10) satisfies the condition

$$\mathcal{L}\left[\{A'_{\mu}\},\{\psi'_f\},\{\bar{\psi}'_f\};x'\right]\, d^4x' = \Big(\mathcal{L}\left[\{A_{\mu}\},\{\psi_f\},\{\bar{\psi}_f\};x\right] + \partial_{\nu}\,\delta\Omega^{\nu}(x)\Big)d^4x,$$

where $\delta\Omega$ is some infinitesimal four-vector. Then a current defined by

$$\begin{aligned} J^{\mu}_{\text{cons}} \;=\; & \mathcal{L}\,\delta x^{\mu} + \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}A^{a}_{\nu})}\Big(\delta A^{a}_{\nu} - (\partial_{\rho}A^{a}_{\nu})\,\delta x^{\rho}\Big) \\ & + \sum_{f=1}^{N_f}\Bigg[\frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\psi_f^{\alpha C})}\Big(\delta\psi_f^{\alpha C} - (\partial_{\rho}\,\psi_f^{\alpha C})\,\delta x^{\rho}\Big) \\ & \qquad + \frac{\partial\mathcal{L}}{\partial(\partial_{\mu}\bar{\psi}_f^{\alpha C})}\Big(\delta\bar{\psi}_f^{\alpha C} - (\partial_{\rho}\,\bar{\psi}_f^{\alpha C})\,\delta x^{\rho}\Big)\Bigg] - \delta\Omega^{\mu} \end{aligned} \tag{2.19}$$

is conserved, i.e.,

$$\partial_{\mu}J^{\mu}_{\text{cons}} = 0 \tag{2.20}$$

under the assumption that the equations of motion hold. Note that it is always possible to modify the current by four-vectors of the form

$$J^{\mu}_{(0)} = \partial_{\nu}\Gamma^{\mu\nu}, \qquad \Gamma^{\mu\nu} = -\Gamma^{\nu\mu} \tag{2.21}$$

or by vectors that vanish when the equations of motion are in force. Defining now a charge by

$$\mathcal{Q} = \int J^{0}_{\text{cons}}\, d^3\mathbf{x}$$

and assuming that the spatial components J^{k}_{cons} vanish sufficiently rapidly at infinity, it follows from equation (2.20) and the divergence theorem that $\mathcal{Q}$ is a constant of motion. In a quantised theory it is required that Green's functions involving the current J^{μ}_{cons} are divergence-free, and the quantum analogues of equation (2.20) are known as Ward identities. However, it is impossible to find a regularisation procedure that would preserve all the symmetries of the classical theory. As a result, some classical symmetries are broken by quantum anomalies.

2.4.2 Gauge invariance

As stated previously, the Lagrangian (2.10) is invariant under local gauge transformations defined by equations (2.2) and (2.8). Taking the exponential parameters ε^a in (2.7) to be infinitesimal constants, we get a global transformation of the form (2.18) with

$$\begin{cases} \delta x^\mu = 0 \\ \delta A^a_\mu(x) = ig\,[A_\mu(x), \varepsilon]^a \\ \delta \psi_f^{\alpha C}(x) = -ig\,\varepsilon^a\,(T_a)^C{}_D\,\psi_f^{\alpha D}(x). \end{cases}$$

Applying formula (2.19) we then see that the corresponding conserved current is the colour-electric current J^μ defined in (2.16). The colour-magnetic current $\widetilde{J}^\mu$ of equation (2.17) is equally conserved, but there is no symmetry of the action corresponding to it. The conservation of $\widetilde{J}^\mu$ follows trivially from the fact that the current is of the form (2.21) with $\Gamma^{\mu\nu} = *F^{\mu\nu}$. The colour electric and magnetic charges are given by

$$Q^a = \int J^{0a}\,d^3\mathbf{x}, \qquad \widetilde{Q}^a = \int \widetilde{J}^{0a}\,d^3\mathbf{x}.$$

In Yang–Mills theories gauge invariance is both a blessing and a curse. On one hand, it is vital for the renormalisability of quantised theories, and therefore one would like to retain it manifestly at the quantum level as long as possible. On the other hand, gauge angles are nonphysical variables and they must be eliminated at some stage. In fact, it is not the fields (A_μ, ψ_f) themselves that are physical objects but their gauge orbits

$$\begin{aligned} \mathcal{O}(A_\mu, \psi_f) \;=\; \Big\{ \Big(A^{(\omega)}_\mu(x), \psi^{(\omega)}_f(x) \Big) \;\Big|\; & A^{(\omega)}_\mu(x) = \omega(x) A_\mu(x)\,\omega^{-1}(x) + \frac{i}{g}\Big(\partial_\mu\,\omega(x)\Big)\omega^{-1}(x), \\ & \psi^{(\omega)}_f(x) = \omega_f(x)\,\psi_f(x), \qquad \omega(x) \in G \Big\}. \end{aligned}$$

The space of orbits is usually parametrised by fixing the gauge, that is, by selecting a unique representative from each orbit. This can be done with the help of a gauge condition, which is a non-gauge-invariant equation of the form

$$F_a\big[\{A_\mu\}, \{\psi_f\}; x\big] = 0, \qquad a = 1, \ldots, \dim G. \tag{2.22}$$

The gauge must be attainable, i.e., for every field configuration (A_μ, ψ_f) there must exist a gauge-transformed configuration $(A^{(\omega)}_\mu, \psi^{(\omega)}_f)$ that satisfies the gauge condition $F_a\big[\{A^{(\omega)}_\mu\}, \{\psi^{(\omega)}_f\}; x\big] = 0$. Uniqueness means that only one field configuration in each orbit is allowed to satisfy the condition (2.22).

Surprisingly enough, attainability and uniqueness are so restrictive requirements that it is almost impossible to find a gauge fulfilling them both. Usually it is uniqueness that fails and the gauge condition is satisfied by two or more gauge-equivalent configurations [46]. There is a theorem by Singer [47] which states that no continuous gauge fixing is possible in the set of gauge potentials defined over S^4. The sphere S^4 is a compactified version of Euclidean $\mathbb{R}^4$ where the infinity limit is regarded as a point on the sphere. $\mathbb{R}^4$ then becomes a chart for one hemisphere of S^4. The coordinates of the other hemisphere are given by the inversion

$$\begin{cases} x'^\mu = \frac{x^\mu}{r^2}, \qquad r^2 = \delta_{\rho\sigma}\,x^\rho x^\sigma \\ A'_\mu(x') = \chi(x)\,(r^2\,\delta_\mu{}^\nu - 2\,\delta_{\mu\rho}\,x^\rho x^\nu)\,A_\nu(x)\,\chi^{-1}(x) + \frac{i}{g}\big(\partial_\mu\,\chi(x)\big)\chi^{-1}(x), \end{cases} \tag{2.23}$$

where the G-valued mapping χ is called a transition function. Compactification of $\mathbb{R}^4$ now amounts to requiring that A'_μ be smooth enough at the point $x' = 0$. Similarly, if a local gauge transformation matrix $\omega(x)$ has been defined on $\mathbb{R}^4$, the corresponding matrix on the hemisphere containing infinity reads

$$\omega'(x') = \chi(x)\,\omega(x)\,\chi^{-1}(x).$$

The transition function χ is characterised by the homotopy group $\pi_3(G)$. If the group is trivial, the transition function can be deformed to unity. Otherwise there will be a twist labelled by some nontrivial element of $\pi_3(G)$. (See Refs. [48] and [49] for further details.) Singer's theorem does not forbid the existence of ambiguity-free gauges on $\mathbb{R}^4$ or in Minkowski space. For example, the orthogonal gauge

$$x^\mu A_\mu(x) = 0$$

is unique up to global gauge transformations [50]. An apparent alternative to gauge-fixing would be the formulation of Yang–Mills theory in terms of gauge-invariant variables. However, the construction of such variables requires specifying the gauge angles that are going to be removed (see Ref. [51] and Paper III of this thesis). In other words, one must still choose a gauge.

2.4.3 Poincaré invariance

Poincaré transformations encompass Lorentz boosts, rotations and translations. When operating on space-time coordinates, they take the form

$$x'^\mu = \Lambda^\mu{}_\nu\, x^\nu + b^\mu, \tag{2.24}$$

where the Lorentz transformations Λ are defined by the property that they preserve the Minkowski metric, i.e.,

$$g_{\mu\nu}\,\Lambda^\mu{}_\rho \Lambda^\nu{}_\sigma = g_{\rho\sigma}.$$

It is customary to say that the gauge potential A_μ now transforms as a one-form, but this does not have to be the case [29,52]. Namely, local gauge invariance implies that we are free to attach a gauge transformation to any other symmetry. The Poincaré transformation of A_μ can thus be defined as

$$A'_\mu(x') = \omega_{\Lambda,b}(x)\,\Lambda_\mu{}^\nu A_\nu(x)\,\omega^{-1}_{\Lambda,b}(x) + \frac{i}{g}\Big(\partial_\mu\,\omega_{\Lambda,b}(x)\Big)\omega^{-1}_{\Lambda,b}(x), \tag{2.25}$$

where the matrix $\omega_{\Lambda,b}$ is now a function of the Poincaré transformation parameters. This formula is particularly useful in proving the Lorentz covariance of gauge-fixed fields. For example, the temporal gauge $A_0 = 0$ seemingly violates Lorentz covariance, but if one applies formula (2.25) to the gauge-fixed potential with

$$\omega_\Lambda(t,\mathbf{x}) = \omega_\Lambda(0,\mathbf{x})\,\overrightarrow{T}\exp\left(ig\,\Lambda_0{}^k\int_0^t A_k(s,\mathbf{x})\,ds\right),$$

the result is a Lorentz transformation that preserves the gauge $A'_0 = 0$. (Time ordering $\overrightarrow{T}$ means that the components of the gauge potential must be in ascending order with respect to their time arguments.)

In order to define Lorentz transformations for spinors we must confine ourselves to one connected component of the Lorentz group, which is normally taken to be the subgroup of proper ortochronous transformations ($\det\Lambda = 1$, $\Lambda^0{}_0 \geq 1$). Such transformations can be written as exponentials

$$\Lambda^{\mu}{}_{\nu} = \left[\exp\left(-\frac{i}{2}\,\epsilon^{\rho\sigma} M_{\rho\sigma}\right)\right]^{\mu}{}_{\nu}, \tag{2.26}$$

where the parameters $\epsilon^{\rho\sigma}$ are antisymmetric and the generators $M_{\rho\sigma}$ are matrices defined by

$$(M_{\rho\sigma})^{\mu}{}_{\nu} = i\,(g_{\rho}{}^{\mu}\, g_{\sigma\nu} - g_{\sigma}{}^{\mu}\, g_{\rho\nu})\,.$$

The spinor representation corresponding to (2.26) is then

$$S^{\alpha}{}_{\beta}(\Lambda) = \left[\exp\left(-\frac{i}{2}\,\epsilon^{\rho\sigma}\,\Sigma_{\rho\sigma}\right)\right]^{\alpha}{}_{\beta},$$

where

$$(\Sigma_{\rho\sigma})^{\alpha}{}_{\beta} = \frac{i}{4}\,[\gamma_{\rho},\gamma_{\sigma}]^{\alpha}{}_{\beta}.$$

The Poincaré transformation formula for spinors is now written as

$$\psi'_f(x') = \omega_{\Lambda,b}(x) S(\Lambda)\psi_f(x). \tag{2.27}$$

Representations for space and time reversals depend on the particular form of the γ-matrices. Adopting the choice (2.9), space reversal can be defined by

$$S(\mathbb{P}) = \eta_P\gamma^0, \qquad |\eta_P| = 1, \qquad \mathbb{P} = \mathrm{diag}(1,-1,-1,-1).$$

Time reversal is a more peculiar transformation in which

$$S(\mathbb{T}) = \eta_T\gamma_5\gamma_2\gamma_0, \qquad |\eta_T| = 1, \qquad \mathbb{T} = \mathrm{diag}(-1,1,1,1)$$

with

$$\gamma_5 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix},$$

and formula (2.25) is replaced by

$$A'_{\mu}(x') = -\,\omega_T(x)\,\mathbb{T}_{\mu}{}^{\nu} A_{\nu}(x)\,\omega_T^{-1}(x) + \frac{i}{g}\Big(\partial_{\mu}\,\omega_T(x)\Big)\omega_T^{-1}(x).$$

(See Refs. [53,54] for further details.) It is straightforward to verify that the action is invariant under the combined transformations (2.24), (2.25) and (2.27). Poincaré invariance must also survive quantisation as it is one of the most important symmetries in physics.

Let us now find the Noether current corresponding to translations. If b^{μ} is taken to be infinitesimal, equations (2.24), (2.25) and (2.27) reduce to the form (2.18) with

$$\begin{cases} \delta x^{\mu} = b^{\mu} \\ \delta A^{a}_{\mu}(x) = \big(\nabla_{\mu}(A)\, b^{\nu} w_{\nu}(x)\big)^{a} \\ \delta\psi_f^{\alpha C}(x) = -ig\, b^{\nu} w^{a}_{\nu}(x)\,(T_a)^{C}{}_{D}\,\psi_f^{\alpha D}(x), \end{cases}$$

where w^a_ν specifies the associated gauge transformation by

$$\omega_b(x) = \mathbb{1} - ig\, b^\nu w^a_\nu(x)\, T_a + \mathcal{O}(b^2).$$

Formula (2.19) gives the conserved current

$$\begin{aligned} T^{\mu\nu}(w) &= F^{\mu\rho}_a \big(\nabla_\rho(A) w^\nu - \partial^\nu A_\rho\big)^a - \frac{1}{4}\, g^{\mu\nu} F^a_{\rho\sigma}\, F^{\rho\sigma}_a + \sum_f \bar\psi_f \gamma^\mu D^\nu(w) \psi_f \\ &\quad - g^{\mu\nu} \sum_f \bar\psi_f \big(\gamma^\rho D_\rho(A) + m_f\big)\psi_f, \end{aligned} \tag{2.28}$$

which is recognised as the energy-momentum tensor. Its components $T^{0\mu}$ can be regarded as densities of the total four-momentum, i.e.,

$$H = \int T^{00}\, d^3\mathbf{x}, \qquad P^k = \int T^{0k}\, d^3\mathbf{x}.$$

Note that it is still possible to alter the definition of $T^{\mu\nu}$ by adding a term of the form (2.21). This option will be exploited later.

Infinitesimal Lorentz transformations take the form (2.18) with

$$\begin{cases} \delta x^\mu = \epsilon^\mu{}_\nu x^\nu \\ \delta A^a_\mu(x) = \epsilon_\mu{}^\nu A^a_\nu(x) + \frac{1}{2}\, \epsilon^{\rho\sigma} \big(\nabla_\mu(A) L_{\rho\sigma}(x)\big)^a \\ \delta \psi^{\alpha C}_f(x) = -\frac{i}{2}\, \epsilon^{\rho\sigma} (\Sigma_{\rho\sigma})^\alpha{}_\beta\, \psi^{\beta C}_f(x) - \frac{i}{2}\, g\, \epsilon^{\rho\sigma} L^a_{\rho\sigma}(x)\, (T_a)^C{}_D\, \psi^{\alpha D}_f(x), \end{cases}$$

where the associated gauge transformation reads

$$\omega_\epsilon(x) = \mathbb{1} - \frac{i}{2}\, g\, \epsilon^{\rho\sigma} L^a_{\rho\sigma}(x)\, T_a + \mathcal{O}(\epsilon^2).$$

Using the energy-momentum tensor (2.28), the conserved current can be written as follows:

$$\begin{aligned} \mathcal{M}^{\mu\rho\sigma}(L) &= x^\rho\, T^{\mu\sigma}(0) - x^\sigma\, T^{\mu\rho}(0) + A^{\rho a} F^{\mu\sigma}_a - A^{\sigma a} F^{\mu\rho}_a - F^{\mu\nu}_a \big(\nabla_\nu(A) L^{\rho\sigma}\big)^a \\ &\quad - \sum_f \bar\psi_f \gamma^\mu \left(i\, \Sigma^{\rho\sigma} + ig\, L^{\rho\sigma}\right) \psi_f. \end{aligned} \tag{2.29}$$

The components $\mathcal{M}^{0kl}$ are related to the total angular momentum $\mathbb{J}^k$ through

$$\mathbb{J}^k = \int \frac{1}{2}\, \varepsilon^k{}_{lm}\, \mathcal{M}^{0lm}\, d^3\mathbf{x},$$

and the components $\mathcal{M}^{0k0}$ are densities of the so-called boost vector

$$\mathbb{K}^k = \int \mathcal{M}^{0k0}\, d^3\mathbf{x}.$$

The currents (2.28) and (2.29) can now be combined to form a symmetric energy-momentum tensor [55]. In fact, symmetry is required of a tensor that is going to be coupled to general relativity. The trick is to notice that the current $\mathcal{M}^{\mu\rho\sigma}$ splits into two parts,

$$\mathcal{M}^{\mu\rho\sigma}(L) = x^\rho\, T^{\mu\sigma}(w) - x^\sigma\, T^{\mu\rho}(w) + \Gamma^{\mu\rho\sigma},$$

$$
\begin{aligned}
\Gamma^{\mu\rho\sigma} &= A^{\rho a}F_a^{\mu\sigma} - A^{\sigma a}F_a^{\mu\rho} - \sum_f \bar\psi_f\gamma^\mu\, i\,\Sigma^{\rho\sigma}\psi_f \\
&\quad -F_a^{\mu\nu}\big(\nabla_\nu(A)L^{\rho\sigma}\big)^a - x^\rho F_a^{\mu\nu}\big(\nabla_\nu(A)w^\sigma\big)^a + x^\sigma F_a^{\mu\nu}\big(\nabla_\nu(A)w^\rho\big)^a \\
&\quad -\sum_f \bar\psi_f\gamma^\mu\, ig\,(L^{\rho\sigma} + x^\rho w^\sigma - x^\sigma w^\rho)\,\psi_f,
\end{aligned}
$$

where the first two terms refer to orbital angular momentum and the first line of $\Gamma^{\mu\rho\sigma}$ to spin. The energy-momentum tensor is now symmetrised by forming the combination

$$
T^{\mu\nu}(w) + \frac{1}{2}\,\partial_\rho\,(\Gamma^{\rho\mu\nu} + \Gamma^{\mu\nu\rho} + \Gamma^{\nu\mu\rho})\,,
$$

where the second term is seen to be of the form (2.21). Removing terms that vanish when the equations of motion are in force, we arrive at the expression

$$
\begin{aligned}
\theta^{\mu\nu} &= \theta_0^{\mu\nu} + \frac{1}{2}\,\partial_\rho\,\partial_\sigma\Big(F_a^{\mu\sigma}\,(x^\rho w^\nu - x^\nu w^\rho - L^{\nu\rho})^a\Big) \\
&\quad +\frac{1}{2}\,\partial_\rho\,\partial_\sigma\Big(F_a^{\nu\sigma}\,(x^\rho w^\mu - x^\mu w^\rho - L^{\mu\rho})^a\Big), \\
\theta_0^{\mu\nu} &= -F_a^{\mu\rho}\,(F_\rho{}^\nu)^a - \frac{1}{4}\,g^{\mu\nu}F^a_{\rho\sigma}\,F_a^{\rho\sigma} \\
&\quad +\frac{1}{4}\sum_f \bar\psi_f\left(\gamma^\mu\overrightarrow{D}^\nu(A) - \overleftarrow{D}^{\dagger\nu}(A)\gamma^\mu + \gamma^\nu\overrightarrow{D}^\mu(A) - \overleftarrow{D}^{\dagger\mu}(A)\gamma^\nu\right)\psi_f \\
&\quad -g^{\mu\nu}\sum_f \bar\psi_f\left(\frac{1}{2}\,\gamma^\rho\overrightarrow{D}_\rho(A) - \frac{1}{2}\,\overleftarrow{D}^\dagger_\rho(A)\gamma^\rho + m_f\right)\psi_f,
\end{aligned}
$$

known as the Belinfante tensor. The surface terms could also be transformed away by adding counterterms of the form (2.21). This tensor is now both conserved and symmetric, i.e.,

$$
\partial_\mu\theta^{\mu\nu} = 0, \qquad \theta^{\mu\nu} = \theta^{\nu\mu},
$$

but its trace $\theta^\mu{}_\mu$ does not necessarily vanish. With the help of $\theta^{\mu\nu}$ it is also possible to alter the definition (2.29) by introducing

$$
\mathcal{M}_B^{\mu\rho\sigma} = x^\rho\theta^{\mu\sigma} - x^\sigma\theta^{\mu\rho}.
$$

It turns out that $\mathcal{M}^{\mu\rho\sigma}$ and $\mathcal{M}_B^{\mu\rho\sigma}$ differ only by terms of the form (2.21) and by terms which vanish when the equations of motion hold.

2.4.4 Conformal symmetries

If the fermions in the Lagrangian (2.10) are massless, the classical theory is invariant under scalings and special conformal transformations. Scalings are defined by

$$
\begin{cases}
x'^\mu = e^{-s}x^\mu \\
A'_\mu(x') = \omega_s(x)\,e^s A_\mu(x)\,\omega_s^{-1}(x) + \frac{i}{g}\big(\partial_\mu\omega_s(x)\big)\omega_s^{-1}(x) \\
\psi_f'^{\alpha C}(x') = \omega_s(x)^C{}_D\,e^{(3/2)s}\,\psi_f^{\alpha D}(x)
\end{cases}
$$

with

$$
\omega_s(x) = \exp\big(-igs\,v^a(x)\,T_a\big),
$$

and the corresponding Noether current reads

$$J^\mu_s = -x_\nu\, T^{\mu\nu}(0) + F^{\mu\nu}_a \big(A_\nu + \nabla_\nu(A)v\big)^a - \sum_f \bar\psi_f \gamma^\mu \left(\frac{3}{2} - ig\, v^a\, T_a\right)\psi_f.$$

As before, it is possible to give a modified but essentially equivalent definition

$$J^\mu_{s,B} = -x_\nu\, \theta^{\mu\nu}_0 + \partial_\nu \left(F^{\mu\nu}_a v^a\right).$$

The divergence of this current is proportional to the fermion masses,

$$\partial_\mu J^\mu_s = \sum_f m_f\, \bar\psi_f \psi_f,$$

showing that scale symmetry is explicitly broken by massive fermions. In addition to this, scale symmetry is also broken anomalously at the quantum level due to the emergence of a fundamental mass scale. In fact, this symmetry must be broken — if unbroken, it would render the mass spectrum of the theory either continuous or zero, contrary to observations [56].

Special conformal transformations are parametrised by a four-vector c. The transformation equations are

$$\begin{cases} x'^\mu = \frac{x^\mu - x^2 c^\mu}{1 - 2\,c\cdot x + c^2 x^2} \\ A'_\mu(x') = \omega_c(x)\Big[(1 - 2\,c\cdot x + c^2x^2)\, g_\mu{}^\nu + 2\,(x_\mu - x^2 c_\mu)\, c^\nu \\ \qquad\qquad -2\,(c_\mu - 2\,c\cdot x\, c_\mu + c^2 x_\mu)\, x^\nu\Big] A_\nu(x)\, \omega_c^{-1}(x) + \frac{i}{g}\big(\partial_\mu \omega_c(x)\big)\omega_c^{-1}(x) \\ \psi'^{\alpha C}_f(x') = \omega_c(x)^C{}_D\, (1 - 2\,c\cdot x + c^2x^2)\, [(1 - c\cdot x)\mathbb{1} - 2i\, x_\rho c_\sigma\, \Sigma^{\rho\sigma}]^\alpha{}_\beta\, \psi^{\beta D}_f(x), \end{cases}$$

where

$$\omega_c(x) = \exp\left(-ig\, c^\mu k^a_\mu(x)\, T_a\right).$$

The current given by (2.19) can again be modified so that it takes a simplified form

$$K^{\mu\nu}_B = -(2\,x^\nu x_\rho - x^2 g^\nu{}_\rho)\, \theta^{\mu\rho}_0 - \partial_\rho \left(F^{\mu\rho}_a\, k^{\nu a}\right).$$

At the classical level this symmetry is broken explicitly by fermion masses,

$$\partial_\mu K^{\mu\nu}_B = 2\,x^\nu \sum_f m_f\, \bar\psi_f \psi_f,$$

but in a quantised theory it is additionally broken by the same anomaly that leads to the violation of scale invariance. It should also be noted that the Minkowskian counterpart of the inversion (2.23)

$$\begin{cases} x'^\mu = \frac{x^\mu}{x^2} \\ A'_\mu(x') = \chi(x)\,(x^2\, g_\mu{}^\nu - 2\,x_\mu x^\nu)\, A_\nu(x)\, \chi^{-1}(x) + \frac{i}{g}\big(\partial_\mu \chi(x)\big)\chi^{-1}(x) \\ \psi'^{\alpha C}_f(x') = \chi(x)^C{}_D\, x^2 x_\mu (\gamma^\mu)^\alpha{}_\beta\, \psi^{\beta D}_f(x) \end{cases}$$

defines a symmetry transformation of the Lagrangian (2.10) with massless fermions. As this transformation is discrete, there is no conserved quantity attached to it.

2.4.5 Fermionic symmetries

The presence of fermions increases the number of classical symmetries obeyed by the action (2.10). Some of these symmetries are exact, while some are only good approximations which would become exact if the fermion masses or their differences tended to zero. The exact symmetries consist of global U(1) phase transformations

$$\psi_f'(x) = e^{-i\,\phi_f}\psi_f(x)$$

which are responsible for the conservation of the flavour current

$$j_f^\mu = i\,\bar{\psi}_f\gamma^\mu\psi_f.$$

The corresponding charge is the flavour number

$$n_f = \int \psi_f^\dagger\psi_f\, d^3\mathbf{x}.$$

Experiments indicate that flavour numbers must also be conserved at the quantum level.

Since the Lagrangian (2.10) treats both fermion chiralities equally, it possesses an additional U(1) symmetry of the form

$$\psi_f'(x) = e^{-i\,\tilde{\phi}_f\gamma_5}\,\psi_f(x) \tag{2.30}$$

provided that the corresponding flavours are light. The resulting axial current

$$j_{f,A}^\mu = i\,\bar{\psi}_f\gamma^\mu\gamma_5\psi_f$$

is broken by explicit mass terms,

$$\partial_\mu j_{f,A}^\mu = 2i\,m_f\,\bar{\psi}_f\gamma_5\psi_f,$$

but also by a quantum anomaly [30,31]. The anomalous divergence of $j_{f,A}^\mu$ takes the form of a second Chern class, indicating that the nonconservation of axial flavour numbers is related to the topology of the gauge potential. Indeed, this connection has made it possible to explain by instantons the breaking of the classical U(1) symmetry where all the light quarks are multiplied by a flavour-independent phase of the form (2.30). As instantons give nonzero contributions to the second Chern class, they break this undesired symmetry through the anomaly of the current [43]. An anomaly also appears in the covariant divergence of the Lie algebra -valued current

$$J_{f,A}^\mu = ig\left(\bar{\psi}_f\gamma^\mu\gamma_5\,T_a\,\psi_f\right)T^a.$$

Classically it holds that

$$\nabla_\mu(A)J_{f,A}^\mu = 2ig\,m_f\left(\bar{\psi}_f\gamma_5\,T_a\,\psi_f\right)T^a,$$

but an anomalous term will alter this relation at the quantum level [56–58].

If the masses of some fermion flavours are equal, the action remains invariant under rotations of these flavours. If there are N' such flavours, the corresponding symmetry consists of global SU(N') rotations of the form

$$\psi_f'(x) = \sum_{f'=1}^{N'}\left(e^{-i\,\beta^a\mathcal{T}_a}\right)_f^{\ f'}\,\psi_{f'}(x), \qquad f = 1,\ldots N', \tag{2.31}$$

where the $\mathcal{T}_a$'s are the generators of SU(N'). In QCD a symmetry of this kind is approximately present, since the lightest quarks u and d have masses with the same order of magnitude. Sometimes even the third lightest quark s is added into the picture, and the approximate symmetry then becomes that of SU(3) rotations. The current associated with the transformation (2.31) is

$$\mathcal{V}_a^\mu = \sum_{f=1}^{N'} \sum_{f'=1}^{N'} i\, \bar{\psi}_f \gamma^\mu \left(\mathcal{T}_a\right)_f{}^{f'} \psi_{f'},$$

and its divergence reads

$$\partial_\mu \mathcal{V}_a^\mu = \sum_{f=1}^{N'} \sum_{f'=1}^{N'} i\, (m_f - m_{f'})\, \bar{\psi}_f \left(\mathcal{T}_a\right)_f{}^{f'} \psi_{f'}.$$

The corresponding charge forms a generalised flavour isospin vector

$$I_a^{\mathcal{V}} = \int \sum_{f=1}^{N'} \sum_{f'=1}^{N'} \psi_f^\dagger \left(\mathcal{T}_a\right)_f{}^{f'} \psi_{f'}\, d^3\mathbf{x},$$

and it is approximately a constant of motion. This symmetry must remain anomaly-free after quantisation.

Axial flavour rotations

$$\psi'_f(x) = \sum_{f'=1}^{N'} \left(e^{-i\gamma_5 \tilde{\beta}^a \mathcal{T}_a}\right)_f{}^{f'} \psi_{f'}(x), \qquad f = 1, \ldots N' \tag{2.32}$$

also define an approximate symmetry transformation of light fermions. The Noether current is

$$\mathcal{A}_a^\mu = \sum_{f=1}^{N'} \sum_{f'=1}^{N'} i\, \bar{\psi}_f \gamma^\mu \gamma_5 \left(\mathcal{T}_a\right)_f{}^{f'} \psi_{f'},$$

giving rise to a generalised axial isospin vector

$$I_a^{\mathcal{A}} = \int \sum_{f=1}^{N'} \sum_{f'=1}^{N'} \psi_f^\dagger\, \gamma_5 \left(\mathcal{T}_a\right)_f{}^{f'} \psi_{f'}\, d^3\mathbf{x}.$$

The divergence of this current is proportional to the fermion masses,

$$\partial_\mu \mathcal{A}_a^\mu = \sum_{f=1}^{N'} \sum_{f'=1}^{N'} i\, (m_f + m_{f'})\, \bar{\psi}_f \gamma_5 \left(\mathcal{T}_a\right)_f{}^{f'} \psi_{f'},$$

and no anomalous terms are expected to appear at the quantum level. However, in QCD this symmetry is broken spontaneously, since the observed particle spectrum does not obey the symmetry derived from (2.32).

2.5 Path-integral quantisation

Quantisation of Lagrangian systems is done using the path-integral method. Given a gauge condition of the form (2.22), the generating functional of Green's functions is given by the formula [59,60]

$$\begin{aligned} Z[J,\xi_f,\bar\xi_f] &= N\int\Big(\prod_{x,\mu,a} dA^a_\mu(x)\Big)\Big(\prod_{x,\alpha,C,f} d\psi^{\alpha C}_f(x)\, d\bar\psi^{\alpha C}_f(x)\Big)\prod_{x,a}\delta\Big(F_a[\{A_\mu\},\{\psi_f\};x]\Big) \\ &\quad\times\Big(\mathrm{Det}\,\mathcal{F}\Big)\exp\bigg(i\int\Big[\mathcal{L}+J^\mu_a A^a_\mu+\sum_f\Big(\bar\xi_{f,\alpha C}\,\psi^{\alpha C}_f+\bar\psi^{\alpha C}_f\xi_{f,\alpha C}\Big)\Big]d^4x\bigg), \end{aligned} \tag{2.33}$$

where $\mathcal{F}$ is the Faddeev–Popov matrix

$$\begin{aligned} \mathcal{F}_{ab}(x,y) &= \frac{\delta F_a[\{A^{(\omega)}_\mu\},\{\psi^{(\omega)}_f\};x]}{\delta\varepsilon^b(y)}, \qquad \omega(y)=\exp\big(-ig\,\varepsilon^b(y)\,T_b\big) \\ &= -\Big[h_b{}^c\,\partial^{(y)}_\mu - g\,f_d{}^c{}_b\,A^d_\mu(y)\Big]\frac{\delta F_a(x)}{\delta A^c_\mu(y)} \\ &\quad -ig\sum_f\Big(\psi^{\alpha D}_f(y)(T_b)^C{}_D\,\frac{\delta}{\delta\psi^{\alpha C}_f(y)}F_a(x)-\bar\psi^{\alpha D}_f(y)(T_b)^C{}_D\,\frac{\delta}{\delta\bar\psi^{\alpha C}_f(y)}F_a(x)\Big). \end{aligned}$$

The fermion fields $(\psi_f,\bar\psi_f,\xi_f,\bar\xi_f)$ are now regarded as anticommuting Grassmann variables. The derivatives of Z with respect to the sources $(J,\xi_f,\bar\xi_f)$ yield Green's functions, i.e., vacuum-vacuum amplitudes of the form

$$G^{(n)}(x_1,\cdots,x_n)=\langle 0|\overleftarrow{T}\,\Big(\widehat{O}_1(x_1)\cdots\widehat{O}_n(x_n)\Big)\,|0\rangle,$$

provided that a Wick rotation is performed on the time axis of the action. $\widehat{O}_i$ is a generic symbol for operators corresponding to the fields $(A^a_\mu,\psi^{\alpha C}_f,\bar\psi^{\alpha C}_f)$, and the notation $\overleftarrow{T}$ indicates that the operators are arranged in descending time order.

Nobody knows how to evaluate the path integral (2.33) exactly. In practice one must do with a perturbation expansion around some exactly calculable special case, which is usually taken to be the free theory with $g=0$. In this case the integrand of Z can be made Gaussian, and the Green functions then decompose into formal Taylor series in powers of g. Unfortunately these series contain infinite loop corrections, and they have to be regularised before the theory makes sense. Several regularisation methods are available nowadays, but the most popular one is dimensional regularisation, which was developed in Ref. [11]. The procedure starts by modifying the dimensionality of space-time from 4 to $d=4-\varepsilon$. As a result, the dimensions of the fields change and the coupling constant becomes dimensionful. The divergent loop integrals reside in the vertex functions generated by the effective action

$$\begin{aligned} \Gamma[A_c,\psi_{f,c},\bar\psi_{f,c}] &= \frac{1}{i}\log Z[J_c,\xi_{f,c},\bar\xi_{f,c}] \\ &\quad -\int\Big[(J_c)^\mu_a\,(A_c)^a_\mu+\sum_f\Big((\bar\xi_{f,c})_{\alpha C}\,(\psi_{f,c})^{\alpha C}+(\bar\psi_{f,c})^{\alpha C}(\xi_{f,c})_{\alpha C}\Big)\Big]d^4x, \end{aligned}$$

where the sources $(J_c, \xi_{f,c}, \bar{\xi}_{f,c})$ are expressed as functions of the external fields $(A_c, \psi_{f,c}, \bar{\psi}_{f,c})$ by solving the equations

$$\begin{aligned}
(A_c)^a_\mu(x) &= \frac{1}{i}\,\frac{\delta}{\delta\,(J_c)^\mu_a(x)}\log Z[J_c,\xi_{f,c},\bar{\xi}_{f,c}] \\
(\psi_{f,c})^{\alpha C}(x) &= \frac{1}{i}\,\frac{\delta}{\delta\,(\bar{\xi}_{f,c})_{\alpha C}(x)}\log Z[J_c,\xi_{f,c},\bar{\xi}_{f,c}] \\
(\bar{\psi}_{f,c})^{\alpha C}(x) &= -\frac{1}{i}\,\frac{\delta}{\delta\,(\xi_{f,c})_{\alpha C}(x)}\log Z[J_c,\xi_{f,c},\bar{\xi}_{f,c}].
\end{aligned}$$

Functional derivatives of Γ with respect to the fields $(A_c, \psi_{f,c}, \bar{\psi}_{f,c})$ give the vertex functions $\widetilde{\Gamma}^{(n)}(x_1,\cdots,x_n)$, which on the other hand are sums of connected one-particle-irreducible Feynman graphs. Dimensional regularisation now turns the vertex functions into finite expressions which can be split into regular and singular parts according to their behaviour in the limit $d \to 4$. Once the singular parts have been identified, they can be absorbed into redefinitions of the parameters and the fields as follows:

$$A^a_\mu = Z_A^{1/2} A^a_{\mu,r}, \qquad \psi_f = Z_\psi^{1/2}\psi_{f,r}, \qquad g = Z_g\, g_r \mu^{2-d/2}, \qquad m_f = Z_m Z_\psi^{-1} m_{f,r}.$$

Here μ is an arbitrary mass parameter, which must be introduced for the sake of dimensional consistency. Gauge-fixing parameters and ghost fields are also redefined as above. The singular coefficients Z_i are determined so that they cancel the loop divergences in every order of perturbation theory. When expressed in terms of the renormalised parameters and fields, denoted by the subscript r above, the theory remains well-defined in the limit $d \to 4$. If massless particles such as gluons are present, there will also be infrared divergences, but they are easier to cure than ultraviolet divergences.

As mentioned before, the mass μ breaks scale invariance in the renormalised theory. This breaking in turn has quantitative effects on the renormalised Green functions. The bare and renormalised momentum space vertex functions are related through

$$\begin{aligned}
&\widetilde{\Gamma}^{(n)}\big(\,p_1,\cdots,p_{n-1};g,\{m_f\},d\,\big) = Z_A^{-\frac{n_A}{2}}\, Z_\psi^{-\frac{n_\psi}{2}}\,\widetilde{\Gamma}^{(n)}_r\big(\,p_1,\cdots,p_{n-1};g_r,\{m_{f,r}\},\mu,d\,\big), \\
&\quad n_A + n_\psi = n,
\end{aligned}$$

where n_A indicates the number of A-fields and n_ψ the total number of ψ_f- and $\bar{\psi}_f$-fields in the definition of $\widetilde{\Gamma}^{(n)}$. Taking derivatives with respect to μ and passing to the limit $d \to 4$ yields the Callan–Symanzik equation [61,62]

$$\begin{aligned}
&\left[\mu\frac{\partial}{\partial\mu} + \beta(g_r)\frac{\partial}{\partial g_r} + \sum_f \gamma_{m_f}(g_r)\, m_{f,r}\,\frac{\partial}{\partial m_{f,r}} - n_A\,\gamma_A(g_r) - \,n_\psi\gamma_\psi(g_r)\right] \\
&\times\widetilde{\Gamma}^{(n)}_r\big(\,p_1,\cdots,p_{n-1};g_r,\{m_{f,r}\},\mu,d=4\,\big) = 0,
\end{aligned}$$

where

$$\begin{aligned}
\beta(g_r) &= \mu\frac{dg_r}{d\mu} \\
\gamma_{m_f}(g_r) &= \frac{\mu}{m_{f,r}}\frac{dm_{f,r}}{d\mu}
\end{aligned} \tag{2.34}$$

$$\begin{aligned}\gamma_A(g_r) &= \frac{1}{2}\,\mu\frac{d}{d\mu}\log Z_A \\ \gamma_\psi(g_r) &= \frac{1}{2}\,\mu\frac{d}{d\mu}\log Z_\psi.\end{aligned}$$

Note that the functions β and γ are independent of the fermion masses in the so-called MS (or $\overline{\mathrm{MS}}$) renormalisation scheme [63]. Equations (2.34) determine the scaling behaviour of the variables and the parameters as follows: a vertex function with scaled momenta is given by the formula

$$\widetilde{\Gamma}^{(n)}_r\big(\,e^s p_1,\cdots,e^s p_{n-1};g_r(0),\{m_{f,r}(0)\},\mu\,\big) = f(s)\,\widetilde{\Gamma}^{(n)}_r\big(\,p_1,\cdots,p_{n-1};g_r(s),\{m_{f,r}(s)\},\mu\,\big),$$

where the running parameters are solutions to the differential equations

$$\begin{aligned}\frac{dg_r(s)}{ds} &= \beta\big(g_r(s)\big) \\ \frac{dm_{f,r}(s)}{ds} &= m_{f,r}(s)\,\big[\gamma_{m_f}\big(g_r(s)\big)-1\big] \\ \frac{1}{f}\,\frac{df(s)}{ds} &= 4-n_A-\frac{3}{2}\,n_\psi-n_A\,\gamma_A\big(g_r(s)\big)-n_\psi\gamma_\psi\big(g_r(s)\big)\end{aligned}$$

with the initial condition $f(0)=1$. The functions (2.34) are thus responsible for deviations from the classical scaling behaviour. The β-function of QCD was calculated in Ref. [13], and for a Lagrangian of the form (2.10) it reads

$$\beta(g_r) = -\frac{g_r^3}{4\pi^2}\left(\frac{11}{12}C_1-\frac{1}{3}N_fC_2\right)+\mathcal{O}(g_r^5),$$

where C_1 and C_2 are constants determined by the generators of the gauge group through the relations

$$\begin{aligned}f_{ac}{}^d f_{bd}{}^c &= -C_1\,\delta_{ab} \\ \mathrm{Tr}(T_aT_b) &= C_2\,\delta_{ab}.\end{aligned}$$

To lowest order the running coupling constant then reads

$$g_r(s) = \frac{g_r(0)}{\sqrt{1+s\,\frac{g_r(0)^2}{2\pi^2}\left(\frac{11}{12}C_1-\frac{1}{3}N_fC_2\right)}}.$$

In QCD the group constants satisfy

$$\frac{11}{12}C_1-\frac{1}{3}N_fC_2 = \frac{7}{4} > 0,$$

indicating that the running coupling constant tends to zero in the high-energy limit $s\to\infty$. This is the property of asymptotic freedom, which means that quarks and gluons behave as free particles at small spatial separations. Contrary to this, $g_r(s)$ grows rapidly at low energies and blows up at a finite value of s. This singularity is probably due to a perturbative approximation of the β-function, but for physical reasons the coupling constant is indeed expected to grow large in the low-energy regime so that colour confinement may result.

Perturbative path-integral quantisation is a powerful method for calculating transition amplitudes, but it has its limitations. First, the perturbation expansion in powers of g does not converge. Even if it did so, nonanalytic expressions in g would still be beyond the scope of the expansion. This does not mean that perturbation theory would become useless in the strong-coupling regime, but rather that one should derive an expansion around a Lagrangian that is relevant in the low-energy regime. The free theory is obviously suitable only in the weak-coupling limit. Furthermore, in the strong-coupling regime one would be interested in determining the energy spectrum and the physical states of the theory, because that would make it possible to compare the predictions of QCD with known experimental properties of hadrons. The path-integral formalism is not at its best in finding the state vectors, although there exists a procedure for extracting wave functions from transition amplitudes. In fact, canonical quantisation in terms of a functional Schrödinger equation would be a more suitable way to study the properties of the states. As a quantisation method the canonical approach is also more fundamental than the path-integral method, since there are examples such as the nonlinear σ-model where a straightforward application of path integrals gives incorrect results [64]. Moreover, path integrals are not mathematically rigorous objects because their integration measures cannot be defined unambiguously. In practical calculations path integrals are first converted into ordinary integrals by discretising space-time. Once the integrations have been performed, the discretisation is removed with a suitable limiting procedure. Unfortunately different discretisations give different results, thereby making a universal definition of the integration measure impossible [65,66]. In a theory with a well-defined Hamiltonian this problem can be circumvented by deriving path-integral formulas from the canonical formalism, and in fact this was the method originally used by Feynman in his derivation of the quantum-mechanical path integral. Finally, as the path-integral formalism involves classical fields only, the structure of quantum operators is kept hidden. For all these reasons it would be most desirable to formulate quantum Yang–Mills theory also in the canonical framework.

Chapter 3

Hamiltonian formalism

3.1 Motivation

Hamiltonian formalism has proved to be a reliable pathway to quantum physics. If it is possible to formulate a classical system in terms of canonically conjugate variables, then a quantised system is obtained by replacing the classical variables with quantum operators whose commutators (or anticommutators) are derived from the classical Poisson brackets by the heuristic substitution

$$\{\,\cdot\,,\,\cdot\,\} \to \frac{1}{i\hbar}\,[\,\cdot\,,\,\cdot\,]\,.$$

Of course, this is only a rule of thumb which must be applied with care. Namely, there are an infinite number of different quantum operators satisfying the same commutation relations, and these operators in general are physically inequivalent. One classical system may thus give rise to several different quantum systems among which one then has to find the system that is realised in nature. Even when the particular form of the fundamental quantum operators has been decided, there will still be ordering ambiguities in the definitions of composite operators. Due to these problems one cannot thus regard the canonical formalism as an easier method of quantisation than the path-integral formalism. One might even say that the advantage of the canonical approach lies in the fact that it makes all quantisation ambiguities inescapable, whereas the path-integral approach hides them behind the seemingly elegant but in reality vacuous definition of the integration measure. To be fair, it must now be noted that the measure problem is also present in the Hamiltonian formalism, because the normalisation of state vectors in the Schrödinger picture is determined by a path integral. The Schrödinger formalism also requires regularisation and renormalisation, since for example multiple functional derivatives defined at the same point often lead to infinite expressions. As the Schrödinger picture is not manifestly Lorentz-covariant, its perturbative renormalisation becomes more awkward than in the path-integral formalism [67,68]. However, an intrinsic renormalisation procedure, independent of the Fock vacuum, seems to be better suited for Schrödinger-quantised operators [69]. The functional integral and canonical operator approaches to quantisation should therefore be regarded as complementary methods sharing the same physical problems and ambiguities. Which formalism to use depends on the nature of the problem in question. In QCD at low energies one is mainly interested in the properties of hadronic states, and the functional Schrödinger formalism would then seem to be the most appropriate framework for studying this issue.

3.2 The Gauss law problem

The first step towards canonical quantisation consists of converting a classical Lagrangian system into a Hamiltonian one. The standard procedure for a Lagrangian of the form

$$L = \int \mathcal{L}\left[\{A_\mu\}, \{\psi_f\}, \{\bar{\psi}_f\}; x\right] d^3\mathbf{x}$$

begins with defining canonical momenta by the formulas

$$\begin{aligned} \pi_{f,\alpha C}(\mathbf{x}) &= \frac{\delta L}{\delta\, \partial_0 \psi_f^{\alpha C}(\mathbf{x})} \\ \bar{\pi}_{f,\alpha C}(\mathbf{x}) &= \frac{\delta L}{\delta\, \partial_0 \bar{\psi}_f^{\alpha C}(\mathbf{x})} \\ \Pi_a^\mu(\mathbf{x}) &= \frac{\delta L}{\delta\, \partial_0 A_\mu^a(\mathbf{x})}. \end{aligned}$$

The resulting equations are then solved for the generalised velocities, and after that a Hamiltonian is constructed through the Legendre transformation

$$H = \int \Big[\Pi_a^\mu\, \partial_0 A_\mu^a + \sum_f \left(\pi_{f,\alpha C}\, \partial_0 \psi_f^{\alpha C} + \bar{\pi}_{f,\alpha C}\, \partial_0 \bar{\psi}_f^{\alpha C}\right)\Big] d^3\mathbf{x} \;-\; L. \tag{3.1}$$

With the Lagrangian (2.10) this procedure fails, because the resulting canonical momenta read

$$\pi_{f,\alpha C}(\mathbf{x}) = i\, \psi_{f,\alpha C}^\dagger(\mathbf{x}) \tag{3.2a}$$

$$\bar{\pi}_{f,\alpha C}(\mathbf{x}) = 0 \tag{3.2b}$$

$$\Pi_a^\mu(\mathbf{x}) = -F_a^{\mu 0}(\mathbf{x}) = -\partial^0 A_a^\mu(\mathbf{x}) + \partial^\mu A_a^0(\mathbf{x}) - g\, f^{bc}{}_a A_b^\mu(\mathbf{x})\, A_c^0(\mathbf{x}), \tag{3.2c}$$

indicating in particular that $\Pi_a^0 = 0$. Clearly these equations are impossible to solve for the velocities $(\partial_0 A_0, \partial_0 \psi_f, \partial_0 \bar{\psi}_f)$. The difficulties with fermions are relatively easy to cure since they originate from the fact that the Lagrangian is linear in the derivatives of ψ_f. Such first-order Lagrangians have been analysed in Refs. [70,71] with the result that it is indeed possible to express them in a Hamiltonian form. One can even adopt the naive approach where a Hamiltonian system is formed by combining the definitions (3.1) and (3.2a). The problem with a vanishing Π^0 is more fundamental and cannot be solved by simple manipulations. In order to examine this problem, let us employ the definitions (3.2) to construct a Hamiltonian

$$H = \int \left(\frac{1}{2}\, \Pi_k^a\, \Pi_a^k + \frac{1}{4}\, F_{kl}^a\, F_a^{kl} + \sum_f \pi_f \gamma^0 \left(\gamma^k D_k(A) + m_f\right) \psi_f - G_a\, A_0^a\right) d^3\mathbf{x}, \tag{3.3}$$

where

$$G_a := \Big(\nabla^k(A)\, \Pi_k\Big)_a + \sum_f ig\, (\pi_f\, T_a\, \psi_f) \tag{3.4}$$

and the index k runs over spatial coordinates only. If we now choose the pairs (A_k^a, Π_a^k) and $(\psi_f^{\alpha C}, \pi_{f,\alpha C})$[1] as canonically conjugate variables, the Hamiltonian equations of motion

[1]Note that $\psi_f^{\alpha C}$ and $\pi_{f,\alpha C}$ are still related by equation (3.2a). This constraint would be avoided in real-valued variables

$$\begin{cases} q_f^{\alpha C} = \frac{1}{\sqrt{2}}(\psi_f^{\alpha C} - i\pi_f^{\alpha C}) \\ p_f^{\alpha C} = \frac{1}{\sqrt{2}}(-i\psi_f^{\alpha C} + \pi_f^{\alpha C}). \end{cases}$$

read

$$\partial_0 A^a_k(\mathbf{x}) = \frac{\delta H}{\delta \Pi^k_a(\mathbf{x})}, \qquad \partial_0 \Pi^a_k(\mathbf{x}) = -\frac{\delta H}{\delta A^k_a(\mathbf{x})}$$
$$\partial_0 \psi^{\alpha C}_f(\mathbf{x}) = \frac{\delta H}{\delta \pi_{f,\alpha C}(\mathbf{x})}, \qquad \partial_0 \pi_{f,\alpha C}(\mathbf{x}) = -\frac{\delta H}{\delta \psi^{\alpha C}_f(\mathbf{x})}.$$

These equations reproduce correctly the Dirac equation and the dynamical part of the Yang–Mills equations, i.e.,

$$\begin{aligned} &\nabla_\nu(A) F^{k\nu} = j^k \\ &\big(\gamma^\mu D_\mu(A) + m_f\big)\psi_f = 0, \end{aligned} \tag{3.5}$$

but Gauss's law is lost. In the canonical variables Gauss's law would correspond to the equation

$$G_a(\mathbf{x}) = 0,$$

where the G_a's are those defined in equation (3.4). Although the actual law is missing, the canonical equations of motion still state that

$$\big(\nabla_0(A) G\big)_a = 0 \tag{3.6}$$

under the assumption that A_0 is independent of all the other variables. This result means that the generators G_a are covariantly constant with respect to time in the Hamiltonian dynamics. Setting the generators to zero will then not give rise to further constraints on the canonical variables. This property will play an important role later, when specific methods are devised for implementing the Gauss law.

The difficulties with the Hamiltonian formulation of Yang–Mills theory can be understood from the point of view of gauge invariance. There cannot be a unique Hamiltonian that would govern the time evolution of all the variables, because the dynamical equations can always be modified by arbitrary time-dependent gauge transformations. This arbitrariness manifests itself in the Hamiltonian (3.3) through the presence of A_0. Since none of the Yang–Mills equations involves the time derivative of A_0, it is possible to regard A_0 as an external field that determines the gauge degrees of freedom in the dynamics of the remaining variables. This dynamical gauge freedom can be fixed by assigning A_0 some prescribed value, but there still remains the freedom of performing time-independent gauge transformations in the Hamiltonian system. These transformations are generated by the composite fields (3.4) in the sense that their Poisson brackets with the canonical variables yield time-independent gauge transformation formulas in infinitesimal form:

$$\begin{aligned} \{G_a(\mathbf{x}), A^b_k(\mathbf{y})\} &= -h_a{}^b\, \partial^{(x)}_k \delta(\mathbf{x}-\mathbf{y}) - g\, f_{ca}{}^b\, A^c_k(\mathbf{y})\, \delta(\mathbf{x}-\mathbf{y}) \\ \{G_a(\mathbf{x}), \Pi^k_b(\mathbf{y})\} &= -g\, f^c{}_{ab}\, \Pi^k_c(\mathbf{y})\, \delta(\mathbf{x}-\mathbf{y}) \\ \{G_a(\mathbf{x}), \psi^{\alpha C}_f(\mathbf{y})\} &= -ig\, (T_a)^C{}_D\, \psi^{\alpha D}_f(\mathbf{y})\, \delta(\mathbf{x}-\mathbf{y}) \\ \{G_a(\mathbf{x}), \pi_{f,\alpha C}(\mathbf{y})\} &= ig\, (T_a)^D{}_C\, \pi_{f,\alpha D}(\mathbf{y})\, \delta(\mathbf{x}-\mathbf{y}). \end{aligned}$$

Apparently the physical degrees of freedom must be invariant under these gauge transformations. This calls for methods to construct gauge-invariant canonical variables.

3.3 Lagrangian approach

One might try to solve Gauss's law and fix the dynamical gauge freedom already at the level of the Lagrangian. This would then give rise to such a new Lagrangian that the relations between generalised velocities and canonical momenta would be invertible. The Lagrangian form of Gauss's law reads

$$\nabla_k(A)\nabla^k(A)A^0 - \nabla_k(A)\,\partial^0 A^k = j^0 \tag{3.7}$$

with the fermionic matter density j^0 given in (2.13). It was proposed in Ref. [72] that this equation could be solved in the generalised Coulomb gauge

$$\nabla_k(A)\,\partial^0 A^k = 0, \tag{3.8}$$

which would result in A^0 satisfying the covariant Poisson equation

$$\nabla_k(A)\nabla^k(A)A^0 = j^0. \tag{3.9}$$

The attainability of the gauge (3.8) and the solvability of equation (3.9) have been proved in Paper II of this thesis under fairly mild assumptions regarding the smoothness and the asymptotic behaviour of the fields involved. Sufficient asymptotic conditions for making the solutions unique are given in Ref. [73]. For practical applications it would be desirable to solve the gauge condition (3.8) for the potential A_k so that the parameters of the solution could then be adopted as unconstrained variables. While this looks like a difficult task, it is also possible to apply the method of Lagrange multipliers, where one adds to the Lagrangian (2.10) a term that reproduces the constraint (3.8),

$$\mathcal{L}' = \mathcal{L} + C^a(x)\left(\nabla_k(A)\,\partial^0 A^k(x)\right)_a. \tag{3.10}$$

It is understood here that A^0 is given by equation (3.9) as a function of the remaining variables but independent of any generalised velocities. Berezin has argued that one can now perform a Legendre transformation to canonical variables and after that determine the multiplier field C from the constraint equation [74]. Unfortunately this argument applies correctly to holonomic constraints only [75], yielding unsatisfactory results with nonholonomic constraints such as equation (3.8). Probably this is a reflection of well-known problems with the construction of a proper action principle for nonholonomic constraints [76]. In order to see what happens let us try to apply Berezin'n procedure to the modified Lagrangian (3.10). The canonical momenta corresponding to the gauge potential are then given by

$$\Pi^k_a(\mathbf{x}) = \left(\nabla^k(A)A^0 - \partial^0 A^k + \nabla^k(A)C\right)_a, \tag{3.11}$$

and for the fermionic momenta equation (3.2a) is reproduced. Combining formulas (3.8), (3.9) and (3.11) it is seen that the multiplier field satisfies the equation

$$\nabla_k(A)\nabla^k(A)C = \nabla_k(A)\Pi^k - j^0. \tag{3.12}$$

Given that this condition holds, the Hamiltonian can be written as

$$\begin{aligned} H \;=\; \int\Bigg[&\frac{1}{2}\Big(\Pi_k - \nabla_k(A)C + \nabla_k(A)A_0\Big)^a\Big(\Pi^k - \nabla^k(A)C + \nabla^k(A)A_0\Big)_a + \frac{1}{4}\,F^a_{kl}\,F^{kl}_a \\ &+\frac{1}{2}\Big(\nabla_k(A)A_0\Big)^a\Big(\nabla^k(A)A_0\Big)_a + \sum_f \pi_f\gamma^0\left(\gamma^k D_k(A) + m_f\right)\psi_f\Bigg]\,d^3\mathbf{x}. \end{aligned} \tag{3.13}$$

Although Gauss's law is now in force by construction, the original dynamical equations (3.5) cannot be obtained from the canonical equations of motion unless [72]

$$\nabla^k(A)\,\partial_0 C + 2ig\,[\partial_0 A^k, C] = 0. \tag{3.14}$$

This condition also emerges at the Lagrangian level if one varies the density (3.10). There it is resolved by considering infinitesimal gauge variations of the potential around configurations which satisfy the condition (3.8), i.e.,

$$\delta A_k(x) = \nabla_k(A)\,\varepsilon(x), \qquad \nabla_k(A)\,\partial_0 A^k(x) = 0.$$

Since the modified action must be stationary against coordinate field variations of all forms, including gauge variations at the endpoints of the time interval in question, the multiplier field must satisfy the equation

$$\nabla_k(A)\nabla^k(A)C = 0$$

and converge to zero at spatial infinity. According to Ref. [73] this implies that $C = 0$ and the consistency condition (3.14) is then met. However, equation (3.12) now reveals that the canonical variables are not free, because they are subject to the constraint

$$\nabla_k(A)\Pi^k - j^0 = 0,$$

which is recognised as the Hamiltonian form of Gauss's law. Although Gauss's law was solved explicitly at the Lagrangian level through the determination of A_0, it reappears in the canonical formalism due to the fact that the gauge condition (3.8) was left unsolved. In view of this metamorphosis of constraints it seems that the procedure above has merely been a rederivation of the Hamiltonian setup presented in Section 3.2. Indeed, if one assigns A_0 the form given by (3.9), it is then easily seen that the Hamiltonians (3.3) and (3.13) are equivalent up to boundary terms. However, the Gauss law generators no longer satisfy equation (3.6), because A_0 now depends on the canonical variables. It turns out that the G_a's are actually constants of motion, i.e.,

$$\partial_0\, G_a = 0.$$

The generalised Coulomb gauge thus bears some resemblance to the temporal gauge $A_0 = 0$, although the Hamiltonians corresponding to these gauges differ by an interaction term, when there are fermions in the theory. However, a glance at formula (3.3) shows that the form of A_0 has no physical significance, since the only term with a dependence on A_0 vanishes in the Gauss law limit $G_a \to 0$.

The difficulties with the explicit construction of an unconstrained Lagrangian apparently stem from the complicated expression (3.7) for Gauss's law in terms of the gauge potential. A simpler form would arise in an alternative formalism at the cost of increasing the number of variables. This formalism is based on a Lagrangian of the form

$$\begin{aligned}\mathcal{L} \;&=\; -\frac{1}{2}\,E^a_k E^k_a - E^k_a\big(\nabla_k(A)A_0 - \partial_0 A_k\big)^a - \frac{1}{4}\,F^a_{kl}\,F^{kl}_a \\ &\quad - \sum_f \bar\psi_f\big(\gamma^\mu D_\mu(A) + m_f\big)\psi_f,\end{aligned} \tag{3.15}$$

where the components of the colour-electric field E_k are now regarded as independent

variables. Comparing this expression to the Lagrangian (2.10), it is seen that both of them reproduce the same equations of motion. The only difference lies in the status of the equation

$$E_k = -\partial_k A_0 + \partial_0 A_k + ig[A_0, A_k],$$

which is a definition in the traditional formalism, but an equation of motion in the first-order approach. One can thus solve Gauss's law in the form

$$\nabla_k(A) E^k = j^0 \tag{3.16}$$

without having to assume any connection between the colour-electric field and the gauge potential. Once an explicit solution has been obtained, the resulting unconstrained Lagrangian can then be converted into a Hamiltonian system by applying the techniques given in Refs. [70,71] for general first-order Lagrangians. Note that the Lagrangian (3.15) can be expressed as

$$\begin{aligned}\mathcal{L} &= -\frac{1}{2} E^a_k E^k_a + E^k_a \, \partial_0 A^a_k - \frac{1}{4} F^a_{kl} \, F^{kl}_a - \sum_f \bar{\psi}_f \big(\gamma^0 \partial_0 + \gamma^k D_k(A) + m_f\big) \psi_f \\ &\quad + A^a_0 \big(\nabla_k(A) E^k - j^0\big)_a - \partial_k \left(E^k_a A^a_0\right),\end{aligned}$$

which suggests that A_0 disappears from the equations of motion in a system where Gauss's law is in force. This happens even without reference to any particular gauge.

In Paper I of this thesis a method for solving the first-order Gauss law (3.16) is presented in the case of the gauge group SU(3). In the beginning the colour-electric field is split into longitudinal and transverse parts,

$$E_k = E^L_k + E^T_k,$$

and the longitudinal component is expressed as a covariant gradient of a field that satisfies the covariant Poisson equation with the fermionic source, i.e.,

$$\begin{aligned} E^L_k &= \nabla_k(A)\Phi, \\ \nabla_k(A)\nabla^k(A)\Phi &= j^0. \end{aligned}$$

It then remains to parametrise the transverse part in such a way that the equation

$$\nabla^k(A) E^T_k = 0$$

holds as an identity. In Ref. [77] it was suggested that this equation could be solved by mimicking the well-known result of Abelian theories stating that divergence-free vector fields can be written as curls. Motivated by this, one takes an ansatz of the form

$$E^T_k = \varepsilon_k{}^{lm} \, \nabla_l(A) C_m$$

and finds out that in a non-Abelian theory the fields C_k are subject to the constraint

$$ig[B^k, C_k] = 0, \tag{3.17}$$

where B^k denotes the colour-magnetic field ($B^k = *F^{0k}$). If the gauge group is SU(2), the generic solution of this constraint is simply

$$C_k(x) = \alpha_{kl}(x) B^l(x),$$

where the coefficient fields α_{kl} are symmetric,

$$\alpha_{kl}(x) = \alpha_{lk}(x).$$

Unfortunately this simplicity seems to be a property of SU(2) only. An extension of this method to SU(3) has been performed successfully in Paper I, but it turns out that the solution of equation (3.17) becomes much more complicated than in the SU(2) case. One of the reasons for this arising complexity is the fact that in SU(3) the Cartan subalgebra is two-dimensional, and for every matrix there thus exists a two-dimensional space of matrices commuting with it. Nevertheless, the results of Paper I yield a solution to the Gauss law (3.16), and in principle this solution could now be used to develop Hamiltonian formalism according to the principles of Refs. [70,71]. In practice, however, one expects a complicated colour-electric field to produce an equally complicated Hamiltonian.

3.4 Constrained quantum theory

3.4.1 Dirac quantisation

An elegant-looking functional formulation of Yang–Mills theory is obtained if one quantises the Hamiltonian (3.3) in the temporal gauge $A_0 = 0$, neglecting Gauss's law at first [56,69]. Components of the gauge potential and their conjugate momenta then become quantum operators which obey the equal-time commutation rule

$$\left[\widehat{A}^a_k(\mathbf{x}), \widehat{\Pi}^l_b(\mathbf{y})\right] = i\, g_k{}^l\, h^a{}_b\, \delta(\mathbf{x}-\mathbf{y}).$$

The usual realisation of this relation is to define the momentum operator as a derivative

$$\widehat{\Pi}^k_a(\mathbf{x}) = -i\frac{\delta}{\delta A^a_k(\mathbf{x})} \tag{3.18}$$

operating on functionals of A_k. A representation for the fermionic anticommutation relation

$$\left\{\widehat{\psi}^{\alpha C}_f(\mathbf{x}), \widehat{\pi}^{\beta D}_{f'}(\mathbf{y})\right\} = i\,\delta_{ff'}\,\delta^{\alpha\beta}\,\delta^{CD}\,\delta(\mathbf{x}-\mathbf{y})$$

is given similarly by [78]

$$\widehat{\pi}_{f,\alpha C}(\mathbf{x}) = i\frac{\delta}{\delta\psi^{\alpha C}_f(\mathbf{x})}, \tag{3.19}$$

where ψ_f is now a complex Grassmann field and the functional derivative operates on the left. Due to the presence of fermionic variables some basic definitions in the functional Schrödinger picture need to be modified. A state vector is still a functional of the form

$$\Psi = \Psi\left[\{A^a_k\}, \{\psi^{\alpha C}_f\}\right],$$

but its dual reads [69]

$$\begin{aligned}\overline{\Psi}\big[\{A^a_k\}, \{\psi^{\alpha C}_f\}\big] \;&=\; \int\Big(\prod_{\mathbf{x},\alpha,C,f} d\chi^*_{f,\alpha C}(\mathbf{x})\Big)\exp\left(\int\sum_f \chi^*_{f,\alpha C}(\mathbf{x})\,\psi^{\alpha C}_f(\mathbf{x})\,d^3\mathbf{x}\right)\\ &\quad\times\Psi\big[\{A^a_k\}, \{\chi_{f,\alpha C}\}\big]^*.\end{aligned}$$

It is assumed here that some ordering of variables has been specified in the integration measure. An inner product between two states is defined formally as a path integral

$$\langle\Phi|\Psi\rangle = \int \Big(\prod_{\mathbf{x},k,a} dA_k^a(\mathbf{x})\Big)\Big(\prod_{\mathbf{x},\alpha,C,f} d\psi_f^{\alpha C}(\mathbf{x})\Big)\overline{\Phi}\big[\{A_k^a\},\{\psi_f^{\alpha C}\}\big]\Psi\big[\{A_k^a\},\{\psi_f^{\alpha C}\}\big]. \qquad (3.20)$$

Note that these definitions are in harmony with the classical constraint (3.2a), because it now holds that

$$\langle\Psi|\,\psi_{f,\alpha C}(\mathbf{x})\,|\Phi\rangle^* = \langle\Phi|\,\frac{\delta}{\delta\psi_f^{\alpha C}(\mathbf{x})}\,|\Psi\rangle,$$

i.e.,

$$\widehat{\psi}^\dagger_{f,\alpha C}(\mathbf{x}) = \frac{\delta}{\delta\psi_f^{\alpha C}(\mathbf{x})} = -i\,\widehat{\pi}_{f,\alpha C}(\mathbf{x}). \qquad (3.21)$$

The Hamiltonian is obtained from the classical expression (3.3) by moving the fermionic momentum operator to the right,

$$\begin{aligned}\widehat{H}_0 \;=\; \int\Big[&-\frac{1}{2}\,\frac{\delta^2}{\delta A_k^a(\mathbf{x})\,\delta A_a^k(\mathbf{x})} + \frac{1}{4}\,F_{kl}^a(\mathbf{x})F_a^{kl}(\mathbf{x}) \\ &-i\sum_f\Big(\gamma^0\left(\gamma^k D_k(A) + m_f\right)\psi_f(\mathbf{x})\Big)^{\alpha C}\,\frac{\delta}{\delta\psi_f^{\alpha C}(\mathbf{x})}\Big]\,d^3\mathbf{x}. \qquad (3.22)\end{aligned}$$

The Gauss law generators are constructed similarly, the result being

$$\widehat{G}_a(\mathbf{x}) = -i\Big(h_a{}^c\,\partial_k^{(x)} - g\,f_b{}^c{}_a A_k^b(\mathbf{x})\Big)\frac{\delta}{\delta A_k^c(\mathbf{x})} + g\sum_f (T_a)^C{}_D\,\psi_f^{\alpha D}(\mathbf{x})\,\frac{\delta}{\delta\psi_f^{\alpha C}(\mathbf{x})}. \qquad (3.23)$$

It does not seem possible to make the $\widehat{G}_a$'s vanish as operators. Instead, the implementation of Gauss's law can be achieved by requiring that physical state vectors be annihilated by these generators, i.e.,

$$\widehat{G}_a(\mathbf{x})\,\Psi_{\rm phys} = 0. \qquad (3.24)$$

On the other hand, operating on a state with the $\widehat{G}_a$'s corresponds to evaluating the variation of the state functional under an infinitesimal gauge transformation. One would now be tempted to conclude that Gauss's law requires the physical state functionals to be gauge invariant, but the situation is not that simple. This is because the nontrivial topology of local fixed-time gauge transformations may constitute an obstruction to exponentiating the infinitesimal constraint equation (3.24). The topology arises in the case of gauge transformations that tend to a direction-independent limit at spatial infinity [29,79]. Such transformations can be regarded as gauge group -valued mappings defined on the sphere S^3 instead of $\mathbb{R}^3$. They are classified by the homotopy group $\pi_3(G)$, which for most SU(N) and SO(N) groups is nontrivial. Usually the elements of $\pi_3(G)$ are labelled by integers, also known as winding numbers. Gauge transformations with a nonzero winding number cannot be deformed to identity, and as a result, they cannot be obtained by exponentiating the Gauss law generators. Transformations of this kind do not therefore necessarily leave physical state functionals invariant. However, physical observables must be invariant even under nontrivial gauge transformations, and this implies that these transformations can only alter state functionals by a phase factor. One then

concludes that in a theory where $\mathbb{R}^3$ is compactified to S^3, the interpretation of Gauss's law is the following [29]:

$$\Psi_{\rm phys}\big[\{A_k^{(\omega_n)}\},\{\psi_f^{(\omega_n)}\}\big] = e^{-in\theta}\,\Psi_{\rm phys}\big[\{A_k\},\{\psi_f\}\big], \tag{3.25}$$

where n is the winding number of the gauge transformation ω_n and θ is the so-called vacuum angle. It has been shown in Ref. [29] that one could also define a theory with exactly gauge-invariant state functionals, but the θ-angle would then enter the Hamiltonian as if it were constructed from a Lagrangian where the topological term (2.5) is present. Finally it should be stressed that the whole interpretation of Gauss's law is sensitive to quantisation ambiguities. For example, the operator ordering in formula (3.23) was determined by the requirement that $\widehat{G}_a$ be an infinitesimal generator of gauge transformations. In an alternative realisation of the fermionic anticommutation rules, presented in Ref. [80], the operators $\widehat{G}_a$ turn out not to be equal with the gauge transformation generators. Either this means that certain projection operators must be introduced [81] or the representation is anomalous.

Working in the functional Schrödinger picture consists of solving the eigenvalue equation

$$\widehat{H}_0\,\Psi_{\rm phys} = E\,\Psi_{\rm phys} \tag{3.26}$$

together with the gauge invariance condition (3.25). However, due to this condition the inner product (3.20) will diverge for physical state vectors. The problem is cured by adding a gauge-fixing term and a corresponding Faddeev–Popov determinant into the integration measure as in (2.33) to make sure that the integration is actually performed over gauge orbits. The biggest problem ahead is still the regularisation and renormalisation of the Hamiltonian (3.22) and the generators (3.23). No method is known which would remove all infinities once and for all, but at least a perturbative approach seems to be possible. Fock space solutions to the Schrödinger equation are known in the noninteracting case ($g = 0$) [69,78], and this in principle allows one to renormalise the theory order by order. So far this procedure has not been carried through for the Hamiltonian (3.22), but in an interacting theory of scalars it seems to work properly [67,68]. On the other hand, in Refs. [69,80] it is argued that the renormalisation of Schrödinger picture operators should be performed intrinsically, without reference to the Fock vacuum. The idea is to regulate finite transformations generated by the operator in question and thereby deduce the renormalised form of the infinitesimal generator. Regardless of which method one prefers, renormalisation may sometimes lead to an inescapable anomaly which at worst invalidates the whole quantum theory. For instance, doubts have been cast on the validity of the representation (3.18) for the momentum operators [56]. In non-Abelian theories involving chiral fermions the Gauss law also acquires an anomaly [82], but this does not concern QCD, since the coupling to gauge bosons is purely vectorial there.

Despite the apparent elegance of equations (3.26) and (3.25) no-one has managed to work out the physical particle spectrum of QCD from them. One disadvantage of this constrained formalism is the fact that the fields appearing in the Hamiltonian are not truly physical variables because they are redundant in number. An unconstrained formalism can be derived by making a unitary transformation to variables where the nonphysical degrees of freedom, i.e., those violating Gauss's law, are easy to isolate and remove. When written in terms of the remaining variables, the state functionals then automatically satisfy Gauss's law. Several methods have been developed for performing such a reduction [51,83–90], but unfortunately the resulting Hamiltonians are always rather complicated.

Probably due to this complexity these unconstrained quantum formalisms have not led to a breakthrough in the Schrödinger picture either.

3.4.2 BRST formalism

In certain quantum systems the implementation of constraints in the form of equation (3.24) is a too stringent requirement, because it will overreduce the physical Hilbert space. The quantum BRST formalism [91–95] is a generalisation of Dirac quantisation which does not have this shortcoming. The central idea in the BRST method is to attach an additional pair of fermionic ghost variables to every constraint of the system. The constraint generators and the ghosts are then combined suitably to form a fermionic BRST charge operator. The charge operator is nilpotent, and the physical Hilbert space can thus be identified with the cohomology of the charge operator, when it acts on ghost-free states in the extended system.

In the Hamiltonian framework of Section 3.2 one starts with the canonical pairs (A^a_k, Π^k_a), $(\psi^{\alpha C}_f, \pi_{f,\alpha C})$ and usually extends the phase space by adding the pair (A^a_0, Π^0_a). There will then also be a pair of fermionic ghost fields $(\eta^a, \overline{\mathcal{P}}_a)$ attached to Gauss's law and another pair $(\overline{\eta}_a, \mathcal{P}^a)$ corresponding to the constraint $\Pi^0_a = 0$. Fields with a negative ghost number are denoted by a bar. At the quantum level all the variables become operators which can be realised in a way similar to Section 3.4.1. If state functionals are taken to have the form

$$\Psi = \Psi\left[\{A^a_\mu\}, \{\psi^{\alpha C}_f\}, \{\eta^a\}, \{\mathcal{P}^a\}\right], \tag{3.27}$$

the momentum operators $\widehat{\Pi}^k_a(\mathbf{x})$ and $\widehat{\pi}_{f,\alpha C}(\mathbf{x})$ are represented as in (3.18) and (3.19), while the remaining operators are expressed as

$$\widehat{\Pi}^0_a(\mathbf{x}) = -i\frac{\delta}{\delta A^a_0(\mathbf{x})}, \qquad \widehat{\overline{\mathcal{P}}}_a(\mathbf{x}) = i\frac{\delta}{\delta \eta^a(\mathbf{x})}, \qquad \widehat{\overline{\eta}}_a(\mathbf{x}) = i\frac{\delta}{\delta \mathcal{P}^a(\mathbf{x})}.$$

Left differentiation is adopted for ghosts, and the ghost coordinates η^a are taken to be real and the antighost momenta $\mathcal{P}^a$ imaginary, i.e.

$$\eta^a(\mathbf{x})^* = \eta^a(\mathbf{x}), \qquad \mathcal{P}(\mathbf{x})^* = -\mathcal{P}(\mathbf{x}).$$

With these conventions the dual of a state vector is defined as

$$\begin{aligned}\overline{\Psi}\big[\{A^a_\mu\}, \{\psi^{\alpha C}_f\}, \{\eta^a\}, \{\mathcal{P}^a\}\big] \;=\; & \int\Big(\prod_{\mathbf{x},\alpha,C,f} d\chi^*_{f,\alpha C}(\mathbf{x})\Big) \exp\left(\int \sum_f \chi^*_{f,\alpha C}(\mathbf{x})\, \psi^{\alpha C}_f(\mathbf{x})\, d^3\mathbf{x}\right) \\ & \times \Psi\big[\{A^a_\mu\}, \{\chi_{f,\alpha C}\}, \{\eta^a\}, \{\mathcal{P}^a\}\big]^*,\end{aligned}$$

and the inner product between two states then reads

$$\begin{aligned}\langle\Phi|\Psi\rangle \;=\; & \int\Big(\prod_{\mathbf{x},\mu,a} dA^a_\mu(\mathbf{x})\Big)\Big(\prod_{\mathbf{x},\alpha,C,f} d\psi^{\alpha C}_f(\mathbf{x})\Big)\Big(\prod_{\mathbf{x},a} d\eta^a(\mathbf{x})\Big)\Big(\prod_{\mathbf{x},a} d\mathcal{P}^a(\mathbf{x})\Big) \\ & \times \overline{\Phi}\big[\{A^a_\mu\}, \{\psi^{\alpha C}_f\}, \{\eta^a\}, \{\mathcal{P}^a\}\big]\, \Psi\big[\{A^a_\mu\}, \{\psi^{\alpha C}_f\}, \{\eta^a\}, \{\mathcal{P}^a\}\big].\end{aligned} \tag{3.28}$$

This inner product is not positive-definite, but with it the relation (3.21) holds and the ghost momenta become anti-Hermitean operators.

There is a lot of freedom in choosing the explicit form of the BRST charge operator, but probably the simplest expression is the following:

$$\widehat{\Omega} = \int \Big(\eta^a(\mathbf{x})\, \widehat{G}_a(\mathbf{x}) + \frac{1}{2}\, g\, f_{ab}{}^c\, \eta^a(\mathbf{x}) \eta^b(\mathbf{x}) \widehat{\overline{\mathcal{P}}}_c(\mathbf{x}) - i\, \mathcal{P}^a(\mathbf{x})\, \widehat{\Pi}^0_a(\mathbf{x}) \Big)\, d^3\mathbf{x}.$$

The Gauss law generators $\widehat{G}_a$ are defined in (3.23), and it follows from their commutator algebra that the BRST charge is nilpotent, i.e.

$$\widehat{\Omega}^2 = 0.$$

The constraints can thus be implemented cohomologically by requiring that physical states be BRST-closed, that is,

$$\widehat{\Omega}\, \Psi_{\text{phys}} = 0. \tag{3.29}$$

BRST-exact states can be written in the form

$$\Psi = \widehat{\Omega}\, \Phi \qquad \text{for some } \ \Phi, \tag{3.30}$$

and they are identified with null states. This identification is justified by the observation

$$\langle\, \Psi_{\text{phys}} |\, \widehat{\Omega}\, \Phi\, \rangle = \langle\, \widehat{\Omega}^\dagger \Psi_{\text{phys}} |\, \Phi\, \rangle = \langle\, \widehat{\Omega}\, \Psi_{\text{phys}} |\, \Phi\, \rangle = 0,$$

which shows that BRST-exact states are orthogonal to physical states in the inner product (3.28). In mathematical terminology the physical Hilbert space is now contained in the BRST cohomology classes

$$\text{Ker}\ \widehat{\Omega}\, /\, \text{Im}\, \widehat{\Omega}$$

consisting of BRST-closed functionals which are equivalent if their difference is BRST-exact:

$$\Psi_{\text{phys}} \sim \Psi_{\text{phys}} + \widehat{\Omega}\, \Phi, \qquad \widehat{\Omega}\, \Psi_{\text{phys}} = 0.$$

BRST-closedness is only a necessary property of physical states, and it does not exclude the presence of ghosts. Defining the ghost number operator as

$$\widehat{\mathcal{G}} = -i \int \Big(\eta^a(\mathbf{x})\, \widehat{\overline{\mathcal{P}}}_a(\mathbf{x}) + \mathcal{P}^a(\mathbf{x})\, \widehat{\overline{\eta}}_a(\mathbf{x}) \Big)\, d^3\mathbf{x},$$

one should thus add the requirement that physical states have ghost number zero,

$$\widehat{\mathcal{G}}\, \Psi_{\text{phys}} = 0. \tag{3.31}$$

Together the conditions (3.29) and (3.31) yield exactly the same physical states as Dirac's constrained quantisation method. Indeed, expanding an arbitrary functional of the form (3.27) in a power series in the ghost fields (η^a, $\mathcal{P}^a$) and applying condition (3.31), it is easily seen that physical state functionals are independent of the ghosts,

$$\Psi_{\text{phys}} = \Psi_{\text{phys}} \left[\{A^a_\mu\}, \{\psi^{\alpha C}_f\} \right].$$

Condition (3.29) and the linear independence of the ghost states involved then reproduces the constraints

$$\widehat{G}_a(\mathbf{x})\, \Psi_{\text{phys}} = 0, \qquad \widehat{\Pi}^0_a(\mathbf{x})\, \Psi_{\text{phys}} = -i\, \frac{\delta\, \Psi_{\text{phys}}}{\delta A^a_0(\mathbf{x})} = 0,$$

which are equivalent with those of Section 3.4.1. None of the physical states can be BRST-exact, because that would require the functional Φ in (3.30) to have a negative ghost number, which is impossible. BRST physical states are thus in one-to-one correspondence with those of Section 3.4.1. As in Dirac quantisation, the inner product (3.28) is ill-defined for physical states and an analogous regularisation is needed.

It still remains to determine the BRST Hamiltonian appearing in the Schrödinger equation

$$\widehat{H}\Psi = E\Psi \tag{3.32}$$

of the extended system. Apparently one has to quantise the classical Hamiltonian (3.3) and modify the result according to the requirements of BRST invariance. In general, an operator $\widehat{O}$ is said to be BRST-closed if

$$\big[\widehat{O},\widehat{\Omega}\big]_{\pm} = 0, \qquad \big[\widehat{O},\widehat{\Omega}\big]_{\pm} = \begin{cases} \{\widehat{O},\widehat{\Omega}\} & \text{for fermionic } \widehat{O} \\ [\widehat{O},\widehat{\Omega}] & \text{for bosonic } \widehat{O}. \end{cases}$$

BRST-exact operators can be written in the form

$$\widehat{O} = \big[\widehat{K},\widehat{\Omega}\big]_{\pm} \qquad \text{for some operator } \widehat{K},$$

and it then follows from the Jacobi identity and the nilpotency of $\widehat{\Omega}$ that BRST-exact operators are automatically BRST-closed. As with states, operators corresponding to physical observables are contained in the BRST cohomology classes, where

$$\widehat{O}_{\text{phys}} \sim \widehat{O}_{\text{phys}} + \big[\widehat{K},\widehat{\Omega}\big]_{\pm}, \qquad \big[\widehat{O}_{\text{phys}},\widehat{\Omega}\big]_{\pm} = 0.$$

Returning to the Hamiltonian (3.3), it turns out that a suitable BRST-closed quantum extension reads

$$\widehat{H} = \widehat{H}_0 - \int \Big(A_0^a(\mathbf{x})\,\widehat{G}_a(\mathbf{x}) + g\,f_{ab}{}^c A_0^a(\mathbf{x})\eta^b(\mathbf{x})\widehat{\overline{\mathcal{P}}}_c(\mathbf{x}) + i\,\mathcal{P}^a(\mathbf{x})\,\widehat{\overline{\mathcal{P}}}_a(\mathbf{x}) \Big)\, d^3\mathbf{x},$$

where $\widehat{H}_0$ is the temporal gauge Hamiltonian (3.22). It is easy to see that the Hamiltonians $\widehat{H}$ and $\widehat{H}_0$ coincide on physical states, and in fact they are BRST-equivalent operators, since

$$\widehat{H} = \widehat{H}_0 + \big\{\widehat{K},\widehat{\Omega}\big\} \qquad \text{with } \widehat{K} = i\int A_0^a(\mathbf{x})\,\widehat{\overline{\mathcal{P}}}_a(\mathbf{x})\,d^3\mathbf{x}.$$

Equations (3.29), (3.31) and (3.32) thus reproduce the corresponding equations (3.24) and (3.26) of Dirac quantisation. Suitable regularisation and renormalisation procedures will be required to redefine singular operators and states, but the BRST formalism has the advantage that local gauge invariance becomes easier to preserve because it is contained in one single global generator.

The BRST method has become popular in the quantisation of Yang–Mills theory due to its mathematical elegance and its ability to merge local gauge symmetry into the BRST charge. Moreover, the excess of variables allows one to interpolate between different ghost-free formulations of the theory and to check whether they are equivalent. With the help of ghosts it is also possible to preserve locality and manifest Lorentz covariance at the quantum level in a way that is consistent with the requirements of unitarity. However, if one wishes to formulate Yang–Mills theory in terms of unconstrained gauge-invariant variables by explicitly solving Gauss's law, the BRST formalism then means stepping into the opposite direction, because the number of redundant variables is increased in it.

3.5 Unconstrained classical theory

It was mentioned earlier that quantisation is a nonunique procedure where different choices may lead to physically inequivalent quantum theories. One such choice is the order in which quantisation and the implementation of Gauss's law are performed. It is well known that not all classical canonical transformations have unitary counterparts at the quantum level [96]. For example, it has been argued that the vacuum θ-angle of classical gluodynamics can be transformed away [97], while this is certainly not true in the quantum theory. Since we do not know beforehand which quantum formalism describes nature, it is worthwhile to experiment with different possibilities and to work out their consequences. One way to achieve this goal is to start by implementing Gauss's law already at the classical level. The classical theory is then certainly correct and one can freely investigate which quantisation prescription yields results that agree with observations.

3.5.1 Dirac's formalism

The mainstream of work on constrained systems is based on Dirac's general theory [98, 95,99]. This formalism can be applied to the Hamiltonian formulation of the Lagrangian system (2.10) in a fairly straightforward manner. In Dirac's terminology the failure in trying to solve equations (3.2) for the generalised velocities is expressed by saying that the theory contains primary constraints of the form

$$\Phi^i(\mathbf{x}) = 0,$$

where

$$\Phi^1_{f,\alpha C}(\mathbf{x}) = \pi_{f,\alpha C}(\mathbf{x}) - i\,\psi^\dagger_{f,\alpha C}(\mathbf{x}) \tag{3.33a}$$

$$\Phi^2_{f,\alpha C}(\mathbf{x}) = \bar\pi_{f,\alpha C}(\mathbf{x}) \tag{3.33b}$$

$$\Phi^3_a(\mathbf{x}) = \Pi^0_a(\mathbf{x}). \tag{3.33c}$$

One then defines a Hamiltonian by combining formulas (3.1) and (3.2), implementing the constraints with Lagrange multipliers. Up to boundary terms, the result reads

$$\begin{aligned} H \;=\; \int\Bigg(&\frac{1}{2}\,\Pi^a_k\,\Pi^k_a + \frac{1}{4}\,F^a_{kl}\,F^{kl}_a + \sum_f \bar\psi_f\left(\gamma^k D_k(A) + m_f\right)\psi_f - G_a\,A^a_0 \\ &+ \lambda^a_3\,\Pi^0_a + \sum_f\Big[\lambda^{\alpha C}_{1,f}\left(\pi_f + \bar\psi_f\gamma^0\right)_{\alpha C} + \lambda^{\alpha C}_{2,f}\,\bar\pi_{f,\alpha C}\Big]\Bigg)\,d^3\mathbf{x}, \end{aligned} \tag{3.34}$$

where G_a is now written as

$$G_a = \Big(\nabla^k(A)\,\Pi_k\Big)_a - \sum_f ig\left(\bar\psi_f\gamma^0\,T_a\,\psi_f\right).$$

At this stage all the pairs (A^a_μ, Π^μ_a), $(\psi^{\alpha C}_f, \pi_{f,\alpha C})$ and $(\bar\psi^{\alpha C}_f, \bar\pi_{f,\alpha C})$ are regarded as canonical variables. In view of the canonical equations of motion governed by the Hamiltonian (3.34) one can interpret the multipliers λ_i as arbitrary values for the velocities which could not be determined from equations (3.2). The next step is to check whether the Hamiltonian equations of motion reproduce the original Euler–Lagrange equations (2.11)

– (2.12) and whether they preserve the constraints (3.33) in time. It turns out that the requirement of preserving the fermionic constraints (3.33a) – (3.33b) merely fixes the corresponding Lagrange multipliers by

$$\begin{aligned}\lambda_{1,f}^{\alpha C} &= \Big[\gamma^0\Big(\gamma^k D_k(A) + m_f\Big)\psi_f - ig\,A_0\,\psi_f\Big]^{\alpha C} \\ \lambda_{2,f}^{\alpha C} &= \Big[\bar\psi_f\Big(\overleftarrow{D}_k^\dagger(A)\gamma^k - m_f\Big)\gamma^0 + ig\,\bar\psi_f A_0\Big]^{\alpha C},\end{aligned}$$

and this results in the reproduction of Dirac's equation for the field ψ_f and its adjoint. As the constraints are now compatible with the dynamical equations, they can be implemented safely. The pairs $(\psi_f^{\alpha C}, i\,\psi_{f,\alpha C}^\dagger)$ are thus canonically conjugate variables, and the reduced Hamiltonian reads

$$H = \int\left(\frac{1}{2}\,\Pi_k^a\,\Pi_a^k + \frac{1}{4}\,F_{kl}^a\,F_a^{kl} + \sum_f i\,\psi_f^\dagger\gamma^0\left(\gamma^k D_k(A) + m_f\right)\psi_f - G_a\,A_0^a + \lambda_3^a\,\Pi_a^0\right)d^3\mathbf{x}. \tag{3.35}$$

Needless to say, the constraint (3.33c) is more troublesome to handle than its fermionic counterparts. Since

$$\partial_0\,\Pi_a^0(\mathbf{x}) = G_a(\mathbf{x}),$$

the preservation of this constraint requires Gauss's law to hold. In Dirac's terminology this is phrased by saying that there appears a secondary constraint

$$\Phi_a^4(\mathbf{x}) = G_a(\mathbf{x}) = 0. \tag{3.36}$$

No further constraints arise, because the Gauss law generators are covariant constants as in equation (3.6), and the implementation of Gauss's law is therefore consistent with the Hamiltonian dynamics. One is left with the Hamiltonian (3.35) and the constraints (3.33c) and (3.36). In Dirac's theory all constraints are classified according to their Poisson bracket relations. To be exact, it is understood that the brackets are first evaluated without assuming the constraints to hold, taking them into account only in the final result. In this framework the constraints (3.33c) and (3.36) are said to be first class, because their Poisson brackets with each other vanish on the constraint surface. The presence of such constraints is regarded as a manifestation of gauge invariance,[2] and one is allowed to impose additional gauge constraints in order to fix the arbitrary degrees of freedom. Let us denote these constraints by

$$\begin{aligned}\chi_a^3(\mathbf{x}) &= 0 \\ \chi_a^4(\mathbf{x}) &= 0.\end{aligned}$$

The functional form of the χ's must satisfy two requirements. First, the Poisson bracket matrix

$$\begin{aligned}&M_{ab}^{ij'}(\mathbf{x},\mathbf{y}) = -M_{ba}^{j'i}(\mathbf{y},\mathbf{x}) = \Big\{\Phi_a^i(\mathbf{x}), \chi_b^{j'}(\mathbf{y})\Big\}, \\ &M_{ab}^{ij}(\mathbf{x},\mathbf{y}) = \Big\{\Phi_a^i(\mathbf{x}), \Phi_b^j(\mathbf{y})\Big\}, \\ &M_{ab}^{i'j'}(\mathbf{x},\mathbf{y}) = \Big\{\chi_a^{i'}(\mathbf{x}), \chi_b^{j'}(\mathbf{y})\Big\}, \qquad i,j,i',j' = 3,4\end{aligned} \tag{3.37}$$

[2] In Dirac's theory the notion of gauge is broader than in Yang–Mills theory. See Refs. [95,99].

must be invertible, and second, the preservation of the additional constraints must only fix the multiplier λ_3 in the Hamiltonian (3.35) without giving rise to further constraints.[3] In Dirac's language, the original first-class system is converted into a system with second-class constraints. One can then use the whole system of constraints to eliminate the nonphysical degrees of freedom.

If the construction of physical variables becomes difficult, it is also possible to formulate the dynamical equations in terms of the original variables by employing the Dirac bracket. This bracket is defined by the expression

$$\begin{aligned}\left\{f,g\right\}_D &= \left\{f,g\right\} - \int\Big[\Big\{f,\Phi_a^i(\mathbf{x})\Big\}\left(M^{-1}\right)_{ij}^{ab}(\mathbf{x},\mathbf{y})\Big\{\Phi_b^j(\mathbf{y}),g\Big\} \\ &\quad +\Big\{f,\Phi_a^i(\mathbf{x})\Big\}\left(M^{-1}\right)_{ij'}^{ab}(\mathbf{x},\mathbf{y})\Big\{\chi_b^{j'}(\mathbf{y}),g\Big\} \\ &\quad +\Big\{f,\chi_a^{i'}(\mathbf{x})\Big\}\left(M^{-1}\right)_{i'j}^{ab}(\mathbf{x},\mathbf{y})\Big\{\Phi_b^j(\mathbf{y}),g\Big\} \\ &\quad +\Big\{f,\chi_a^{i'}(\mathbf{x})\Big\}\left(M^{-1}\right)_{i'j'}^{ab}(\mathbf{x},\mathbf{y})\Big\{\chi_b^{j'}(\mathbf{y}),g\Big\}\Big]\,d^3\mathbf{x}\,d^3\mathbf{y},\end{aligned}$$

where Poisson brackets appear on the right-hand side and M^{-1} denotes the inverse of the matrix (3.37). The Dirac bracket is more appropriate for constrained systems than the Poisson bracket because it is compatible with the constraints, i.e., for arbitrary phase space functionals it holds that

$$\Big\{f,\Phi_a^i(\mathbf{x})\Big\}_D = 0, \qquad \Big\{f,\chi_a^{i'}(\mathbf{x})\Big\}_D = 0.$$

Moreover, the Hamiltonian equations of motion are given by the Dirac bracket

$$\partial_0 f = \{f,H\}_D$$

when the constraints are in force. When quantising the system, the commutation (or anticommutation) relations should be derived from the Dirac bracket.

Although Dirac's formalism covers a large variety of physical systems, it is not satisfactory in every respect. For instance, the formalism may seem conceptually a little confusing, since one must make a distinction between weak and strong equations depending on whether one inserts a constraint inside Poisson brackets or not. Furthermore, as suggested in Refs. [70,71], with first-order systems Dirac's formalism can be replaced by other methods which go more straightforwardly to the point. As regards the implementation of constraints, it is also regrettable that Dirac's method fails to give an explicit canonical structure to the physical sector of the phase space. It has been proved that the Dirac bracket is equivalent with a Poisson bracket evaluated in the physical subspace [100], but still no general prescription for constructing canonical variables in the physical sector emerges from the formalism. Dirac's theory must therefore be complemented with some constructive method before an unconstrained canonical formalism can be obtained. Such Dirac-motivated reductions have been performed successfully for Yang–Mills theory [101,102], but as always, the results are complicated.

[3]Note, however, that the equations of motion can be modified by attaching a Lagrange multiplier to the Gauss law constraint (3.36).

3.5.2 Abelianisation

An unconstrained canonical formulation of Yang–Mills theory can be derived from the Hamiltonian (3.3) by dividing the phase space into physical and nonphysical sectors. The nonphysical subspace consists not only of the Gauss law generators G_a but also of their canonical conjugates. This can be understood intuitively by noting that the dynamical equations of these conjugate variables depend on A_0, which is an arbitrary gauge-dependent field in the Hamiltonian. Roughly speaking, one can then conclude that the conjugates of the G_a's determine the gauge degrees of freedom in the system. In an Abelian theory the nonphysical variables can be isolated by performing a canonical transformation which takes the G_a's to some new canonical momenta and the conjugates of the G_a's to coordinates corresponding to these momenta. The implementation of Gauss's law is then achieved simply by setting the momenta to zero, ignoring their conjugate coordinates. However, in non-Abelian theories the canonical structure of the nonphysical sector is more intricate due to the nontrivial Poisson bracket algebra of the generators,

$$\{G_a(\mathbf{x}), G_b(\mathbf{y})\} = -g\, f_{ab}{}^c\, G_c(\mathbf{y})\delta(\mathbf{x}-\mathbf{y}). \tag{3.38}$$

Since all canonical momenta form a set with vanishing Poisson brackets, it is recognised that not all components G_a can become momentum-like variables.

Abelianisation is a method of turning the Gauss law generators into an equivalent set of variables with vanishing Poisson brackets. The idea is rather simple: the generators are related to the new momenta through multiplication by an invertible matrix,

$$P_a^1(\mathbf{x}) = \int M_a{}^b(\mathbf{x},\mathbf{y})\, G_b(\mathbf{y})\, d^3\mathbf{y}. \tag{3.39}$$

The matrix M must be chosen in such a way that the momenta are in involution with each other, i.e.,

$$\left\{P_a^1(\mathbf{x}), P_b^1(\mathbf{y})\right\} = 0.$$

Even though the most general form of M is difficult to determine from these requirements, a relatively simple solution was found in Ref. [103]. When written in terms of the inverse of M, this solution reads

$$(M^{-1})^{ab}(\mathbf{x},\mathbf{y}) = -\left\{G^a(\mathbf{x}), \chi_1^b(\mathbf{y})\right\}, \tag{3.40}$$

where χ_1 is a function of the canonical coordinates A_k and ψ_f only. Moreover, χ_1 can be chosen as the canonical conjugate of the momentum P_a^1, i.e.,

$$Q_1^a(\mathbf{x}) = \chi_1^a(\mathbf{x}). \tag{3.41}$$

In the physical picture χ_1 is interpreted as a gauge-defining variable, and M^{-1} then becomes the Faddeev–Popov matrix corresponding to this gauge choice. Whether this matrix is invertible everywhere or whether Gribov ambiguities occur depends on the uniqueness of the gauge.

Equations (3.39) – (3.41) fall in the category of extended point transformations, which means that the new momenta depend linearly and homogeneously on the old momenta, and the new coordinates are independent of the old momenta. Presumably one can complete the change of variables thus initiated by adhering to similar transformations.

The generating functional takes the form

$$F_2\big[\{A_k^a\},\{\psi_f^{\alpha C}\},\{P_a^k\},\{p_{f,\alpha C}\}\big] = \int\Big(\chi_l^a\big[\{A_k\},\{\psi_f\}\big]P_a^l + \sum_f \xi_f^{\alpha C}\big[\{A_k\},\{\psi_f\}\big]p_{f,\alpha C}\Big)\,d^3\mathbf{x},$$

where P_a^k and $p_{f,\alpha C}$ stand for momenta in the new set. The transformation equations then read

$$\begin{aligned}
Q_k^a(\mathbf{x}) &= \frac{\delta F_2}{\delta P_a^k(\mathbf{x})} = \chi_k^a(\mathbf{x}) \\
q_f^{\alpha C}(\mathbf{x}) &= \frac{\delta F_2}{\delta p_{f,\alpha C}(\mathbf{x})} = \xi_f^{\alpha C}(\mathbf{x}) \\
\Pi_a^k(\mathbf{x}) &= \frac{\delta F_2}{\delta A_k^a(\mathbf{x})} = \int\left(\frac{\delta\chi_l^b(\mathbf{y})}{\delta A_k^a(\mathbf{x})}\,P_b^l(\mathbf{y}) + \sum_f \frac{\delta\xi_f^{\alpha C}(\mathbf{y})}{\delta A_k^a(\mathbf{x})}\,p_{f,\alpha C}(\mathbf{y})\right)d^3\mathbf{y} \\
\pi_{f,\alpha C}(\mathbf{x}) &= \frac{\delta F_2}{\delta\,\psi_f^{\alpha C}(\mathbf{x})} = \int\left(\frac{\delta\chi_l^b(\mathbf{y})}{\delta\,\psi_f^{\alpha C}(\mathbf{x})}\,P_b^l(\mathbf{y}) + \sum_{f'} \frac{\delta\xi_{f'}^{\beta D}(\mathbf{y})}{\delta\,\psi_f^{\alpha C}(\mathbf{x})}\,p_{f',\beta D}(\mathbf{y})\right)d^3\mathbf{y}.
\end{aligned} \tag{3.42}$$

Upon inserting these expressions into formula (3.4) one gets

$$G_a(\mathbf{x}) = -\int\Big(\big\{G_a(\mathbf{x}),\chi_k^b(\mathbf{y})\big\}\,P_b^k(\mathbf{y}) + \sum_f\big\{G_a(\mathbf{x}),\xi_f^{\alpha C}(\mathbf{y})\big\}\,p_{f,\alpha C}(\mathbf{y})\Big)\,d^3\mathbf{y},$$

but on the other hand, equations (3.39) – (3.40) indicate that

$$G_a(\mathbf{x}) = -\int\big\{G_a(\mathbf{x}),\chi_1^b(\mathbf{y})\big\}\,P_b^1(\mathbf{y})\,d^3\mathbf{y}.$$

Comparison of these two formulas leads to the conclusion that

$$\big\{G_a(\mathbf{x}),\chi_k^b(\mathbf{y})\big\} = 0, \qquad \big\{G_a(\mathbf{x}),\xi_f^{\alpha C}(\mathbf{y})\big\} = 0, \qquad k = 2,3,$$

i.e., the functionals χ_2, χ_3 and ξ_f must be invariant under topologically trivial gauge transformations. Following Ref. [51] and Paper III of this thesis, such gauge-invariant functionals can be constructed from gauge-transformed coordinates $A_k^{(\omega)}$ and $\psi_f^{(\omega)}$, where ω is determined by the property

$$\chi_1^a\big[\{A_k^{(\omega)}\},\{\psi_f^{(\omega)}\}\big] = 0.$$

In other words, ω takes the original coordinates into the gauge $\chi_1 = 0$. The explicit form of the functionals χ_k and ξ_f depends on the particular gauge choice.

Abelianisation is an elegant method for performing the desired reduction of the Hamiltonian (3.3), and it has been used in Ref. [104] to define unconstrained variables for pure Yang–Mills theory with the gauge group SU(2). The fact that Abelianisation can be attained by extended point transformations may also be advantageous, because it is widely presumed that these transformations have unitary counterparts at the quantum level. On the other hand, extended point transformations usually lead to nonlocal Hamiltonians, and they also make the invertibility of the system (3.42) less straightforward to verify.

3.5.3 Function group method

A novel method for implementing Gauss's law in classical Hamiltonian Yang–Mills theory is introduced in Paper III of this thesis. It tackles the problem of transforming the algebra (3.38) into a canonical form, but from a different perspective than Abelianisation. No attempt is made to turn the algebra involutive, but rather it is examined whether the Gauss law generators can be written in terms of new canonical variables in such a way that the algebra (3.38) is reproduced. The motivation of this approach goes back to an ancient paper by Lie [105], where the notion of a function group was presented. Lie defined these groups as sets of phase space functions whose Poisson brackets close on the set.[4] He also proved that the elements of every function group can be parametrised with an equal number of canonical variables. When these ideas are applied to the algebra (3.38), it is immediately recognised that the Gauss law generators form a function group. According to Lie's theorem they can be parametrised with a set of new canonical variables such that the dimension of this set equals the dimension of the gauge algebra.

Although Lie's theorem guarantees the existence of a canonical parametrisation, it does not tell how to construct one. In Paper III it is argued that the procedure should begin by defining the maximum number of independent momentum variables in terms of the Gauss law generators. The structure of these momenta can be deduced from the properties of the gauge algebra, if one identifies the G_a's with the generators of the algebra and the Poisson bracket with the commutator. The problem then reduces to finding the maximum number of mutually involutive, functionally independent elements from the enveloping algebra of the fundamental representation. Clearly the Casimir operators and the Cartan subalgebra generators belong to this set, but also the Casimir operators of some subalgebras must be taken into account. Once the canonical momenta have been identified, the remaining coordinates are then defined so that the algebra (3.38) emerges as a result of the fundamental Poisson brackets of the new variables. In the case of the group SU(2), which is covered in Paper III, such a parametrisation can be expressed as follows:

$$\begin{aligned} G_1 &= \sqrt{p_1^2 - p_2^2}\,\cos(g\,q_2) \\ G_2 &= -\sqrt{p_1^2 - p_2^2}\,\sin(g\,q_2) \\ G_3 &= p_2. \end{aligned} \tag{3.43}$$

Of course, this parametrisation is not unique, but its inverse

$$\begin{aligned} p_1 &= \sqrt{G_1^2 + G_2^2 + G_3^2} \\ p_2 &= G_3 \\ q_2 &= -\frac{2}{g}\arctan\left(\frac{\sqrt{G_1^2 + G_2^2} - G_1}{G_2}\right) \end{aligned} \tag{3.44}$$

displays nicely the algebraic features discussed above. In particular, p_1^2 is recognised as the Casimir operator of SU(2) and p_2 as a Cartan subalgebra generator. The canonical conjugate of p_1 is not determined at this stage.

[4] Actually function groups come close to the present notion of Lie algebras.

Completing the transformation initiated by equations (3.44) is a lengthy procedure. If ξ_i denotes a generic canonical variable in the new set, different from the two pairs already introduced, the fundamental Poisson brackets between ξ_i and the variables (q_1, p_1, q_2, p_2) then necessarily vanish. In view of the parametrisation (3.43) this means that ξ_i satisfies the conditions

$$\{G_a(\mathbf{x}), \xi_i(\mathbf{y})\} = 0, \qquad \frac{\delta \xi_i(\mathbf{y})}{\delta p_1(\mathbf{x})} = 0. \tag{3.45}$$

As before, the first condition is interpreted as a requirement of invariance under topologically trivial gauge transformations. Since there are already three non-gauge-invariant variables (q_1, q_2, p_2) in the new set, they determine the gauge angles that have to be removed in the construction of gauge-invariant variables. As explained in Paper III, the elimination of q_2 and p_2 is achieved by transforming the original variables into a gauge where only the T_1-component of the Gauss law generators is nonvanishing. This gauge happens to fall in the class of Abelian gauges in which some homogeneously transforming object is rotated into the direction of a chosen Cartan subalgebra in colour space [16]. After the rotation there still remains the freedom of performing Abelian gauge transformations in the subalgebra. In the SU(2) case under study the residual Abelian gauge invariance is that of U(1) rotations in the direction of T_1. These rotations are generated by the momentum p_1 and parametrised by its conjugate q_1. In Paper III this U(1) gauge freedom is eliminated symbolically at first, without specifying q_1 explicitly. The result is a set of intermediate variables which depend on p_1 and whose Poisson brackets are complicated.

The elimination of p_1 requires fixing the U(1) gauge explicitly. When combined with the Poisson brackets of the intermediate variables, the gauge choice yields a system of functional differential equations, which determine the p_1-dependence of the variables. The integration constants of this system then satisfy both of the requirements (3.45), and accordingly, they can be regarded as defining a transformation to a new set of variables. The Poisson bracket relations of these new variables in turn reveal that the set actually forms a function group. However, there are redundant variables in the set, since the transformed fields have the property that the corresponding Gauss law generators lie in the T_1-direction. Unlike the Dirac formalism, this property is not a constraint but a functional identity which can be taken inside Poisson brackets, demonstrating that there are no weak equations in the function group formalism. Although this newest set of variables can rather easily be parametrised with the correct number of canonical variables, there is still one peculiarity in the process. Namely, the gauge-invariant variables exhibit an internal U(1) symmetry, which must also be eliminated by proper gauge-fixing. This done, the resulting variables finally determine the desired extension of the transformation (3.44).

The final transformation connects two sets of variables

$$(A^a_k, \Pi^k_a) \quad \Longleftrightarrow \quad \begin{cases} (q_i, p_i), & i = 1, 2, 3 \\ (Q^1_1, P^1_1) & \\ (Q^a_2, P^2_a), & a = 1, 3 \\ (Q^a_3, P^3_a), & a = 1, 2, 3, \end{cases}$$

where the pairs (q_i, p_i) of the new set correspond to gauge angles and gauge transformation generators. The pairs $(Q^{a'}_k, P^k_{a'})$, are obtained from the original variables (A^a_k, Π^k_a) by a gauge transformation which annihilates certain momentum components. The corresponding coordinates are solved for from a functional identity similar to Gauss's law.

The transformation equations read

$$\begin{aligned} A^a_k &= \left(\Omega^T\right)^a{}_b \left(Q^b_k + \frac{1}{g}\,\delta_{k1}\,(O_4)^b{}_2 \frac{p_1}{P^1_1}\sqrt{1-\left(\frac{p_3}{p_1}\right)^2} - \frac{1}{g}\,\delta_{k2}\,\delta^b{}_2 \frac{p_3}{P^2_3}\right) \\ &\quad -\frac{1}{2g}\,\varepsilon_{bc}{}^a \left(\Omega^T\,\partial_k\Omega\right)^{cb} \\ \Pi^k_a &= \left(\Omega^T\right)_a{}^b\, P^k_b, \end{aligned} \tag{3.46}$$

where Ω and O_4 are orthogonal matrices parametrised by the pairs (q_i, p_i) and

$$\begin{aligned} Q^2_1 &= -\frac{1}{gP^1_1}\sum_{k=2}^{3}\left(\partial_k P^k_3 - g\,\varepsilon_b{}^c{}_3 Q^b_k\,P^k_c\right) + \frac{1}{g}\,p_3\,\frac{P^2_1}{P^1_1 P^2_3} \\ Q^3_1 &= \frac{1}{gP^1_1}\Big(\partial_3 P^3_2 - \sum_{k=2}^{3} g\,\varepsilon_b{}^c{}_2 Q^b_k\,P^k_c\Big) \\ Q^2_2 &= \frac{1}{gP^2_3}\left(\partial_k P^k_1 - g\,\varepsilon_b{}^c{}_1\,Q^b_3\,P^3_c\right) \\ P^1_2 &= P^1_3 \;=\; P^2_2 \;=\; 0. \end{aligned} \tag{3.47}$$

Note that in a compactified theory the gauge transformation matrix Ω may be topologically nontrivial. On the other hand, the results of Ref. [97] suggest that topology probably plays no role at the classical level.

In the new variables Gauss's law corresponds to setting

$$p_1 = p_2 = p_3 = 0. \tag{3.48}$$

It can be seen from the transformation equations that the coordinates conjugate to these momenta become ambiguous in the Gauss law limit. Hence they must be discarded as nonphysical variables. The dynamics of the physical variables $(Q^{a'}_k, P^k_{a'})$ can be derived by inserting the expressions (3.46) into the Hamiltonian (3.3) (without fermions) and then imposing the Gauss law (3.48). The resulting physical Hamiltonian reads

$$H_{\rm phys} = \int\left(\frac{1}{2}\,P^k_a P^a_k + \frac{1}{4}\,\Phi^a_{kl}\,\Phi^{kl}_a\right) d^3\mathbf{x},$$

where

$$\Phi^a_{kl} = \partial_l Q^a_k - \partial_k Q^a_l + g\,\varepsilon_{bc}{}^a Q^b_k\,Q^c_l$$

and the definitions (3.47) are applied in connection with the requirement (3.48). It may be a little surprising that the Hamiltonian is local, but this is actually a consequence of the fact that the gauge condition of Paper III has been formulated in terms of canonical momenta. The dependence of the Hamiltonian on the coupling constant is also somewhat surprising, since $H_{\rm phys}$ decomposes into a finite series

$$H_{\rm phys} = \frac{1}{2g^2}H^{(0)} + \frac{1}{g}\,H^{(1)} + H^{(2)} + g\,H^{(3)} + \frac{g^2}{2}H^{(4)}, \tag{3.49}$$

which is singular in the Abelian limit $g \to 0$. This singularity stems from the fact that in the function group method Abelian Gauss law generators are always parametrised

with momentum-like variables only, whereas conjugate coordinates must necessarily be introduced in the non-Abelian case. The canonical momenta of an Abelian theory are also gauge-invariant and therefore impossible to use as ingredients of a gauge.

The function group method is probably not the easiest way of passing to an unconstrained Hamiltonian formulation of Yang–Mills theory, but it brings to light the algebraic features of the construction process. Moreover, it has a connection to the Abelian projection method, which has been conjectured to give some insight into the colour confinement problem [16]. Although the function group method suffers from the usual problem of unconstrained formalisms, namely, the complexity of the Hamiltonian, it must be recognised that the weak-coupling term $H^{(0)}$ in the expansion (3.49) can be quantised easily because it depends only on the canonical momenta. The entire Hamiltonian is more difficult to quantise due to ordering problems and fourth-order terms involving coordinate fields. The question of quantisation is important for the physical applicability of the Hamiltonian, but one should also derive extensions of Paper III to more general Lie groups and to fermionic degrees of freedom. The generalisation principles are easy to state, but more difficult to carry out. For instance, in the case of SU(3) the canonical momenta appearing in the parametrisation of the Gauss law generators should be constructed from the following elements of the enveloping algebra:

$$\begin{cases} f_1 = G_a\, G^a \\ f_2 = d^{abc}\, G_a G_b G_c \\ f_3 = G_1^2 + G_2^2 + G_3^2 \\ f_4 = G_3 \\ f_5 = G_8. \end{cases}$$

There are two Casimir operators, one sub-Casimir and two Cartan subalgebra generators. The remaining three coordinate-like parameters should then be defined so that the algebra (3.38) will result from the fundamental Poisson brackets. Equipped with thus parametrised Gauss law generators, one should then follow the principles of Paper III. Fermions are probably easier to add, because the gauge algebra (3.38) is the same regardless of their presence. Fermions thus come into play at the second stage of the method, where gauge-invariant variables are constructed. Apparently it suffices to express in the fundamental representation all the gauge transformations appearing in the procedure and to make sure that the final set of variables possesses a canonical structure.

Chapter 4

Conclusions

We have seen that Gauss's law can be implemented in the Hamiltonian formalism in various ways. In spite of the diversity that these methods exhibit, they all share the common feature that a number of ignorable variables must be specified by choosing a gauge. Depending on the nature of the gauge, some approaches are more suitable than others. For instance, Abelianisation combined with extended point transformations is suitable for gauges which are expressed in terms of coordinate-like variables in an extended Hamiltonian system, whereas the function group method applies well to gauges formulated with momentum-like variables. Dirac's formalism and the Lagrangian approach seem to cover a wider class of gauges, but the price to pay is the lack of an explicit canonical structure. Another feature which is common to all unconstrained formalisms is the complexity of the physical Hamiltonian. This is an inevitable consequence of the fact that the Gauss law does not seem to have very simple solutions. The physical Hamiltonians also contain singularities at points where the selected gauge choice encounters a Gribov ambiguity. The analysis of these singularities plays an important role, particularly when gauges related to the Abelian projection method are employed.

With all the methods presented in Chapter 3, the construction of physical variables becomes a solvable problem, although not an easy one. However, quantisation is still a considerably more difficult task, because no unambiguous prescription exists for performing it. Among all possible representations for the canonical commutation or anticommutation relations one is supposed to find a representation where the operator algebra remains anomaly-free after renormalisation. The most important classical symmetries must also be preserved, while certain symmetries such as scale invariance and axial U(1) symmetry must acquire an anomaly. It is also possible that one representation of the canonical operator algebra covers only one sector of the whole quantum theory, and several representations are thus required depending on the number of distinct sectors in the theory. For instance, the constrained formalism of Section 3.4.1 seems to describe the weak-coupling regime well, but it is questionable whether it applies to the strong-coupling regime. Usually quantisation problems are studied by the method of trial and error, that is, by choosing some quantisation prescription and then finding out whether it leads to a consistent theory which also agrees with observations. At present it seems that there is still a lot of work to do in this field before a universally defined quantum Yang–Mills theory is at hand.

Appendix A

Lie groups and algebras

Lie groups are topological groups which have the structure of an analytic manifold in such a way that all group operations are analytic. A matrix Lie group is a group whose elements are invertible finite square matrices. Depending on the definition of the group in question, there will be additional requirements that the group elements have to satisfy. As Yang–Mills theory is solely concerned with matrix groups, the discussion will be limited to them. More general treatments can be found in Refs. [106–108].

A d_r-dimensional matrix representation of a Lie group G is a mapping

$$D_r : G \to \mathrm{GL}(d_r, \mathbb{C})$$

that takes elements of G to invertible $d_r \times d_r$ -matrices in such a way that the group structure of G is preserved, i.e.,

$$D_r(\omega_1)D_r(\omega_2) = D_r(\omega_1\omega_2), \qquad \omega_1, \omega_2 \in G.$$

The d_r-dimensional vector space where the matrices $D_r(\omega)$ operate is called the representation space. Whether this space is real or complex depends on the representation in question. If the representation matrices are the same as those appearing in the definition of G, the representation is said to be fundamental. In the adjoint representation the representation space is the Lie algebra of G, and the mapping D_{adj} is determined by the property

$$X \xmapsto{D_{\mathrm{adj}}(\omega)} \omega X \omega^{-1}, \qquad X \in \mathcal{G},$$

where $\mathcal{G}$ denotes the Lie algebra. The right-hand side refers to the ordinary matrix product.

The Lie algebra of a Lie group is the tangent space at the unit element of the group. Hence there is a neighbourhood in which elements of a matrix group can be written as

$$\omega = \exp\left(iX^a T_a\right), \tag{A.1}$$

where the X^a's are real or complex coordinates and the index a runs over the dimension of the group. The matrices T_a are the basis vectors of the Lie algebra $\mathcal{G}$, and the factor i is a peculiarity of physical conventions — in mathematical literature it is omitted. Lie algebras can also be defined without reference to groups as vector spaces equipped with a Lie bracket. The Lie bracket is a $\mathcal{G}$-valued bilinear form which is antisymmetric in its arguments and satisfies the Jacobi identity. In matrix Lie algebras the bracket is defined in terms of the matrix commutator in such a way that relations of the form

$$[T_a, T_b] = i f_{ab}{}^c T_c \tag{A.2}$$

hold for all basis elements. The coefficients ${f_{ab}}^c$ are known as structure constants of the algebra. Note that the factor i is again omitted in mathematical conventions.

A d_r-dimensional matrix representation of a Lie algebra $\mathcal{G}$ is a linear mapping ρ_r which takes elements of $\mathcal{G}$ to $d_r \times d_r$ -matrices in a way that preserves the commutator:

$$\rho_r\left(i[X,Y]\right) = i\left[\rho_r(X), \rho_r(Y)\right], \qquad X, Y \in \mathcal{G}.$$

Since the mapping is linear, it holds that

$$\rho_r(X) = \rho_r(X^a T_a) = X^a \rho_r(T_a),$$

and the generators $\rho_r(T_a)$ of the representation obey the algebra (A.2), i.e.,

$$[\rho_r(T_a), \rho_r(T_b)] = i{f_{ab}}^c\, \rho_r(T_c). \tag{A.3}$$

The representation is called fundamental if the representation matrices are given by the Lie algebra $\mathcal{G}$ itself. As with groups, the adjoint representation is one in which the Lie algebra is the representation space. The generators of this representation are defined in terms of the structure constants as

$$\Big(\rho_{\mathrm{adj}}(T_a)\Big)^c_{\ b} = i{f_{ab}}^c,$$

and the algebra (A.3) then follows from the Jacobi identity of the fundamental representation. Any group representation can locally be obtained by exponentiating the corresponding representation of the algebra, that is, for group elements of the form (A.1) the relation

$$D_r\Big(\exp\left(iX^a T_a\right)\Big) = \exp\Big(iX^a \rho_r(T_a)\Big)$$

holds true.

If an inner product is defined in the algebra $\mathcal{G}$, it is given by a positive-definite Hermitean metric tensor h,

$$(X, Y) := h_{ab}(X^a)^* Y^b.$$

If the algebra is real, h is positive-definite and symmetric. The metric tensor and its inverse are used to raise and lower indices,

$$X_a := h_{ab} X^b, \qquad X^a = h^{ab} X_b, \qquad h^{ac} h_{cb} = {\delta^a}_b.$$

A Lie algebra is called compact if it is real and has a positive-definite metric that satisfies

$$(X, i[Y,Z]) = -\left(i[Y,X], Z\right) \qquad \text{for all} \quad X, Y, Z \in \mathcal{G}. \tag{A.4}$$

Note that the notion of compactness does not refer to topology in this context. Lie groups can be compact in the topological sense, and the definition of a compact Lie algebra has been tailored so that algebras of compact Lie groups satisfy it. In such a case the inner product is invariant under the adjoint action of the group, i.e.,

$$\left(\omega X \omega^{-1}, \omega Y \omega^{-1}\right) = (X, Y) \qquad \text{for all} \quad \omega \in G,$$

and equation (A.4) is then the infinitesimal form of this relation.

If a subspace of a Lie algebra is closed under the Lie bracket, it is called a subalgebra. A subalgebra $\mathcal{S}$ is an ideal if

$$i\left[Y, X\right] \in \mathcal{S} \qquad \text{for all} \quad Y \in \mathcal{S}, \quad X \in \mathcal{G}.$$

If the algebra $\mathcal{G}$ contains no Abelian ideals other than the null space $\{0\}$, it is called semisimple. Whether an algebra is semisimple or not can be deduced from the structure constants with the help of the Killing form

$$K(X,Y) = k_{ab}X^aY^b, \qquad k_{ab} := f_{ac}{}^d f_{bd}{}^c.$$

The connecting theorem is due to Cartan and Weyl:

A Lie algebra $\mathcal{G}$ is semisimple if and only if its Killing form is non-degenerate, i.e., if the tensor k is invertible. Moreover, a real Lie algebra $\mathcal{G}$ is semisimple and compact if and only if the Killing form is negative-definite.

An immediate consequence of this theorem is the fact that in compact semisimple Lie algebras the Killing form can be used to define the metric tensor as

$$h_{ab} = -ck_{ab},$$

where c is some positive constant.

In the representation theory of Lie algebras and in Paper III of this thesis it is necessary to find a maximal set of matrices that commute with every element of the algebra or at least with each other. The former are called Casimir operators, and they belong to the enveloping algebra of $\mathcal{G}$, which is generated by all sums and products of the algebra elements. In semisimple algebras one of these operators is given by

$$C_2 = k^{ab}T_aT_b,$$

where k^{ab} stands for the inverse of the Killing form. The number and form of the other Casimir operators depends on the detailed structure of the algebra. As regards maximal sets of matrices that commute with each other, they can be constructed by collecting together independent Casimir operators, including certain subalgebra Casimirs, and the generators of a so-called Cartan subalgebra. A Cartan subalgebra $\mathcal{H}$ of a semisimple Lie algebra $\mathcal{G}$ is an Abelian subalgebra of maximal dimension such that the commutator matrices

$$(\mathrm{ad}\, X)^a{}_b := i[X,T_b]^a = -f_{cb}{}^a X^c.$$

are simultaneously diagonalisable for all elements $X \in \mathcal{H}$. The dimension of the Cartan subalgebra is called the rank of the algebra $\mathcal{G}$. It is known that in semisimple algebras the number of independent Casimir operators is equal to the rank of the algebra.

Appendix B

Functional analysis

Paper II of this thesis is a mathematical study on the existence and smoothness of solutions to the covariant Poisson equation on $\mathbb{R}^3$. The method is based on functional analysis, and the following sections are devoted to the presentation of some basic definitions in this field. The focus is on concepts employed in Paper II. More traditional applications of functional analysis to partial differential equations can be found in Refs. [109,110]. It is assumed that the reader is familiar with Lebesgue integration and elementary topology.

B.1 Banach and Hilbert spaces

Linear functional analysis deals with vector spaces where an inner product or at least a norm is defined. It is understood that the norm is always positive-definite and in particular that only the null vector has zero norm. Let us denote this norm by $||\cdot||$. A sequence (x_n) is said to be a Cauchy-sequence if for every $\varepsilon > 0$ there exists a number $n_\varepsilon \in \mathbb{N}$ such that

$$||x_n - x_m|| < \varepsilon \qquad \text{for all} \quad n, m \geq n_\varepsilon.$$

A space where all Cauchy sequences converge is called complete. Banach spaces are defined as normed vector spaces which in addition are complete. Similarly, Hilbert spaces are complete inner product spaces. Every Hilbert space is therefore also a Banach space, but the converse does not hold.

A linear mapping $f : X \to \mathbb{R}$ defined on a normed space X is said to be a bounded linear functional on X if there exists a positive constant C such that the inequality

$$|f(x)| \leq C\,||x||$$

holds for all elements $x \in X$. Every linear functional is bounded if and only if it is continuous.

A subset S of a normed vector space X is said to be dense if every element of X can be approximated by elements of S, i.e., if for every $x \in X$ and $\varepsilon > 0$ there exists an element $s \in S$ such that

$$||x - s|| < \varepsilon.$$

A Banach or Hilbert space is called separable if it is possible to construct a countable basis, that is, a set of linearly independent elements $\{\Psi_n\}$ such that every vector can be approximated with arbitrary accuracy by linear combinations of the basis elements.

Paper II of this thesis uses extensions of the spaces L^p and L^p_{loc} defined on $\mathbb{R}^3$. The L^p spaces with $1 \leq p < \infty$ consist of Lebesgue measurable real functions f for which the integral

$$||f||_p := \left(\int_{\mathbb{R}^3} |f(\mathbf{x})|^p \, d^3\mathbf{x} \right)^{1/p}$$

is defined and finite. If the Lebesgue measure of $\mathbb{R}^3$ is denoted by μ, then measurable real functions which are essentially bounded on $\mathbb{R}^3$, i.e., functions for which

$$||f||_\infty := \operatorname*{ess\,sup}_{\mathbf{x}\in\mathbb{R}^3} |f(\mathbf{x})| := \inf \Big\{ \alpha \in \mathbb{R} \,\Big|\, \mu\left(\{ \mathbf{x} \in \mathbb{R}^3 \,|\, |f(\mathbf{x})| > \alpha \} \right) = 0 \Big\}$$

is finite belong to the space L^∞. Similarly, if $|f|^p$ is integrable on compact subsets of $\mathbb{R}^3$, f belongs to the local space L^p_{loc}. Functions which are essentially bounded on compact subsets of $\mathbb{R}^3$ constitute the space L^∞_{loc}. The global L^p spaces, including L^∞, are Banach spaces with the norms given above, provided that one identifies functions which agree almost everywhere. Actually L^2 is even a Hilbert space, because its norm is induced by the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^3} f(\mathbf{x})\, g(\mathbf{x})\, d^3\mathbf{x}.$$

Extending the above definitions to Lie algebra -valued mappings, one obtains the spaces $L^p(\mathbb{R}^3, \mathcal{G})$ and $L^p_{\mathrm{loc}}(\mathbb{R}^3, \mathcal{G})$. The Lie algebra $\mathcal{G}$ is assumed to be compact, and its inner product is determined by a positive-definite metric tensor h_{ab}. If $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalues of h, it holds for all elements $X \in \mathcal{G}$ that

$$\lambda_{\min} \sum_{a=1}^{d} (X^a)^2 \leq |X|^2 \leq \lambda_{\max} \sum_{a=1}^{d} (X^a)^2 \,, \qquad d = \dim \mathcal{G},$$

indicating that $\mathcal{G}$ is topologically equivalent with $\mathbb{R}^d$. Measurability and integrability of $\mathcal{G}$-valued mappings

$$F : \mathbb{R}^3 \to \mathcal{G}$$

are then defined in the same fashion as for $\mathbb{R}^d$-valued mappings. The extended L^p norms are

$$||F||_p = \left(\int_{\mathbb{R}^3} |F(\mathbf{x})|^p \, d^3\mathbf{x} \right)^{1/p}$$

$$||F||_\infty = \operatorname*{ess\,sup}_{\mathbf{x}\in\mathbb{R}^3} |F(\mathbf{x})|,$$

and the spaces $L^p(\mathbb{R}^3, \mathcal{G})$ consist of measurable mappings for which the corresponding norms are defined and finite. The local spaces $L^p_{\mathrm{loc}}(\mathbb{R}^3, \mathcal{G})$ are defined similarly, requiring only integrability of $|F|^p$ or essential boundedness of $|F|$ on compact subsets of $\mathbb{R}^3$. The global spaces $L^p(\mathbb{R}^3, \mathcal{G})$ are Banach spaces, and $L^2(\mathbb{R}^3, \mathcal{G})$ is a Hilbert space, when equipped with the inner product

$$\langle F, G \rangle = \int_{\mathbb{R}^3} \big(F(\mathbf{x}), G(\mathbf{x}) \big)\, d^3\mathbf{x}.$$

The L^p norms satisfy a number of important inequalities. First, there is the triangle inequality

$$||f + g||_p \leq ||f||_p + ||g||_p \qquad \text{for all} \quad f, g \in L^p, \qquad 1 \leq p \leq \infty,$$

also known as Minkowski's inequality. The inner product of L^2 satisfies the Schwarz inequality

$$|\langle f, g\rangle| \leq ||f||_2 \, ||g||_2 \qquad \text{for all} \quad f, g \in L^2.$$

This inequality can be generalised, and the result is the Hölder inequality

$$\int_{\mathbb{R}^3} |fg| \, d^3\mathbf{x} \leq ||f||_p \, ||g||_q,$$

where $f \in L^p$, $g \in L^q$ and

$$\frac{1}{p} + \frac{1}{q} = 1, \qquad 1 \leq p, q \leq \infty.$$

B.2 Sobolev spaces

There are many ways of defining Sobolev spaces, but roughly speaking one can regard their elements as extensions of continuous mappings. Let C^n stand for the space of n times continuously differentiable real functions on $\mathbb{R}^3$, and let $C^n(\mathbb{R}^3, \mathcal{G})$ be the corresponding space of $\mathcal{G}$-valued mappings. The notation C^∞ refers to mappings which have continuous derivatives to all orders, and omission of the superscript n means that the mappings are merely continuous. The support of a mapping is the closure of all points where the mapping is nonzero:

$$\text{spt}F := \overline{\left\{\mathbf{x} \in \mathbb{R}^3 \,\middle|\, F(\mathbf{x}) \neq 0\right\}}.$$

The spaces $C_c^n(\mathbb{R}^3, \mathcal{G})$ consist of mappings that belong to $C^n(\mathbb{R}^3, \mathcal{G})$ and whose support is a compact subset of $\mathbb{R}^3$. The first step towards Sobolev spaces consists of supplying the C^n spaces with suitable norms. In Paper II of this thesis it is shown that the expression

$$\begin{aligned} < \Phi, \Psi >_1 &= \int_{\mathbb{R}^3} \frac{1}{w^{1-\sigma}} \big(\nabla_k(A)\Phi, \nabla^k(A)\Psi\big) \, d^3\mathbf{x}, \\ w(\mathbf{x}) &= \left(1 + |\mathbf{x}|^2\right)^{1/2}, \quad 0 < \sigma \leq 1 \end{aligned} \tag{B.1}$$

defines an inner product on $C_c^1(\mathbb{R}^3, \mathcal{G})$, when $A_k \in L^2_{\text{loc}}(\mathbb{R}^3, \mathcal{G})$. The parameter σ is a fixed number which will later control the asymptotic properties of the Sobolev spaces. At first sight it seems that the proposed norm

$$||\Phi||_{1,2} := \langle \Phi, \Phi \rangle_1^{1/2}$$

could vanish for nonzero mappings. However, there exists a positive constant C such that the inequality

$$\left(\int_{\mathbb{R}^3} \frac{1}{w^{3-\sigma}} |\Phi|^2 \, d^3\mathbf{x}\right)^{1/2} \leq C \left(\int_{\mathbb{R}^3} \frac{1}{w^{1-\sigma}} \big(\nabla_k(A)\Phi, \nabla^k(A)\Phi\big) \, d^3\mathbf{x}\right)^{1/2} \tag{B.2}$$

holds in $C_c^1(\mathbb{R}^3, \mathcal{G})$. The expression (B.1) thus induces a proper norm.

Although $C_c^1(\mathbb{R}^3, \mathcal{G})$ is now an inner product space, it is not a Hilbert space. Since the norm $||\cdot||_{1,2}$ and the left-hand side of (B.2) are actually weighted L^2 norms, it is realised that Cauchy sequences in $C_c^1(\mathbb{R}^3, \mathcal{G})$ converge to elements in certain weighted L^2 spaces. More precisely, if (Φ_i) is a Cauchy sequence in $C_c^1(\mathbb{R}^3, \mathcal{G})$, then there exists a mapping Ψ such that

$$w^{-(3-\sigma)/2} \, \Psi \in L^2(\mathbb{R}^3, \mathcal{G})$$

and

$$\left(\int_{\mathbb{R}^3} \frac{1}{w^{3-\sigma}} |\Phi_i - \Psi|^2 \, d^3\mathbf{x}\right)^{1/2} \xrightarrow[i\to\infty]{} 0.$$

The first order Sobolev space $H_1(\mathbb{R}^3, \mathcal{G})$ is now defined as an extension of $C_c^1(\mathbb{R}^3, \mathcal{G})$ which contains the limits of all Cauchy sequences. In other words,

$$H_1(\mathbb{R}^3, \mathcal{G}) = \overline{C_c^1(\mathbb{R}^3, \mathcal{G})},$$

where the closure is taken in the norm $||\cdot||_{1,2}$. By construction, $H_1(\mathbb{R}^3, \mathcal{G})$ is a Hilbert space when one identifies mappings that agree almost everywhere.

If Ψ is an element of $H_1(\mathbb{R}^3, \mathcal{G})$ and (Φ_i) a Cauchy sequence converging to Ψ in the norm $||\cdot||_{1,2}$, it is seen that the sequence of covariant derivatives $(\nabla_k(A)\Phi_i)$ converges in the weighted L^2 norm induced by the right-hand side of (B.1). The limits of these derivative sequences are defined as weak covariant derivatives of Ψ, denoted by $\nabla_k(A)\Psi$. The weak derivatives are square integrable in the sense that

$$w^{-(1-\sigma)/2} \, \nabla_k(A)\Psi \in L^2(\mathbb{R}^3, \mathcal{G}),$$

and they fall in the category of distributional derivatives, which are characterised by the property that the relation

$$\int_{\mathbb{R}^3} \big(\Phi, \nabla_k(A)\Psi\big) \, d^3\mathbf{x} = -\int_{\mathbb{R}^3} \big(\nabla_k(A)\Phi, \Psi\big) \, d^3\mathbf{x}$$

holds for all mappings $\Phi \in C_c^\infty(\mathbb{R}^3, \mathcal{G})$. In particular, Ψ does not have to be differentiable in the ordinary sense.

Weak derivatives are useful in finding distributional solutions to the covariant Poisson equation

$$\nabla_k(A)\nabla^k(A)Z = F, \tag{B.3}$$

where both Z and F are $\mathcal{G}$-valued mappings. If the weak covariant derivatives $\nabla_k(A)Z$ and their commutators with A_k are at least locally integrable, the left-hand side of the Poisson equation (B.3) can be interpreted weakly as a requirement that the relation

$$\int_{\mathbb{R}^3} \big(\widetilde{\Phi}, \nabla_k(A)\nabla^k(A)Z\big) \, d^3\mathbf{x} = -\int_{\mathbb{R}^3} \big(\nabla_k(A)\widetilde{\Phi}, \nabla^k(A)Z\big) \, d^3\mathbf{x},$$

holds for all $\widetilde{\Phi} \in C_c^\infty(\mathbb{R}^3, \mathcal{G})$. Distributional solutions of (B.3) therefore satisfy the equation

$$-\int_{\mathbb{R}^3} \big(\nabla_k(A)\widetilde{\Phi}, \nabla^k(A)Z\big) \, d^3\mathbf{x} = \int_{\mathbb{R}^3} \big(\widetilde{\Phi}, F\big) \, d^3\mathbf{x} \qquad \text{for all } \ \widetilde{\Phi} \in C_c^\infty(\mathbb{R}^3, \mathcal{G}).$$

If the mappings $\widetilde{\Phi}$ are written as

$$\widetilde{\Phi}(\mathbf{x}) = \frac{1}{w(\mathbf{x})^{1-\sigma}}\Phi(\mathbf{x}), \qquad \Phi \in C_c^\infty(\mathbb{R}^3, \mathcal{G}),$$

the weak Poisson equation takes a weighted form,

$$\begin{aligned}\int_{\mathbb{R}^3} \frac{1}{w^{1-\sigma}} \big(\nabla_k(A)\Phi, \nabla^k(A)Z\big) \, d^3\mathbf{x} \, - \, (1-\sigma) \int_{\mathbb{R}^3} \frac{x_k}{w^{3-\sigma}} \big(\Phi, \nabla^k(A)Z\big) \, d^3\mathbf{x} \\ = -\int_{\mathbb{R}^3} \frac{1}{w^{1-\sigma}} \big(\Phi, F\big) \, d^3\mathbf{x}.\end{aligned} \tag{B.4}$$

The existence of a solution can now be inferred from the Lax–Milgram theorem [109,110]:

Let H be a real Hilbert space with norm $||\cdot||$ and let B denote a bilinear mapping $B : H \times H \to \mathbb{R}$. Assume that there exist constants α, $\beta > 0$ such that

$$|B(Y,Z)| \leq \alpha\,||Y||\,||Z|| \qquad \textit{for all } Y, Z \in H$$
$$B(Z,Z) \geq \beta\,||Z||^2.$$

If f is a bounded linear functional on H, then there exists in H a unique element Z such that

$$B(Y,Z) = f[Y] \qquad \textit{for all } Y \in H.$$

It is shown in Paper II that the left-hand side of equation (B.4) defines a bilinear mapping $B(\Phi, Z)$ with the required properties on the Hilbert space $H_1(\mathbb{R}^3, \mathcal{G})$. The right-hand side of (B.4) in turn is seen to define a bounded linear functional on $H_1(\mathbb{R}^3, \mathcal{G})$ when

$$w^{(1+\sigma)/2} F \in L^2(\mathbb{R}^3, \mathcal{G}).$$

The Lax–Milgram theorem then states that there is an element $Z \in H_1(\mathbb{R}^3, \mathcal{G})$ such that equation (B.4) holds for all mappings $\Phi \in H_1(\mathbb{R}^3, \mathcal{G})$ and in particular for those belonging to $C_c^\infty(\mathbb{R}^3, \mathcal{G})$. In other words, Z is a distributional solution to the covariant Poisson equation.

B.3 Continuous embeddings

It was mentioned earlier that weak solutions of partial differential equations are not necessarily even differentiable in the ordinary sense. One then wonders what are the conditions under which a weak solution is actually a classical solution with continuous derivatives. This question can be studied with the help of so-called Sobolev embeddings. Generally, if a Banach space Y is a subset of another Banach space X in such a way that the norms of these spaces satisfy an inequality of the form

$$||y||_X \leq C\,||y||_Y \qquad \text{for all } \; y \in Y$$

with some positive constant C, then it is said that Y is continuously embedded in X. If one now considers Sobolev spaces with derivatives of sufficiently high order in the definition of the norm, it turns out that these Sobolev spaces can be continuously embedded in spaces consisting of continuously differentiable and bounded mappings. This means that every element of a higher order Sobolev space is differentiable (modulo sets of measure zero) up to a certain order.

Sobolev spaces of order n can be defined as extensions of the spaces $C_c^n(\mathbb{R}^3, \mathcal{G})$. As shown in Paper II, the expression

$$< \Phi, \Psi >_n = \sum_{p=1}^{n} \int_{\mathbb{R}^3} w^{(2p-3)(1-\sigma)} \big(\nabla_{k_1}(A) \cdots \nabla_{k_p}(A)\Phi, \nabla^{k_1}(A) \cdots \nabla^{k_p}(A)\Psi \big)\, d^3\mathbf{x}$$

defines an inner product on $C_c^n(\mathbb{R}^3, \mathcal{G})$, if the gauge potential satisfies certain local smoothness assumptions. Denoting the corresponding norm by

$$||\Phi||_{n,2} := \langle \Phi, \Phi \rangle_n^{1/2},$$

the Sobolev space $H_n(\mathbb{R}^3, \mathcal{G})$ is defined as a closure of $C_c^n(\mathbb{R}^3, \mathcal{G})$ in this norm, i.e.,

$$H_n(\mathbb{R}^3, \mathcal{G}) = \overline{C_c^n(\mathbb{R}^3, \mathcal{G})}.$$

Elements of $H_n(\mathbb{R}^3, \mathcal{G})$ possess square integrable weak derivatives up to nth order:

$$w^{(p-3/2)(1-\sigma)} \nabla_{k_1}(A) \cdots \nabla_{k_p}(A) Z \in L^2(\mathbb{R}^3, \mathcal{G}), \qquad p = 1, \ldots, n.$$

In Paper II it is proved that these spaces are embedded in the spaces $C_B^{n-2}(\mathbb{R}^3, \mathcal{G})$, which consist of mappings $\Phi \in C^{n-2}(\mathbb{R}^3, \mathcal{G})$ that are bounded in the norm

$$||\Phi||_{n-2,\infty} = \max_{0 \leq p \leq n-2} \sup_{\mathbf{x} \in \mathbb{R}^3} \left\{ w(\mathbf{x})^{p(1-\sigma)} \Big(\nabla_{k_1}(A) \cdots \nabla_{k_p}(A) \Phi(\mathbf{x}), \nabla^{k_1}(A) \cdots \nabla^{k_p}(A) \Phi(\mathbf{x}) \Big)^{1/2} \right\}.$$

At this stage it is assumed that the gauge potential is continuous and $A_k \in C^{n-3}(\mathbb{R}^3, \mathcal{G})$ for $n \geq 3$ so that the covariant derivatives in the above definition yield continuous mappings. The inclusion

$$H_n(\mathbb{R}^3, \mathcal{G}) \subset C_B^{n-2}(\mathbb{R}^3, \mathcal{G})$$

is then a continuous embedding. It remains to determine the conditions under which a weak solution of the covariant Poisson equation actually belongs to higher order Sobolev spaces. This can be done by deriving bounds for the norms $||\cdot||_{n,2}$ in terms of lower order norms and weighted L^2 derivative norms of the mapping F in (B.4). As a rule, these estimates are difficult to work out, and Paper II is no exception in this respect.

Bibliography

[1] A. Jaffe and E. Witten, *Quantum Yang–Mills Theory* (Clay Mathematics Institute Millennium Problem, 2000).

[2] *Selected Papers on Quantum Electrodynamics*, edited by J. Schwinger (Dover, New York, 1958).

[3] C. N. Yang and R. L. Mills, Phys. Rev. **96** (1954) 191.

[4] R. Utiyama, Phys. Rev. **101** (1956) 1597.

[5] P. W. Higgs, Phys. Lett. **12** (1964) 132; Phys. Rev. Lett. **13** (1964) 508; Phys. Rev. **145** (1966) 1156.

[6] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321.

[7] S. L. Glashow, Nucl. Phys. **22** (1961) 579.

[8] S. Weinberg, Phys. Rev. Lett. **19** (1967) 1264.

[9] A. Salam, in *Elementary Particle Theory*, proceedings of the 8th Nobel symposium, Lerum, Sweden, 1968, edited by N. Svartholm (Almqvist & Wiksell, Stockholm, 1968), p. 367.

[10] G. 't Hooft, Nucl. Phys. B **33** (1971) 173; **35** (1971) 167.

[11] G. 't Hooft and M. Veltman, Nucl. Phys. B **44** (1972) 189; **50** (1972) 318.

[12] H. Fritzsch, M. Gell-Mann and H. Leutwyler, Phys. Lett. B **47** (1973) 365.

[13] D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30** (1973) 1343; H. D. Politzer, *ibid.* 1346.

[14] S. Weinberg, Phys. Rev. D **8** (1973) 605.

[15] S. Mandelstam, Phys. Lett. B **53** (1975) 476.

[16] G. 't Hooft, Nucl. Phys. B **190** (1981) 455.

[17] M. Creutz, Phys. Rev. D **21** (1980) 2308.

[18] Ph. de Forcrand, J. Stat. Phys. **43** (1986) 1077.

[19] K. D. Born, E. Laermann, N. Pirch, T. F. Walsh and P. M. Zerwas, Phys. Rev. D **40** (1989) 1653.

[20] N. Seiberg and E. Witten, Nucl. Phys. B **426** (1994) 19; **430** (1994) 485 (Erratum); **431** (1994) 484.

[21] H. Georgi and S. L. Glashow, Phys. Rev. Lett. **32** (1974) 438.

[22] H. Georgi, in *Particles and Fields*, AIP Conference Proceedings, edited by C. Carlson (American Institute of Physics, New York, 1975), Vol. 23, p. 575.

[23] J. Wess and B. Zumino, Nucl. Phys. B **78** (1974) 1.

[24] A. Salam and J. Strathdee, Phys. Lett. B **51** (1974) 353.

[25] S. Ferrara and B. Zumino, Nucl. Phys. B **79** (1974) 413.

[26] T. K. Leen, Annals Phys. **147** (1983) 417.

[27] D. J. Toms, Phys. Rev. D **27** (1983) 1803; M. Omote and S. Ichinose, *ibid.* 2341.

[28] N. Seiberg and E. Witten, JHEP **9** (1999) 032.

[29] R. Jackiw, Rev. Mod. Phys. **52** (1980) 661.

[30] S. L. Adler, Phys. Rev. **177** (1969) 2426.

[31] J. S. Bell and R. Jackiw, Nuovo Cimento A **60** (1969) 47.

[32] I. Segal, J. Funct. Anal. **33** (1979) 175.

[33] J. Ginibre and G. Velo, Commun. Math. Phys. **82** (1981) 1.

[34] Y. Choquet-Bruhat and D. Christodoulou, Ann. Sci. École Norm. Sup. (4) **14** (1981) 481.

[35] D. M. Eardley and V. Moncrief, Commun. Math. Phys. **83** (1982) 171; *ibid.* 193.

[36] S. Coleman, Commun. Math. Phys. **55** (1977) 113.

[37] G. 't Hooft, Nucl. Phys. B **79** (1974) 276.

[38] A. M. Polyakov, JETP Lett. **20** (1974) 194.

[39] T. T. Wu and C. N. Yang, Phys. Rev. D **12** (1975) 3843.

[40] H. Nielsen and P. Olesen, Nucl. Phys. B **61** (1973) 45.

[41] B. Julia and A. Zee, Phys. Rev. D **11** (1975) 2227.

[42] A. A. Belavin, A. M. Polyakov, A. S. Schwartz and Yu. S. Tyupkin, Phys. Lett. B **59** (1975) 85.

[43] G. 't Hooft, Phys. Rev. D **14** (1976) 3432.

[44] E. Noether, Nachr. v. d. Ges. d. Wiss. zu Göttingen (1918) 235.

[45] E. L. Hill, Rev. Mod. Phys. **23** (1951) 253.

[46] V. N. Gribov, Nucl. Phys. B **139** (1978) 1.

[47] I. M. Singer, Commun. Math. Phys. **60** (1978) 7.

[48] M. Nakahara, *Geometry, Topology and Physics* (IOP, Bristol, 1990).

[49] M. Daniel and C. M. Viallet, Rev. Mod. Phys. **52** (1980) 175.

[50] C. Cronström, Phys. Lett. B **90** (1980) 267.

[51] M. Creutz, I. J. Muzinich and T. N. Tudron, Phys. Rev. D **19** (1979) 531.

[52] C. Cronström, *Lectures on Gauge Theory*, Internal Report HU-TFT-81-3 (Research Institute for Theoretical Physics, University of Helsinki, 1981).

[53] S. Weinberg, *The Quantum Theory of Fields I* (Cambridge University Press, 1995).

[54] G. C. Branco, L. Lavoura and J. P. Silva, *CP Violation* (Oxford University Press, 1999).

[55] F. J. Belinfante, Physica **6** (1939) 887.

[56] R. Jackiw, in *Current Algebra and Anomalies* (World Scientific, Singapore, 1985).

[57] S. Weinberg, *The Quantum Theory of Fields II* (Cambridge University Press, 1996).

[58] K. Fujikawa, Phys. Rev. Lett. **42** (1979) 1195.

[59] R. P. Feynman, Rev. Mod. Phys. **20** (1948) 367.

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

[61] C. G. Callan, Phys. Rev. D **2** (1970) 1541.

[62] K. Symanzik, Commun. Math. Phys. **18** (1970) 227.

[63] G. 't Hooft, Nucl. Phys. B **61** (1973) 455.

[64] I. S. Gerstein, R. Jackiw, B. W. Lee and S. Weinberg, Phys. Rev. D **3** (1971) 2486.

[65] R. H. Cameron, J. Math. and Phys. **39** (1960) 126.

[66] M. Noga, Phys. Rev. D **38** (1988) 3158.

[67] K. Symanzik, Nucl. Phys. B **190** (1981) 1.

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

[69] R. Jackiw, in *Field Theory and Particle Physics*, proceedings of the 5th Jorge André Swieca Summer School, Campos do Jordão, Brazil, 1989, edited by O. J. P. Éboli, M. Gomes and A. Santoro (World Scientific, Singapore, 1990), p. 78.

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

[71] C. Cronström and M. Noga, Nucl. Phys. B **428** (1994) 449.

[72] C. Cronström, Acta Phys. Slov. **50** (2000) 369.

[73] C. Cronström, J. Math. Phys. **46** (2005) 073503.

[74] F. Berezin, Uspehi Mat. Nauk **29** (1974) 183.

[75] C. Cronström, private communication.

[76] H. Rund, *The Hamilton–Jacobi Theory in the Calculus of Variations* (Van Nostrand, London, 1966).

[77] P. Majumdar and H. S. Sharatchandra, Phys. Rev. D **58** (1998) 067702.

[78] T. Barnes and G. I. Ghandour, Czech. J. Phys. B **29** (1979) 256; Nucl. Phys. B **166** (1980) 125.

[79] M. Creutz and T. N. Tudron, Phys. Rev. D **16** (1977) 2978.

[80] R. Floreanini and R. Jackiw, Phys. Rev. D **37** (1988) 2206.

[81] C. Kiefer and A. Wipf, Annals Phys. **236** (1994) 241.

[82] L. D. Faddeev, Phys. Lett. B **145** (1984) 81.

[83] J. Goldstone and R. Jackiw, Phys. Lett. B **74** (1978) 81.

[84] V. Baluni and B. Grossman, Phys. Lett. B **78** (1978) 226.

[85] M. B. Halpern, Phys. Rev. D **19** (1979) 517.

[86] F. Lenz, H. W. L. Naus and M. Thies, Annals Phys. **233** (1994) 317.

[87] M. Bauer, D. Z. Freedman and P. E. Haagensen, Nucl. Phys. B **428** (1994) 147.

[88] P. E. Haagensen and K. Johnson, Nucl. Phys. B **439** (1995) 597.

[89] H. Nachbagauer, Phys. Rev. D **52** (1995) 3672.

[90] L. Chen, M. Belloni and K. Haller, Phys. Rev. D **55** (1997) 2347.

[91] C. Becchi, A. Rouet and R. Stora, Commun. Math. Phys. **42** (1975) 127; Annals Phys. **98** (1976) 287.

[92] I. V. Tyutin, Lebedev Institute preprint No. 39 (1975).

[93] E. S. Fradkin and G. A. Vilkovisky, Phys. Lett. B **55** (1975) 224.

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

[95] M. Henneaux and C. Teitelboim, *Quantization of Gauge Systems* (Princeton University Press, 1992)

[96] M. Moshinsky and T. H. Seligman, Annals Phys. **114** (1978) 243.

[97] A. M. Khvedelidze, D. M. Mladenov, H.-P. Pavel and G. Röpke, Eur. Phys. J. C **24** (2002) 137.

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

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

[100] T. Maskawa and H. Nakajima, Progr. Theor. Phys. **56** (1976) 1295.

[101] A. G. Izergin, V. E. Korepin, M. A. Semenov-Tian-Shansky and L. D. Faddeev, Theor. Math. Phys. **38** (1979) 1.

[102] R. Marnelius, Nucl. Phys. B **194** (1982) 433.

[103] S. Hwang, Nucl. Phys. B **351** (1991) 425.

[104] A. M. Khvedelidze and H.-P. Pavel, Phys. Rev. D **59** (1999) 105017.

[105] S. Lie, Math. Ann. **8** (1875) 215.

[106] D. H. Sattinger and O. L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry and Mechanics* (Springer, New York, 1986)

[107] H. Bacry, *Leçons sur la Théorie des Groupes et les Symétries des Particules Élémentaires* (Gordon & Breach, Paris, 1967).

[108] A. O. Barut and R. Rączka, *Theory of Group Representations and Applications* (Polish Scientific Publishers, Warsaw, 1977).

[109] L. C. Evans, *Partial Differential Equations* (American Mathematical Society, Providence, 1998).

[110] R. C. McOwen, *Partial Differential Equations* (Prentice Hall, Upper Saddle River, 1996).