Commun. Math. Phys. 160, 431–456 (1994)

Communications in
Mathematical
Physics


# Constrained Quantisation, Gauge Fixing and the Gribov Ambiguity

**David McMullan**

School of Mathematics and Statistics, University of Plymouth, Drake Circus, Plymouth, Devon PL48AA, U.K. Email: D.MCMULLAN@plymouth.ac.uk

Received: 28 December 1992/in revised form: 23 April 1993

**Abstract:** The role played by gauge fixing in the description of superselection sectors for a simple quantum mechanical system is analysed. By viewing this as a theory with constraints, it is shown that the possibility of having inequivalent gauge fixing conditions (Gribov's ambiguity) signals the existence of inequivalent reductions to a physical quantum theory, and hence superselection sectors. This point of view is contrasted with the more traditional one that identifies superselection sectors with inequivalent quantisations. It is argued that emphasising the role of gauge fixing (along with the Gribov problem) will allow for a more direct extension of these ideas to quantum field theory and, in particular, gauge theories.

## 1. Introduction

Classically a Yang–Mills theory is characterised by its coupling constant. After quantising, though, an additional parameter is needed, measuring the vacuum angle. We can think of this angle $\theta$ as labelling various superselection sectors; it then being a matter of experimentation (through the analysis of CP-violating effects) to fix its physical value. Initially $\theta$ arose from a semi-classical analysis [1]. However it is now clear (as discussed, for example, in [2]) that the emergence of such an angle reflects a general property of quantising when the configuration space[1] of the system has non-contractible loops (the $\theta$ angle then emerging from the representation theory of the first homotopy group, $\pi_1$, of this space).

The existence of such non-trivial loops in Yang–Mills theory follows from the fact that the group of all gauge transformations (in 3-dimensional space) is disconnected, and indeed has components labelled by the integers. Thus it is the set $\hat{\mathbb{Z}}$, of all irreducible unitary representations of the group of integers $\mathbb{Z}$, that determines the superselection sector: as is well known, $\hat{\mathbb{Z}} \simeq SO(2)$ – hence we get an angle.

[1] The configuration space in question being the space of physically inequivalent potentials $\mathscr{A}/\mathscr{G}$ (with $\mathscr{G}$ the group of gauge transformation) as opposed to the (extended configuration) space $\mathscr{A}$ of all Yang–Mills potentials – which is topologically quite trivial

This analysis is directly analogous to the identification of the circle, $S^1$, as the coset space $\mathbb{R}/\mathbb{Z}$ – the group of integers here playing the role of the components of the group of gauge transformations. Indeed, the quantum theory of a particle moving on a circle also has superselection sectors labelled by an angle; which can have consequences analogous to the observed Aharonov–Bohm effect [3].

Having seen the physical significance of the $\pi_1$-effects, it is natural to enquire about the consequences for quantisation of the rest of the rich topological structures found in the Yang–Mills configuration space. Excluding the exciting things that can happen when coupling to matter, the only noticeable effect attributed to them is the Gribov problem [4]. This arises when we try to fix the gauge using a continuous function. Then a consequence of the non-triviality of the configuration space is that such a guage fixing cannot be constructed globally [5] – we must be content with local gauge fixing.

In the context of perturbation theory this appears to present no real problem as one would expect to be able to stay within the region where the gauge fixing is well defined. How one extends beyond this region is then far from clear, and still open to debate. The majority view seems to be that, by being slightly more careful, this potential problem can be seen as no more than a technical blemish (see, for example, the comments in [2]). However, others have thought differently. Indeed Gribov tried to relate confinement in QCD to the existence of a maximal region where gauge fixing works (the Gribov horizon), and this point of view has been taken up more recently by Zwanziger [6]. The full significance of this analysis, though, is not yet clear.

So the current state of affairs seems to be that it is only through the $\pi_1$-effects that superselection sectors enter Yang–Mills theory – all the other nonlinearities only adding technical problems to a reasonably clear class of quantisations.

If true, this would be a surprising result since we know of many quantum mechanical systems which have no $\pi_1$-structure and yet have interesting superselection sectors. The simplest example of this is the free particle moving on the two sphere, $S^2$ ($\pi_1(S^2) = 0$). One approach[2] to the quantisation of this system involves applying Mackey's theory of induced representation [8] to the two sphere viewed as the coset space $SO(3)/SO(2)$. Hence the superselection sectors are labelled by an integer describing the irreducible unitary representations of the $SO(2)$ subgroup. More generally, for motion on the coset space $G/H$, with $G$ and $H$ not too wild, the superselection sectors are labelled by the dual space $\hat{H}$ of equivalence classes of irreducible unitary representations of the little group $H$.

The physical configuration space $\mathscr{A}/\mathscr{G}$ for Yang–Mills theory seems close enough to this form for use to try and apply Mackey's results to describe all, or at least some other, superselection sectors of the theory. Indeed, this is essentially how the $\theta$ vacua is understood. However, we now notice that the technical restrictions on $G$ and $H$ are far too restrictive to be systematically applied to field theory, where even locally compact spaces have to be sacrificed. Such mathematical niceties do not usually perturb physicists from applying the constructions anyway – just to see what happens. The real problem, though, for such a program is that this description of superselection sectors relies on natural structures to be found in quantum mechanics (invariant measures, full description of $\hat{H}$, concrete Hilbert spaces of

[2] Another approach is to use geomentic quantisation [7]: then it is the structure of $H^2(S^2, \mathbb{Z})$ that is responsible of the superselection sectors

states, etc.), which are really more than we can expect to find in a quantum field theory. What we would like to see is an alternative description of this type of quantisation that involves structures which can be more readily extended to field theory.

The aim of this paper is to initiate such a reformulation. The main conclusion of this analysis will be that it is precisely the Gribov type of problem with gauge fixing that is responsible for the superselection sectors found in these quantum mechanical models, thus suggesting that we should re-examine our response to the Gribov problem found in Yang–Mills theory.

Quite a lot of ground needs to be covered before we can arrive at this conclusion. The main starting point for the analysis will be to view the motion on the coset space $G/H$ as motion on $G$ subject to the constraints which generate the action of $H$ on $G$. This is clearly motivated by the standard description of Yang–Mills theory as dynamics on the extended configuration space $\mathscr{A}$ subject to the Gauss law constraints, which generate the action of the gauge transformations on $\mathscr{A}$.

If we now follow Dirac's [9] analysis of such a system, then we should first quantise on the extended configuration space $G$ – identifying the states as $L^2$ functions on $G$, then reduce to the physical theory by identifying the physical states as those annihilated by the quantum constraints. For these systems, if $H$ is compact, both of these steps are well defined and one recovers a quantum theory – but only one! There are no superselection sectors.

It is possible to recover the superselection sectors by modifying, by hand, the quantum constraints. For the example of a particle on $S^2$ this involves replacing the classical first class constraint $r_3 = 0$ (the notation is explained in Sect. 3) by the quantum constraint $r_3 = n$. For different integers $n$ one gets different reductions which reproduce the known superselection sectors. The extension of this argument to the general $G/H$ situation is not too difficult [10], although one now has to cope with the fact that the quantum constraints will become a mixture of first and second class ones.

This modified version of Dirac's procedure does lead to a satisfactory constrained description of Mackey's results. However, it is not a constrained account of why superselection sectors arise.

In [11] it was proposed that the superselection sectors could be recovered by modifying the first step in Dirac's analysis. So, rather than using the simplest, or naive, quantisation on $G$, it was suggested that we first should identify $G$ with the coset $(G \times H)/\tilde{H}$, where $\tilde{H} \simeq H$ is a diagonal subgroup of $G \times H$. Appealing to Mackey's analysis of such systems allows us to deduce that there will be superselection sectors labelled by $\hat{\tilde{H}} \simeq \hat{H}$. The physical states are then recovered by projecting with the constraints as before. The motivation for this construction is two-fold: First the naive quantisation, as described in [11], takes as basic the generators of the left action of $G$ on itself; whereas the constraints, which generate a right action, are treated as derived objects. Hence there is a *potential* ordering problem associated with the constraints, which can be avoided on the extended space $(G \times H)/\tilde{H}$ (in Sect. 3.1 we shall see, though, that this potential problem does not actually arise in practice). Secondly, keeping with the naive quantisation does not yield the expected superselection sectors – hence, in the view of [11], the naive quantisation must be wrong.

Now it must be stressed that, using Mackey's results, *we do already know how to quantise on the coset space $G/H$ – the whole point of this exercise then is to recover* *these results from a constrained analysis on* $G$. Thus, from the constrained dynamical perspective, the above proposal also cannot be considered as an attractive solution to this problem. Rather, if we are trying to understand how to quantise on the coset space $G/H$, and we wish to identify $G$ with $(G \times H)/\tilde{H}$, then we should also treat this second description as constrained motion on $G \times H$. As we only get one quantisation this way, we will then be forced to view $G \times H$ as the coset $(G \times H \times \tilde{H})/\tilde{\tilde{H}}$. Viewing this as constrained motion on $G \times H \times \tilde{H}$ will again force us to bigger and bigger spaces, with no hope of getting the superselection sectors.

It might be argued that this repeated identification of a coset space as a constrained system is rather extreme, and is causing more problems than it solves. But we must remember that the only real justification for this reappraisal of the known quantum results is so that the techniques developed can be applied to real, physically relevant, field theories. There we really do just want to use the naive quantisation on the extended Yang–Mills configuration space $\mathscr{A}$ – it would be highly unattractive if we are forced to identify $\mathscr{A}$ with some larger space with a more exotic quantisation.

So we will adopt the working philosophy that a constrained analysis is deemed successful if it satisfies the following general guidelines:

(1) Classically it involves an extended phase space and a reduction procedure to the true degrees of freedom;
(2) Quantisation involves the naive quantisation on the extended phase space, that is, the fact that there are constraints is not taken into account;
(3) Physical quantum states are recovered using a direct extension to the quantum theory of the classical reduction used in (1);
(4) The states described in (3) agree with the states obtained by directly quantising on the true degrees of freedom, if this is possible[3].

The discussion above shows that Dirac's approach is not successful in this sense since (4) does not hold for the coset spaces. Also, as will be argued in Sect. 2, it fails on account of (3) and also (2) if non-compact groups are allowed. The proposal of [11] is also unsuccessful on these accounts, especially since it does not conform to step (2). However, we make no claim that an unsuccessful approach is not useful. In particular, the method whereby the constraints are modified by hand (hence not keeping to step (3)) can be extended to Yang–Mills theory with surprising consequences [10].

It is not at all clear that there is any successful constrained formalism – the aim of this paper is to show that there is, at least for the particle moving on $S^2 = SO(3)/SO(2)$. The more general coset example will be presented elsewhere as it involves more subtle, group theoretic arguments that are important, but obscure the essential new methods needed to tackle this problem.

The plan of this paper is as follows: After this introduction, in Sect. 2, a detailed discussion of how states are described in various formulations of constrained dynamics will be presented. This will start with Dirac's approach applied to a simple first-class system. The difficulties encountered in translating the classical

---

3 For Yang–Mills theory this comparison is not readily available, so we are then forced to essentially define the quantum theory as that which emerges from the constrained analysis. To have any confidence in this we need to ensure that, at least any applications where (4) can be tested, the constrained formalism is successful

characterization of physical states to the quantum theory will be reviewed and various unsatisfactory aspects to the whole construction emphasised. After that, the description of such a first-class system using ghost variables and BRST-symmetry will be presented. This will involve a careful account of how classical states on a super phase space should be defined, and isolated. It will be seen that, upon quantisation, many of the problems encountered with Dirac's approach are now not apparent. But, there is now a new type of problem related to how the quantum states should be defined. In the end we will be forced to conclude that even this approach is flawed.

The interesting thing about the shortcomings of these two standard methods is that they, in some sense, complement each other. This motivates a new approach to constrained dynamics that combines the best of both worlds by using ghosts not just for the constraints but also for the gauge fixing. Thus the BRST-charge will be replaced by a new, non-abelian charge; allowing for a natural definition of both the classical and quantum states, along with a systematic way to isolate the physical ones in both contexts.

Armed with this new method for analysing constrained systems, in Sect. 3 the example of a particle moving on $SO(3)$, subject to the constraint associated to the right action of $SO(2)$, will be discussed in detail. The interesting thing about this example is that it has a Gribov type problem when gauge fixing is used. A consequence of which will be that the methods developed in Sect. 2 now lead to inequivalent reductions in the quantum theory, reflecting the inequivalent classes of gauge fixing terms available in the model. Finally we shall conclude with some remarks on the possible morals to be learnt from this analysis, and the consequences for field theory.

## 2. States and Constraints

The objective here is to highlight some of the general problems encountered in isolating the physical states in a constrained analysis. This discussion will be presented through the analysis of two simple systems. The first has extended configuration space $\mathbb{R}^{n+1}$ and constraint $p_0 = 0$ (non-compact example), while the second has extended configuration space $\mathbb{R}^n \times S^1$ and constraint $p_\theta = 0$ (compact example). In both cases the physical configuration space is $\mathbb{R}^n$. We use the notation that in both cases the coordinates on the extended phase space, $P_{\text{ext}} = T^* Q_{\text{ext}}$, are $(q^a, p_a)$, $a = 0, \ldots, n$; with $(q^0, p_0)$ identified with $(\theta, p_\theta)$ in the compact case. The coordinates on the physical phase space $P_{\text{phy}} = T^*\mathbb{R}^n = \mathbb{R}^{2n}$ are then $(q^\alpha, p_\alpha)$, $\alpha = 1, \ldots, n$. Classical states will generically be denoted by $\omega$, while $\psi$ will be used for quantum ones.

*2.1. States and the Dirac Formalism.* In both classical and quantum mechanics an essential distinction is made between states and observables. Classically a state[4] is a point of a phase space, and the observables are the functions on the phase space. Quantum mechanically the states are elements of a Hilbert space and the observables are (self-adjoint) operators on this space. The notion of a state and an observable in both cases are closely connected – the states being a class of dual objects to a $(C^*)$-algebra of observables.

[4] We will take all states to be pure in this paper

It is useful to play down the geometric nature of a classical state, and rather represent the state corresponding to the point $(\bar{q}, \bar{p})$ by the $\delta$-functions

$$\omega(\bar{q}, \bar{p}) = \delta(q^a - \bar{q}^a)\delta(p_a - \bar{p}_a) . \tag{2.1}$$

Then, with respect to the Liouville measure $d\mu(q, p)$, we have $\int \omega(\bar{q}, \bar{p})\, d\mu(q, p) = 1$, and the expectation value of any classical observable, $f(q, p)$, in this state is

$$\begin{aligned}\langle f \rangle_{\omega(\bar{q}, \bar{p})} &:= \int \omega(\bar{q}, \bar{p}) f(q, p)\, d\mu(q, p) \\ &= f(\bar{q}, \bar{p}) ,\end{aligned} \tag{2.2}$$

the value of the function at the point $(\bar{q}, \bar{p})$.

The quantum states $\psi$ for this system are then the normalised elements of the Hilbert space $L^2(Q_{\text{ext}})$, of wave functions on the (extended) configuration space. The expectation value of an observable $\hat{f}$ is then $\langle \hat{f} \rangle_\psi = \langle \psi | \hat{f} | \psi \rangle = \int \bar{\psi}(q) \hat{f} \psi(q)\, d\mu(q)$, where $d\mu(q)$ is the measure on the configuration space arising from the volume form on the space.

In order to isolate the physical classical states from the set of extended states (2.1), it is not enough to simply impose the constraint $\langle p_0 \rangle_\omega = 0$, as this only implies $\bar{p}_0 = 0$. On top of this we also need to use a gauge fixing condition. Choosing $q^0$ as the gauge fixing function, the physical states can be described by

$$\omega_{\text{phy}} = \{\omega(\bar{q}, \bar{p}) \colon \langle p_0 \rangle_\omega = \langle q^0 \rangle_\omega = 0\} . \tag{2.3}$$

Thus, in this gauge, we can write

$$\omega_{\text{phy}} = \delta(p_0)\delta(q^0)\delta(q^\alpha - \bar{q}^\alpha)\delta(p_\alpha - \bar{p}_\alpha) . \tag{2.4}$$

Note that since the classical states are identified with specific distributions on the phase space, we can also write condition (2.3) as

$$p_0 \omega_{\text{phy}} = q^0 \omega_{\text{phy}} = 0 . \tag{2.5}$$

In the Dirac approach to quantising constrained systems, the classical constraints become operator identities on the quantum states. But, from this point of view we cannot now directly transcribe the classical identification of states (2.5) into the quantum theory, since the uncertainty principle tells us that there are no states (or generalised states) which satsify both $\hat{p}_0 | \psi \rangle = 0$ and $\hat{q}^0 | \psi \rangle = 0$. In order to circumvent this problem one usually reasons as follows: The classical states are defined on the whole of the phase space, while the quantum ones are just wave functions over the configuration space. So, in the passage to the quantum theory half of the classical states have already been dealt with—thus there is no need for the gauge fixing in the quantum theory. All we need to do, then, is to define the physical states by the Dirac condition

$$\psi_{\text{phy}} = \{\psi \colon \hat{p}_0 \psi = 0\} . \tag{2.6}$$

At first sight such an identification of the physical states seems to work; giving credence to the view that gauge fixing has no role in isolating the quantum states. Indeed, (2.6) tells us that the physical states have no dependence on the unphysical $q^0$-direction; they are just functions of the physical $q$'s.

The problem, though, with this identification is that one needs to ensure that there are solutions to (2.6) in $L^2(Q_{\text{ext}})$. For the non-compact example this is not the case – wave functions constant in the $q^0$-direction cannot be normalised using the measure $d\mu(q)$ on the extended configuration space. To resolve this, while keeping the identification (2.6) of physical states, we need to modify the measure, i.e., $d\mu(q) \to d\mu(q)\delta(q^0)$, consequently the constraints will no longer be self-adjoint, and we have been forced into a highly non-standard quantisation.

Things are better in the compact example since solutions to (2.6) do exist – so we have a sensible set of states. However, in this example we are viewing the physical configuration space as the coset space $(\mathbb{R}^n \times S^1)/S^1$ (in much the same way as $G$ was viewed as $(G \times H)/\tilde{H}$ in the introduction) hence we would expect superselection sectors labelled by $S_O(2) = \mathbb{Z}$. The Dirac analysis, though, has only recovered the trivial sector.

The conclusion of this short account of Dirac's approach is that it might be useful for some specific applications (such as the trivial sector of systems with compact symmetry groups), but it is not successful in the sense discussed in the introduction.

*2.2. States and Ghost Variables.* A powerful and popular approach to constrained dynamics involves enlarging the extended phase space to a super phase space. The new fermionic coordinates, for our examples, being a ghost variable $\eta$ and its conjugate $\rho$ ($\eta\eta = \rho\rho = 0$, $\eta\rho = -\rho\eta$, $\{\eta, \rho\} = \{\rho, \eta\} = 1$). We shall postpone until Sect. 2.3 the discussion of how these new variables are used in a constrained analysis; here we shall concentrate on how to define states, both classically and quantum mechanically, in this graded context.

The states (2.1) have various nice properties: they pick out the points of the phase space; they are normalised to one; they are characters for the algebra[5] of functions on the phase space, i.e., a state determines a map $\omega(f) := \langle f \rangle_\omega$, from the functions to the reals such that

$$\omega(f)\omega(g) = \omega(fg) . \tag{2.7}$$

Some care is needed in extending these properties of states to the graded situation [12]. If we focus on the role of states as recovering the points of the (super) manifold then a sensible definition of a *graded state*, concentrated at the point $(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho})$, is

$$\omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) = \delta(q^a - \bar{q}^a)\delta(p_a - \bar{p}_a)\delta(\rho - \bar{\rho})\delta(\eta - \bar{\eta}) . \tag{2.8}$$

The delta functions for the odd variables are simply given by $\delta(\eta - \bar{\eta}) = \eta - \bar{\eta}$, with a similar expression for the conjugate ghost. Hence we have $\int \omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) d\mu(q, p, \eta, \rho) = 1$, where $d\mu(q, p, \eta, \rho) = d\mu(q, p)\, d\eta\, d\rho$. Acting on the function $\mathscr{F} = f_0 + f_1\eta + f_2\rho + f_3\eta\rho$, with the $f_i$'s functions on the bosonic phase space, we get

$$\begin{aligned}\langle \mathscr{F} \rangle_\omega &:= \int \mathscr{F}(q, p, \eta, \rho)\omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) d\mu(q, p, \eta, \rho) \\ &= f_0(\bar{q}, \bar{p}) + f_1(\bar{q}, \bar{p})\bar{\eta} + f_2(\bar{q}, \bar{p})\bar{\rho} + f_3(\bar{q}, \bar{p})\bar{\eta}\bar{\rho} .\end{aligned} \tag{2.9}$$

This evaluation of the function, though, has not given us a real number but rather an element of the graded extension of the reals needed to define a super manifold (see the discussion in [12]). We note, though, that the graded states for which $\bar{\eta} = 0$

---

[5] The class of functions being the continuous functions which tend to zero at infinity; this is then a $C^*$-algebra with the sup norm

and $\bar{\rho} = 0$, are states in the sense described by (2.7). From (2.9) we see that if $\omega$ is such a restricted graded state then $\langle \mathscr{F} \rangle_\omega = f_0(\bar{q}, \bar{p})$, a trivial extension of the states given by (2.2).

For the quantum states it is natural to use wave functions of both the bosonic configuration variables and the fermionic one. Taking $\Psi(q, \eta) = \psi_0(q) + \psi_1(q)\eta$, and using the product measure $d\mu(q, \eta) = d\mu(q)d\eta$, we find that

$$\langle \Psi, \Psi \rangle = \int (\bar{\psi}_0 \psi_1 + \bar{\psi}_1 \psi_0)\, d\mu(q) \, . \tag{2.10}$$

To extract any useful information from this we require that the integrand is a probability density function. There are various possible ways to ensure this; the most natural of which is to require that both $\psi_0$ and $\psi_1$ are wave functions on the configuration space, and that they are equal:

$$\psi_0 = \psi_1 \, , \tag{2.11}$$

then $\langle \Psi, \Psi \rangle = 2\langle \psi_0, \psi_0 \rangle$. Other, more involved, possibilities will be discussed in the next section.

*2.3. The* BRST-*Charge and Physical States.* In order to extract from the graded states (2.8) the physical ones we can simply treat the ghost variables as some extra constraints and, following (2.3), require that

$$\omega_{\text{phy}} = \{\omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) \colon \langle q^0 \rangle_\omega = \langle p_0 \rangle_\omega = \langle \eta \rangle_\omega = \langle \rho \rangle_\omega = 0\} \, . \tag{2.12}$$

Thus, though, is not too exciting, and will suffer the problems encountered earlier in the Dirac approach when extending to the quantum theory.

An important property of ghost variables, and indeed the whole reason why they were introduced in the first place, is that we can encode the conditions (2.12) in a much more succinct way.

The BRST-charge $\mathscr{Q}$ for this system is defined to be

$$\mathscr{Q} = p_0 \eta \, . \tag{2.13}$$

This charge generates the BRST transformation on states through the action of its Hamiltonian vector field

$$\delta = -p_0 \frac{\partial}{\partial \rho} - \eta \frac{\partial}{\partial q^0} \, . \tag{2.14}$$

If we require that $\delta\omega = 0$, then we see from (2.8) that $\bar{p}_0 = \bar{\eta} = 0$. So we have encoded half of the conditions needed in (2.12) to isolate the physical states. The normal expectation is that the cohomological structure associated with the BRST charge, $\delta^2 = 0$, suffices to characterise the rest of the conditions needed to describe the physical states. However, this is not the case [12], and an additional charge is needed which explicitly depends on the gauge fixing.

The (symplectic) dual, $\bar{\mathscr{Q}}$, to the BRST charge is given by [12, 13],

$$\bar{\mathscr{Q}} = q^0 \rho \, . \tag{2.15}$$

Its action on states is given by the Hamiltonian vector field $\bar{\delta}$, where

$$\bar{\delta} = -q^0 \frac{\partial}{\partial \eta} + \rho \frac{\partial}{\partial p_0} \, . \tag{2.16}$$

The condition $\bar{\delta}\omega = 0$ then implies that $\bar{q}^0 = \bar{\rho} = 0$. We thus see that the classical physical states can be identified with those graded states that are invariant under both the BRST and dual-BRST transformations, i.e.

$$\omega_{\rm phy} = \{\omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) \colon \delta\omega = \bar{\delta}\omega = 0\} \ . \tag{2.17}$$

The action of the BRST transformation, and its dual, on the quantum states is given by[6]

$$\hat{\mathcal{Q}} = -i\eta\frac{\partial}{\partial q^0} \ , \tag{2.18}$$

and

$$\hat{\bar{\mathcal{Q}}} = -iq^0\frac{\partial}{\partial \eta} \ . \tag{2.19}$$

From (2.17) we are tempted to define

$$\Psi_{\rm phy} = \{\Psi \colon \hat{\mathcal{Q}}\Psi = \hat{\bar{\mathcal{Q}}}\Psi = 0\} \ , \tag{2.20}$$

that is, directly incorporate the gauge fixing condition into the isolation of the quantum states. The surprising thing is that now we can make sense of this identification, in contrast to Dirac's approach.

Indeed, (2.20) is equivalent to the conditions

$$\frac{\partial\psi_0}{\partial q^0} = 0 \quad \text{and} \quad q^0\psi_1 = 0 \ . \tag{2.21}$$

In the non-compact case neither of these have any solutions unless distributions are allowed. Admitting such functions allow us to identify the physical wave functions with those generalised (super) wave functions of the form

$$\Psi_{\rm phy} = \psi_0 + \delta(q^0)\psi_0\eta \ , \tag{2.22}$$

where $\dfrac{\partial\psi_0}{\partial q^0} = 0$. These are normalisable states which satisfy (2.20), however, condition (2.11) is violated since we have $\psi_1 = \delta(q^0)\psi_0$.

Even in the compact case, where distributions do not at first seem to be forced upon us, condition (2.11) must be violated in order to find non-trivial solutions to (2.20). Alternatively, [14], the anti-BRST charge $\mathcal{Q}_{\rm anti}$ can be used instead of $\bar{\mathcal{Q}}$ in this situation, where

$$\mathcal{Q}_{\rm anti} = p_0\rho \ . \tag{2.23}$$

This charge has ghost number minus one and, in general, is constructed out of the BRST-charge under the interchange of the ghosts with their conjugates – which we note is a canonical transformation. Classically $\mathcal{Q}$ and $\mathcal{Q}_{\rm anti}$ do not fully isolate the physical states since requiring anti-BRST invariance will only imply $\bar{p}_0 = \bar{\rho} = 0$; $\bar{q}^0$

---

[6] As usual in the Schrödinger representation we take $\hat{q} = q$, $\hat{p} = -i\dfrac{\partial}{\partial q}$, and with our convention on the ghost Poisson bracket the same is true for the ghosts; $\hat{\eta} = \eta$, $\hat{\rho} = -i\dfrac{\partial}{\partial \eta}$. We note that the conjugate ghost is taken to be anit-Hermitian

is still unrestricted. However, for compact symmetry groups, solving $\hat{\mathscr{Q}}\Psi = \hat{\mathscr{Q}}_{\text{anti}}\Psi = 0$, with condition (2.11) on states, does work giving

$$\Psi_{\text{phy}} = \psi_0 + \psi_0\eta\ , \qquad (2.24)$$

where $\dfrac{\partial\psi_0}{\partial q^0} = 0$. However, just as we saw for the Dirac analysis, this description has just picked out one sector from all the expected superselection sectors, and the method used for isolating the physical quantum states is not the same as that used in the classical theory – gauge fixing has been avoided.

So our conclusions for the use of ghosts, along with $\mathscr{Q}$ and $\bar{\mathscr{Q}}$, as a constrained formalism is that one can get further than in the standard Dirac approach – since now the states are isolated by much the same conditions, (2.17) and (2.20), in both the classical and quantum theory. However, in the quantum theory, the class of wave functions has had to be extended due to the constraints: the desired relation (2.11) only holds in the compact case when $\mathscr{Q}_{\text{anti}}$ is used instead of $\bar{\mathscr{Q}}$, indeed then $\hat{\mathscr{Q}}\Psi = 0$ implies that $\hat{\mathscr{Q}}_{\text{anti}}\Psi = 0$. We also note that the states used do not have a definite ghost number.

*2.4. A Second Class Approach.* We have seen that by introducing ghost variables the classical role of gauge fixing in the isolation of physical states can be extended to the quantum theory. The only caveat being that the class of wave functions needed has to be extended to include distributions and condition (2.11) has to be dropped. Indeed we have seen that condition (2.11) on quantum states is closely connected to the anit-BRST charge. However, this charge does not use gauge fixing and does not isolate any normalisable states in the non-compact case.

These uncertainties in how one should define states clearly violates the spirit of condition (2) for a successful quantisation, and its generalisation to more complex systems is quite involved [15]. If this was the best we could do them we might be willing to overlook these technical shortcomings. However, we shall now see that we can do better, even for the non-compact case.

Given that ghosts allow gauge fixing to act on states in the quantum theory, the divide between (first class) constraints and gauge fixing, which was so essential in Dirac's approach, now seems artificial. Indeed, since the BRST charge was constructed solely out of the constraints, it was precisely this divide that accounts for the fact that the anti-BRST charge contains no gauge fixing information. In order to treat constraints and gauge fixing on a more equal footing we combine them to form a second class set – then add ghosts for them all, [13].

The second class analogue of the BRST charge (2.13), using $q^0$ as the gauge fixing condition, is defined to be

$$\mathscr{S} = p_0\eta^1 + q^0\eta^2\ . \qquad (2.25)$$

Here $\eta^i$, $(i = 1, 2)$, are the ghost variables, with conjugates $\rho_i$. In contrast to the BRST charge, which was abelian ($\{\mathscr{Q}, \mathscr{Q}\} = 0$), we now have a non-abelian charge, $\{\mathscr{S}, \mathscr{S}\} = -2\eta^1\eta^2$, reflecting the second class nature of the constraints.

The graded states are now given by

$$\omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) = \delta(q^a - \bar{q}^a)\delta(p_a - \bar{p}_a)\delta(\rho_i - \bar{\rho}_i)\delta(\eta^i - \bar{\eta}^i)\ . \qquad (2.26)$$

The second class charge $\mathscr{S}$ acts on these states through its Hamiltonian vector field $\kappa$, where

$$\kappa = -p_0 \frac{\partial}{\partial \rho_1} - q^0 \frac{\partial}{\partial \rho_2} - \eta^1 \frac{\partial}{\partial q^0} + \eta^2 \frac{\partial}{\partial p_0} . \tag{2.27}$$

Then $\kappa\omega = 0$ implies that $\bar{p}_0 = \bar{q}^0 = \bar{\eta}^1 = \bar{\eta}^2 = 0$. To deal with the conjugate ghosts, we extend the argument presented at the end of Sect. 2.3, and define the anti-$\mathscr{S}$ charge to be

$$\mathscr{S}_{\text{anti}} = p_0\rho_1 + q^0\rho_2 . \tag{2.28}$$

Then $\kappa_{\text{anti}}\omega = 0$ implies that $\bar{p}_0 = \bar{q}^0 = \bar{\rho}^1 = \bar{\rho}^2 = 0$. Thus the classical physical states are described in this second class formulation by

$$\omega_{\text{phy}} = \{\omega(\bar{q}, \bar{p}, \bar{\eta}, \bar{\rho}) \colon \kappa\omega = \kappa_{\text{anti}}\omega = 0\} . \tag{2.29}$$

Due to the extra ghosts, the structure of the quantum states is now richer than that encountered earlier. Working in a polarisation[7] where the wave functions are functions of the ghost variables $\eta^1$ and $\rho_2$, we see that a general wave function $\Psi(q, \eta^1, \rho_2)$ can be decomposed into states of ghost number zero, $\Psi_0$, and states of ghost number one, $\Psi_1$, and minus one, $\Psi_{-1}$. We write

$$\Psi_0 = \psi_0(q) + \psi_1(q)\eta^1\rho_2 , \tag{2.30}$$

and

$$\Psi_1 = \psi_3(q)\eta^1, \quad \Psi_{-1} = \psi_4(q)\rho_2 , \tag{2.31}$$

where the $\psi_n$'s are square integrable wave functions on the extended, bosonic configuration space. Then, following the discussion presented after (2.10) where we now use the measure $d\mu(q, \eta^1, \rho_2) = d\mu(q)d\eta^1 d\rho_2$, we see that the simplest way to ensure that these states give sensible probabilistic results is to require that

$$\psi_0 = -\psi_1 \quad \text{and} \quad \psi_3 = \psi_4 . \tag{2.32}$$

We note that the ghost number zero states $\Psi_0 = (1 - \eta^1\rho_2)\psi_0$ can be characterised as those states (2.30) which satisfy $K\Psi_0 = 0$, where the ghost number one operator $K$ is given by

$$K = \hat{\eta}^1 - i\hat{\eta}^2 . \tag{2.33}$$

The quantum version of the second class charge (2.25) is

$$\hat{\mathscr{S}} = -i\eta^1 \frac{\partial}{\partial q^0} - iq^0 \frac{\partial}{\partial \rho_2} . \tag{2.34}$$

Then $\hat{\mathscr{S}}\Psi = 0$ implies

$$\frac{\partial \psi_4}{\partial q^0} = q^0\psi_4 = 0 , \tag{2.35}$$

and

$$\frac{\partial \psi_0}{\partial q^0} - q^0\psi_1 = 0 . \tag{2.36}$$

---

7 We note, though, that the conclusions of the analysis presented here do not depend on this choice of polarisation, although for other choices the resulting physical states may not have ghost number zero

The first conditions, along with (2.32), imply that $\psi_3 = \psi_4 = 0$. Condition (2.36) becomes, after using (2.32),

$$\frac{\partial \psi_0}{\partial q^0} + q^0 \psi_0 = 0 \,, \tag{2.37}$$

which implies that

$$\psi_0(q) = e^{-1/2\,(q^0)^2}\psi_{\rm phy} \,, \tag{2.38}$$

where $\psi_{\rm phy}$ is a wave function on the physical configuration space. Note that due to the Gaussian factor in (2.38), $\psi_0$ is a wave function on the extended configuration space, even in the non-compact case. In fact, these states satisfy the weak condition on physical states that $\langle\psi_0|\hat{p}_0|\psi_0\rangle = \langle\psi_0|\hat{q}^0|\psi_0\rangle = 0$. Just as for the first class system, conditions (2.32) automatically imply that the states satisfying $\hat{\mathscr{S}}\Psi = 0$ will also satisfy $\hat{\mathscr{S}}_{\rm anti}\Psi = 0$.

$$\begin{aligned}\Psi_{\rm phy} &:= \{\Psi(q, \eta^1, \rho_2)\colon \hat{\mathscr{S}}\Psi = \hat{\mathscr{S}}_{\rm anti}\Psi = 0\} \,, \\ &= \{\Psi(q, \eta^1, \rho_2)\colon \hat{\mathscr{S}}\Psi = 0\} \,, \end{aligned} \tag{2.39}$$

where the components of the wave functions $\Psi$ satisfy conditions (2.32). From the discussion above we see that such states can be identified with the ghost number zero states which satisfy $\hat{\mathscr{S}}\Psi_0 = 0$. Indeed we have seen that

$$\Psi_{\rm phy} \simeq (1 - \eta^1\rho_2)e^{-1/2\,(q^0)^2}\psi_{\rm phy} \,. \tag{2.40}$$

In order to understand how superselection sectors can emerge in this formulation, and indeed to give a fuller description of these methods even when such sectors are not expected, we now discuss how the second class charges $\mathscr{S}$ and $\mathscr{S}_{\rm anti}$ are constructed for a more general gauge fixing condition. How gauge fixing can lead to superselection sectors will then be discussed in the next section.

Recall that the prescription for constructing the BRST-charge appropriate to a system with a general set of first-class constraints, $\phi_i$, is that one requires the charge to be of the form

$$\mathscr{Q} = \phi_i\eta^i + \text{higher order ghost terms} \,. \tag{2.41}$$

The additional terms being determined (almost uniquely) by the conditions that $\mathscr{Q}$ has ghost number one and is abelian: $\{\mathscr{Q}, \mathscr{Q}\} = 0$. Given such a charge it then follows that the BRST-charge needed for any equivalent set of constraints can be obtained by applying even canonical transformations to the original charge. The ability to do this is central to the success of these methods in the quantum theory [15].

For a second class set of constraints $(p_0, f)$, where now $f = f(q)$ is some[8] gauge fixing condition for the constraint $p_0$, i.e.,

$$\{f, p_0\} = \frac{\partial f}{\partial q^0} \neq 0 \,, \tag{2.42}$$

[8] For simplicity we only deal with a gauge fixing condition that is a function of the configuration space variables

we might be tempted to simply define the second class charge to be $\mathscr{S} = p_0\eta^1 + f\eta^2$, and construct its anti-version by the normal recipe of interchanging ghosts with their conjugates, i.e., take $\mathscr{S}_{\text{anti}} = p_0\rho_1 + f\rho_2$. This seems to be as good a prescription as any since now the addition of any higher order ghost terms, as in (2.41), cannot be fixed in any obvious way due to the non-abelian structure of the charge. However, this proposal does not have the essential property that changes in the gauge fixing term can be induced by canonical transformations, hence, following [15], there would be no unitary equivalence in the quantum theory relating the physical theory in different (but closely related) gauges.

Instead, we define the second class charge appropriate to the second class system $(p_0, f)$ to be

$$\begin{aligned}\mathscr{S} &= p_0\eta^1 + \{f, p_0\}f\eta^2 - \{\ln\{f, p_0\}, p_0\}\eta^1\eta^2\rho_2\,,\\ &\equiv p_0\eta^1 + \frac{\partial f}{\partial q^0}f\eta^2 - \left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}\eta^1\eta^2\rho_2\,. \end{aligned}\tag{2.43}$$

To show that this charge does have the property that (small) changes in the gauge fixing condition, $f \to f + \Delta f$, can be induced by canonical transformations, we write $\Delta f = \varepsilon\dfrac{\partial f}{\partial q^0}g$, where $g(q)$ is some arbitrary function. Any even function $h$ will induce the canonical transformation $\mathscr{S} \to \hat{\mathscr{S}} = \mathscr{S} + \varepsilon\{\mathscr{S}, h\}$. If we take

$$h = gp_0 + \frac{\partial g}{\partial q^0}(\eta^1\rho_1 + \eta^2\rho_2)\,,\tag{2.44}$$

then a straightforward calculation shows that the resulting $\tilde{\mathscr{S}}$ is given by (2.43) but with the gauge fixing term $f$ replaced by $f + \Delta f$.

We now need to construct the appropriate anti-charge to (2.43). How this is done depends on the status we assign to such an object. If we simply define the anti-charge to be that function, of ghost number minus one, *derived* from the second class charge (2.43) under the interchange of ghosts with their conjugates, then it is clear that we should take

$$\mathscr{S}_{\text{anti}} = p_0\rho_1 + f\frac{\partial f}{\partial q^0}\rho_2 - \left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}\eta^2\rho_1\rho_2\,.\tag{2.45}$$

The dependence of this charge on the gauge fixing term is then *defined* to be that derived from $\mathscr{S}$'s dependence. We do not separately apply the canonical transformation generated by (2.44) to this charge.

If, however, we do not wish to view the anti-charge as a purely derived object, but rather identify it with a function whose dependence on the gauge fixing function $f$ is such that under the canonical transformation generated by (2.44) it becomes a function of the gauge fixing function $f + \Delta f$, then we are forced to take

$$\mathscr{S}'_{\text{anti}} = \left(\frac{\partial f}{\partial q^0}\right)^{-2}\left(p\rho_1 + f\frac{\partial f}{\partial q^0}\rho_2 + \left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}\eta^2\rho_1\rho_2\right).\tag{2.46}$$

Clearly the relationship between $\mathcal{S}$ and $\mathcal{S}'_{\text{anti}}$ is not the naive one of interchanging ghosts with their conjugates. Rather $\mathcal{S}'_{\text{anti}}$ is constructed out of $\mathcal{S}$ under the canonical transformation

$$\begin{aligned} q &\to q\,, \\ p &\to p - 2\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}(\eta^1\rho_1 + \eta^2\rho_2)\,, \\ \eta^i &\to \left(\frac{\partial f}{\partial q^0}\right)^{-2}\rho_i\,, \\ \rho_i &\to \left(\frac{\partial f}{\partial q^0}\right)^{2}\eta^i\,. \end{aligned} \tag{2.47}$$

For the present we keep both forms of the anti-charge.

Using the new charges (2.43) and (2.45) or (2.46), it is clear that the physical classical states, as defined by (2.29), are those graded states for which $\langle p_0\rangle = \langle f\rangle = \langle \eta^i\rangle = \langle \rho_i\rangle = 0$.. Thus we see that the function $f$ is being used as a gauge fixing condition, as expected.

In the quantum theory we need to be slightly more careful due to the need to order terms in $\mathcal{S}$, $\mathcal{S}_{\text{anti}}$ and especially $\mathcal{S}'_{\text{anti}}$. Keeping $\mathcal{S}$ self adjoint we have

$$\begin{aligned} \hat{\mathcal{S}} = &\left(-i\frac{\partial}{\partial q^0} + \frac{i}{2}\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}\right)\eta^1 - if\frac{\partial f}{\partial q^0}\frac{\partial}{\partial \rho_2} \\ &- i\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}\eta^1\rho_2\frac{\partial}{\partial \rho_2}\,. \end{aligned} \tag{2.48}$$

Acting on the states $\Psi$ we find that $\hat{\mathcal{S}}\Psi = 0$ implies that $\psi_4 = 0$ (hence $\psi_3 = 0$) and, using (2.32),

$$\left(\frac{\partial}{\partial q^0} + f\frac{\partial f}{\partial q^0} - \frac{1}{2}\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\,\partial q^0}\right)\psi_0 = 0\,. \tag{2.49}$$

We now need to check what this condition implies for both $\hat{\mathcal{S}}_{\text{anti}}\Psi$ and $\hat{\mathcal{S}}'_{\text{anti}}\Psi$. Clearly we would hope that in both cases we get zero.

Keeping $\hat{\mathcal{S}}_{\text{anti}}$ anti-Hermitian we find

$$\begin{aligned} \hat{\mathcal{S}}_{\text{anti}} = &\left(-\frac{\partial}{\partial q^0} - \frac{1}{2}\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\partial q^0}\right)\frac{\partial}{\partial \eta^1} + f\frac{\partial f}{\partial q^0}\rho_2 \\ &+ \left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0\partial q^0}\rho_2\frac{\partial}{\partial \rho_2}\frac{\partial}{\partial \eta^1}\,. \end{aligned} \tag{2.50}$$

Acting on the states we directly recover Eq. (2.49), hence $\hat{\mathcal{S}}\Psi = 0$ implies that $\hat{\mathcal{S}}_{\text{anti}}\Psi = 0$.

For $\mathscr{S}'_{\text{anti}}$ more care is needed since we also have to order the first term. The net result of which is that

$$\hat{\mathscr{S}}'_{\text{anti}} = -i\left(\frac{\partial f}{\partial q^0}\right)^{-2}\left(-i\frac{\partial}{\partial q^0} + i\frac{3}{2}\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0 \partial q^0}\right)\frac{\partial}{\partial \eta^1}$$
$$+ f\left(\frac{\partial f}{\partial q^0}\right)^{-1}\rho_2 - \left(\frac{\partial f}{\partial q^0}\right)^{-3}\frac{\partial^2 f}{\partial q^0 \partial q^0}\rho_2 \frac{\partial}{\partial \rho_2}\frac{\partial}{\partial \eta^1}\,. \tag{2.51}$$

Thus acting on $\Psi_0$ we get

$$\hat{\mathscr{S}}_{\text{anti}}\Psi_0 = \left(\frac{\partial f}{\partial q^0}\right)^{-2}\left(\frac{\partial}{\partial q^0} + f\frac{\partial f}{\partial q^0} - \frac{1}{2}\left(\frac{\partial f}{\partial q^0}\right)^{-1}\frac{\partial^2 f}{\partial q^0 \partial q^0}\right)\psi_0 \rho_2\,. \tag{2.52}$$

Hence, from (2.42), if $\hat{\mathscr{S}}_{\text{anti}}\Psi = 0$ then $\hat{\mathscr{S}}_{\text{anti}}\Psi = 0$ as required.

An important point to note is that the requirement that there are states satisfying the physicality condition (2.49) will, in general, impose restrictions on the type of gauge fixing allowed. In particular, for the compact example expression (2.40) for the physical states is not an allowed state since the $\psi_0$ part is not an element of $L^2(S^1 \times \mathbb{R}^n)$ as it is not periodic in $q^0$. This also reflects the fact that the function $q^0$ is not a continuous function on the circle, thus we should really use functions of the form $f = e^{iq^0} - 1$ for the gauge fixing in this situation. A more interesting quantum restriction on the allowed gauge fixing term will be discussed in the next section.

In conclusion we have seen that (modulo any account of the superselection sectors) a successful constraint formalism can be developed if ghosts are introduced for both the constraint and the gauge fixing. By construction, if the gauge fixing condition can be smoothly deformed into $q^0 = 0$, then the solutions to (2.49) will be unitarily equivalent to the states (2.38). We have also seen that this result does not depend on which definition of anti-charge is used. Clearly, though, definition (2.45) is simpler than (2.46).

Actually the analysis presented here is only really half of the story – we have yet to discuss the isolation of physical observables in this second class approach. Such an account will be presented in Sect. 3 within the specific example considered there, since we shall see that in the quantum theory the choice of gauge fixing has some bearing on the choice of basic observables.

## 3. A Model with a Gribov Problem

The aim of this section is to spell out, in some detail, how the methods developed in Sect. 2 can be applied to a non-trivial, compact system. The example considered is that of motion on the Lie group $SO(3)$, subject to the constraint associated with the right action of $SO(2)$. The non-triviality of this example being the inability to construct a global gauge fixing condition for this action. For completeness we shall start with a brief review of the dynamics on a configuration space which is a Lie group, and its quantisation. Then the constrained formalism developed in Sect. 2 will be applied to the specific example of $G = SO(3)$ and the various quantum reductions described in detail. Throughout this section we denote the Lie algebra associated with the (compact, semi-simple) Lie group $G$ by $\mathfrak{g}$, and its dual by $\mathfrak{g}^*$. The structure functions of the Lie group $G$ will be denoted by $f^a_{bc}$.

*3.1. Dynamics and Quantisation on G.* The arena for a Hamiltonian description of the dynamics on $G$, and in particular for $SO(3)$, is the phase space $T^*G$. As with all such cotangent bundles this space comes equipped with a canonical symplectic form, from which the Poisson brackets can be calculated [16]. In order to give a fuller description of these structures we exploit the fact that $T^*G$ is actually a trivial bundle over $G$, indeed: $T^*G = \mathfrak{g}^* \times G$. There are two such identifications possible, reflecting the left or right action of $G$ on itself. For reasons to be apparent later, we work in the trivialisation adapted to the left action of $G$ on itself.

If $T_a$ is a basis of $\mathfrak{g}$, and $\theta^a$ the dual basis of $\mathfrak{g}^*$, then the point $l_a\theta^a \in \mathfrak{g}^*$ will be denoted by $l$. The above trivialisation of $T^*G$ them amounts to identifying $(l, x) \in \mathfrak{g}^* \times G$ with $l_a\theta^a_{\mathrm{MC}} \in T^*G$, where $\theta^a_{\mathrm{MC}}$ are the components of the right-invariant Maurer–Cartan form $\theta_{\mathrm{MC}} = T_a\theta^a_{\mathrm{MC}}$ on $G$. (Recall that if $R_a$ is the right invariant vector field on $G$ whose value at the identity is $T_a$, then $\langle\theta_{\mathrm{MC}}, R_a\rangle = T_a$ and $d\theta^a_{\mathrm{MC}} = \frac{1}{2}f^a_{bc}\theta^b_{\mathrm{MC}} \wedge \theta^c_{\mathrm{MC}}$.)

In these coordinates the Liouville form is $\theta(l, x) = l_a\theta^a_{\mathrm{MC}}$, so that the symplectic form is given by

$$\Omega = -d\theta = -dl_a \wedge \theta^a_{\mathrm{MC}} - \frac{1}{2}f^a_{bc}l_a\theta^b_{\mathrm{MC}} \wedge \theta^c_{\mathrm{MC}} . \tag{3.1}$$

Given a function $f$ on $T^*G$, its Hamiltonian vector field, $X_f$, is defined by

$$i_{X_f}\Omega = df = \frac{\partial f}{\partial l_a}dl_a + R_a(f)\theta^a_{\mathrm{MC}} . \tag{3.2}$$

From which we can deduce that

$$X_f = -\left(R_a(f) + \frac{\partial f}{\partial l_b}f^c_{ba}l_c\right)\frac{\partial}{\partial l_a} + \frac{\partial f}{\partial l_a}R_a . \tag{3.3}$$

Thus the Poisson bracket $\{f_1, f_2\} := X_{f_1}(f_2)$, between functions on $T^*G$, is given by

$$\{f_1, f_2\} = \frac{\partial f_1}{\partial l_a}R_a(f_2) - R_a(f_1)\frac{\partial f_2}{\partial l_a} - f^c_{ab}l_c\frac{\partial f_1}{\partial l_a}\frac{\partial f_2}{\partial l_b} . \tag{3.4}$$

Using (3.3), we see that the Hamiltonian vector field associated with the momentum variables $l_a$ are simply

$$X_{l_a} = R_a - f^c_{ab}l_c\frac{\partial}{\partial l_b} , \tag{3.5}$$

from which we can deduce that the $l_a$'s are the generators for the symplectic lift to $T^*G$ of the left action of $G$ on itself (which we recall induces the right invariant vector fields on $G$). Indeed $\{l_a, l_b\} = -f^c_{ab}l_c$, as expected for such a left action.

As well as the left action of $G$ on itself, there is also the right action. The symplectic lift of this action to $T^*G$ will be generated by functions, $r_a$, linear in momenta, i.e.,

$$r_a = \Lambda^b_a l_b , \tag{3.6}$$

for some smooth set of functions $\Lambda^b_a$ on $G$. We have from (3.3),

$$X_{r_a} = \Lambda^b_a R_b - (R_d(\Lambda^b_a)l_b + \Lambda^b_a f^c_{bd}l_c)\frac{\partial}{\partial l_d} . \tag{3.7}$$

Since the left and right actions of $G$ commute, we know that $\{r_a, l_d\} = 0$. This is equivalent to the condition $X_{r_a}(l_d) = 0$, which, from (3.7), implies that

$$R_d(\Lambda^b_a) = f^b_{dc}\Lambda^c_a \,. \tag{3.8}$$

Consequently we can deduce that $\{\Lambda^b_a, l_b\} = -R_b(\Lambda^b_a) = 0$. Hence, in constrast to the statement found in [11], we see that if we follow a route to quantisation whereby the left momenta $l_a$ is taken as fundamental, then there will be no factor ordering ambiguities in representing the quantum operators corresponding to the right momenta $r_a$, i.e, since both $\Lambda^b_a$ and $\hat{l}_b$ will be taken to be self adjoint operators, we can define the self adjoint operator $\hat{r}_a$ to be $\hat{r}_a = \Lambda^b_a\hat{l}_b = \hat{l}_b\Lambda^b_a$.

For the dynamics we opt for the simplest possible form for the Hamiltonian; so we take

$$H = l_a l_a \,. \tag{3.9}$$

As is well known, the classical trajectories then correspond to the free motion of a particle, with respect to the Cartan metric, on $G$ (to be more precise, it is with respect to a trivially rescaled metric to avoid factors of $-\frac{1}{2}$).

In order to quantise we must identify what the quantum states are, and how the basic observables are to be represented as operators on these states. Exploiting the group structure of $G$ allows for a very simple passage to the quantum theory (the so-called naive quantisation of [11]). The wave functions are taken to be elements of $L^2(G)$. Then for the position functions, $f \in C^\infty(G)$, we take

$$(\hat{f}\psi)(x) = f(x)\psi(x) \,, \tag{3.10}$$

while for the left momentum we define $\hat{l}_a$ by

$$\begin{aligned}(\hat{l}_a\psi)(x) &= \frac{d}{dt}\psi(e^{-tT_a}x)|_{t=0} \,, \\ &\equiv \pi_L(T_a)\psi(x) \,, \end{aligned} \tag{3.11}$$

where $\pi_L$ is the left-regular representation of $G$ on $L^2(G)$.

The quantum operator corresponding to the right momentum is then given by $\hat{r}_a = \Lambda^b_a\hat{l}_b$, while that for the Hamiltonian is simply given by $\hat{H} = \hat{l}_a\hat{l}_a$. We stress that there are no ordering ambiguities in either of these definitions.

*3.2. Constrained Dynamics on* $SO(3)$. Elements of the rotation group $SO(3)$ can be labelled by the Euler angles $\alpha$, $\beta$ and $\gamma$, where $0 \leqq \alpha < 2\pi$, $0 \leqq \beta \leqq \pi$ and $0 \leqq \gamma < 2\pi$. Indeed we can write $g \in SO(3)$ as

$$g = R(\alpha, \beta, \gamma) = e^{-i\alpha J_3}e^{-i\beta J_2}e^{-i\gamma J_3} \,, \tag{3.12}$$

where the $J_a$'s are the standard generators of the Lie algebra $so(3)$. We note, though, that the relationship between these angles and rotations is not one to one. Hence they do not form a globally valid system of coordinates for the rotation group [17].

We take as our classical constraint the condition that

$$r_3 = 0 \,. \tag{3.13}$$

This constraint generates the right action of an $SO(2)$ subgroup of $SO(3)$. Hence the reduced configuration space can be identified with the coset space

$SO(3)/SO(2) \simeq S^2$: the two sphere. The group $SO(3)$ is thus the total space of an $SO(2)$-principal bundle over the two sphere. As is well known, this non-trivial principal bundle has no global cross sections; hence there will be a Gribov problem when one tries to fix a gauge. However, even though no good gauge fixing can be found for this system, we will still need some locally valid way to fix the gauge in order to carry out the quantisation. The important question is then to determine how the associated reduction depends on the particular candidate for such a gauge fixing term.

The gauge fixing condition will generically be taken to be of the form $f(\alpha, \beta, \gamma) = 0$. However, directly working with the Euler angles will lead to some difficulties since we will also need to calculate Poisson brackets like $\{f, l_a\}$ and we need to find an explicit expression for the matrix $\Lambda_a^b$ used in expression (3.8). To deal with these problems we introduce a local set of Riemann normal (or canonical) coordinates; where the element $g = \exp(-ix^a J_a)$ of $SO(3)$ has coordinates $x^a$.

In terms of these coordinates we know that there must exist functions $L_a^b$ and $R_a^b$ such that we can write the left ($L_a$) and right ($R_a$) invariant vector fields on $SO(3)$ as

$$L_a = L_a^b \frac{\partial}{\partial x^b}, \quad R_a = R_a^b \frac{\partial}{\partial x^b} . \tag{3.14}$$

To find these functions we exploit the fact (see, for instance, Ref. 17, problem III.8) that

$$(L^{-1})_a^b = \left( \frac{1 - \exp(-x \cdot f)}{x \cdot f} \right)_a^b , \tag{3.15}$$

where $(x \cdot f)_a^b := x^c f^b_{\ ca} = x^c \varepsilon_{bca}$. It then follows that[9]

$$(L^{-1})_a^b = \delta_c^b \frac{\sin x}{x} + \frac{x^b x_a}{x^2}\left(1 - \frac{\sin x}{x}\right) + \frac{x^c}{x^2} f^b_{\ ac}(1 - \cos x) . \tag{3.16}$$

Here we have set $x = \sum_a \sqrt{(x^a)^2}$ and, as discussed earlier, we are using a rescaled Cartan metric so that $x^a = x_a$. We can thus deduce that

$$L_b^a = \delta_b^a \frac{x}{2} \cot\frac{x}{2} + \frac{x^a x_b}{x^2}\left(1 - \frac{x}{2}\cot\frac{x}{2}\right) - \frac{1}{2} x^c f^a_{\ bc} . \tag{3.17}$$

A similar analysis shows that the matrix $R_b^a$ is given by

$$R_b^a = \delta_b^a \frac{x}{2} \cot\frac{x}{2} + \frac{x^a x_b}{x^2}\left(1 - \frac{x}{2}\cot\frac{x}{2}\right) + \frac{1}{2} x^c f^a_{\ bc} . \tag{3.18}$$

If $p_a$ is the momentum conjugate to $x^a$: $\{x^a, p_a\} = \delta_b^a$, then its Hamiltonian vector field is simply given by $X_{p_a} = -\partial/\partial x^a$. The above analysis then shows that we can locally write the left momentum as

$$l_b = -R_b^a p_a . \tag{3.19}$$

[9] We note that there is a difference in sign, in the final term, between this expression and that found in [17]. As a test for the correctness of this form we note that as $x \to 0$, (3.15) becomes $\delta_a^b - \frac{1}{2} x^c f^b_{\ ca}$ as does (3.16)

The momenta $r_a$ is, by definition, the generator of the right action of $SO(3)$, hence from (3.7) and (3.8) we know that its Hamiltonian vector field is simply $L_a$. Thus we can deduce that

$$r_b = -L^a_b p_a . \tag{3.20}$$

Using (3.19) and (3.20) we see that we can identify the matrix $\Lambda^a_b$, which relates the left and right momenta (3.6), as

$$\begin{aligned} \Lambda^a_b &= (R^{-1})^a_c L^c_b \\ &= \delta^a_b \cos x + \frac{x^a x_b}{x^2}(1-\cos x) - \frac{x^c}{x} f^a_{bc} \sin x . \end{aligned} \tag{3.21}$$

This is also equal to the matrix $(\exp(x \cdot f))^a_b$, and hence we recover the standard result that $L_a = (\mathrm{Ad}\, g)_* R_a$.

Using the canonical coordinates $x^a$ and the left momenta $l_a$, we see that the basic Poisson brackets on $T^*SO(3)$ are:

$$\begin{aligned} \{x^a, x^b\} &= 0 , \\ \{x^a, l_b\} &= -R^a_b , \\ \{l_a, l_b\} &= -f^c_{ab} l_c . \end{aligned} \tag{3.22}$$

Using these fundamental brackets the Poisson brackets of the various momentum variables with the Euler angles can be derived since we have the following relations between the various coordinate systems on $SO(3)$:

$$\begin{aligned} \alpha &= -\arctan\frac{x_2}{x_1} - \arctan\left(\frac{x_3}{x}\tan\frac{x}{2}\right) , \\ \beta &= 2\arcsin\left(\frac{\sqrt{x_1^2 + x_2^2}}{x}\sin\frac{x}{2}\right) , \\ \gamma &= \arctan\frac{x_2}{x_1} - \arctan\left(\frac{x_3}{x}\tan\frac{x}{2}\right) . \end{aligned} \tag{3.23}$$

Using these equations, along with (3.22) and (3.18)–(3.20), one can confirm that $\{\gamma, r_3\} = 1$, $\{\alpha, l_3\} = 1$ and $\{\alpha, r_3\} = \{\beta, r_3\} = \{\beta, l_3\} = \{\gamma, l_3\} = 0$. One can also now show that:

$$\{\alpha, l_1\} = -\cos\alpha \cot\beta, \quad \{\beta, l_1\} = -\sin\alpha, \quad \{\gamma, l_1\} = \frac{\cos\alpha}{\sin\beta} , \tag{3.24}$$

$$\{\alpha, l_2\} = -\sin\alpha \cot\beta, \quad \{\beta, l_2\} = \cos\alpha, \quad \{\gamma, l_2\} = \frac{\sin\alpha}{\sin\beta} . \tag{3.25}$$

Given a gauge fixing condition $f = f(\alpha, \beta, \gamma)$ (we postpone until the final subsection any discussion about what type of gauge fixing terms we should use) we can now follow the analysis presented in Sect. 2.4; hence we add ghost variables and construct the second class charge

$$\mathscr{S} = r_3\eta^1 + f\{f, r_3\}\eta^2 - \{f, r_3\}^{-1}\{\{f, r_3\}, r_3\}\eta^1\eta^2\rho_2 . \tag{3.26}$$

In Sect. 2 the kinematical role of this charge, along with the anti-charge (2.44) or (2.45), in isolating the physical states was discussed at length. Before analysing the consequences of inequivalent gauge fixing on the reduction, we now need to address the important problem of how the physical dynamics is described using these charges, and in particular, how the Hamiltonian (3.9) is related to a physical one. This discussion will also supply sensible conditions to impose on the type of gauge fixing considered.

*3.3. Classical and Quantum Observables.* Classically the observables are identified with the functions on a phase space. Hence, in this constrained system, the physical observables are simply the functions on the true degrees of freedom. Given any function $f$ on the extended phase space (including the ghost variables), its expectation value in a physical state (2.29) will be equivalent to the value of some physical observable; the problem we need to resolve is to determine what class of functions on the extended phase space are describing the same physical observable within this second class approach.

There is, in fact, an equivalence relation we can define on functions such that two functions are equivalent if their value in the physical states (2.29) are equal. Clearly if $f = h + \{\mathscr{S}, v_1\} + \{\mathscr{S}_{\rm anti}, v_2\}$ then $f$ is equivalent to $h$ since $\langle\{\mathscr{S}, v_1\}\rangle_{\omega_{\rm phy}} = \langle v_1\rangle_{K\omega_{\rm phy}} = 0$ for all functions $v_1$ (similarly $\langle\{\mathscr{S}_{\rm anti}, v_2\}\rangle_{\omega_{\rm phy}} = 0$). We thus identify the physical observables with these equivalence classes of functions.

From the analysis presented in Sect. 2.4, it is straightforward to see that this equivalence relation is no more than the usual idea of weak equivalence [9] when one has the constraints $(r_3, f, \eta^i, \rho_i)$. Indeed we have, for example, the identification that

$$r_3 = -\frac{1}{2}\{\mathscr{S}, \rho_1\} - \frac{1}{2}\{\mathscr{S}_{\rm anti}, \eta^1\} \tag{3.27}$$

so that $\langle r_3\rangle_{\omega_{\rm phy}} = 0$, i.e. the constraint $r_3$ is equivalent to the zero function. Given such an equivalence class description of the physical observables, it is natural to look for a representative of the physical observables, on the extended phase space, which preserves the identification (2.29) of the physical states. Rather than doing this for a general observable, let us focus on the specific system at hand.

Upon reduction to the true degrees of freedom, the left momenta will become the generators of the Killing vector fields on the two sphere. Hence, as we have seen, we can associate a classical physical observable with them, and indeed with the whole equivalence class $[l_a]$ of weakly equivalent functions. In this second class formulation we choose the representative $l_a^*$ of the equivalence class $[l_a]$ to be that momenta such that

$$\{l_a^*, \mathscr{S}\} = \{l_a^*, \mathscr{S}'_{\rm anti}\} = 0\,, \tag{3.28}$$

and

$$\{l_a^*, l_b^*\} = -f^c_{ab} l_c^*\,. \tag{3.29}$$

Clearly such an observable will be compatible with the identification (2.29) of the physical states.

One finds that $l_a^*$ is uniquely given, in this gauge, by

$$l_a^* = l_a - \{f, l_a\}\{f, r_3\}^{-1} r_3 - \{\{f, l_a\}\{f, r_3\}^{-1}, r_3\}(\eta^1\rho_2 + \eta^2\rho_2)\,. \tag{3.30}$$

We recognise the non-ghost part of $l_a^*$ as simply the $*$-observable associated with $l_a$ in this second class system [9]. We also note that this representative transforms as it should under the canonical transformation generated by (2.43).

Hence, in this classical description, we take as the representative of the class $[H]$ of physical Hamiltonians

$$H^* = l_a^* l_a^* \,. \tag{3.31}$$

Then from (3.28) we see that $\{H^*, \mathscr{S}\} = \{H^*, \mathscr{S}'_{\mathrm{anti}}\} = 0$.

If we restrict the class of gauge fixing so that $f = f(\gamma)$, then $\{f, l_a\} = \{\gamma, l_a\}\{f, r_3\}$. Hence the ghost term in (3.30) vanishes since $\{\{\gamma, l_a\}, r_3\} = 0$. In this case we will also satisfy the condition $\{l_a^*, \mathscr{S}_{\mathrm{anti}}\} = 0$.

In the quantum theory we now need to follow the same basic strategy adopted above to isolate the physical observables. However, this does not mean that we can directly transcribe the prescription (3.30) into the quantum theory. The reason being that $l_a^*$ is not canonically equivalent to $l_a$, but rather weakly equivalent; which in turn was justified by the equivalence class structure imposed on functions by the second class charges. Now it will still be true that the quantised second class charges will impose an equivalence class structure on operators, but now it is more restrictive. In particular, from the definition of physical states (2.39), we will have $\hat{f}$ equivalent to $\hat{f} + \hat{h}_1[\hat{\mathscr{S}}, \hat{\mathscr{S}}_{\mathrm{anti}}] + \hat{h}_2 N_g$, where $\hat{h}_1$ is any operator which commutes with both $\hat{\mathscr{S}}$ and $\hat{\mathscr{S}}_{\mathrm{anti}}$, $\hat{h}_2$ is an arbitrary ghost number zero operator, and $N_g$ is the ghost number operator $\hat{\eta}^1\hat{\rho}_1 + \hat{\eta}^2\hat{\rho}_2$. A consequence of this is that now the constraint $\hat{r}_3$ is not equivalent to the zero operator since, from (2.49) and using the restriction that $f = f(\gamma)$, we have

$$\hat{r}_3 \Psi_{\mathrm{phy}} = \left( i f \frac{\partial f}{\partial \gamma} - \frac{i}{2}\left(\frac{\partial f}{\partial \gamma}\right)^{-1} \frac{\partial^2 f}{\partial \gamma\, \partial \gamma} \right) \Psi_{\mathrm{phy}} \,, \tag{3.32}$$

which will not, in general, be equal to zero! In particular, if the gauge fixing cannot be chosen to be canonically related to a globally well defined one (as in the non-compact example with $f = q^0$, where the expectation value of $\hat{p}_0$ was zero (2.40)) then there is no reason to expect this to be zero. Thus in the quantum theory the operators $\hat{l}_a$ and $\hat{l}_a^*$ are not necessarily equivalent.

Restricting attention to the left momenta; we want to find an operator equivalent, in the above sense, to $\hat{l}_a$ such to the resulting operator preserves the identification of physical states (2.39). Now we have that, acting on $\Psi_0$,

$$[\hat{l}_a, \hat{\mathscr{S}}] = \{l_a, \gamma\} \left( \frac{i}{2} \{\{f, r_3\}^{-1} \{\{f, r_3\}, r_3\}, r_3\} \eta^1 + \{f\{f, r_3\}, r_3\} \hat{\eta}^2 \right), \tag{3.33}$$

where we use the notation that $\{l_a, \gamma\}$, etc, denotes the multiplicative operator corresponding to the configuration space function $\{l_a, \gamma\}$. If the gauge fixing condition is chosen such that

$$\frac{1}{2}\{f, r_3\}^{-1}\{\{f, r_3\}, r_3 = f\{f, r_3\} - C \,, \tag{3.34}$$

where $C$ is a constant, then

$$[\hat{l}_a, \hat{\mathscr{S}}] = \{l_a, \gamma\}\{f\{f, r_3\}, r_3\}(i\eta^1 + \hat{\eta}^2) \,, \tag{3.35}$$

which is zero on $\Psi_{\mathrm{phy}}$ since from Eq. (2.30) we see that the ghost terms are just $iK$.

Hence, for the class of gauges which satisfy (3.34), we find that the operators $\hat{l}_a$ are physical observables, and hence so will be $\hat{H} = \hat{l}_a \hat{l}_a$. We also note that if (3.34) holds then from (3.32) we have

$$\hat{r}_3 \Psi_{\mathrm{phy}} = iC \Psi_{\mathrm{phy}} , \tag{3.36}$$

that is, the physical states are eigenstates of the momentum constraint. On $L^2(SO(3))$ the possible eigenstates of $\hat{r}_3$ are parameterised by the integers. Thus it would be interesting to see if we can find gauge fixing terms such that (3.34) can be solved with $C = in$, for all integers $n$.

*3.4. The Superselection Sectors for $S^2$.* Having set up all the machinery for quantising this constrained system we now need to address the thorny issue of what type of gauge fixing we should use. The first problem we have to face is that the ability to construct good gauge fixing terms is a *global* issue, however, most of the detailed constructions presented in this section have been *local* in character: that is, we are either using canonical coordinates or Euler angles – both of which are only defined locally on $SO(3)$. In terms of these coordinates it would appear that this system is only slightly more complicated than the compact example discussed in Sect. 2. So following the discussion after Eq. (2.52), we would be tempted to use the unambiguously good gauge fixing condition $f = e^{i\gamma} - 1$, which is continuous function of $\gamma$ and has the unique solution $\gamma = 0$ – what more could one ask of a gauge fixing term?

The problem is that this function cannot be extended to a continuous function on the whole of $SO(3)$ while still preserving these properties – this is the Gribov problem.

There seems to be two possible strategies we could follow now. The first is that we only work with functions that are globally well defined, and thus make manifest the fact that there are no good gauge fixing terms to be singled out; the problem with this, though, is that they will in general not be simply functions of $\gamma$ in our local description. Alternatively, and the approach we shall follow, we can just work with the local description in terms of, say, the Euler angles; but take into account the global results by realising that the gauge fixing $e^{i\gamma} - 1$ is no better than $e^{-in\gamma} - 1$, for any non-zero integer $n$.

These two points of view are not totally disconnected. Indeed we can think of these simple functions of $\gamma$ as being equivalent (in the sense implied by the canonical transformation generated by (2.44)) to the local expression for globally well defined gauge fixing functions which have $n$ distinct solutions in the region of $SO(3)$ described by the Euler angles.

Classically gauge fixing with functions of the form $e^{-in\gamma} - 1$ will have the effect of introducing multiple copies of the true degrees of freedom, which is inconvenient but in no way disastrous. In each such copy we still have $r_3 = 0$ and hence the dynamics is unaltered. It is also clear that one cannot change the number of copies by a canonical transformation. Hence the distinct, canonically inequivalent, gauge fixing terms will be characterised by the integer $n$ introduced above.

For the quantum theory we see that these inequivalent gauge fixing classes will lead to inequivalent reductions and hence superselection sectors. Indeed, from (3.34) we need to solve the nonlinear ordinary differential equation

$$\frac{d^2 f}{d\gamma^2} - 2f \frac{df}{d\gamma}\frac{df}{d\gamma} + 2C \frac{df}{d\gamma} = 0 . \tag{3.37}$$

Setting $\frac{df}{d\gamma} = e^{-2C\gamma} W$ we find that, in order to solve (3.37), we must have $W \propto e^{f^2}$. Hence, in terms of the representative functions discussed above, we must have $C = in$ and $\int e^{-f^2} df = e^{-i2n\gamma} - 1$. That is $\mathrm{Erf}(f) = e^{-i2n\gamma} - 1$, where $\mathrm{Erf}(f)$ is the error function of $f$. So, for the simplest form of the physical states, we need to choose the gauge fixing function in the $n$th sector to be the function

$$f = \mathrm{Erf}^{-1}(e^{-i2n\gamma} - 1) . \tag{3.38}$$

We note that because we are on $SO(3)$, we cannot use gauge fixing terms of the form $f = \mathrm{Erf}^{-1}(e^{-in\gamma} - 1)$ since that would imply that $iC$ could be an half integer, which could not then be a solution to (3.36). On $SU(2)$, though, such a possibility is allowed and we get a larger class of allowed gauge fixing terms.

Hence, in this gauge, the physical states in the $n$th sector are given by wave functions of the form

$$\Psi^n_{\text{phy}} = e^{in\gamma}\psi_0(\alpha, \beta, 0)(1 - \eta^1\rho_2) . \tag{3.39}$$

As is well known, such states for different $n$ are unitarily inequivalent; reflecting the inequivalence of the gauge fixing used to define them.

From the construction of the naive quantisation (3.11), and using the relations (3.24)–(3.25), the left momentum can be represented on the states $\Psi^n_{\text{phy}}$ by the operators

$$\begin{aligned}
\hat{l}^n_1 &= i\cos\alpha\cot\beta\frac{\partial}{\partial\alpha} + i\sin\alpha\frac{\partial}{\partial\beta} + n\frac{\cos\alpha}{\sin\beta} , \\
\hat{l}^n_2 &= i\sin\alpha\cot\beta\frac{\partial}{\partial\alpha} - i\cos\alpha\frac{\partial}{\partial\beta} + n\frac{\sin\alpha}{\sin\beta} , \\
\hat{l}^n_3 &= -i\frac{\partial}{\partial\alpha} .
\end{aligned} \tag{3.40}$$

The Hamiltonian in the different sectors is thus given by

$$\begin{aligned}
\hat{H}^n &= \hat{l}^n_a\hat{l}^n_a, \\
&= -\csc^2\beta\left(\frac{\partial^2}{\partial\alpha^2} - 2in\cos\beta\frac{\partial}{\partial\alpha} - n^2\right) - \frac{\partial^2}{\partial\beta^2} - \cot\beta\frac{\partial}{\partial\beta} .
\end{aligned} \tag{3.41}$$

We note that although $n = 0$ is not strictly allowed within this second class approach, it is a valid choice since then we are essentially not using a gauge fixing and the second class charge has reduced to the BRST-charge. Hence we recover the trivial sector of the theory.

To get a better insight into the structure of the different reductions we follow the usual route [18] of writing the wave functions (3.39), defined as wave functions on a superspace, as wave functions on the two sphere. Using spherical coordinates $(\phi, \theta)$ on $S^2$; we thus need a unitary mapping from the states (3.38) to the wave functions $\psi(\phi, \theta)$ in $L^2(S^2)$. To facilitate this we recall [19] that there is a canonical measurable section from $S^2$ to $SO(3)$ given by setting $x_3 = 0$. Then $x_1$ and $x_2$ will become the normal coordinates on $S^2$, i.e.,

$$x_1 = -\theta\sin\phi, \quad x_2 = \theta\cos\phi . \tag{3.42}$$

From (3.23) we see that this is equivalent to setting $\alpha = \phi$, $\beta = \theta$ and $\gamma = -\phi$. So we see that this is precisely the section $s_+$ used in [18].

Trivially extending the argument presented in [18], we now define the unitary mapping $T$, from the states $\Psi^n_{\text{phy}}$ to $L^2(S^2)$, by

$$(T\Psi^n_{\text{phy}})(\phi, \theta) = 4\pi\psi^n_0(\phi, \theta, -\phi) , \tag{3.43}$$

where $\Psi^n_{\text{phy}} = (1 - \eta^1\rho_2)\psi^n_0$. The inverse to this is simply given by

$$(T^*\psi)(\alpha, \beta, \gamma, \eta^1, \rho_2) = \frac{1}{4\pi} e^{in(\alpha+\gamma)}\psi(\alpha, \beta)(1 - \eta^1\rho_2) . \tag{3.44}$$

Under this mapping the momentum operators $\hat{l}_a$ become (under the substitution $\alpha \to \phi$, $\beta \to \theta$, $\partial/\partial\alpha \to \partial/\partial\phi + in$ and $\partial/\partial\beta \to \partial/\partial\theta$)

$$\hat{\ell}^n_1 = i\cos\alpha\cot\theta\frac{\partial}{\partial\phi} + i\sin\phi\frac{\partial}{\partial\theta} + n\frac{\cos\phi}{\sin\theta}(1 - \cos\theta) ,$$

$$\hat{\ell}^n_2 = i\sin\phi\cot\theta\frac{\partial}{\partial\phi} - i\cos\phi\frac{\partial}{\partial\theta} + n\frac{\sin\phi}{\sin\theta}(1 - \cos\theta) ,$$

$$\hat{\ell}^n_3 = -i\frac{\partial}{\partial\phi} + n . \tag{3.45}$$

Introducing polar coordinates $y_1 = \sin\theta\cos\phi$, $y_2 = \sin\theta\sin\phi$ and $y_3 = \cos\theta$, then we recognise these operators to be the expression for the angular momentum operators

$$\hat{\ell}^n = \mathbf{y} \wedge (\hat{\mathbf{p}} + n\mathbf{A}) + \mathbf{y} , \tag{3.46}$$

where $\hat{p}_a = -i\partial/\partial y^a$ is the usual momentum operator and the vector potential $\mathbf{A}$ is given by

$$\mathbf{A} = \frac{(1 - \cos\theta)}{\sin\theta}\hat{e}_\phi . \tag{3.47}$$

These are precisely the expressions for the angular momentum when a Dirac monopole of charge $n$ is at the centre of the two sphere of unit radius [20]. From the discussion following Eq. (3.38) we see that if we had used the identification $S^2 = SU(2)/U(1)$, then we would recover the Dirac monopole coupling, but now with charge $n/2$.

In summary we have seen in this model that the Gribov problem in fixing a gauge leads to inequivalent reductions, and hence quantisations on $S^2$. Infinitesimal changes in the gauge fixing term used lead to unitarily related quantum theories using the results of (2.43). These reductions precisely agree with the results obtained by directly quantising on the two-sphere using either group theoretic [21, 19], $C^*$-algebraic [18] or geometric [7] techniques. As discussed in the introduction, such sectors can also be *described* within a constrained formalism that allows for the modification of the constraints by hand [10]: however, such an approach does not give a constrained *account* of why such sectors arise. Thus we see that the second class approach initiated in this paper is a successful constrained formalism.

## 4. Conclusions

One of the most surprising aspects of the process of quantisation is the possibility of having superselection sectors. In this paper it has been shown that by introducing a constrained formulation of the theory, these inequivalent quantisations can be understood as arising from the difficulties encountered in constructing an unambiguous gauge fixing term. So at least for the model considered, the Gribov ambiguity is seen to be responsible for the rich class of possible quantisations.

The bulk of the agreement presented here, and in some sense the hidden agenda behind the presentation, has been to show that gauge fixing is essential in an operator approach to constrained systems: even for those with no Gribov problem. This focus on the operator methods is due to the fact the almost all accounts of the emergence of superselection sectors are described within this language [7, 8, 18, 19, 21]. This should be contrasted with the path integral approach to constrained systems where the need for gauge fixing is manifest [22], yet the possibility of superselection sectors is never discussed beyond the possible $\pi_1$-effects. However, even there one must take into account the fact that inequivalent gauge fixing terms will lead to different theories. A full account of how these superselection sectors can emerge from a path integral analysis is an interesting topic and will be presented elsewhere.

Although this paper has just concentrated on a simple model, it is still instructive to reflect on how the results of this exercise could be interpreted in the wider, field theoretic, context. If one accepts the possibility that there are superselection sectors (i.e., that nature does not conspire to only use the naive quantisation of a system) then this example has shown that the Gribov ambiguity is not simply a technical annoyance, but rather an essential ingredient for a full description of the theory. Also knowledge of the actual sector you are in is important for perturbation theory in order to determine what the vacuum structure of the theory is. (Is there a monopole or not?) One can only speculate as to the relevance of superselection sectors to problems like confinement in QCD. However the structures suggested here are to be seen in addition to the consequences of having the Gribov horizon [6]. Clearly a full understanding of the structure of QCD must take into account all the relevant possibilities available.

*Acknowledgements.* I wish to thank Martin Lavelle and especially Izumi Tsutsui for many helpful conversations.

## References

1. Jackiw, R., Rebbi, C.: Phys. Rev. Lett. **37**, 172 (1976)
2. Jackiw, R.: In Relativity, Groups and Topology II. Les Houches 1983, Dewitt, B.S., Stora, R. (eds.) Amsterdam: North–Holland 1984
3. Aharanov, Y., Bohm, D.: Phys. Rev. **115**, 485 (1959)
4. Gribov, V.N.: Nucl. Phys. **B139**, 1 (1978)
5. Singer, I.: Commun. Math. Phys. **60**, 7 (1978)
6. Zwanziger, D.: Nucl. Phys. **B321**, 591 (1989); Nucl. Phys. **B323**, 513 (1989)
7. Woodhouse, N.: Geometric Quantization. Oxford: Clarendon Press 1980
8. Mackey, G.W.: Induced Representations. New York: Benjamin 1968
9. Dirac, P.A.M.: Lectures on Quantum Mechanics. New York: Yeshiva 1964
10. McMullan, D., Tsutsui, I.: BPST instanton and spin from inequivalent quantizations. Phys. Lett. B (to appear)

11. Landsman, N.P., Linden, N.: Nucl. Phys. **B371**, 415 (1992)
12. McMullan, D.: Commun. Math. Phys. **149**, 161 (1992)
13. McMullan, D.: Nucl. Phys. **B363**, 451 (1991)
14. van Holten, J.W.: Nucl. Phys. **B339**, 158 (1990)
15. McMullan, D., Paterson, J.: Phys. Lett. **B202**, 358 (1988); J. Math. Phys. **30**, 477, 487 (1989)
16. Abraham, R., Marsden, J.E.: Foundations of Mechanics (2nd. ed.). Reading, MA: Benjamin/Cummings 1978
17. Choquet-Bruhat, Y., DeWitt-Morette, C., Deillard-Bleick, M.: Analysis, manifolds and physics (revised edition). Amsterdam: North-Holland 1982
18. Landsman, N.P.: Rev. Math. Phys. **2**, 45, 73 (1991)
19. Landsman, N.P., Linden, N.: Nucl. Phys. **B365**, 121 (1991)
20. Biedenharn, L.C., Louck, J.D.: The Racah–Wigner algebra in quantum physics. Encycl. Math. Appl. Vol. 9, Reading, MA: Addison-Wesley 1991
21. Isham, C.J.: In: Relativity, Groups and Topology II. Les Houches 1983, Dewitt, B.S. Stora, R. Amsterdam: North-Holland 1984
22. Faddeev, L.D.: In: Methods in Fields Theory. Les Houches 1975, Balian, R., Zinn-Justin, J. (eds.) Amsterdam: North-Holland 1976

Communicated by G. Felder