On the implementation of constraints through projection operators



2001 J. Phys. A: Math. Gen. 34 1019

(http://iopscience.iop.org/0305-4470/34/5/307)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **34** (2001) 1019–1036 www.iop.org/Journals/ja PII: S0305-4470(01)17489-5

# On the implementation of constraints through projection operators

**Achim Kempf and John R Klauder**

Institute for Fundamental Theory, Departments of Physics and Mathematics, University of Florida, Gainesville, FL 32611, USA

E-mail: kempf@phys.ufl.edu and klauder@phys.ufl.edu

Received 22 September 2000, in final form 6 December 2000

**Abstract**

Quantum constraints of the type $Q|\psi_{phys}\rangle = 0$ can be straightforwardly implemented in cases where $Q$ is a self-adjoint operator for which zero is an eigenvalue. In that case, the physical Hilbert space is obtained by projecting onto the kernel of $Q$, i.e. $H_{phys} = \ker Q = \ker Q^*$. It is, however, non-trivial to identify and project onto $H_{phys}$ when zero is not in the point spectrum but instead is in the continuous spectrum of $Q$, because then $\ker Q = \emptyset$.

Here, we observe that the topology of the underlying Hilbert space can be harmlessly modified, namely, loosely speaking, in the direction perpendicular to the constraint surface. Consequently, $Q$ becomes non-self-adjoint, which then allows us to conveniently obtain $H_{phys}$ as the proper Hilbert subspace $H_{phys} = \ker Q^*$ on which one can project as usual. In the simplest case, the necessary change of topology amounts to passing from an $L^2$ Hilbert space to a Sobolev space.

PACS numbers: 0370, 0240, 0250N, 0365

## 1. Introduction

Numerous classical dynamical systems are distinguished by the presence of constraints, which, in a phase space formulation, act to restrict the system to the *constraint hypersurface*, a submanifold of the original phase space with a positive co-dimension. In practice, this restriction is accomplished by one or more real constraint functions, $\phi_\alpha(p, q) = 0$, $\alpha = 1, 2, \ldots, A$, $A < \infty$, which are non-dynamical (i.e. no time derivatives) and serve to constrain the system to the constraint hypersurface. When dealing with concrete examples it is useful to further categorize a system of constraints (along with their associated Hamiltonian) into classes (first and second), as well as other subdivisions (closed, open, irreducible, reducible, regular, irregular, etc); these categories are well described in the literature and are not reviewed here [1]. For the purposes of the present paper, it is sufficient to focus on the several constraint functions themselves, and we need not be too concerned about any specific subclassification of the set of constraints.

  

Our principal interest lies in quantization, or more particularly, with the quantum theory of constraints. We assume that as quantum operators, the constraints $\Phi_\alpha$, $\alpha = 1, \ldots, A$, are represented by self-adjoint operators determined in some fashion from the classical constraint functions by some consistent but unspecified quantization procedure. Just as the classical constraints act to restrict the system to a subset of the original classical phase space, it is the role of the quantum constraints, in like manner, to restrict the system to a subset of the original quantum mechanical phase space. Since the 'quantum mechanical phase space' is a Hilbert space, such a restriction is ideally imposed by the several constraint conditions $\Phi_\alpha\,|\psi\rangle_{phys} = 0$, $\alpha = 1, \ldots, A$. As a linear equation, it follows that the vectors $|\psi\rangle_{phys}$ form a linear space, and since the constraint operators are self-adjoint, the given space is closed, hence a subspace $H_{phys}$, the presumed *physical Hilbert space*, in the original Hilbert space, $\mathcal{H}$. For evident reasons we focus attention on those cases where $\dim(H_{phys}) > 0$.

The foregoing scenario does indeed hold for certain families of constraints, namely, in cases where the constraint operators $\Phi_\alpha$ each have a set of simultaneous eigenvectors with eigenvalue zero. When that is the case, we may also consider the *single* constraint operator $X^2 \equiv \sum_{\alpha=1}^{A} \Phi_\alpha^2$, regarded as a self-adjoint operator, and observe that $X^2\,|\psi\rangle_{phys} = 0$ is completely equivalent to the several equations $\Phi_\alpha\,|\psi\rangle_{phys} = 0$, $\alpha = 1, \ldots, A$. An example of this kind of constraint situation is given by $\Phi_\alpha = J_\alpha$, $\alpha = 1, 2, 3$, where the operators $J_\alpha$ satisfy the Lie algebra for $SO(3)$ (or $SU(2)$) and the condition $\sum J_\alpha^2\,|\psi\rangle_{phys} = 0$ corresponds to a restriction to spherically symmetric states. We also observe that we can also set $H_{phys} \equiv \mathbb{E}H$, where $\mathbb{E} = \mathbb{E}^* = \mathbb{E}^2$ is a *projection operator*, which in the present case is $\mathbb{E} = \mathbb{E}(\sum J_\alpha^2 = 0)$. For comparison purposes with what is to come we also note, equivalently, that

$$\mathbb{E} = \mathbb{E}\Big(\sum J_\alpha^2 \leqslant c\hbar^2\Big) = \mathbb{E}\Big(\sum J_\alpha^2 = 0\Big) \tag{1}$$

where $c \leqslant 2$ for $SO(3)$ (or $0 \leqslant c < \frac{3}{4}$ for $SU(2)$). The discussion and example of the present paragraph refer to the ideal situation regarding the constraint operators.

More generally, the constraint operators do *not* fulfil the ideal (Dirac) criteria given above. In fact, it frequently happens that the set of constraint operators have no non-zero eigenvector with eigenvalue zero. This situation may arise in two fundamentally different ways. One of these ways refers to cases where $X^2 \equiv \sum_{\alpha=1}^{A} \Phi_\alpha^2$ has a discrete spectrum (in the vicinity of zero) which *does not include zero*. An example of this situation is given by the two classical constraints $\phi_1 = p = 0$ and $\phi_2 = q = 0$ (for the same degree of freedom), which then become the quantum constraints $\Phi_1 = P$ and $\Phi_2 = Q$, two operators which satisfy the Heisenberg commutation relation $[Q, P] = \mathrm{i}$ (with $\hbar = 1$). The ideal equations, $P\,|\psi\rangle_{phys} = 0$ and $Q\,|\psi\rangle_{phys} = 0$, imply (modulo domain issues) that $[Q, P]\,|\psi\rangle_{phys} = \mathrm{i}\,|\psi\rangle_{phys} = 0$, namely, that $|\psi\rangle_{phys} = 0$ and thus $H_{phys} = \emptyset$, which is unacceptable. (This is a typical case of second-class constraints.) To avoid this situation, we replace the ideal conditions by the choice $H_{phys} \equiv \mathbb{E}H$, where in the present case $\mathbb{E} = \mathbb{E}(P^2 + Q^2 \leqslant \hbar) = |0\rangle\langle 0|$, the projection operator onto the harmonic oscillator ground state, thus leading to a one-dimensional $H_{phys}$. More generally, we accommodate this kind of situation by the criterion that $H_{phys} = \mathbb{E}H$, where

$$\mathbb{E} = \mathbb{E}(X^2 \leqslant \delta(\hbar)^2) \tag{2}$$

and $\delta(\hbar)$ is an $\hbar$-dependent regularization parameter to be chosen on a case-by-case basis. The choice of a quadratic combination of constraints is not written in stone, and a discussion of alternative choices is presented elsewhere [2].

The second manner in which the idealized situation may fail arises when the operator $X^2 \equiv \sum_{\alpha=1}^{A} \Phi_\alpha^2$ has its zero in the *continuous spectrum*. An example of this situation is given by the single classical constraint $\phi = q = 0$, which then becomes the quantum constraint

$\Phi = Q$. Other examples could arise from two degrees of freedom for which, in an obvious notation, e.g. (a) $\Phi_1 = Q_1$, $\Phi_2 = Q_2$, or (b) $\Phi_1 = Q_1$, $\Phi_2 = P_2$, or (c) $\Phi = Q_1^2 + Q_2^2 - 1$, etc. In these cases $X^2 \equiv \sum_{\alpha=1}^{A} \Phi_\alpha^2$ has its zero in the continuum. Consequently, $\mathbb{E}(X^2 = 0) \equiv 0$, which we deem to be unacceptable. In place of this idealized condition we once again choose

$$\mathbb{E} = \mathbb{E}(X^2 \leqslant \delta(\hbar)^2) \tag{3}$$

where $\delta(\hbar) > 0$. Observe, in the present case, that for any $\delta > 0$, it follows that $H_{phys} = \mathbb{E}H$ is an *infinite-dimensional* (regularized) physical Hilbert space. Although it is actually possible to work with this regularized space, say in cases where $\delta$ is extremely small, for example, $\delta = 10^{-1000}$, in some natural units, it is analytically preferable if we are able to take the limit $\delta \to 0$. However, this limit cannot be taken in any straightforward fashion since

$$\lim_{\delta\to 0} \langle\lambda|\mathbb{E}(X^2 \leqslant \delta^2)|\phi\rangle \equiv 0 \tag{4}$$

for any pair $|\lambda\rangle$, $|\phi\rangle \in H$, whenever the zero of $X^2$ lies in the continuum.

The problem of quantum constraints with their zero in the continuum is well known and has often been studied in the literature. Briefly summarized, such problems have been studied, for example, by

(a) introducing Gel'fand triplets [3];
(b) using an operator-valued inner product [4];
(c) specialized algebraic representations [5]; or
(d) rescaled limits within suitable subspaces [6].

In this context, see also [7] and the classic works by Dirac [8].

As one version of type (d) above, we imagine working in a *representation* in which $X$ is diagonalized (as $x$), and therefore

$$\langle\lambda|\mathbb{E}(X^2 \leqslant \delta^2)|\phi\rangle = \int\!\!\int_{-\delta}^{\delta} \lambda^*(x, y)\, \phi(x, y)\, \mathrm{d}x\, \mathrm{d}\sigma(y) \tag{5}$$

where the variable $y$ corresponds to any degeneracy. In this form it is clear, as $\delta \to 0$, that the right-hand side vanishes. To overcome this situation, let us first *restrict* attention to a subset of functions, e.g.

$$\mathcal{D} \equiv \{\chi(x, y) : \ \chi(x, y) \in \{\text{polynomial}(x, y)\, \mathrm{e}^{-(x^2+y^2)}\}\} \tag{6}$$

and consider $\lambda = \lambda_0 \in \mathcal{D}$ and $\phi = \phi_0 \in \mathcal{D}$, so that

$$\langle\lambda_0|\mathbb{E}(X^2 \leqslant \delta^2)|\phi_0\rangle \equiv \int\!\!\int_{-\delta}^{\delta} \lambda_0^*(x, y)\, \phi_0(x, y)\, \mathrm{d}x\, \mathrm{d}\sigma(y). \tag{7}$$

In the present case it follows that

$$\begin{aligned}\lim_{\delta\to 0} (2\delta)^{-1} \langle\lambda_0|\mathbb{E}(X^2 \leqslant \delta^2)|\phi_0\rangle &= \lim_{\delta\to 0} (2\delta)^{-1} \int\!\!\int_{-\delta}^{\delta} \lambda_0^*(x, y)\, \phi_0(x, y)\, \mathrm{d}x\, \mathrm{d}\sigma(y)\\ &= \int \lambda_0^*(0, y)\, \phi_0(0, y)\, \mathrm{d}\sigma(y)\\ &\equiv (\lambda_0, \phi_0).\end{aligned}$$

This final expression defines a sesqui-linear form characterizing a pre-Hilbert space. Completion of the pre-Hilbert space in the usual fashion (i.e. inclusion of limits of Cauchy sequences in the inner-product-induced norm $\|\phi_0\| \equiv \sqrt{(\phi_0, \phi_0)}$, plus the identification of elements as equivalence classes of functions as necessary) leads to the *true* physical Hilbert space, $H_{phys}$, in which the constraint condition $X^2 = 0$ is finally satisfied.

In the previous discussion it is noteworthy that we chose a *representation* for the constraint operators in order to achieve a successful rescaling and limit as the regularization was removed. Other schemes may avoid the regularization and its subsequent removal, but they all require a *representation* to be introduced.

A natural question then arises as to whether or not it is possible to devise a procedure to enforce a quantum constraint whose zero lies in the continuum in an *abstract fashion*, i.e. is it possible to impose such a constraint *without* introducing any specific representation whatsoever? It is this question that we address in the remainder of this paper, and it is noteworthy that the answer to that question is in the affirmative.

The key to obtaining this affirmative answer is worth noting. The traditional view of obtaining a physical Hilbert subspace in the case of zero in the continuous spectrum is to *change the Hilbert space*, essentially, by imposing the constraint as a distribution. The new method introduced in this paper *changes the underlying topology* of the existing (pre-)Hilbert space and then is able to impose the constraint with a zero in the continuum in the same way one imposes a constraint with a discrete zero!

In this section, we have used the symbols $P$ and $Q$ in their traditional Heisenberg and Schrödinger senses. Note well that the same symbols $P$ and $Q$ appear in the subsequent sections, but they are used in a much wider sense. In particular, $Q$ is used as a *generic constraint operator with its zero in the continuum*, while $P$ is used as a *maximally symmetric operator* satisfying $[Q, P] = \mathrm{i}\mathbb{1}$ on a suitable domain *which need not even be dense*.

## 2. The continuous spectrum and topology

Let us begin with the observation that, for topological reasons, in the case of zero being in the continuous spectrum of $Q$, there cannot exist a projection from the original Hilbert space $H$ to a physical Hilbert subspace $H_{phys}$.

To see this, let us consider the simplest case, namely where the spectrum of the constraint operator $Q$ is non-degenerate, and is given by an interval, say $I = [0, \infty)$. In the spectral representation of $Q$, the original Hilbert space $H$ is therefore the space $H = L^2(I)$ of square-integrable functions $f \in H$ over the interval $I$.

We would like to find a physical Hilbert subspace $H_{phys}$ corresponding to '$Q = 0$'. It should be a one-dimensional vector space of 'function values at 0'. We could then identify this space with $\mathbb{C}$. To this end, let us consider the linear functional $\psi$ defined by the property that it maps every continuous function $f \in H$ onto its value at zero, i.e.

$$\psi : f \to f(0). \tag{8}$$

Intuitively, one might assume that $\psi$ is the desired projection onto a one-dimensional 'physical' subspace of function values at zero.

However, $\psi$ is not a projection. The problem is that, even though we restricted the domain of $\psi$ to continuous representatives of square-integrable functions, $\psi$ is not a continuous map: consider, for example, the sequence $\{g_n(x)\}_{n\in\mathbb{N}}$ of continuous square-integrable functions

$$g_n(x) := A\,\mathrm{e}^{-nx^2} \tag{9}$$

for an arbitrary non-zero constant $A$. In the topology of the Hilbert space $H$, namely the norm topology induced by the $L^2$ scalar product of $H$, the sequence clearly converges to the null vector in $H$, i.e. to the function $g(x) \equiv 0$. Namely, $\lim_{n\to\infty} g_n = g$ in the sense that $\lim_{n\to\infty} \|g - g_n\|^2 = \lim_{n\to\infty} \int_{-\infty}^{\infty} \mathrm{d}x\, A^* A \exp(-2nx^2) = 0$. On the other hand, $\lim_{n\to\infty} \psi(g_n) = A$, which proves the discontinuity.

This means the following: we can take any continuous square-integrable function $f \in H$ and add a null sequence such as $g_n(x)$. Then, in the limit we recover the vector $f \in H$. However, crucially, as far as the original topology of the Hilbert space $H$ is concerned, we can in this way give $\psi(f)$ any arbitrary value that we wish, i.e. $\psi$ is ill-defined.

Intuitively, the underlying reason for the ill-definedness of $\psi$ is of course that even continuous functions $f \in H$ can be arbitrarily sharply peaked over a point, say zero, meaning that functions $f \in H$ could possess arbitrarily different values at zero, while their $L^2$-distance is arbitrarily small, i.e. while such functions can still be arbitrarily close within the topology of $H$.

Thus, the situation is that $\psi$ is non-continuous and therefore also unbounded. Since projectors are of course bounded, $\psi$ is not a projection! Another way to see this is that by Riesz' theorem, $\psi$, being non-continuous, is not contained in (the dual of) $H$, and can therefore not be projected onto.

The case of real physical interest is of course where the spectrum of $Q$ is degenerate at zero, and where one therefore expects the physical Hilbert space to be higher dimensional. However, it is clear that the fundamental problem, namely that $\psi$ is not a projector, persists to all non-trivial cases—as long as we stick to the original topology of $H$. Our aim is therefore to suitably modify the topology of the original Hilbert space in order to be able to project onto the physical subspace.

## 3. Hints from the theory of distributions and Sobolev spaces

Before we describe our solution to the general problem, let us recall some basic facts from the theory of distributions and Sobolev spaces, facts which can be used directly to solve the problem for the simple non-degenerate example which we just discussed. Our general method will be motivated by this example.

We begin by recalling how, in spite of our arguments above, the map $\psi$ can be made into a continuous and bounded map, in fact, into a projection. The price to be paid is that the *topology* of the original Hilbert space must be changed. Intuitively, the topology of the function space needs to be changed in such a manner that functions which differ only by sharp peaks (which make arbitrarily little difference in the $L^2$ topology) are separated in the new topology.

The basic observation underlying the theory of Sobolev spaces is that a sufficiently improved topology is induced by a simple new scalar product on the function space (see, e.g., [9]). The new scalar product is arranged to be sensitive to the rate of change of functions. In the norm topology induced by this scalar product even those functions can be separated which differ only by spikes that are so sharp that the $L^2$ topology cannot distinguish them (e.g. those differing by sequences such as the sequence $\{g_n(x)\}$ of equation (9)).

In particular, while the scalar product of $H$ is of course

$$\langle f_1 | f_2 \rangle = \int_I \mathrm{d}x\, f_1^*(x) f_2(x) \tag{10}$$

the scalar product of the Sobolev Hilbert space $H^1$ is defined, using distributional derivatives, as

$$(f_1 | f_2) = \int_I \mathrm{d}x \left[ f_1^*(x) f_2(x) + \frac{\mathrm{d}f_1^*(x)}{\mathrm{d}x} \frac{\mathrm{d}f_2(x)}{\mathrm{d}x} \right]. \tag{11}$$

Note that throughout we will use the notation $\langle\, | \,\rangle$ for the scalar product in the initial Hilbert space and $(\, | \,)$ for the new scalar product.

Let us check whether the functional $\psi$ is a continuous functional over $H^1$: to this end, we recall, by Riesz' representation theorem, that all continuous linear functionals over a Hilbert space can be identified with the Hilbert space vectors themselves, via the scalar product action. Thus, if $\psi$ is indeed a continuous functional over $H^1$, then we should now be able to identify a representation of $\psi$ as a vector in $H^1$! In fact,

$$\psi(x) = \mathrm{e}^{-x} \tag{12}$$

is the representation of $\psi$ in $H^1$. It is normalized with respect to the scalar product, equation (11), of $H^1$. The reader may check, using a brief calculation, that indeed $(\psi, f) = f(0)$ for all $f \in H^1$. Thus, $\psi$, which maps continuous functions onto their values at a point, or more precisely

$$\Pi := |\psi) \otimes (\psi| \tag{13}$$

becomes a projection, as desired. It is also clear that this could not have been achieved without a change in topology.

**Remark: higher Sobolev spaces.** We should also mention that our particular choice of change of topology—to move from $H$ to $H^1$—is not the only possible one to achieve our goal of making $\psi$ a projection: to see this, let us be more precise about the properties of the functions in $H^1$ as opposed to those in $H$. For a function $f(x)$ to be in $H^1$ it must not only be square integrable: $f(x)$ must also be the indefinite integral of some square-integrable function '$\mathrm{d}f(x)/\mathrm{d}x$'. We recall that a function $f(x)$ is an indefinite integral exactly if it is absolutely continuous (on any countable set of non-overlapping intervals). Thus, moving from the Hilbert space $H = L^2$ to the Hilbert space $H^1$ improves the continuity of the functions $f(x)$ to absolute continuity. However, the derivatives $\mathrm{d}f(x)/\mathrm{d}x$ of functions in $H^1$ need only be in $L^2$, not $H^1$. Thus, in $H^1$ it is still not possible to project $\mathrm{d}f(x)/\mathrm{d}x$ onto its value at a fixed point. In fact, $\mathrm{d}f(x)/\mathrm{d}x$ need only be defined almost everywhere. As here we are only interested in projecting onto the values of functions and not their derivatives this is not a problem for our purposes.

Nevertheless, for completeness, let us recall how the topology could be arranged to also improve the behaviour of the functions' derivatives. As the theory of Sobolev spaces shows, it is sufficient to this end to use higher-derivative operators in the scalar product, i.e. to use, for example, the scalar product $(f_1, f_2) := \int_I \mathrm{d}x \sum_{j=0}^{n} \frac{\mathrm{d}^j f_1^*(x)}{\mathrm{d}x^j} \frac{\mathrm{d}^j f_2(x)}{\mathrm{d}x^j}$ for some $n > 1$ and its induced norm topology. In this way one obtains the higher Sobolev spaces $H^n$. Obviously, $H \supset H^1 \supset H^2 \supset H^3 \ldots$. In the function space $H^n$, we can project the derivatives $\mathrm{d}^{m-1}f/\mathrm{d}x^{m-1}$ of functions $f(x)$, up to $m = n - 1$, onto their values at any given point. It is clear that, while we could use any higher Sobolev space, $H^1$ does suffice for our purposes here.

## 4. The new method

Let us now consider the general problem of projecting onto the physical subspace in the case of zero being in the continuous spectrum of a self-adjoint constraint operator $Q$ whose spectrum may be arbitrarily degenerate. We would like to generalize the procedure used when zero is only in the point and not in the continuous spectrum of the self-adjoint constraint operator $Q$, in which case one can straightforwardly define the physical subspace as the kernel of $Q$, i.e.

$$H_{phys} := \ker Q = \ker Q^*. \tag{14}$$

In the case where zero is only in the continuous and not in the point spectrum of $Q$, the problem is of course that then $\ker Q = \ker Q^* = \emptyset$, which is obviously not the desired subspace.

Our main idea in this paper is therefore to treat this case by modifying the topology of the initial Hilbert space $H$ in a manner analogous to passing from an $L^2$ Hilbert space to the Sobolev space $H^1$. (In the general case, where zero is also in the point spectrum, one may project out the kernel of $Q$ as usual, before applying our procedure.)

The new Hilbert space, which we will call $\tilde{H}$, will be a subspace of the original Hilbert space when considered as a vector space, and the action of all operators, their commutation relations, etc, therefore remains unchanged. Also $Q$ is of course still the same, as a linear map. Crucially, however, $\tilde{H}$ will be different as a Hilbert space, being equipped with a new scalar product. This scalar product changes the induced norm topology so that the $*$ structure will change and the operator $Q$ will no longer be self-adjoint, i.e. so that $Q \neq Q^*$! As our main finding, we will show that this modification enables us to identify the physical subspace as

$$H_{phys} := \ker Q^*. \tag{15}$$

Since $H_{phys}$ is therefore a proper Hilbert subspace, namely $H_{phys} \subset \tilde{H}$, this means that once one has passed to the new topology, i.e. from $H$ to $\tilde{H}$, one can again implement the constraint by projection.

Explicitly, in order to carry out the programme of suitably modifying the topology, we begin by completing the description of the degree of freedom which is to be constrained: namely, in addition to the constraint operator $Q$, let us also consider a variable $P$ which is conjugate to $Q$, i.e. a maximal symmetric operator $P$ which obeys

$$[Q, P] = \mathrm{i}\mathbb{1}. \tag{16}$$

By definition, therefore, the domain of $P$ is the maximal domain $D_P \in H$ on which this commutation relation holds and on which $P$ is symmetric. We note that while, in general, $P$ will not be self-adjoint on $D_P$; in particular, it is never if $Q$ is positive, the property of being maximally symmetric will suffice. As we will see, it will also not matter for our purposes that the domain $D_P \subset H$ on which the commutation relation, equation (16), holds will in general not even be dense in the original Hilbert space $H$—as it clearly will not be if $Q$ also possesses a point spectrum. Intuitively, this is because for our purposes only the part of the spectrum around zero matters.

In analogy with the Sobolev space example for functions, let us now consider the domain $D_{P^*}$ and let us change the topology on it to obtain a new Hilbert space $\tilde{H}$: namely, as a vector space, we define $\tilde{H}$ to be identical to $D_{P^*}$ while we equip $D_{P^*}$ with a new scalar product which then induces a new norm topology. Denoting the scalar product in the original Hilbert space $H$, and thus on $D_{P^*}$, by $\langle\,|\,\rangle$, we define the new scalar product $(\,|\,)$ on $\tilde{H}$ through

$$(v_1|v_2) := \langle v_1|v_2\rangle + \langle P^* v_1|P^* v_2\rangle. \tag{17}$$

In other words, $\tilde{H}$ is the graph Hilbert space of $P^*$. In the simple case where the spectrum of $Q$ is purely continuous and non-degenerate this Hilbert space is the Sobolev space $\tilde{H} = H^1$.

An advantage of our functional analytic definition is that, unlike in our discussion of the simple example of the non-degenerate spectrum and its Sobolev space, we now no longer need to work in the spectral representation of $Q$. Let us denote the domain of $Q$ in $\tilde{H}$ by $\tilde{D}_Q$ (consisting of all vectors $\phi \in \tilde{H}$ for which $Q\phi$ has a finite norm with respect to the new scalar product given in equation (17)). On $\tilde{D}_Q$ the operator $Q$ is no longer self-adjoint, $Q \neq Q^*$ (and not even symmetric). We then define the physical Hilbert subspace $H_{phys} \subset \tilde{H}$ as the kernel of $Q^*$:

$$H_{phys} := \ker Q^*. \tag{18}$$

This definition has a simple interpretation: we require that if we act with $Q$ on any vector in its domain in $\tilde{H}$ then the resulting vector is orthogonal to all physical vectors. In other words, we define the physical subspace $H_{phys}$ as the orthogonal complement of the range of $Q$ in the Hilbert space $\tilde{H}$.

For added clarity, let us be fully precise regarding the definition of $H_{phys}$ through equation (18). As a vector space, $H_{phys}$ is given by

$$H_{phys} := \left\{ |v\rangle \in \tilde{H} = D_{P^*} \mid \forall w \in \tilde{D}_Q : \langle Qw|v\rangle + \langle P^*Qw|P^*v\rangle = 0 \right\} \qquad (19)$$

i.e. all its vectors are also vectors in $H$, when considered as a vector space. The scalar product in $H_{phys}$ is given by $(\,|\,)$ in equation (17).

To summarize, we begin by completing the picture of the degree of freedom which is to be constrained, namely by augmenting $Q$ by a symmetric operator $P$ obeying the commutation relation $[Q, P] = \mathrm{i}\mathbb{1}$ on its maximal domain $D_P$ in the original Hilbert space $H$. Second, we change the scalar product and consequently the induced topology on the domain $D_{P^*} \subset H$ to obtain the graph Hilbert space, $\tilde{H}$, of $P^*$ with the scalar product equation (11). Third, we identify the physical subspace $H_{phys}$ as the proper Hilbert subspace $H_{phys} = \ker Q^*$ of the Hilbert space $\tilde{H}$. Thus, after passing from $H$ to $\tilde{H}$ we can again project onto a physical subspace which is a proper Hilbert subspace.

**Remark: the special case of positive $Q$.** If $Q$ is a positive self-adjoint operator with zero being in the purely continuous spectrum, i.e. when zero is actually a boundary of a piece of the continuous spectrum, then there is additional structure which can be used.

Namely, consider in this case $D_{P^*}$ modulo $D_P$ in the new topology, i.e. consider $\tilde{H}_B := D_{P^*} \ominus D_P$. By von Neumann's theory of self-adjoint extensions of symmetric operators, the domains of $P^*$ and $P$ differ exactly by the space of boundary functionals, and we have the von Neumann formula

$$D_{P^*} \ominus D_P = N^+ \oplus N^- \qquad (20)$$

where $N^\pm$ are the deficiency spaces of $P$. Since we assume that zero is a boundary of the spectrum we can therefore conclude that $H_{phys} \subset N^+ \oplus N^-$. Consequently, we can say that all physical vectors must be a linear combination of vectors in the kernels of either the operator $(P^* + \mathrm{i})$ or $(P^* - \mathrm{i})$. In concrete representations, this fact can yield useful differential equations.

## 5. Examples

Let us illustrate how the new method works with simple examples in non-relativistic quantum mechanics. We will choose constraint operators $Q$ which are functions only of the position operators $Q = Q(\boldsymbol{x})$. In this way, the associated classical constraint surface, namely the zero set of $Q = 0$, read as a classical equation, will be obvious. We will show that our new method projects on the Hilbert space of functions over the constraint manifold.

### *5.1. Example: projecting onto a point on the line*

As a first example, let us reconsider the simple case of a constraint $Q = 0$ for a self-adjoint operator $Q$ whose spectrum is continuous at zero and non-degenerate. The non-degeneracy means of course that we expect the physical Hilbert space to be one dimensional. This example will basically be the same as that in which we discussed Sobolev spaces. In order to make this case slightly non-trivial let us choose the spectrum of $Q$ as

$$I = [0, 1] \cup \{2, 3, \ldots\}. \qquad (21)$$

Thus, in the spectral representation of $Q$, our starting Hilbert space $H$ is $H = L^2(0,1) \oplus l^2$, with the scalar product

$$\langle \phi_1 | \phi_2 \rangle = \int_0^1 \mathrm{d}q\, \phi_1^*(q)\phi_2(q) + \sum_{n=2}^{\infty} \phi_1^*(n)\phi_2(n). \tag{22}$$

We begin by considering a maximally symmetric operator which obeys $[Q, P] = \mathrm{i}\mathbb{1}$, namely $P = -\mathrm{i}\partial_q$ on its domain $D_P$. Functions in $D_P$ are absolutely continuous, square integrable, vanish at the interval boundary, and their almost-everywhere defined derivative is also square integrable. We note that, while $D_P$ is not dense in $H$, since it lacks the eigenspaces to the discrete eigenvalues, this will not matter when we project onto '$Q = 0$'.

We now equip $D_P$ with the new scalar product, equation (11), to obtain $\tilde{H}$:

$$(\phi_1 | \phi_2) = \langle \phi_1 | \phi_2 \rangle + \langle P^*\phi_1 | P^*\phi_2 \rangle \tag{23}$$

$$= \int_0^1 \mathrm{d}q\, \left\{ \phi_1^*(q)\phi_2(q) + \left(\partial_q \phi_1(q)\right)^* \partial_q \phi_2(q) \right\}. \tag{24}$$

Next, we define the (Sobolev) Hilbert space $\tilde{H}$ as the graph Hilbert space of $P^*$, i.e. as the domain of $P^*$, with the scalar product equation (24).

Now we are ready to calculate the physical subspace $H_{phys} = \ker Q^*$. Explicitly, according to equation (19), the condition for vectors $\psi \in \tilde{H}$ to be in the physical domain $H_{phys}$ now reads

$$\int_0^1 \mathrm{d}q\, \left\{ \phi^*(q) q \psi(q) + \left(\partial_q q \phi^*(q)\right) \partial_q \psi(q) \right\} = 0 \qquad \forall \phi \in \tilde{D}_Q. \tag{25}$$

This yields the condition, again for all $\phi \in D_Q$,

$$\int_0^1 \mathrm{d}q\, \left\{ \phi^*(q) q \psi(q) \;-\; q\phi^*(q)\partial_q^2 \psi(q) \right\} + \left[ q\phi^*(q)\partial_q \psi(q) \right]_0^1 = 0 \tag{26}$$

which means that

$$q(\partial_q^2 - 1)\psi(q) = 0 \tag{27}$$

and

$$\partial_q \psi(q)|_{q=1} = 0. \tag{28}$$

There is only one solution to equations (27) and (28), up to normalization:

$$\psi(q) = \mathrm{e}^{-q} + \mathrm{e}^{q-2} \tag{29}$$

This vector spans the one-dimensional physical Hilbert subspace $H_{phys} \subset \tilde{H}$.

Let us check whether $\psi$ indeed projects functions onto their values at zero (up to the normalization constant). Let $\phi$ be any vector in $\tilde{H}$. Indeed,

$$(\psi | \phi) = \int_0^1 \left\{ \left(\mathrm{e}^{-q} + \mathrm{e}^{q-2}\right) \phi(q) + \left(\partial_q \left(\mathrm{e}^{-q} + \mathrm{e}^{q-2}\right)\right) \partial_q \phi(q) \right\} \tag{30}$$

$$= \left[ \left(-\mathrm{e}^{-q} + \mathrm{e}^{q-2}\right) \phi(q) \right]_0^1 \tag{31}$$

$$= c\phi(0) \tag{32}$$

with the normalization constant $c = (1 - \mathrm{e}^{-2})$.

Finally, let us note that since in this case $Q$ is positive, and zero is therefore a boundary of the continuous spectrum, we could have also found this solution by using von Neumann's theory as mentioned above.

Namely, the only boundary vectors, i.e. the only vectors in $H$ obeying either $(P^*\pm \mathrm{i})|\psi\rangle = 0$, i.e. the only normalizable solutions obeying either of $(-\mathrm{i}\partial_q \pm \mathrm{i})\psi(q)$, are $\mathrm{e}^{\pm q}$. Thus, we could have narrowed down the search for the physical subspace, knowing that it has to lie within the boundary space $H_B$ spanned by these two vectors, as it does of course, being spanned by $\psi$ in equation (29).

### 5.2. Example: similarly with isospinors

Let us now consider an example of a constraint operator $Q$ whose spectrum at zero is finitely degenerate, so that we can expect the corresponding physical, i.e. constrained, Hilbert space to be multi-dimensional.

To this end, let us consider the kinematics of a quantum mechanical particle which possesses isospin and which lives, say, on the positive half-line. In its Hilbert space $H$ the scalar product of wavefunctions then reads

$$\langle\phi|\phi'\rangle = \sum_{i=1}^{N}\int_0^\infty \mathrm{d}x\, \phi_i^*(x)\,\phi_i'(x) \tag{33}$$

We wish to constrain the particle from the bulk to its boundary by choosing a constraint operator $Q$, which acts as $Q\psi_i(x) = x\psi_i(x)$. By suitably imposing '$Q = 0$' we intend to project wavefunctions onto the $N$-dimensional isospinor space at $x = 0$ as the physical subspace. Within the original topology of $H$ this is not possible because $\ker Q = \ker Q^* = \emptyset$. Following our general method, we therefore introduce the symmetric operator $P = -\mathrm{i}\partial_x$, obeying $[Q, P] = \mathrm{i}\mathbb{1}$ on its domain. We can now equip this domain, or more precisely, the domain $D_{P^*}$, with the new scalar product equation (17)

$$(\phi|\phi') = \sum_{i=1}^{N}\int_0^\infty \mathrm{d}x\,\left\{\phi_i^*(x)\phi_i'(x) + (\partial_x\phi_i^*(x))\partial_x\phi_i'(x)\right\} \tag{34}$$

to obtain the new Hilbert space $\tilde{H}$ which possesses the improved topology. In $\tilde{H}$ the operator $Q$ is no longer self-adjoint and we identify $H_{phys} = \ker Q^*$. According to equation (19), the condition for vectors $\psi \in \tilde{H}$ to be in the physical subspace, i.e. in $\ker Q^*$, now reads

$$\sum_{i=1}^{N}\int_0^\infty \mathrm{d}x\,\left\{\phi_i^*(x)x\psi_i(x) + (\partial_x x\phi_i^*(x))\partial_x\psi_i(x)\right\} = 0 \qquad \forall\,\phi \in \tilde{D}_Q \tag{35}$$

which yields

$$x(\partial_x^2 - 1)\psi_i(x) = 0. \tag{36}$$

The solution space is spanned by the wavefunctions $\psi_i^{(n)}(x) = \delta_{n,i}\mathrm{e}^{-x}$ $(n = 1, \ldots, N)$, i.e. those vectors and their linear combinations represent $H_{phys}$ in $\tilde{H}$. The projector $\Pi$ onto the physical subspace is $\Pi = \sum_{n=1}^{N}|\psi^{(n)})\otimes(\psi^{(n)}|$.

Let us verify that the scalar product of the $\psi^{(n)}$ with an arbitrary wavefunction $\phi_i(x)$ projects onto the isospinor space at zero. Indeed,

$$(\psi^{(n)}|\phi) = \sum_{i=1}^{N}\int_0^\infty \mathrm{d}x\,\left\{\delta_{n,i}\mathrm{e}^{-x}\phi_i(x) + \partial_x\mathrm{e}^{-x}\delta_{n,i}\partial_x\phi_i(x)\right\} \tag{37}$$

$$= \left[-\mathrm{e}^{-x}\phi_n(x)\right]_0^\infty \tag{38}$$

$$= \phi_n(0). \tag{39}$$

Clearly, our treatment of this example also applies straightforwardly in the case $N = \infty$, to obtain an infinite-dimensional physical sub-Hilbert space spanned by the $\{\phi^{(n)}\}$. This means that the classical constraint 'manifold' which here was a set of $N$ discrete points, could also be taken to be an infinite set of discrete points.

### 5.3. Example: projecting onto the boundary of the half-plane

Let us now consider the case where the classical constraint manifold is actually a continuous manifold. To this end we consider the example of a particle which lives, say, in a two-dimensional space on the half-plane defined by $x_1 \geqslant 0$, i.e. in its Hilbert space the scalar product of wavefunctions reads

$$\langle\phi|\phi'\rangle = \int_{-\infty}^{\infty} \mathrm{d}x_2 \int_0^{\infty} \mathrm{d}x_1\, \phi^*(x_1, x_2)\phi'(x_1, x_2). \tag{40}$$

We choose $Q = x_1$, in order to constrain the particle from the two-dimensional bulk to its one-dimensional boundary at $x_1 = 0$. Clearly, $Q$ is positive, self-adjoint and possesses the half-axis $x_1 \geqslant 0$ as its infinitely degenerate spectrum.

To employ our method, we use the symmetric operator $P := -\mathrm{i}\partial_{x_1}$. As required, it obeys $[Q, P] = \mathrm{i}\mathbb{1}$. Using $P$, the new scalar product $(\,|\,)$ reads

$$(\phi|\phi') = \int_{-\infty}^{\infty} \mathrm{d}x_2 \int_0^{\infty} \mathrm{d}x_1 \left\{\phi^*\phi' + (\partial_{x_1}\phi^*)\partial_{x_1}\phi'\right\} \tag{41}$$

which yields for our purposes the topologically improved Hilbert space $\tilde{H}$ over $D_{P^*}$. According to equation (19), physical vectors $\psi \in H_{phys} = \ker Q^*$ now obey

$$\int_{-\infty}^{\infty} \mathrm{d}x_2 \int_0^{\infty} \mathrm{d}x_1 \left\{\phi^* x_1\psi + (\partial_{x_1} x_1\phi^*)\partial_{x_1}\psi\right\} = 0 \qquad \forall\, \phi \in \tilde{D}_Q \tag{42}$$

which means

$$x_1(\partial_{x_1}^2 - 1)\psi = 0. \tag{43}$$

Thus, the physical vectors $\psi \in H_{phys}$ are represented in $\tilde{H}$ as wavefunctions $\psi(x_1, x_2) = f(x_2)\,\mathrm{e}^{-x_1}$, where $f(x_2)$ is an arbitrary square-integrable function. Let us verify explicitly that the scalar product of an arbitrary vector $\phi \in \tilde{H}$ with a physical vector $\psi \in H_{phys} \subset \tilde{H}$ is the integral over the constraint surface. Indeed,

$$\begin{aligned}(\psi|\phi) &= \int_{-\infty}^{\infty} \mathrm{d}x_2 \int_0^{\infty} \mathrm{d}x_1 \left\{f^*(x_2)\,\mathrm{e}^{-x_1}\phi(x_1, x_2) + (\partial_{x_1} f^*(x_2)\,\mathrm{e}^{-x_1})\partial_{x_1}\phi(x_1, x_2)\right\} \\ &= \int_{-\infty}^{\infty} \mathrm{d}x_2 \left[-\mathrm{e}^{-x_1} f^*(x_2)\phi(x_1, x_2)\right]_0^{\infty} \\ &= \int_{-\infty}^{\infty} \mathrm{d}x_2\, f^*(x_2)\phi(0, x_2). \end{aligned} \tag{44}$$

### 5.4. Example: projecting onto a cylinder in $\mathbb{R}^3$

For a less trivial example, let us now consider quantum mechanics in three spacetime dimensions, choosing as the constraint operator

$$Q = (X_1^2 + X_2^2 - R^2). \tag{45}$$

We then of course expect the constrained, physical Hilbert space to be the Hilbert space of square-integrable functions over the cylinder with radius $R$ around the $x^3$-axis.

In the original Hilbert space, $H$, the scalar product of wavefunctions reads, choosing cylindrical coordinates,

$$\langle\phi_1|\phi_2\rangle = \int_{-\infty}^{\infty} \mathrm{d}x_3 \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\infty} \mathrm{d}r\, r\phi_1^*\, \phi_2 \tag{46}$$

while $Q$ acts as

$$Q\phi(r,\varphi,x_3) = (r^2 - R^2)\,\phi(r,\varphi,x_3). \tag{47}$$

As the first step, we introduce the operator $P := -\mathrm{i}\frac{1}{2r}\partial_r$ on its domain $D_P$ of absolutely continuous square-integrable functions $\phi$ which vanish at $r = 0$ and for which the almost-everywhere defined derivative $-\mathrm{i}\frac{1}{2r}\partial_r\phi$ is also square integrable. Second, we equip $D_{P^*}$ with the new scalar product, equation (11), and its induced topology, to obtain the Hilbert space $\tilde{H}$. Third, we can now calculate the physical domain $H_{phys} = \ker Q^*$. According to equation (19), a state $\psi$ is in the physical subspace exactly if for all $\phi \in \tilde{D}_Q$ it obeys

$$0 = \int_{-\infty}^{\infty} \mathrm{d}x_3 \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\infty} \mathrm{d}r\, r \left(\phi^*(r^2 - R^2)\psi + \left[\frac{1}{2r}\partial_r(r^2 - R^2)\phi^*\right]\left[\frac{1}{2r}\partial_r\psi\right]\right) \tag{48}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x_3 \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\infty} \mathrm{d}r\, r \left(\phi^*(r^2 - R^2)\psi - (r^2 - R^2)\phi^*\frac{1}{2r}\partial_r\left[\frac{1}{2r}\partial_r\psi\right]\right)$$
$$+ \int_0^{2\pi} \mathrm{d}\varphi \int_{-\infty}^{\infty} \mathrm{d}x_3 \left[\frac{1}{2}(r^2 - R^2)\phi^*\frac{1}{2r}\partial_r\psi\right]_0^{\infty}. \tag{49}$$

Thus, physical states must obey the boundary condition at $r = 0$, for all $\phi \in \tilde{D}_Q$,

$$\lim_{r\to 0^+} \int_0^{2\pi} \mathrm{d}\varphi \int_{-\infty}^{\infty} \mathrm{d}x_3\, \frac{1}{2}(r^2 - R^2)\phi^*\frac{1}{2r}\partial_r\psi = 0 \tag{50}$$

since the boundary term at infinity vanishes due to the square integrability. Also, for all positive $r$ with $r \neq R$ the physical states must obey the differential equation:

$$r(r^2 - R^2)\psi = r(r^2 - R^2)\frac{1}{2r}\partial_r\frac{1}{2r}\partial_r\psi. \tag{51}$$

The solutions to this differential equation are of the form

$$\psi(r,\varphi,x_3) = f(\varphi,x_3)\big(A\mathrm{e}^{-r^2} + B\mathrm{e}^{r^2}\big). \tag{52}$$

Let us begin by considering the part of the solution in the region $r < R$: we know that $\psi$ must be continuous everywhere, and this is non-trivial at the origin, $r = 0$: namely, for $\psi$ to be continuous at the origin one needs either $\psi(r = 0,\varphi,x_3) = 0$, or that $\psi(r = 0,\varphi,x_3)$ is independent of $\varphi$. Thus, we obtain two different possible behaviours of physical states in the region $r < R$: the first set of physical states vanishes at the origin, i.e.

$$\psi(r,\varphi,x_3) = f(\varphi,x_3)\big(\mathrm{e}^{-r^2} - \mathrm{e}^{r^2}\big). \tag{53}$$

Here, *a priori*, $f(\varphi,x_3)$ is some arbitrary square-integrable function. However, solutions must also obey the boundary condition equation (50). Since all $\psi$ either vanish at $r = 0$, or are

$\varphi$-independent, equation (50) takes the form

$$0 = \lim_{r\to 0^+} \int_0^{2\pi} \mathrm{d}\varphi \, \frac{1}{2r} \partial_r \psi \tag{54}$$

$$= \lim_{r\to 0^+} \int_0^{2\pi} \mathrm{d}\varphi \, f(\varphi, x_3)\big(-\mathrm{e}^{-r^2} - \mathrm{e}^{r^2}\big). \tag{55}$$

Thus,

$$0 = -2 \int_0^{2\pi} \mathrm{d}\varphi \, f(\varphi, x_3) \tag{56}$$

which means that in solutions of the form given in equation (53) the $f(\varphi, x_3)$ are indeed arbitrary square-integrable function, with the exclusion of the $\varphi$-zero-modes.

However, the zero-modes are not lost: the solutions which we have just obtained are only the set of solutions which are continuous at $r = 0$ by virtue of vanishing there. As we mentioned above, there is a second set of physical states, namely those which are continuous at $r = 0$. These do not depend on the variable $\varphi$, i.e. they are of the form

$$\psi(r, \varphi, x_3) = g(x_3)\big(A\mathrm{e}^{-r^2} + B\mathrm{e}^{r^2}\big). \tag{57}$$

For these solutions the boundary condition equation (50) reads

$$0 = \lim_{r\to 0^+} \int_0^{2\pi} \mathrm{d}\varphi \, \frac{1}{2r} \partial_r \psi \tag{58}$$

$$= \lim_{r\to 0^+} \int_0^{2\pi} \mathrm{d}\varphi \, g(x_3)\big(-A\mathrm{e}^{-r^2} + B\mathrm{e}^{r^2}\big) \tag{59}$$

$$= -2\pi g(x_3)(A - B) \tag{60}$$

which implies that $B = A$. Thus, the physical states which carry the zero-modes of the angular degree of freedom $\varphi$ are represented in the region $r < R$ by the functions

$$\psi(r, \varphi, x_3) = g(x_3)\big(\mathrm{e}^{-r^2} + \mathrm{e}^{r^2}\big). \tag{61}$$

Clearly, for the part of the solution in the region $r > R$, square integrability requires solutions of the form of equation (52) with $B = 0$. The solutions of the differential equation for $r < R$ and for $r > R$ are glued together by the requirement that every physical state should be continuous at $r = R$. We therefore obtain that the physical states are spanned by states of one of two forms. Either,

$$\psi(r, \varphi, x_3) = f(\varphi, x_3) \begin{cases} (\mathrm{e}^{r^2} - \mathrm{e}^{-r^2}) & r < R \\ \mathrm{e}^{-r^2}\,(\mathrm{e}^{2R^2} - 1) & r \geqslant R \end{cases} \tag{62}$$

where $f$ is an arbitrary square-integrable function of $x_3$ and $\varphi$, without, however, the zero-modes in $\varphi$, because of equation (56). The rest of the physical states are the zero-modes in $\varphi$, and are represented as functions of the form

$$\psi(r, \varphi, x_3) = g(x_3) \begin{cases} (\mathrm{e}^{-r^2} + \mathrm{e}^{r^2}) & r < R \\ \mathrm{e}^{-r^2}\,(1 + \mathrm{e}^{2R^2}) & r \geqslant R. \end{cases} \tag{63}$$

The physical Hilbert space $H_{phys} \subset \tilde{H}$ is spanned by these functions, and is equipped with the scalar product of $\tilde{H}$, as given in equation (17).

In order to project an arbitrary function in $\tilde{H}$ down to $H_{phys}$ we need to take scalar products of arbitrary functions $\phi$ in $\tilde{H}$ with functions in the physical Hilbert subspace $H_{phys}$. Let us check that, as desired, this scalar product reduces to an integral over the product of two functions over the surface of the cylinder $r = R$:

$$(\phi, \psi) = \int_{-\infty}^{\infty} \mathrm{d}x_3 \int_0^{2\pi} \mathrm{d}\varphi \int_0^{\infty} \mathrm{d}r\, r \left( \phi^* \psi + \left[ \frac{1}{2r} \partial_r \phi^* \right] \left[ \frac{1}{2r} \partial_r \psi \right] \right) \tag{64}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x_3 \int_0^{2\pi} \mathrm{d}\varphi \left\{ \int_0^{R} \mathrm{d}r\, r \left( \phi^* \psi - \phi^* \frac{1}{2r} \partial_r \left[ \frac{1}{2r} \partial_r \psi \right] \right) \right. \tag{65a}$$

$$\left. + \int_R^{\infty} \mathrm{d}r\, r \left( \phi^* \psi - \phi^* \frac{1}{2r} \partial_r \left[ \frac{1}{2r} \partial_r \psi \right] \right) \right\} \tag{65b}$$

$$+ \int_{-\infty}^{\infty} \mathrm{d}x_3 \int_0^{2\pi} \mathrm{d}\varphi \left\{ \left[ \frac{1}{2} \phi^* \frac{1}{2r} \partial_r \psi \right]_{0_+}^{R_-} + \left[ \frac{1}{2} \phi^* \frac{1}{2r} \partial_r \psi \right]_{R_+}^{\infty} \right\} \tag{65c}$$

where, to be precise, in line (65$c$) the evaluations at the interval boundaries are limits taken from within the interval. The terms in lines (65$a$) and (65$b$) vanish because all physical states $\psi$ obey equation (51). In order to evaluate line (65$c$) we decompose $\phi$ into its $\varphi$-zero-modes $\phi_0$ and the rest, $\phi_1$:

$$\phi(r, \varphi, x_3) = \phi_0(r, x_3) + \phi_1(r, \varphi, x_3) \tag{66}$$

where $\phi_1(r = 0, \varphi, x_3) = 0$. Recall that we found that while the physical subspace consists of all integrable functions, say $h(\varphi, x_3)$ over $x_3$ and over $\varphi$-space, its representation as functions in the original Hilbert space realizes the $\varphi$-zero-modes somewhat specially. Namely, if we decompose an arbitrary $h(\varphi, x_3)$ into their $\varphi$-zero-modes and the rest as

$$h(\varphi, x_3) = f(\varphi, x_3) + g(x_3) \tag{67}$$

then these are represented as

$$\psi(r, \varphi, x_3) = \begin{cases} f(\varphi, x_3)\, (\mathrm{e}^{r^2} - \mathrm{e}^{-r^2}) + g(x_3)(\mathrm{e}^{r^2} + \mathrm{e}^{-r^2}) & r < R \\ f(\varphi, x_3)\, \mathrm{e}^{-r^2}(\mathrm{e}^{2R^2} - 1) + g(x_3)\, \mathrm{e}^{-r^2}(1 + \mathrm{e}^{2R^2}) & r \geqslant R. \end{cases} \tag{68}$$

We then read off:

$$\frac{1}{2r} \partial_r \psi|_{0_+} = 2 f(\varphi, x_3) \tag{69}$$

$$\frac{1}{2r} \partial_r \psi|_{R_-} = f(\varphi, x_3)(\mathrm{e}^{R^2} + \mathrm{e}^{-R^2}) + g(x_3)(\mathrm{e}^{R^2} - \mathrm{e}^{-R^2}) \tag{70}$$

$$\frac{1}{2r} \partial_r \psi|_{R_+} = f(\varphi, x_3)(-\mathrm{e}^{R^2} + \mathrm{e}^{-R^2}) - g(x_3)(\mathrm{e}^{R^2} + \mathrm{e}^{-R^2}). \tag{71}$$

Using

$$\phi_1(0, \varphi, x_3) = 0 \tag{72}$$

$$\int_0^{2\pi} \mathrm{d}\varphi\, \phi_1^*(r, \varphi, x_3) g(x_3) = 0 \tag{73}$$

and

$$\int_0^{2\pi} \mathrm{d}\varphi\, \phi_0^*(r, x_3) f(\varphi, x_3) = 0 \tag{74}$$

we therefore find that

$$(\phi,\psi) = \frac{1}{2}\int_0^{2\pi}\int_{-\infty}^{\infty} \mathrm{d}\varphi\,\mathrm{d}x_3 \left\{ \left(\phi_0^*(r,x_3)+\phi_1^*(r,\varphi,x_3)\right)\frac{1}{2r}\partial_r\psi|_{r=R_-} \right.$$

$$- \left(\phi_0^*(r,x_3)+\phi_1^*(r,\varphi,x_3)\right)\frac{1}{2r}\partial_r\psi|_{r=0_+}$$

$$\left. - \left(\phi_0^*(r,x_3)+\phi_1^*(r,\varphi,x_3)\right)\frac{1}{2r}\partial_r\psi|_{r=R_+}\right\} \tag{75}$$

$$= \mathrm{e}^{R^2}\int_0^{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\varphi\,\mathrm{d}x_3 \left\{\phi_0^*(R,x_3)g(x_3)+\phi_1^*(R,\varphi,x_3)f(\varphi,x_3)\right\} \tag{76}$$

$$= \mathrm{e}^{R^2}\int_0^{2\pi}\int_{-\infty}^{\infty}\mathrm{d}\varphi\,\mathrm{d}x_3\,\phi^*(R,\varphi,x_3)h(\varphi,x_3) \tag{77}$$

as it should be, with the factor $\mathrm{e}^{R^2}$ being an overall normalization constant.

### 5.5. *Example: projecting onto a line in* $\mathbb{R}^3$

Let us briefly also consider the case of projecting not onto a cylinder, but onto a one-dimensional line in $\mathbb{R}^3$. To this end, we consider, similar to the previous example:

$$Q = x_1^2 + x_2^2 \tag{78}$$

and again $P = -\mathrm{i}\frac{1}{2r}\partial_r$, where $r^2 = x_1^2 + x_2^2$. Within our formalism, the change of the dimensionality of the constraint manifold is automatically taken care of: the differential equation obeyed by the physical states, equation (51), must now hold for all positive $r > 0$. Thus, normalizability shows that all solutions must be of the form $\psi(r,\varphi,x_3) = f(\varphi,x_3)\,\mathrm{e}^{-r^2}$. Since these functions do not vanish at the origin, $r = 0$, continuity at $r = 0$ requires independence of $\varphi$:

$$\psi(r,\varphi,x_3) = g(x_3)\,\mathrm{e}^{-r^2}. \tag{79}$$

While for $R > 0$ these solutions were ruled out by the boundary condition equation (50), here equation (50) is trivially obeyed due to its prefactor $r$.

Let us check whether the scalar product of any physical state $\psi$ with an arbitrary state $\phi$ reduces to the integral of the product of the two functions over the $x_3$-axis. Indeed,

$$(\phi|\psi) = \int_{-\infty}^{\infty}\mathrm{d}x_3\int_0^{2\pi}\mathrm{d}\varphi\left[\frac{1}{2}\phi^*(r,\varphi,x_3)\frac{1}{2r}\partial_r\psi\right]_0^{\infty} \tag{80}$$

$$= \int_{-\infty}^{\infty}\mathrm{d}x_3\int_0^{2\pi}\mathrm{d}\varphi\,\frac{1}{2}\phi^*(0,\varphi,x_3)g(x_3) \tag{81}$$

$$= \pi\int_{-\infty}^{\infty}\mathrm{d}x_3\,\phi^*(0,x_3)g(x_3) \tag{82}$$

where in the last step we used that the functions $\phi$ do not depend on $\varphi$ at $r = 0$. Thus, the scalar product of arbitrary functions in $\tilde{H}$ with a function from the physical subspace indeed yields the integral of the product of the two functions over the $x_3$-axis, as expected.

This example also shows how the formalism takes care of a dimensional reduction: the loss of the $\varphi$-degree of freedom is expressed by the representatives of physical states possessing only zero-modes of the $\varphi$-degree of freedom. In other words, restricting the function space over a cylinder to only its zero-modes in the angle $\varphi$ is the same as reducing the function space to a function space over a line rather than the cylinder.

### 5.6. Example: an impossible constraint

It is perhaps instructive to briefly discuss how an 'impossible' constraint yields an empty physical subspace. To this end, let us consider again in $\mathbb{R}^3$ the constraint

$$Q = x_1^2 + x_2^2 + R^2 \tag{83}$$

for some positive $R > 0$. In this case, physical states must obey the differential equation (51) for all positive $r > 0$ and therefore, by the above arguments, the only possibility are functions of the form given in equation (79). However, unlike in the case $R = 0$, now the boundary condition equation (50) is non-trivial, namely yielding $g(x_3) = 0$. This rules out any solutions, thereby yielding $H_{phys} = \emptyset$, as should be the case.

## 6. Non-uniqueness of $Q$ and $P$, and the topology of the constraint manifold

As we saw already in section 3, our method cannot be unique: while the topology of the Sobolev space $H^1$ is good enough to allow the projection onto a function's value at a point, higher Sobolev spaces $H^n$ could be used as well. Analogously, in our more general method, for a given choice of $Q$ and $P$, the topology of $\tilde{H}$ allows us to project onto the physical Hilbert subspace, while higher Sobolev space analogues of $\tilde{H}$ could be used as well. Interestingly, however, there is even more non-uniqueness in our general method.

Firstly, for a given classical constraint manifold there is a non-uniqueness in the possible choices of the constraint operator $Q$. Secondly, also for a given $Q$ there exists in general a non-uniqueness in the possible choices of a canonical conjugate operator $P$. The question of to what degree the various physical Hilbert subspaces obtained from these different choices are or are not equivalent is subtle and reveals interesting relations to the dimensionality and topology of the constraint manifold.

### 6.1. Non-uniqueness of $Q$ for a given classical constraint

Perhaps surprisingly, even simple classically identical constraints may differ quantum mechanically, even when there are no ordering ambiguities: in order to see this, let us consider the example of the constraint operator $Q_1 = x_1^2 + x_2^2 + x_3^2 - R^2$ in the Hilbert space $H$ of square-integrable functions in $\mathbb{R}^3$. Classically, the constraint manifold, i.e. the set of solutions to $Q_1 = 0$ is a 2-sphere in $\mathbb{R}^3$: the physical Hilbert space which we obtain from $Q_1$ is of course the $L^2$ Hilbert space of functions over this sphere. Similarly, the constraint operator $Q_2 = x_3^2 + (x_1^2 + x_2^2 - R^2)^2$ describes the equator of the sphere described by $Q_1$.

Now, let us consider the constraint operator $Q_3 := Q_1 Q_2$. Classically, its constraint manifold, i.e. the set of points for which $Q_3 = 0$, is still the 2-sphere with radius $R$, i.e. it is identical to the case above. Quantum mechanically, the physical Hilbert space for $Q_2$, however, can now be different from the space of functions over the sphere. This is because $Q_2$ counts the 1-sphere of the equator twice. We might expect that the physical Hilbert space also specifically contains $L^2$ functions over the equator and their scalar product. In this way, the scalar product could consist of evaluating both the two-dimensional integral over the sphere and the one-dimensional integral over the equator of the product of two wavefunctions.

This issue will be investigated further elsewhere. It is clear that whenever such ambiguities arise, i.e. whenever the physical subspaces can be, for example, function spaces over a variety of topologically different manifolds (of varying, and in general possibly not even well defined dimensionality), then this should also have a reflection in the functional analysis of $P$.

### *6.2. Non-uniqueness of the choice of P for a given Q*

For a given choice of $Q$, our condition that $P$ be a maximal symmetric operator obeying $[Q, P] = \mathrm{i}\mathbb{1}$ does not determine the choice of operator $P$ uniquely. Clearly, any two operators $P$ and $P'$ which obey the commutation relation $[Q, P] = [Q, P'] = \mathrm{i}\mathbb{1}$ will differ by a symmetric operator $R$ which commutes with $Q$, i.e. $P' = P + R$ with $[R, Q] = 0$, with the appropriate domains understood. Usually, two such operators $P$ will be connected by a gauge transformation, in the sense that there exists a unitary operator $U$ in $H$, such that $P' = P + R = UPU^{\dagger}$. In this case, as is easily verified, $U$ is also an isometry connecting the corresponding 'generalized Sobolev spaces' $\tilde{H}$ and $\tilde{H}'$, as well as being an isometry connecting the corresponding physical subspaces $H_{phys}$ and $H'_{phys}$. In this case, the choice of $P$ is therefore immaterial.

For example, if two operators $P$ differ by an operator $R$ which is a self-adjoint operator that is independent of the constraint degree of freedom, in the sense that it commutes with $Q$ and $P$, then we find $U = \exp(\mathrm{i}QR)$. Or, if $R$ is a function $k(Q)$ of $Q$ with integral $\int \mathrm{d}q\, k(q) = K(q)$, then we have $U = \exp(\mathrm{i}K(Q))$.

More concretely, let us reconsider the example of section 5.3, where we used $Q = x_1$ to project onto the boundary (the $x_2$-axis) of the half-plane ($x_1 \geqslant 0$). We had chosen $P = -\mathrm{i}\partial_{x_1}$. However, this choice was not unique because obviously, for example, also $P' := -\mathrm{i}\partial_{x_1} + g(x_2)$ obeys $[Q, P'] = \mathrm{i}\mathbb{1}$. The question arises as to whether we obtain the same physical Hilbert space if we use $P'$ rather than $P$. A short calculation shows that, using $P'$, the physical vectors are represented as functions $f(x_2)\,\mathrm{e}^{-x_1(1+\mathrm{i}g(x_2))}$. The difference therefore is only a phase, which vanishes for $x_1 = 0$, i.e. after the projection. Thus, the use of $P'$ yields the same Hilbert space as does the use of $P$.

Similarly, also, for example, $P'' := -\mathrm{i}\partial_{x_1} - \mathrm{i}\partial_{x_2}$ obeys the commutation relation $[Q, P''] = \mathrm{i}\mathbb{1}$. Using $P''$, physical vectors must obey the differential equation $(\partial_{x_1} + \partial_{x_2} + 1)\psi = 0$. Using the coordinates $z := (x_1 + x_2)/2$ and $y := (x_2 - x_1)/2$, we obtain $(\partial_z + 1)\psi = 0$, yielding for the general expression of physical vectors: $\psi = f(2y)\,\mathrm{e}^{-z}$, i.e. $\psi = f(x_2 - x_1)\,\mathrm{e}^{-(x_1+x_2)/2}$. The scalar product of a general vector $\phi \in \tilde{H}$ with a physical vector $\psi \in H_{phys}$ is then

$$
\begin{aligned}
(\psi|\phi) &= \int_0^{\infty} \mathrm{d}x_1 \int_{-\infty}^{\infty} \mathrm{d}x_2 \left\{\psi^*\phi + \left((\partial_{x_1} + \partial_{x_2}) f^*(x_2 - x_1)\,\mathrm{e}^{-(x_1+x_2)/2}\right)(\partial_{x_1} + \partial_{x_2})\phi\right\} \\
&= \int_{-\infty}^{\infty} \mathrm{d}x_2 \left[f^*(x_2 - x_1)\,\mathrm{e}^{-(x_1+x_2)/2}\phi\right]_0^{\infty} \\
&= \int_{-\infty}^{\infty} \mathrm{d}x_2\, f^*(x_2)\,\mathrm{e}^{-x_2/2}\phi(0, x_2) \\
&= \int_{-\infty}^{\infty} \mathrm{d}x_2\, \psi^*(0, x_2)\phi(0, x_2). \qquad (84)
\end{aligned}
$$

Thus, when using $P''$ we still obtain the same physical Hilbert subspace.

Similarly, also $P''' := -\mathrm{i}\partial_{x_1} + g(x_1)$ obeys the commutation relation. Its use leads to physical vectors being represented as functions $\psi = f(x_2)\,\mathrm{e}^{-x_1 - \mathrm{i}\int^{x_1} g(x)\,\mathrm{d}x}$. All we obtain is a new phase factor which vanishes (up to possibly an irrelevant global overall phase) after the projection onto $x_1 = 0$. Thus, we also obtain the same Hilbert space from $P'''$, as long as $g(x_1)$ is integrable.

Even though these examples illustrate the robustness of the projection onto the physical subspace under changes of $P$ we cannot exclude that, in general, $R$ may not be a pure gauge, i.e. it may not be 'integrable' in these simple ways. One may find unitarily non-equivalent maximal representations of the commutation relation $[Q, P] = \mathrm{i}\mathbb{1}$. On the other hand, while *a priori* these representations will lead to non-equivalent generalized Sobolev spaces, some of the resulting representations on the corresponding physical subspaces may still be equivalent after projection.

The investigation of the precise relation between different choices of $Q$ for the same classical constraint, unitarily non-equivalent $P$'s, and the non-uniqueness of the topology of the 'constraint manifold' is probably deep and should be worth pursuing further.

## Acknowledgments

The authors are happy to thank B G Bodmann and S Shabanov for useful comments.

## References

[1] Gitman D M and I V Tyutin 1990 *Quantization of Fields with Constraints* (Berlin: Springer)
Henneaux M and Teitelboim C 1992 *Quantization of Gauge Systems* (Princeton, NJ: Princeton University Press)
[2] Klauder J R 2000 Quantization of constrained systems *Preprint* hep-th/0003297
[3] Hájiĉek P 1986 *J. Math. Phys.* **27** 1800
Kuchar K 1996 Hamiltonian dynamics of gauge systems *Phys. Rev.* D **34** 3031
[4] Landsman N P 1998 *Mathematical Topics Between Classical and Quantum Mechanics* (Heidelberg: Springer)
[5] Grundling H and Hurst C A 1998 Constrained dynamics for quantum mechanics I. Restricting a particle to a surface *J. Math. Phys.* **39** 3091
[6] Klauder J R 1997 Coherent state quantization of constraint systems *Ann. Phys.* **254** 419
[7] Albeverio S, Smolyanov O G and Shavgulidze E T 1998 Some methods of quantization of finite dimensional systems with constraints *Dokl. Math.* **58** 114
[8] Dirac P A M 1964 *Lectures on Quantum Mechanics* (New York: Belfer Graduate School of Science, Yeshiva University)
Dirac P A M 1950 Generalized Hamiltonian dynamics *Can. J. Math.* **2** 129
Dirac P A M 1951 The Hamiltonian form of field dynamics *Can. J. Math.* **3** 1
[9] Al-Gwaiz M A 1992 *Theory of Distributions* (New York: Dekker)