International Journal of Theoretical Physics, Vol. 7, No. 1 (1973), pp. 31–38

The Positivity Conditions In General Relativity†‡

J. N. GOLDBERG and F. S. KLOTZ

Department of Physics, Syracuse University, Syracuse, New York 13210

Received 10 *February* 1972

Abstract

We extend the usual treatment of constraints in the canonical formalism to cases where the restrictions on the canonical variables take the form of inequalities. The result is a general prescription for eliminating such restrictions from the canonical formalism. We apply this prescription to eliminate the positivity conditions in a model theory and in general relativity. In the case of the model we find that elimination of the positivity conditions makes the construction of observables much simpler; in the full theory, however, the spatial constraints introduce complications, and we have not been able to carry out all the calculations explicitly.

1. *Introduction*

The construction of a quantum theory for general relativity through the canonical formalism has been stymied by the lack of a complete non-redundant set of observables (Bergmann, 1962). These observables must be invariants, that is, quantities which have vanishing Poisson brackets with the constraint generators of the invariant transformations of the theory. By introducing an intrinsic coordinate system, Komar has constructed a complete set of observables (Komar, 1958a). These variables, however, are not independent, for they must still satisfy the constraint equations. On the other hand, Komar has also shown that a complete non-redundant set of observables uniquely characterizes the Hamilton–Jacobi functional of general relativity (Komar, 1968b). In this case, however, the explicit construction has not been carried out.

We wish to explore the possibility that the solution of a secondary problem associated with the Hamiltonian formalism of Dirac (1958a, 1958b) may be a significant step in the construction of observables. This is the problem

† This research was supported in part by the National Science Foundation and the Aerospace Research Laboratories.

‡ This paper is based partly on the second author's doctoral dissertation at Syracuse University.

of incorporating into the formalism the 'possitivity conditions', which are most simply expressed by the inequality†

$$g_{ab}\,dx^a\,dx^b > 0 \qquad \forall dx^a \neq 0$$

The positivity conditions require that the initial t = constant hypersurface S be spacelike, i.e., that the 3×3 matrix $\{g_{ab}\}$ be positive definite. They may therefore be regarded as rather complicated nonholonomic constraints in configuration space which arise from the physical interpretation of the canonical co-ordinates. Klauder & Aslaken (1968) have pointed out that such constraints present serious problems in any quantum theory of general relativity. They argue that the q-number analogs of the Dirac variables g_{ab} and p^{ab} cannot serve as the basic co-ordinate and momentum operators in Hilbert space, for, in view of the positivity conditions, the spectra of the operator analogs of the g_{ab} are nonphysical. One would like to be able to formulate the classical field equations in terms of classical variables which identically satisfy the positivity conditions so that this problem could not arise. While such variables will not in general be observables, one might hope to learn something about the observables by simplifying the formalism in this manner. Here we present our first efforts to carry out this program.

We begin with a few general remarks on the role of inequalities in the canonical formalism. The standard phase space formalism is based on the assumption that the canonical variables represent independent degrees of freedom of the physical system. Any restrictions on the independence of these variables, whether holonomic (e.g., constraint equations) or nonholonomic (e.g., inequalities), indicate that in some sense the variables themselves are inappropriate. For these conditions imply that part of the original phase space lacks physical significance; only the subspace they define, the *reduced phase space*, is of physical interest.

For any physically acceptable theory the restrictions which appear must be consistent with the dynamical equations. In other words, the Hamiltonian of the theory must map the reduced phase space onto itself. If this is the case, we can lose no physical information by confining our attention exclusively to the reduced phase space. Whether the constraints are holonomic or nonholonomic, it must be possible to remove them from the formalism by replacing the 'defective' original set of phase space variables with a new set of reduced phase space variables

The positivity conditions in general relativity are not dynamical conditions at all; they merely require that we adhere to the geometrical conventions we have established. In order to state the initial value problem correctly it is necessary to choose the arbitrary functions which appear in the Dirac Hamiltonian so as to insure that the positivity conditions are nowhere violated. We can therefore introduce canonical variables appropriate to the reduced phase space defined by the positivity conditions without losing

† We take Latin indices to run from 1 to 3 and Greek indices to run from 0 to 3. We shall use a signature of $-{+}{+}{+}$ for the space-time metric $g_{\mu\nu}$, so that the intrinsic metric of the initial t = constant hypersurface has the signature $+{+}{+}$.

any physically meaningful solutions to the dynamical equations. This is the general program we shall follow to eliminate the positivity conditions from the Hamiltonian formalism.

We have emphasized above the similarities between nonholonomic and holonomic restrictions on the canonical variables. However, it is important to remember that there remain significant differences between the positivity conditions and constraint equations in the Dirac formalism. For one thing, the positivity conditions do not reduce the number of degrees of freedom in the problem: the physical subspace has exactly the same dimensionality as the original phase space. This is never the case with constraint equations. Furthermore, the positivity conditions do not arise naturally out of the requirements for the internal consistency of the Hamiltonian formalism, as do the constraint equations. The origin of the positivity conditions is geometrical, and therefore independent of the formalism itself. These differences lead to difficulties if one tries to press the analogy between the positivity conditions and the constraint equations too far.

In the following sections we shall eliminate the positivity conditions, first in a simple model theory obtained from the Dirac formalism by imposing certain subsidiary conditions, and then in the full theory of general relativity.

2. *Elimination of the Positivity Conditions in a Model Theory*

Suppose that in the Dirac formalism for general relativity we arbitrarily restrict ourselves to initial data which satisfies, in addition to the constraint equations, the conditions

$$g_{ab,c} \approx 0 \qquad p^{ab}{}_{,c} \approx 0 \tag{2.1}$$
$$g_{ab} \approx g_a \delta_{ab} \qquad p^{ab} \approx \pi^a \delta^{ab} \qquad \text{(no sum over } a)$$

It is not difficult to show that, with the arbitrary functions chosen so that

$$g_{0\nu} = g_0 \delta_{0\nu}, \qquad g_0 < 0, \qquad g_{0,a} = 0 \tag{2.2}$$

the initial conditions (2.1) are propagated by the Dirac Hamiltonian. Consequently there exists a particular subset of solutions to the empty space gravitational field equations which satisfy restrictions (2.1) and (2.2) everywhere; these are the well-known Kasner solutions (Kasner, 1921).

The subsidiary conditions (2.1) and (2.2) reduce the Hamiltonian formalism of general relativity to a much simpler canonical theory with three phase space co-ordinates g_k, three conjugate momenta π^k, and one arbitrary function $(-g_0)^{1/2}$. The simplified theory possesses a single constraint,

$$C \equiv (g_1 \pi^1)^2 + (g_2 \pi^2)^2 + (g_3 \pi^3)^2 - 2g_1 \pi^1 g_2 \pi^2 - 2g_2 \pi^2 g_3 \pi^3 - 2g_3 \pi^3 g_1 \pi^1 \approx 0$$

which reduces the number of independent degrees of freedom to two. The Hamiltonian is

$$H \equiv \tfrac{1}{2}K^{-1}(-g_0)^{1/2}\{(g_1 \pi^1)^2 + (g_2 \pi^2)^2 + (g_3 \pi^3)^2 - 2g_1 \pi^1 g_2 \pi^2 - 2g_2 \pi^2 g_3 \pi^3 - 2g_3 \pi^3 g_1 \pi^1\} \approx 0$$

where $K^2 \equiv g_1 g_2 g_3$, and the positivity conditions take the form

$$g_k > 0 \tag{2.3}$$

Now every function $F(g_a, \pi^a)$ defined on the full six-dimensional phase space coordinatized by the variables g_k, π^k generates a one-parameter family of canonical transformations

$$g_k \to g_k(\lambda)$$
$$\pi^k \to \pi^k(\lambda)$$

where $g_k(\lambda)$ and $\pi^k(\lambda)$ are solutions to the differential equations

$$\frac{dg_k}{d\lambda} = \frac{\partial F}{\partial \pi^k}, \frac{d\pi^k}{d\lambda} = -\frac{\partial F}{\partial g_k}$$

We define *p-observables* as those functions $F(g_k, \pi^k)$ which map the physical subspace defined by the positivity conditions (2.3) onto itself. That is, for p-observable generating functions $\mathcal{O}(g_k, \pi^k)$, the initial conditions

$$g_k(0) \equiv g_k > 0$$

must imply

$$g_k(\lambda) > 0 \qquad \forall \lambda$$

For instance, according to the above definition, the g_k are p-observable because the transformations they generate satisfy

$$\dot{g}_k = 0 \Rightarrow g_k(\lambda) = g_k(0) = g_k > 0$$

Obviously this is a special instance of the general rule that all phase space functions whose brackets with the g_k vanish are p-observable. But the π^k are not p-observable, because they generate the transformations

$$\dot{g}_l = \delta_l^{\ k} \quad g_l(\lambda) = g_l(0) + \lambda \delta_l^{\ k}$$

which give $g_k(\lambda) \leqslant 0$ for $\lambda \leqslant -g_k(0)$. Finally, consider the product $g_1 \pi^1$. We have

$$\dot{g}_1 = g_1 \Rightarrow g_1(\lambda) = g_1(0)\, e^{\lambda} > 0 \qquad \forall \lambda$$
$$\dot{g}_2 = 0 \Rightarrow g_2(\lambda) = g_2(0) > \qquad \forall \lambda$$
$$\dot{g}_3 = 0 \Rightarrow g_3 = g_3(0) > 0 \qquad \forall \lambda$$

Hence $g_1 \pi^1$ is p-observable. Similarly, $g_2 \pi^2$ and $g_3 \pi^3$ are p-observable.

A short discussion of the properties of p-observables, which are in some respects rather surprising, will be found in the Appendix.

The p-observables defined above play a role analogous to that of ordinary observables (or *first-class* quantities) in the presence of constraint equations (Bergmann & Komar, 1962). A complete set of p-observable canonical variables spans the reduced phase space defined by the positivity conditions. These reduced phase space variables must satisfy the positivity conditions identically. For if the positivity conditions restricted the range of any one of them, its canonical conjugate would not be p-observable, contrary to our assumption. Consequently, by introducing a set of canonical variables which are all p-observables, we can eliminate the positivity conditions from the formalism.

The simplest way to carry out this procedure in the present case is to replace the momentum variables π^1, π^2, and π^3 with the p-observable variables $p^1 = g_1\pi^1$, $p^2 = g_2\pi^2$, and $p^3 = g_3\pi^3$, and the co-ordinates g_k with the p-observable canonical conjugates of the p^k, $q_k = \ln g_k$. Conditions (2.3), which may now be written as

$$g_k = e^{q_k} > 0$$

are automatically satisfied and so may be ignored. Since the positivity conditions are preserved by the dynamical equations of our model, no physical information is lost when we replace the restricted g_k, π^k variables with the unrestricted q_k, p^k variables.

The introduction of p-observable canonical variables results in a substantial simplification in the form of the Hamiltonian:

$$H = \tfrac{1}{2}(-g_0)^{1/2} K^{-1}\{(p^1)^2 + (p^2)^2 + (p^3)^2 - 2p^1p^2 - 2p^2p^3 - 2p^3p^1\} \approx 0$$

where $\ln K^2 = q_1 + q_2 + q_3$.

We can simplify the Hamiltonian still further by means of a linear transformation which preserves the p-observable character of the basic canonical variables:

$$\left.\begin{aligned} Q_1 &\equiv \frac{1}{\sqrt{3}}(q_1+q_2+q_3) & P^1 &\equiv \frac{1}{\sqrt{3}}(p^1+p^2+p^3) \\ Q_2 &\equiv \frac{1}{2\sqrt{3}}(q_1+q_2-2q_3 & P^2 &\equiv \frac{1}{\sqrt{3}}(p^1+p^2-2p^3) \\ Q_3 &\equiv \tfrac{1}{2}(q_2+q_1) & P^3 &\equiv p^2-p^1 \end{aligned}\right\} \Rightarrow$$

$$H = \tfrac{1}{2}(-g_0)^{1/2}\exp\left(-\frac{\sqrt{3}}{2}Q_1\right)\{(P^2)^2+(P^3)^2-(P^1)^2\} \approx 0$$

We shall not pursue the above analysis further, for Misner (1969) has discussed this theory, as well as other similar model theories, in great detail. The important point here is the following. The requirement that the basic canonical variables be p-observables has led in a natural way to a new formulation of the model theory which is considerably simpler than the original one. The elimination of the positivity conditions has resulted in the simplification of the constraint equation as well. Consequently the construction of the ordinary observables of the theory is now a much simpler task; indeed, it is almost trivial. As far as the model theory is concerned, at least, the elimination of the positivity conditions constitutes an important step toward the construction of a complete set of nonredundant observables.

3. *Elimination of the Positivity Conditions in General Relativity*

We now wish to apply the above program to the Hamiltonian formalism for general relativity. To do so we must construct a complete set of p-observable canonical variables on the physical subspace of phase space defined by the positivity conditions. This in itself is not very difficult, for many such sets of variables exist. But if we wish to achieve any real simplification by eliminating the positivity conditions, we must take care not to destroy the manifest spatial covariance of the Dirac formalism. Suppose, for instance, that we replace the g_{ab} with the logarithms of the eigenvalues of the 3×3 matrix $\{g_{ab}\}$ together with three additional variables which characterize the unitary transformation that diagonalizes $\{g_{ab}\}$. In this way, by direct analogy with our work on the model, we could eliminate the positivity conditions. But this decomposition of $\{g_{ab}\}$ is not spatially covariant, and as a result it leads to expressions for the spatial constraints $\mathscr{H}_s \approx 0$ which are grotesquely complicated. Alternately, one might set co-ordinate conditions which require the off-diagonal elements of $\{g_{ab}\}$ to vanish, and introduce the logarithms of the diagonal elements of $\{g_{ab}\}$ as canonical variables. Here again the analogy with our work on the model is clear. But in general it will not be possible to propagate the co-ordinate conditions off S, and this approach will break down. In what follows we shall present one method for eliminating the positivity conditions which preserves the manifest spatial covariance of the formalism and so is free of these defects.

We begin by replacing the Dirac configuration space variables g_{ab} with the 3-scalars defined by

$$\gamma^{AB} \equiv \alpha^A{}_{,r}\, \alpha^B{}_{,s}\, e^{rs} \qquad A = 1, 2, 3$$

where the $\alpha^A(g_{rs})$ are three independent 3-scalar functions of the three-dimensional Riemannian tensor. The γ^{AB} are just the contravariant components of the metric tensor of S given in the intrinsic co-ordinate system defined by the α^A. (We assume that the α^A are chosen so that $|\alpha^A{}_{,r}| \neq 0$ at each point of S.) Consequently the positivity conditions now require that the 3×3 matrix of the γ^{AB} be positive definite. The momentum densities conjugate to the γ^{AB} are given by

$$p_{AB} = \int \frac{\delta g'_{ab}}{\delta \gamma^{AB}} p^{ab'}\, d^3 x'$$

At this point there are many different sets of p-observable canonical variables that we may introduce in place of the γ^{AB} and p_{AB}. Perhaps the simplest such set is the one directly analogous to the first set proposed above,

$$Q^{AB} \equiv \{Q\}^{AB} \equiv \{\ln \gamma\}^{AB}$$

$$P_{AB} \equiv \frac{\partial \gamma^{RS}}{\partial Q^{AB}} p_{RS}$$

Here $\{Q\}$ and $\{\gamma\}$ represent 3×3 matrices. The Q^{AB}, as defined above, transform as 3-scalars on S; the P_{AB} transform as spatial scalar densities of weight one. Consequently the spatial constraints assume the form (Klotz, 1972)

$$\mathscr{H}_s = Q^{AB}{}_{,s} P_{AB} \approx 0$$

The remaining constraints depend on the explicit functional form of the intrinsic spatial co-ordinates:

$$\mathscr{H}_L = \mathscr{H}_L(Q^{AB}, P_{AB}) \approx 0$$

It is clear from what we have said previously that the positivity conditions impose no restrictions on the variables Q^{AB}, P_{AB}.

4. *Summary and Conclusion*

In a simple model theory based on general relativity, we have seen that the elimination of the positivity conditions is a considerable help in the construction of the independent observables. It is not clear whether this procedure, when extended to general relativity, is as useful for the construction of observables as it is in the model theory. The spatial covariance of the full theory introduces complications which we did not encounter in the model. While we have eliminated the positivity conditions and simplified the form of the spatial constraints, we have not been able to obtain an explicit expression for $\mathscr{H}_L$. The new canonical variables Q^{AB} and P_{AB} are not, as in the model, simple algebraic functions of the metric tensor, but depend on at least the third derivatives of the g_{ab}. Therefore, the formal simplification we have achieved is deceptive. Although we are farther along than Komar (1958), the remaining task seems to be insurmountable. It appears that the natural starting point for the canonical quantization program for general relativity is the fully reduced phase space defined by both the positivity conditions and the constraint equations. However, we must conclude that there is at present no satisfactory method for constructing the canonical variables of this fully reduced phase space explicitly.

Acknowledgements

The authors want to thank Dr. John Klauder and Professor Peter Bergmann for many stimulating discussions of the positivity conditions and the canonical formalism. One of us (Joshua N. Goldberg) would like to thank the physics department at the Technion for its hospitality while some of these ideas incubated.

APPENDIX

Some Simple Properties of P-Observables

We first consider p-observables within the context of the model theory. If $\mathscr{O}_1$ and $\mathscr{O}_2$ are p-observables, then their sum $\mathscr{O}_1 + \mathscr{O}_2$ must also be

p-observable. For if the transformations generated by $\mathcal{O}_1$ and $\mathcal{O}_2$ individually preserve the conditions

$$g_k > 0 \tag{A.1}$$

then the (product) transformation generated by $\mathcal{O}_1 + \mathcal{O}_2$ must also do so. Furthermore, if $\mathcal{O}_1$ and $\mathcal{O}_2$ are p-observables and

$$[\mathcal{O}_1, \mathcal{O}_2] = 0 \tag{A.2}$$

then $\mathcal{O}_1\mathcal{O}_2$ is p-observable. For suppose that $\mathcal{O}_1$ and $\mathcal{O}_2$ are independent functions of the canonical variables. Then (A.2) implies that by means of an appropriate canonical transformation we can introduce $\mathcal{O}_1$ and $\mathcal{O}_2$ as canonical momenta. The assumption that $\mathcal{O}_1$ and $\mathcal{O}_2$ are individually p-observable implies that independent changes in the conjugate co-ordinates Q^1 and Q^2 preserve (A.1). Consequently the simultaneous changes in Q^1 and Q^2 generated by the product $\mathcal{O}_1\mathcal{O}_2$ must preserve (A.1). That is, $\mathcal{O}_1\mathcal{O}_2$ is p-observable. On the other hand, if $\mathcal{O}_2 = F(\mathcal{O}_1)$, we introduce $\mathcal{O}_1$ as a momentum variable. Since $\mathcal{O}_1$, $\mathcal{O}_2$ and $\mathcal{O}_1\mathcal{O}_2$ generate changes in the conjugate co-ordinate Q^1, the result that $\mathcal{O}_1\mathcal{O}_2$ is p-observable follows immediately from the assumption that $\mathcal{O}_1$ and $\mathcal{O}_2$ are p-observable. This completes the proof of our assertion.

Strangely enough, however, it turns out that the product of two arbitrary p-observable phase space functions is not necessarily p-observable. We have already seen a simple example of this: $(g_1)^{-1}$ and $g_1\pi^1$ are both p-observables but their product $\pi^1 = [(g_1)^{-1}][g_1\pi^1]$ is not. Conversely, it is quite possible to write a p-observable quantity as a product of two functions which are not themselves p-observables. For instance,

$$1 = (\pi^1)(\pi^1)^{-1}$$

Neither π^1 nor its inverse are p-observable, but their product clearly is.

The concept of p-observables can obviously be extended to more general phase space inequalities than (A.1). Since the above arguments are independent of the specific form of conditions (A.1), the properties of p-observables that we have deduced are quite general.

References

Bergmann, P. G. (1962). The general theory of relativity, in *Handbuch der Physik*, Vol. IV. Springer-Verlag, Berlin.

Bergmann, P. G. and Komar, A. B. (1962). Status report on the quantization of the gravitational field, in *Recent Developments in General Relativity*. Pergamon Press, New York.

Dirac, P. A. M. (1958). *Proceedings of the Royal Society* (London), **A246**, 333.

Dirac, P. A. M. (1958). *Proceedings of the Royal Society* (London), **A246**, 326.

Kasner, E. (1921). *American Journal of Mathematics*, **43**, 217.

Klauder, J. and Aslaken, E. (1968). *Journal of Mathematical Physics*, **9**, 206.

Klotz, F. S. (1972). *International Journal of Theoretical Physics*, Vol. 6, No 4, p. 251.

Komar, A. B. (1958). *Physical Review*, **111**, 1182.

Komar, A. B. (1968). *Physical Review*, **170**, 1195.

Misner, C. W. (1969). *Physical Review*, **186**, 1319.