Symmetry and Bifurcations of Momentum Mappings[*]

Judith M. Arms[1], Jerrold E. Marsden[2], and Vincent Moncrief[3]

[1] Department of Mathematics, University of Washington, Seattle, WA 98195, USA
[2] Department of Mathematics, University of California, Berkeley, CA 94720, USA
[3] Department of Physics, Yale University, P.O. Box 2014 Yale Station, New Haven, CT 06520, USA

Abstract. The zero set of a momentum mapping is shown to have a singularity at each point with symmetry. The zero set is diffeomorphic to the product of a manifold and the zero set of a homogeneous quadratic function. The proof uses the Kuranishi theory of deformations. Among the applications, it is shown that the set of all solutions of the Yang-Mills equations on a Lorentz manifold has a singularity at any solution with symmetry, in the sense of a pure gauge symmetry. Similarly, the set of solutions of Einstein's equations has a singularity at any solution that has spacelike Killing fields, provided the spacetime has a compact Cauchy surface.

1. Introduction

A momentum mapping is the conserved quantity associated with a symmetry group acting on phase space. The purpose of this paper is to study the level sets of a momentum mapping and, especially, the zero set. The main results of the paper show that *these level sets have cone-type singularities at any point (in phase space) which itself has some symmetries.*

Level sets of momentum mappings are important in several contexts.

a) The Topology of Hamiltonian Systems with Symmetry

The momentum mapping is conserved by a given Hamiltonian system with symmetry, so knowledge of the level sets and their bifurcations can help in understanding the qualitative features of its flow, as has been emphasized by Smale (1970). In this context, one can reduce a Hamiltonian system with symmetry, looking at the orbit space of a level set of the momentum mapping. This general

procedure, due to Marsden and Weinstein (1974) generalizes the classical elimination of $2k$ variables when k integrals in involution are known as well as Jacobi's elimination of the node in celestial mechanics.

b) Constraints in Classical Relativistic Field Theories

Under certain general conditions, the constraints in Lagrangian field theories can be phrased as saying that an appropriate momentum mapping vanish (see Gotay et al., 1980). This idea first arose in general relativity and (classical Lorentzian) gauge theory. The present paper grew out of our work in general relativity (Fischer et al., 1980) and gauge theory (Arms, 1980). *The study of the space of all solutions to a classical relativistic field theory is thus closely related to the study of the zero set of an associated momentum mapping* [1]. This space of classical solutions plays a key role in perturbation theory about a given solution. If this solution is a point of symmetry for the gauge group generating the constraints, then first order perturbation theory must be supplemented by second order conditions in order to approximate solutions to the nonlinear equations.

   Most of the work on classical field theory assumes that the space of solutions forms a manifold, or has restricted attention to points where that assumption holds. This is true, for example, in Marsden and Weinstein (1974), in the work of Gotay et al. (1978) on the Dirac theory of constraints, in Segal (1978) on Yang-Mills theory and in most perturbation work in field theory. In general relativity however, Brill and Deser (1973) questioned this assumption, and found that it sometimes fails. In fact it fails exactly when there is symmetry, i.e. exactly in the cases of interest, for the known solutions which are perturbed are symmetric. Similar phenomena for more general situations in general relativity were obtained in a series of papers of Choquet-Bruhat et al.; see Fischer et al. (1980) and references therein. For gauge theories similar results are due to Moncrief (1977) and Arms (1979a, 1980).

c) Perturbative Quantum Theory

The breakdown of classical perturbation theory near a solution with symmetry has a quantum analogue that arises if one quantizes the fluctuations about a given symmetric classical background solution. Even at first order, the quantized form of the linearized constraints does not adequately capture the content of the nonlinear constraints. Using the method of Dirac, one must restrict the linearized constraints by certain second order quantum constraints in order to exclude some physically spurious quantum states. An example was worked out by Moncrief (1978); it was shown that without the second order conditions, states which violate the correspondence principle would be allowed.

   A similar phenomenon will occur in the path integral quantization of small fluctuations about a symmetric classical background. One would find it essential to expand certain projections of the constraints (which appear in the classical

---

1   Our work applies to Lorentzian field theories. The results for Euclidean field theories, such as Yang-Mills fields on S^4 or gravitational instantons are different; see Atiyah et al. (1978)

action integral) to higher than linear order to accurately approximate their contribution to the quantum theory.

This paper describes the structure of the singularities in the zero sets of momentum mappings that occur at points of symmetry when the group giving rise to the momentum mapping is compact or admits a local slice. We emphasize zero levels rather than general level sets both for simplicity and because this covers most of the examples of interest. In perturbation theory, it is assumed that the linearized equations in a field theory such as relativity are a good approximation to the original nonlinear equation; i.e., the solution set of the nonlinear equations when linearized, is diffeomorphic to the solution set of the linearized equations. In this paper we show that at points with symmetries the nonlinear zero set is diffeomorphic to the zero set of a homogeneous quadratic form given by the second order perturbation equations. Moreover, this zero set is the product of a manifold of solutions with the same degree of symmetry and a cone of solutions for which one or more of the symmetries has been broken. An important consequence for perturbation theory is that while second order conditions must be imposed on the first order perturbations, there are no additional higher order obstructions to completing the perturbation expansion.

The method of proof involves a function borrowed, via the work of Atiyah et al. (1978) on gauge theory, from the Kuranishi Theorem on deformation of complex structures. However, the proof itself makes no reference to that work. What is directly and extensively used is the machinery of groups of symplectomorphisms (canonical transformations) acting on a symplectic manifold (phase space) which has an additional metric structure. The relevant machinery is reviewed in Sect. 2; for a more leisurely discussion, see Chaps. 3 and 4 of Abraham and Marsden (1978). The main results of the paper are contained in Theorems 1–5 in Sects. 3–6. Section 7 resumes the present discussion and Sect. 8 gives some specific examples.

The critical or bifurcation points that occur are proved to be non-degenerate in a suitable sense (see Theorem 2), so it may be possible to develop a corresponding global Morse theory. [Local Morse theory does in fact appear in the preceding work by Fischer et al. (1979), but those methods are not pursued here.] The present paper deals only with the local structure of the singularities.

2. Background and Notation

We recall some concepts and list the notations that will be used throughout the paper. We let (P, ω) be a given symplectic manifold (possibly infinite dimensional in which case ω is only required to be a weak symplectic form). Let G be a Lie group with Lie algebra $\mathfrak{g}$. Assume G acts symplectically on P; the action is denoted by $(g, x) \mapsto g \cdot x = \Phi_g(x)$. Let

$$J : P \to \mathfrak{g}^*$$

be an Ad*-equivariant momentum mapping. This means two things. First, J satisfies the identity

$$\langle dJ(x) \cdot v, \xi \rangle = \omega_x(\xi_P(x), v), \tag{1}$$

where $\langle \, , \, \rangle$ is the natural pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$, $dJ(x): T_xP \to \mathfrak{g}^*$ is the derivative of J, $v \in T_xP$, the tangent space to P at x, $\xi \in \mathfrak{g}$ and ξ_P is the infinitesimal generator of the group action corresponding to ξ. Second, the following diagram commutes:

$$
\begin{array}{ccc}
P & \xrightarrow{\ \Phi_g\ } & P \\
{\scriptstyle J}\Big\downarrow & & \Big\downarrow{\scriptstyle J} \\
\mathfrak{g}^* & \xrightarrow[\ \mathrm{Ad}_{g^{-1}}^*\ ]{} & \mathfrak{g}^*
\end{array}
$$

where $\mathrm{Ad}_{g^{-1}}^*$ is the co-adjoint action of G on $\mathfrak{g}^*$. For further information on momentum mappings consult, for example, Abraham and Marsden (1978, Chap. 4).

We shall also make the following hypotheses and notations throughout the paper:

1. There is a (weak) Riemannian metric $\langle\!\langle \, , \, \rangle\!\rangle$ on P and a complex structure $\mathbb{J}_x : T_xP \to T_xP$ satisfying $\mathbb{J}^2 = -\,\mathrm{Id}$, $\mathbb{J}$ is symplectic and

$$
\langle\!\langle v_x, w_x \rangle\!\rangle_x = \omega_x(v_x, \mathbb{J}w_x). \tag{2}
$$

The condition that $\mathbb{J}$ is symplectic is equivalent to requiring that $\mathbb{J}$ is $\langle\!\langle \, , \, \rangle\!\rangle$-skew. (We do not assume that P is a Kähler manifold; possibly this assumption, would give additional information.)

2. The group action commutes with $\mathbb{J}$:

$$
T\Phi_g \circ \mathbb{J} = \mathbb{J} \circ T\Phi_g
$$

and hence preserves $\langle\!\langle \, , \, \rangle\!\rangle$ as well. (Conversely, if the action preserves $\langle\!\langle \, , \, \rangle\!\rangle$ it must commute with $\mathbb{J}$.)

3. For $x_0 \in P$, assume that the G orbit of x_0 is a closed submanifold of P. This, together with our assumption that $\langle\!\langle \, , \, \rangle\!\rangle$ is invariant implies that there is a G-slice at x_0 (see Palais, 1957); i.e. there is a submanifold $S_{x_0} \subset P$ containing x_0 and satisfying:

    (i) $g \cdot x_0 = x_0$ implies $g \cdot S_{x_0} = S_{x_0}$,

    (ii) if $g \cdot S_{x_0} \cap S_{x_0} \neq \emptyset$ then $g \cdot x_0 = x_0$, and

    (iii) there is a local cross section $\chi : G/I_{x_0} \to G$ defined in a neighborhood of the identity coset such that the map $S_{x_0} \times G/I_{x_0} \to P$; $(x, u) \mapsto \chi(u) \cdot x$ gives a local diffeomorphism. Here, $I_{x_0} = \{g \in G \mid g \cdot x_0 = x_0\}$ denotes the *isotropy group* of x_0.

The slice S_{x_0} can be chosen as follows. We claim that the $\langle\!\langle \, , \, \rangle\!\rangle$-orthogonal complement of $T_{x_0}(G \cdot x_0)$ is

$$
[T_{x_0}(G \cdot x_0)]^\perp = \ker(dJ(x_0) \circ \mathbb{J}). \tag{3}
$$

Indeed, this follows from (1), (2) and the fact that $T_{x_0}(G \cdot x_0) = \{\xi_P(x_0) \mid \xi \in \mathfrak{g}\}$. Now S_{x_0} can can be chosen in two ways. First of all, assume that P is open in a linear space and relative to this linear structure let S_{x_0} be $\{x_0\}$ + a ball in $[T_{x_0}(G \cdot x_0)]^\perp$ relative to the inner product $\langle\!\langle \, , \, \rangle\!\rangle_{x_0}$ (in infinite dimensions, use a strong metric here). Since the issues are local, the linear structure may be provided by local coordinates, but in several major examples, P is naturally open in a linear space. Secondly, one can choose S_{x_0} to be the $\langle\!\langle \, , \, \rangle\!\rangle$ exponential of such a ball. In any case, we choose $T_{x_0}S_{x_0}$ to be given by (3).

4. The adjoint of $dJ(x_0) : T_{x_0}P \to \mathfrak{g}^*$ is the operator

$$dJ(x_0)^* : \mathfrak{g} \to T_{x_0}P$$

defined by

$$\langle dJ(x_0) \cdot v, \xi \rangle = \langle\!\langle v, dJ(x_0)^* \cdot \xi \rangle\!\rangle .$$

(In infinite dimensions, assume the adjoint exists.)

In terms of adjoints, the defining relation (1) for momentum mappings reads:

$$\xi_P(x_0) = - \mathbb{J} \circ dJ(x_0)^* \cdot \xi . \tag{4}$$

5. For each $x_0 \in P$, assume there is an inner product $(\ ,\ )_{x_0}$ on $\mathfrak{g}^*$ such that $(\ ,\ )_{x_0}$ is invariant under $\mathrm{Ad}^*_{g^{-1}}$ for each g satisfying $gx_0 = x_0$. [In infinite dimensions $(\ ,\ )_{x_0}$ need not be complete.] We define the $(\ ,\ )_{x_0}$ dual of $dJ(x_0)$ to be

$$dJ(x_0)^\dagger : \mathfrak{g}^* \to T_{x_0}P ,$$

where

$$(dJ(x_0) \cdot v, \mu)_{x_0} = \langle\!\langle v, dJ(x_0)^\dagger \cdot \mu \rangle\!\rangle_{x_0} .$$

It follows that

$$\mathfrak{g}^* = \mathrm{range}\, dJ(x_0) \oplus \ker dJ(x_0)^\dagger . \tag{5}$$

[In infinite dimensions assume $dJ(x_0)^\dagger$ or $dJ(x_0)$ is an elliptic operator and apply the Fredholm alternative.]

6. Let $x_0 \in P$ and assume $J(x_0) = 0$. From Ad*-equivariance, it follows that G leaves $J^{-1}(0)$ invariant. In particular, $T_{x_0}(G \cdot x_0) \subset \ker dJ(x_0)$. In other words,

$$\mathrm{range}(- \mathbb{J} \circ dJ(x_0)^*) \subset \ker dJ(x_0) . \tag{6}$$

Thus the two decompositions

$$T_{x_0}P = \ker dJ(x_0) \oplus \mathrm{range}\, dJ(x_0)^*$$

and

$$T_{x_0}P = T_{x_0}(G \cdot x_0) \oplus [T_{x_0}(G \cdot x_0)]^\perp$$
$$= \mathrm{range}(- \mathbb{J} \circ dJ(x_0)^*) \oplus \ker(dJ(x_0) \circ \mathbb{J})$$

can be intersected to produce *Moncrief's decomposition* (see Moncrief, 1975b; Arms et al., 1975):

$$T_{x_0}P = [\mathrm{range}(- \mathbb{J} \circ dJ(x_0)^*)] \oplus [\mathrm{range}(dJ(x_0)^*)]$$
$$\oplus [\ker(dJ(x_0) \circ \mathbb{J}) \cap \ker(dJ(x_0))] . \tag{7}$$

Remark. If $J(x_0) \neq 0$ then analogues of (7) and many results below remain true. If $G_{J(x_0)} \equiv \{ g \in G \mid \mathrm{Ad}^*_{g^{-1}} J(x_0) = J(x_0) \} = G$, then all the results remain valid. If $G_{J(x_0)} \neq G$, we do not know if Theorem 3 below is valid.

We provide a brief glossary of symbols preceeding the bibliography to aid the reader.

3. The Relationship between Regular Points and Symmetries

Let $\mathscr{S}_{x_0}$ be the identity component of the isotropy group I_{x_0} and $\mathfrak{s}_{x_0}$ its Lie algebra; i.e.

$$\mathfrak{s}_{x_0} = \{\xi \in \mathfrak{g} \mid \xi_P(x_0) = 0\}.$$

We will call non-zero elements of $\mathfrak{s}_{x_0}$ *infinitesimal symmetries* of x_0.

Theorem 1. *x_0 is a regular point of J (i.e. $dJ(x_0)$ is surjective) if and only if x_0 has no infinitesimal symmetries.*

Proof. The map $dJ(x_0)$ is surjective iff $\xi \perp \text{range}\, dJ(x_0)$ implies $\xi = 0$ [in the infinite dimensional case assume $dJ(x_0)$ has closed range]. However, $\xi \perp \text{range}\, dJ(x_0)$ means $\langle \xi, dJ(x_0) \cdot v \rangle = 0$ for all v. By (1), this is equivalent to $\xi_P(x_0) = 0$. Thus $dJ(x_0)$ is surjective iff $\mathfrak{s}_{x_0} = \{0\}$. $\square$

Theorem 1 implies that $J^{-1}(0)$ is a manifold in a neighborhood of x_0 if $\mathfrak{s}_{x_0} = \{0\}$. The tangent space at x_0 is

$$T_{x_0} J^{-1}(0) = \ker dJ(x_0).$$

From (4) we notice that

$$\mathfrak{s}_{x_0} = \ker dJ(x_0)^* . \tag{8}$$

To deal with the structure of $J^{-1}(0)$ near a point x_0 at which $\mathfrak{s}_{x_0} \neq \{0\}$, we shall follow the *Liapunov-Schmidt procedure* from bifurcation theory. This proceeds as follows:

Let $\mathbb{P} : \mathfrak{g}^* \to \text{range}\, dJ(x_0)$ be the $(\ ,\ )_{x_0}$ orthogonal projection associated to the decomposition (5). Let $\mathscr{C} = J^{-1}(0)$ and let

$$\mathscr{C}_{\mathbb{P}} = \{x \in P \mid \mathbb{P}J(x) = 0\}.$$

Since $\mathbb{P}J : J \to \text{range}\, dJ(x_0)$ is a submersion at x_0, $\mathscr{C}_{\mathbb{P}}$ is a smooth manifold in an neighborhood of x_0 with

$$T_{x_0} \mathscr{C}_{\mathbb{P}} = \ker dJ(x_0).$$

Define

$$f : \mathscr{C}_{\mathbb{P}} \to \ker dJ(x_0)^\dagger$$

by

$$f(x) = (\text{Id} - \mathbb{P})(J(x)).$$

Thus

$$\mathscr{C} = f^{-1}(0). \tag{9}$$

Note that $f(x_0) = 0$ and $df(x_0) = 0$.

4. Identification of the Degeneracy Space of $d^2 f(x_0)$

In Sects. 5 and 6 we shall prove that $\mathscr{C} \cap S_{x_0}$ can be identified with the zero set of a homogeneous quadratic map. That proof, however, proceeds independently of a

detailed knowledge of the possible degeneracy of $d^2 f(x_0)$. On the other hand, the structure of this set depends on this degeneracy space.

For x in a neighborhood of x_0, let $I_x = \{g \in G | g \cdot x = x\}$ and $\mathscr{S}_x$ be the component of the identity, and let $\mathfrak{s}_x$ be its Lie algebra; thus

$$\mathfrak{s}_x = \{\xi \in g | \xi_P(x) = 0\}. \tag{10}$$

From the slice theorem it follows that there is a neighborhood V of x_0 such that if $x \in V$ then $\mathscr{S}_x$ is conjugate to a subgroup of $\mathscr{S}_{x_0}$. Denote by N_{x_0} those $x \in V$ such that $\mathscr{S}_x$ is conjugate to $\mathscr{S}_{x_0}$ itself, i.e. N_{x_0} consists of elements of the same symmetry type as x_0 (or elements of the same "orbit type"). It is well known that N_{x_0} is a smooth manifold near x_0 (see, for instance, Hermann, 1968). It is clear that $G \cdot x_0 \subset N_{x_0}$, but in many examples N_{x_0} is strictly larger.

Definition. Let $\mathscr{N}_{x_0} = N_{x_0} \cap S_{x_0}$ where S_{x_0} is the G-slice at x_0.

Lemma 1. (a) $\mathscr{N}_{x_0} = \{x \in S_{x_0} | \mathscr{S}_x = \mathscr{S}_{x_0}\} = \{x \in S_{x_0} | \xi_P(x) = 0 \text{ for all } \xi \in \mathfrak{s}_{x_0}\}$.

(b) $\mathscr{N}_{x_0}$ *is a smooth manifold if* S_{x_0} *is chosen sufficiently small and*

$$T_{x_0}\mathscr{N}_{x_0} = \{u \in \ker(dJ(x_0) \circ \mathbb{J}) | \langle d^2 J(x_0)(u, v), \xi \rangle = 0$$

$$\text{for all } v \in T_{x_0}P \text{ and } \xi \in \mathfrak{s}_{x_0}\}.$$

Proof. (a) follows from property (ii) of slices. Now $\mathscr{N}_{x_0}$ is the set of points of fixed symmetry type for the action of $\mathscr{S}_{x_0}$ on S_{x_0}, so it is a smooth manifold; its tangent space is obtained by linearizing the condition $\xi_P(x) = 0$ or equivalently, $\omega_x(\xi_P(x), v) = 0$ for all v; i.e.

$$\langle dJ(x) \cdot v, \xi \rangle = 0 \quad \text{for all} \quad v \in T_x P, \xi \in \mathfrak{s}_{x_0}. \quad \square$$

Remark. The second derivative in Lemma 1(b) is well defined (without using a connection) since the derivative of $\langle J(x), \xi \rangle$ vanishes at x_0 if $\xi \in \mathfrak{s}_{x_0}$.

Lemma 2. $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ *is a manifold (if* S_{x_0} *is chosen sufficiently small) with*

$$T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0}) = \ker dJ(x_0) \cap \ker(dJ(x_0) \circ \mathbb{J}).$$

(The third summand in Moncrief's decomposition (7).)

Proof. By (6)

$$\text{range}(-\mathbb{J} \circ dJ(x_0)^*) \subset \ker dJ(x_0)$$

so taking adjoints,

$$\text{range}(dJ(x_0)^*) \subset \ker(dJ(x_0) \circ \mathbb{J}). \tag{11}$$

Therefore

$$\ker dJ(x_0) + \ker(dJ(x_0) \circ \mathbb{J}) \supset \ker dJ(x_0) + \text{range} \, dJ(x_0)^* = T_{x_0}P.$$

Thus $\mathscr{C}_{\mathbb{P}}$ and S_{x_0} intersect transversally at x_0 so $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ is a manifold near x whose tangent space is the intersection of $T_{x_0}\mathscr{C}_{\mathbb{P}} = \ker dJ(x_0)$ and $T_{x_0}S_{x_0}$ $= \ker(dJ(x_0) \circ \mathbb{J})$. $\quad \square$

Lemma 3. *Let* $\xi \in \mathfrak{g}$, $\eta \in \mathfrak{g}$ *and* $v \in T_{x_0}P$.

(a) *We have the identity*

$$\langle d^2J(x_0) \cdot (\eta_P(x_0), v), \xi \rangle + \langle dJ(x_0) \cdot d\eta_P(x_0) \cdot v, \xi \rangle = \langle dJ(x_0) \cdot v, [\xi, \eta] \rangle.$$

(b) *(Gauge Invariance of* d^2J.*) If* $\xi \in \mathfrak{s}_{x_0}$, *then*

$$\langle d^2J(x_0) \cdot (\eta_P(x_0), v), \xi \rangle = 0$$

if either

(i) $v \in \ker dJ(x_0)$ *or*

(ii) $[\eta, \xi] = 0$.

Proof. By Ad* equivariance,

$$\langle J(gx), \xi \rangle = \langle J(x), \mathrm{Ad}_{g^{-1}} \xi \rangle \tag{12}$$

Differentiating (12) in x,

$$\langle dJ(gx) \cdot T\Phi_g \cdot v, \xi \rangle = \langle dJ(x) \cdot v, \mathrm{Ad}_{g^{-1}} \xi \rangle. \tag{13}$$

Differentiating (13) in g at $g = \mathrm{Id}$ and evaluating at x_0 gives (a). Part (b) follows at once from (a). □

Remark. Strictly speaking, (a) does not make sense without a connection. However, any one may be used, or a coordinate calculation will suffice. Our main interest is getting (b) which makes sense independent of a connection.

Lemma 4. *(*$\mathbb{J}$*-invariance of* $d^2J(x_0)$.*) For* $\xi \in \mathfrak{s}_{x_0}$ *and* $u, v \in T_{x_0}P$ *we have*

$$\langle d^2J(x_0) \cdot (u, v), \xi \rangle = \langle d^2J(x_0) \cdot (\mathbb{J}u, \mathbb{J}v), \xi \rangle.$$

Proof. From $\mathbb{J} \circ T\Phi_g = T\Phi_g \circ \mathbb{J}$ we get $T\mathbb{J} \circ T\xi_P = T\xi_P \circ \mathbb{J}$ (see Abraham and Marsden (1978), Proposition 4.1.28). Let $d\xi_P(x_0) : T_{x_0}P \to T_{x_0}P$ be the intrinsic derivative of ξ_P at the zero x_0. Then we get $\mathbb{J} \cdot d\xi_P(x_0) \cdot u = d\xi_P(x_0) \cdot \mathbb{J}u$. From (1) we get

$$\langle d^2J(x_0) \cdot (u, v), \xi \rangle = \omega_{x_0}(d\xi_P(x_0) \cdot u, v)$$

[each side of this equation is well defined independent of a connection since $\xi_P(x_0) = 0$]. Therefore,

$$\langle d^2J(x_0) \cdot (\mathbb{J}u, \mathbb{J}v), \xi \rangle = \omega_{x_0}(d\xi_P(x_0) \cdot \mathbb{J}u, \mathbb{J}v)$$
$$= \omega_{x_0}(\mathbb{J} \cdot d\xi_P(x_0) \cdot u, \mathbb{J}v).$$

Since $\mathbb{J}$ is symplectic, this becomes

$$\omega_{x_0}(d\xi_P(x_0) \cdot u, v) = \langle d^2J(x_0) \cdot (u, v), \xi \rangle. \quad □$$

Lemma 5. *Consider elements in* $\mathfrak{g}^*$ *with the same symmetry type as* x_0, *i.e. let*

$$\mathfrak{g}_{x_0}^* = \{ v \in \mathfrak{g}^* \mid \mathrm{Ad}_{g^{-1}}^* v = v \quad \text{for all} \quad g \in \mathscr{S}_{x_0} \}$$
$$= \{ v \in \mathfrak{g}^* \mid \langle v, [\xi, \eta] \rangle = 0 \quad \text{for all} \quad \xi \in \mathfrak{s}_{x_0}, \eta \in \mathfrak{g} \}.$$

Then

(a) $J : \mathscr{N}_{x_0} \to \mathfrak{g}_{x_0}^*$ *(i.e.* $J(\mathscr{N}_{x_0}) \subset \mathfrak{g}_{x_0}^*$),

(b) $dJ(x_0) : T_{x_0}\mathscr{N}_{x_0} \to \mathfrak{g}_{x_0}^*$, *and*

(c) $dJ(x_0)^\dagger : \mathfrak{g}_{x_0}^* \to T_{x_0}\mathscr{N}_{x_0}$.

Proof. (a) follows from equivariance and (b) follows from (a). We shall prove (c). By equivariance, for $g \in G$ and $\mu \in \mathfrak{g}^*$,

$$(\mu, J(gx_0))_{x_0} = 0.$$

Differentiating in g:

$$(\mu, dJ(x_0) \cdot \xi_P(x_0))_{x_0} = 0$$

i.e.

$$\langle\!\langle dJ(x_0)^\dagger \mu, \xi_P(x_0) \rangle\!\rangle_{x_0} = 0$$

for all $\xi \in \mathfrak{g}$. Thus $dJ(x_0)^\dagger \mu \perp T_{x_0}(G \cdot x_0)$, so $dJ(x_0)^\dagger \mu \in T_{x_0}S_{x_0}$. It remains to show that

$$\langle d^2 J(x_0)(dJ(x_0)^\dagger \mu, u), \xi \rangle = 0$$

if $\xi \in \mathfrak{s}_{x_0}$, $\mu \in \mathfrak{g}_{x_0}^*$ and $u \in T_{x_0}P$; see Lemma 1(b). By Lemma 4, this is equivalent so showing that

$$\langle d^2 J(x_0)(-\mathbb{J} \circ dJ(x_0)^\dagger \mu, v), \xi \rangle = 0$$

for all $\xi \in \mathfrak{s}_{x_0}$, $\mu \in \mathfrak{g}_{x_0}^*$ and $v \in T_{x_0}P$. By Lemma 3(b)(ii), this will be true if $-\mathbb{J} \circ dJ(x_0)^\dagger \mu = \eta_P(x_0)$ where $[\eta, \xi] = 0$. Let $\eta \in \mathfrak{g}$ correspond to μ via $(\,,\,)_{x_0}$; i.e. $\langle \sigma, \eta \rangle = (\mu, \sigma)_{x_0}$ for all $\sigma \in \mathfrak{g}^*$. Now for $g \in \mathscr{S}_{x_0}$, $(\mu, \sigma)_{x_0} = (\mathrm{Ad}^*_{g^{-1}}\mu, \sigma)_{x_0} = (\mu, \mathrm{Ad}^*_g \sigma)_{x_0}$ [by the assumed invariance of $(\,,\,)_{x_0}$], so $\langle \sigma, \eta \rangle = \langle \mathrm{Ad}^*_g \sigma, \eta \rangle$. Thus $\mathrm{Ad}_g \eta = \eta$, so $[\eta, \xi] = 0$. Also, $dJ(x_0)^\dagger \cdot \mu = dJ(x_0)^* \cdot \eta$ from the definitions. $\square$

Note that (c) is equivalent to saying $dJ(x_0)^*$ maps $\mathfrak{g}_{x_0}$ to $T_{x_0}\mathscr{N}_{x_0}$ where $\mathfrak{g}_{x_0} = \{\eta \in \mathfrak{g} \mid [\eta, \mathfrak{s}_{x_0}] = 0\}$.

Lemma 6. $\mathscr{N}_{x_0} \cap \mathscr{C}_\mathbb{P}$ *is a smooth manifold near x_0 with*

$$T_{x_0}(\mathscr{N}_{x_0} \cap \mathscr{C}_\mathbb{P}) = \ker dJ(x_0) \cap T_{x_0}\mathscr{N}_{x_0}$$

where $T_{x_0}\mathscr{N}_{x_0}$ is given by Lemma 1(b).

Proof. $\mathbb{P}J$ maps $\mathscr{N}_{x_0}$ to $\mathbb{P}\mathfrak{g}_{x_0}^*$ by Lemma 5(a). Since $\mathscr{N}_{x_0} \cap \mathscr{C}_\mathbb{P} = (\mathbb{P}J)^{-1}(0)$, the lemma will be proved if we can show that $\mathbb{P}J$ is a submersion. Now the projection map is given by

$$\mathbb{P} = dJ(x_0) \circ [dJ(x_0)^\dagger \circ dJ(x_0)]^{-1} \circ dJ(x_0)^\dagger,$$

where $dJ(x_0)^\dagger \circ dJ(x_0)$ is regarded as an isomorphism of range $dJ(x_0)^\dagger$ to itself. By Lemma 5, if $\mu \in \mathbb{P}\mathfrak{g}_{x_0}^*$ then we can write

$$\mu = dJ(x_0) \cdot v,$$

where

$$v = [dJ(x_0)^\dagger \circ dJ(x_0)]^{-1} \cdot dJ(x_0)^\dagger v$$

for some $v \in \mathfrak{g}_{x_0}^*$. By Lemma 5(b) and (c), $v \in T_{x_0} \mathcal{N}_{x_0}$. Thus $\mu = \mathbb{P} dJ(x_0) \cdot v$ for $v \in T_{x_0} \mathcal{N}_{x_0}$, so $\mathbb{P}J$ is a submersion at x_0. $\square$

Lemma 7.
$$\mathcal{N}_{x_0} \cap \mathscr{C}_{\mathbb{P}} = \mathcal{N}_{x_0} \cap \mathscr{C}.$$

Proof. Notice that if $\xi \in \mathfrak{s}_{x_0}$ and $x \in \mathcal{N}_{x_0}$ then $dJ(x)^* \xi = 0$. Thus the x-derivative of $\langle J(x), \xi \rangle$ vanishes identically on $\mathcal{N}_{x_0}$. Therefore, $0 = \langle J(x), \xi \rangle$ for $x \in \mathcal{N}_{x_0}$ and $\xi \in \ker dJ(x_0)^*$. Equivalently, $0 = (J(x), v)_{x_0}$ for $x \in \mathcal{N}_{x_0}$ and $v \in \ker dJ(x_0)^\dagger$. But this means that $(I - \mathbb{P})(J(x)) = 0$, so $J(x) = 0$ on $\mathcal{N}_{x_0} \cap \mathscr{C}_{\mathbb{P}}$. $\square$

Thus, solutions of $J(x) = 0$ with the same symmetry type as x_0 form a smooth manifold. The proof of Lemma 7 also establishes the following:

Lemma 8. *Consider the map* $f : \mathscr{C}_{\mathbb{P}} \cap S_{x_0} \rightarrow \ker dJ(x_0)^\dagger$ *defined by*
$$f(x) = (\mathrm{Id} - \mathbb{P})(J(x))$$

(as in Sect. 3). Then $\mathcal{N}_{x_0} \cap \mathscr{C}_{\mathbb{P}}$ *is a manifold of critical points of* f.

The *degeneracy space* of f at a critical point x is defined to be
$$\{u \in T_x(\mathscr{C}_{\mathbb{P}} \cap S_{x_0}) \mid d^2 f(x)(u, v) = 0 \quad \text{for all} \quad v \in T_x(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})\}.$$

Here is our second main result:

Theorem 2. $\mathcal{N}_{x_0} \cap \mathscr{C}$ *is a manifold of non-degenerate critical points for* $f : \mathscr{C}_{\mathbb{P}} \cap S_{x_0} \rightarrow \ker dJ(x_0)^\dagger$ *in the sense that*
    (a) *each* $x \in \mathcal{N}_{x_0} \cap \mathscr{C}$ *satisfies* $df(x) = 0$
and
    (b) *the degeneracy space for* f *at* $x \in \mathcal{N}_{x_0} \cap \mathscr{C}$ *equals* $T_x(\mathcal{N}_{x_0} \cap \mathscr{C})$.

Proof. By Lemmas 7 and 8, (a) holds. Also, we have
$$\langle d^2 f(x_0) \cdot (u, v), \xi \rangle = \langle d^2 J(x_0) \cdot (u, v), \xi \rangle \tag{14}$$

for all $\xi \in \mathfrak{s}_x$, and $u, v \in T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})$. If $u \in T_{x_0}(\mathcal{N}_{x_0} \cap \mathscr{C})$ then (14) vanishes by Lemma 1(b), so u lies in the degeneracy space for f at x_0. Conversely, suppose $u \in T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})$ and u lies in the degeneracy space for f at x_0; thus (14) vanishes for all $\xi \in \mathfrak{s}_{x_0}$ and all $v \in T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})$. If $w_1 \in \text{range}(-\mathbb{J} \circ dJ(x_0)^*)$ then $\langle d^2 J(x_0)(u, w_1), \xi \rangle = 0$ by gauge invariance [Lemma 3(b)(i)]. Now the orthogonal complement of $T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0}) \oplus \text{range}(-\mathbb{J} \circ dJ(x_0)^*)$ is $\text{range}(dJ(x_0)^*)$ by Moncrief's decomposition (7). For $w_2 \in \text{range}(dJ(x_0)^*)$, $\langle d^2 J(x_0)(u, w_2), \xi \rangle = 0$ by Lemma 3(b)(i) and $\mathbb{J}$-invariance. Thus $\langle d^2 J(x_0)(u, w), \xi \rangle = 0$ for all $v \in T_{x_0} P$, so $u \in T_{x_0} \mathcal{N}_{x_0}$. Since all points of $\mathcal{N}_{x_0} \cap \mathscr{C}$ have the same symmetry, the same calculation works at any point $x \in \mathcal{N}_{x_0} \cap \mathscr{C}$. $\square$

The following special case is useful to note:

Corollary. *If* $\dim \ker dJ(x_0)^* = 1$, *then* $\mathcal{N}_{x_0} \cap \mathscr{C}$ *is a nondegenerate critical manifold for* f *and so* $\mathscr{C} \cap S_{x_0}$ *is the product of a cone with* $(\mathcal{N}_{x_0} \cap S_{x_0})$ *and hence* $\mathscr{C}$ *itself is a cone* $\times (\mathcal{N}_{x_0} \cap S_{x_0}) \times (G/I_{x_0})$, *in a neighborhood of* x_0.

The conclusion that $\mathscr{C} \cap S_{x_0}$ is a cone $\times \mathcal{N}_{x_0} \cap S_{x_0}$ follows from the parametrized Morse lemma; cf. Bott (1954).

Example. Let $H : \mathbb{R}^2 \to \mathbb{R}$ be a smooth hamiltonian with a critical point at the origin. Suppose H has all its orbits periodic of the same period. Then H *has a nondegenerate critical point* at $(0, 0)$. This follows from the corollary to Theorem 2 by using $G = S^1$ acting by the flow of X_H. In fact, as we will see in Lemma 17 below, there are symplectic coordinates near $(0, 0)$ in which H is homogeneous quadratic; thus H is a harmonic oscillator.

5. Quadratic Momentum Mappings

Theorem 3 below proves that if J is quadratic then there is a diffeomorphism of a neighborhood of x_0 that maps $\mathscr{C} \cap S_{x_0}$ to the set of zeros of $u \mapsto (I - \mathbb{P}) d^2 J(x_0)(u, u)$ for $u \in T_{x_0} S_{x_0}$. Moreover, this diffeomorphism respects the information derived in Theorem 2 (see Theorems 4 and 4').

The case treated in this section applies to some interesting quadratic examples, such as gauge theories, as we shall see in Sect. 8. The proofs really only require that $(\mathrm{Id} - \mathbb{P}) J$ restricted to S_{x_0} is quadratic. This observation is applied to the constraint equations in general relativity near a spacetime having k spacelike Killing symmetries (the supermomentum constraint is quadratic, but the Hamiltonian constraint is not).

The following elementary example shows that in general, zero sets of quadratic maps need not have conical singularities; thus, even when J is quadratic, it is not obvious that $J^{-1}(0)$ has conical singularities.

Example. Let $F : \mathbb{R}^3 \to \mathbb{R}^2$; $F(x, y, z) = (x^2 + y^2 - z, x^2 - yz)$ and consider $F^{-1}(0, 0)$ near $(0, 0, 0)$. This set does not have a conical singularity, but rather a cusp singularity $z = y^2 + y^3 + \ldots$, $x = y^{3/2} + \ldots$ near $(0, 0, 0)$.

The components of the momentum mapping are linked together in a nontrivial way by Ad*-equivariance. These extra properties actually rule out singularities of higher order than quadratic.

The key to the constructions in this section is the Kuranishi map used by Atiyah et al. (1978). We are grateful to L. Nirenberg and I. Singer for suggesting its consideration[2]. Kuranishi (1965) originally used this map (called F below) in a study of deformations of complex structures.

Our assumptions are as in Sect. 2, except now let us assume that P is (open in) a linear space and J is quadratic; i.e., let $J(x_0) = 0$ and assume that in a neighborhood of $x_0 \in P$,

$$J(x_0 + h) = dJ(x_0) \cdot h + Q(h),$$

where $Q(h) = \frac{1}{2} B(h, h)$, and $B(u, v) = d^2 J(x_0)(u, v)$.

The map $\Delta = dJ(x_0) \circ dJ(x_0)^\dagger$ is an isomorphism of range $dJ(x_0)$ to itself. Let $G = \Delta^{-1} \circ \mathbb{P} : \mathfrak{g}^* \to \text{range } dJ(x_0)$, the "Green's function" for Δ. The crucial map we deal with is

$$F : P \to P$$

$$F(x) = x + dJ(x_0)^\dagger \circ G \circ Q(h), \tag{15}$$

where $h = x - x_0$.

---

2   Keep in mind the fact that the singularities found here do not occur in the Euclidean theory studied by Atiyah et al. (1978)

Lemma 9. *F is a local diffeomorphism of a neighborhood of x_0 to a neighborhood of x_0.*

Proof. $DF(x_0) = \mathrm{Id}$. $\square$

Lemma 10. *F takes a neighborhood of x_0 in $\mathscr{C}_{\mathbb{P}}$ to a neighborhood of x_0 in $\{x_0\} + \ker dJ(x_0)$.*

Proof. We need to show that

$$x \in \mathscr{C}_{\mathbb{P}} \Leftrightarrow F(x) - x_0 \in \ker dJ(x_0).$$

Now

$$dJ(x_0) \cdot (F(x) - x_0) = dJ(x_0) \cdot h + dJ(x_0) \circ dJ(x_0)^\dagger \circ G \circ Q(h).$$

But from the definition of G we have $dJ(x_0) \circ dJ(x_0)^\dagger \circ G = \mathbb{P}$ and so

$$
\begin{aligned}
dJ(x_0)(F(x) - x_0) &= dJ(x_0) \cdot h + \mathbb{P}Q(h) \\
&= \mathbb{P}(dJ(x_0) \cdot h + Q(h)) \\
&= \mathbb{P}(J(x_0 + h)). \quad \square
\end{aligned}
$$

We now explicitly choose S_{x_0} to be the affine slice; i.e. S_{x_0} is a ball in $\{x_0\} + \ker(dJ(x_0) \circ \mathbb{J})$.

Lemma 11. *F maps S_{x_0} to S_{x_0}.*

Proof. For $\xi \in \mathfrak{g}$ and $x \in S_{x_0}$,

$$
\begin{aligned}
\langle dJ(x_0) \circ \mathbb{J}(F(x) - x_0), \xi \rangle &= \langle dJ(x_0) \circ \mathbb{J}(h + dJ(x_0)^\dagger \circ G \circ Q(h)), \xi \rangle \\
&= \langle dJ(x_0) \circ \mathbb{J}(h), \xi \rangle
\end{aligned}
$$

because $\mathrm{range}(\mathbb{J} \circ dJ(x_0)^\dagger) \subset \ker dJ(x_0)$ implies that

$$\langle dJ(x_0) \circ \mathbb{J} \circ dJ(x_0)^\dagger \circ G \circ Q(h), \xi \rangle = 0. \quad \square$$

From Lemmas 10 and 11 we get

Lemma 12. *F restricts to a local diffeomorphism of $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ to*

$$\{x_0\} + \ker dJ(x_0) \cap \ker(dJ(x_0) \circ \mathbb{J}).$$

Note that this proof also shows that $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ is a manifold; cf. Lemma 2.

Lemmas 9–12 do not use the fact that J is quadratic; indeed, $Q(h)$ can be replaced by the remainder $R(h)$ in the Taylor expansion. Now we shall use the fact that J is quadratic. [As we already remarked, it is noteworthy that we only need to assume that $(\mathrm{Id} - \mathbb{P}) \circ J$ is quadratic.] Our main result of this section is as follows.

Theorem 3. *The local diffeomorphism F maps $\mathscr{C} \cap S_{x_0}$ locally $1-1$ onto the cone*

$$
\begin{aligned}
C_{x_0} = \{x_0\} + \{u \in \ker dJ(x_0) \cap \ker(dJ(x_0) \circ \mathbb{J})| \\
\langle d^2 J(x_0)(u, u), \xi \rangle = 0 \quad \text{for all} \quad \xi \in \ker dJ(x_0)^*\}.
\end{aligned}
$$

Thus, locally $J^{-1}(0) = \mathscr{C} \approx C_{x_0} \times G/I_{x_0}$.

Proof. From Lemma 12, we need to show that for $x \in \mathscr{C}_{\mathbb{P}} \cap S_{x_0}$,

$$(\mathrm{Id} - \mathbb{P})Q(F(x) - x_0) = 0 \Leftrightarrow (\mathrm{Id} - P)J(x) = 0. \tag{16}$$

Keep in mind that $(\mathrm{Id} - \mathbb{P})J(x) = (\mathrm{Id} - \mathbb{P})Q(h)$. By Lemma 10, $F(x) - x_0 \in \ker dJ(x_0)$ so letting $\hat{\mathbb{P}}$ be the orthogonal projection onto $\ker dJ(x_0)$, we have

$$\begin{aligned} F(x) - x_0 &= \hat{\mathbb{P}}(F(x) - x_0) \\ &= \hat{\mathbb{P}}(h + dJ(x_0)^\dagger \circ G \circ Q(h)) \\ &= \hat{\mathbb{P}}h. \end{aligned}$$

Therefore,

$$\begin{aligned} (\mathrm{Id} - \mathbb{P})Q(F(x) - x_0) &= (\mathrm{Id} - \mathbb{P})Q(\hat{\mathbb{P}}h) \\ &= (\mathrm{Id} - \mathbb{P})Q(h - (\mathrm{Id} - \hat{\mathbb{P}})h). \end{aligned}$$

Since $(\mathrm{Id} - \mathbb{P})J$ is quadratic, this becomes

$$(\mathrm{Id} - \mathbb{P})Q(F(x) - x_0)$$
$$= (\mathrm{Id} - \mathbb{P})\{Q(h) - B(h, (\mathrm{Id} - \hat{\mathbb{P}})h) + \tfrac{1}{2}B((\mathrm{Id} - \hat{\mathbb{P}})h, (\mathrm{Id} - \hat{\mathbb{P}})h)\}. \tag{17}$$

By $\mathbb{J}$-invariance of $d^2J(x_0)$,

$$(\mathrm{Id} - \mathbb{P})(B(h, (\mathrm{Id} - \hat{\mathbb{P}})h)) = (\mathrm{Id} - \mathbb{P})(B(\mathbb{J}h, \mathbb{J}(\mathrm{Id} - \hat{\mathbb{P}})h)).$$

Now $h \in \ker(dJ(x_0) \circ \mathbb{J})$ since $x \in S_{x_0}$, so $\mathbb{J}h \in \ker(dJ(x_0))$. Also, $\mathbb{J}(\mathrm{Id} - \hat{\mathbb{P}})h \in \mathrm{range}\,\mathbb{J} \circ dJ(x_0)^*$, so $\mathbb{J}(\mathrm{Id} - \hat{\mathbb{P}})h = \eta_P(x_0)$ for some $\eta \in \mathfrak{g}$. Therefore, by gauge invariance [Lemma 3(b)(i)],

$$(\mathrm{Id} - \mathbb{P})B(h, (\mathrm{Id} - \hat{\mathbb{P}})h) = 0.$$

A similar argument together with the fact that $\mathrm{range}(\mathbb{J} \circ dJ(x_0)^*) \subset \ker dJ(x_0)$ shows that

$$(\mathrm{Id} - \mathbb{P})B((\mathrm{Id} - \hat{\mathbb{P}})h, (\mathrm{Id} - \hat{\mathbb{P}})h) = 0.$$

Thus from (17),

$$(\mathrm{Id} - \mathbb{P})Q(F(x) - x_0) = (\mathrm{Id} - \mathbb{P})Q(h),$$

which gives (16). $\quad\square$

Next we show that F maps the degeneracy manifold to the affine space determined by the degeneracy space of the second derivative.

Theorem 4. *F maps $\mathscr{N}_{x_0} \cap \mathscr{C}$ locally $1-1$ onto the affine space*

$$\{x_0\} + \{u \in \ker(dJ(x_0) \circ \mathbb{J}) \cap \ker(dJ(x_0)) \mid \langle d^2J(x_0)(u, v), \xi \rangle = 0$$

for all $\xi \in \ker dJ(x_0)^*$ and all $v \in T_{x_0}P\} = \mathscr{N}_{x_0} \cap \mathscr{C}_{x_0}$.

Proof. We already know that for $x \in S_x \cap \mathscr{C}$, $u = F(x) - x_0 \in \ker(dJ(x_0) \circ \mathbb{J}) \cap \ker(dJ(x_0))$ by Lemmas 10 and 11. Now suppose that $x \in \mathscr{N}_{x_0}$. Then by definition of F and $\mathbb{J}$-invariance,

$$\begin{aligned} \langle d^2J(x_0)(u, v), \xi \rangle &= \langle d^2J(x_0)(h, v), \xi \rangle + \langle d^2J(x_0)(dJ(x_0)^\dagger \circ G \circ Q(h), v), \xi \rangle \\ &= \langle d^2J(x_0)(h, v), \xi \rangle \\ &\quad + \langle d^2J(x_0)(\mathbb{J} \circ dJ(x_0)^\dagger \circ G \circ Q(h), \mathbb{J}v), \xi \rangle. \end{aligned}$$

Now $\mathcal{N}_{x_0}$ is an affine space, so $h \in T_{x_0}\mathcal{N}_{x_0}$. Thus by Lemma 5, $G \circ Q(h) \in \mathfrak{g}_{x_0}^*$. Thus by Lemma 3(b)(ii), the second term vanishes. The first term vanishes by Lemma 1(b). Thus F maps $\mathcal{N}_{x_0} \cap \mathscr{C}$ into the stated space. Since these spaces are tangent at x_0 and F is a local diffeomorphism, the result follows. $\quad\square$

Theorem 4 can be generalized as follows: let $\mathscr{H} \subset \mathscr{S}_{x_0}$ be a Lie subgroup with Lie algebra $\mathfrak{h}$. Define $\mathscr{A}_{\mathscr{H}} = \{x \in P \,|\, dJ(x)^*\xi = 0 \text{ for all } \xi \in \mathfrak{h}\}$, an affine subspace through x_0, let $\mathscr{B}_{\mathscr{H}} = \mathscr{A}_{\mathscr{H}} \cap S_{x_0}$ and let C_{x_0} be the cone in Theorem 3.

Theorem 4′. *We have* $F(\mathscr{A}_{\mathscr{H}}) = \mathscr{A}_{\mathscr{H}}$, $F(\mathscr{B}_{\mathscr{H}}) = \mathscr{B}_{\mathscr{H}}$ *and so* $F(\mathscr{B}_{\mathscr{H}} \cap \mathscr{C}) = \mathscr{B}_{\mathscr{H}} \cap C_{x_0}$ *in a neighborhood of* x_0.

Thus, F preserves symmetry type for any subgroup of $\mathscr{S}_{x_0}$. Theorem 4 is the case $\mathscr{H} = \mathscr{S}_{x_0}$. [Note that as above, we need only assume $(\mathrm{Id} - \mathbb{P})J|S_{x_0}$ is quadratic.]

Proof of Theorem 4′. Since $\mathscr{A}_{\mathscr{H}}$ is affine, $x \in \mathscr{A}_{\mathscr{H}}$ if and only if $\langle d^2J(x_0)(h,v), \xi \rangle = 0$ for all $v \in T_{x_0}P$ and $\xi \in \mathfrak{h}$, where $h = x - x_0$. Then if $u = F(x) - x_0$,
$$\langle d^2J(x_0)(u,v), \xi \rangle = \langle d^2J(x_0)(h,v), \xi \rangle + \langle d^2J(x_0)(dJ(x_0)^\dagger \circ G \circ Q(h), v), \xi \rangle.$$
The first term vanishes by assumption. Since $\mathfrak{h} \subset \mathfrak{s}_{x_0}$, the argument in the proof of Theorem 4 shows that the second term vanishes as well. Thus $F(\mathscr{A}_{\mathscr{H}}) \subset \mathscr{A}_{\mathscr{H}}$. Since $DF(x_0) = \mathrm{Id}$, we have local equality. The remaining assertions follow from Lemma 11 and Theorem 3. $\quad\square$

6. General Momentum Mappings

We shall now obtain the conical structure of $J^{-1}(0)$ without the assumption that J is quadratic. The methods of this section do not rely on those in the previous one. However, the methods here have the disadvantage that they are more difficult to implement in the infinite dimensional case. The examples given in Sect. 8 use the results of the previous section rather than this one. In particular, the method used here relies on the Darboux theorem; a version of this theorem that is useful in infinite dimensional problems is technically quite complicated (see Marsden, 1980, Sect. 1).

The strategy in this section is to show that $j = (\mathrm{Id} - \mathbb{P})J|(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})$ is the momentum map for the action of $\mathscr{S}_{x_0}$. This action has a fixed point at x_0 and the Darboux Theorem can be used to find symplectic coordinates in which the action is linear and hence in which j is quadratic.

Lemma 13. $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ *is a symplectic submanifold of* P.

Proof. In Lemma 2 we showed that $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ is a manifold. It suffices to show that the symplectic form restricted to $T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})$ is non-degenerate. Let $u \in \ker dJ(x_0) \cap \ker(dJ(x_0) \circ \mathbb{J})$ and suppose that
$$\omega_{x_0}(u,v) = \langle\!\langle u, \mathbb{J}v \rangle\!\rangle = 0$$
for all v in the same space. But $\langle\!\langle u, \mathbb{J}w_1 \rangle\!\rangle = 0$ for $w_1 \in \mathrm{Range}\,\mathbb{J} \circ dJ(x_0)^*$ and $\langle\!\langle u, \mathbb{J}w_2 \rangle\!\rangle = 0$ for $w_2 \in \mathrm{Range}\,dJ(x_0)^*$. Thus, by Moncrief's decomposition (7), $\langle\!\langle u, \mathbb{J}v \rangle\!\rangle = 0$ for all $v \in T_{x_0}\mathbb{P}$ and so $u = 0$. $\quad\square$

The preceeding lemma is implicit in Marsden-Weinstein (1974).

Lemma 14. *Let $R \subset P$ be a symplectic submanifold and $H \subset G$ a subgroup leaving R invariant. Then the momentum map of the H action on R is $j = \pi_{\mathfrak{h}} \circ J \circ i_R$ where $i_R : R \to P$ is inclusion, and $\pi_{\mathfrak{h}} : \mathfrak{g}^* \to \mathfrak{h}^*$ is the natural projection. If the momentum map J of the action is Ad^*-equivariant, then so is j.*

This is a straightforward verification from the definition of momentum mappings (1).

Lemma 15. $\mathscr{S}_{x_0}$ *leaves $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ invariant and has a fixed point at x_0.*

Proof. First of all, $\mathscr{S}_{x_0}$ leaves S_{x_0} invariant by property (i) of the slice (see Sect. 2). Next, suppose that $x \in \mathscr{C}_{\mathbb{P}}$; i.e. $\mathbb{P}J(x) = 0$. Then for $g \in \mathscr{S}_{x_0}$, $\mathbb{P}J(gx) = \mathbb{P}\,\text{Ad}^*_{g^{-1}}\,J(x)$. Thus the proof is complete if we can show that $\mathbb{P}\,\text{Ad}^*_{g^{-1}} = \text{Ad}^*_{g^{-1}}\,\mathbb{P}$. But this follows from the facts that $\mathbb{P}$ is an $(\ ,\ )_{x_0}$ orthogonal projection and that $(\ ,\ )_{x_0}$ is $\text{Ad}^*_{g^{-1}}$-invariant for $g \in \mathscr{S}_{x_0}$ (see assumption 5 in Sect. 2). $\quad\square$

Combining Lemmas 13–15, we obtain the following:

Lemma 16. *The momentum mapping for the action of $\mathscr{S}_{x_0}$ on $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ is $j = (I - \mathbb{P})J \,|\, \mathscr{C}_{\mathbb{P}} \cap S_{x_0}$.*

Lemma 17. *There is a symplectic change of coordinates on $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ (the coordinate change has derivative the identity at x_0) in which the action of $\mathscr{S}_{x_0}$ is linear, and hence j is homogeneous quadratic, since $j(x_0) = 0$ and $dj(x_0) = 0$.*

Proof (see Weinstein, 1977, p. 24, last paragraph). Let Ω_0 be the symplectic form on $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ (Lemma 13) and let Ω_1 be the (constant) symplectic form on $T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0})$. Let $\exp : T_{x_0}(\mathscr{C}_{\mathbb{P}} \cap S_{x_0}) \to \mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ be the exponential map associated to the metric $\langle\!\langle\ ,\ \rangle\!\rangle$. Since the action Φ_g preserves $\langle\!\langle\ ,\ \rangle\!\rangle$ then for $g \in \mathscr{S}_{x_0}$,

$$\Phi_g \circ \exp = \exp \circ T\Phi_g(x_0). \tag{18}$$

By the proof of Darboux' theorem (Weinstein, 1977, pp. 22–23) there is a local diffeomorphism f_1 of $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ to itself such that

$$f_1^* \exp_* \Omega_1 = \Omega_0 \tag{19}$$

and $f_1 \circ \Phi_g = \Phi_g \circ f_1$. Let $f_2 = \exp^{-1} \circ f_1$; then by (18)

$$f_2 \circ \Phi_g \circ f_2^{-1} = T\Phi_g(x_0). \tag{20}$$

Also, by (19), f_2 gives a symplectic chart. Thus (20) shows that Φ_g is a linear action in the new coordinates. Since the new coordinates are symplectic, (1) shows that j is quadratic. $\quad\square$

Thus, we have proved the following:

Theorem 5. *There is a local diffeomorphism F of $\mathscr{C}_{\mathbb{P}} \cap S_{x_0}$ that takes $\mathscr{C} \cap S_{x_0}$ locally $1-1$ onto the cone*

$$C_{x_0} = \{u \in \ker dJ(x_0) \cap \ker(dJ(x_0) \circ \mathbb{J}) \,|$$

$$\langle d^2 J(x_0)(u, u), \xi \rangle = 0 \quad \text{for all} \quad \xi \in \ker dJ(x_0)^* \}.$$

Thus $\mathscr{C} \approx C_{x_0} \times (G/I_{x_0})$.

7. Discussion

This section discusses several ways of interpreting and using the results obtained.

a) Bifurcations

Following Smale (1970), we define a *bifurcation point* of a map $f : M \to N$ to be a point $y_0 \in N$ such that $f^{-1}(y)$ changes topological type as y varies in some neighborhood of y_0 [see Marsden (1978) for a discussion of this definition]. If some uniformity condition on the derivative of f is made (e.g. if M is compact) then no regular point can be a bifurcation point (by the implicit function theorem). Theorem 1 shows that for momentum maps, the only candidate bifurcation points are images of points with symmetry group of dimension ≥ 1.

In topology and mechanics one is interested in how the level sets of the momentum map fit together. The reason is simply that any Hamiltonian system leaves these sets invariant, so a knowledge of the topology of those sets can yield information about the flow. Complications in the topology occur precisely at bifurcation points, by definition. In determining this complication, the structure of $J^{-1}(\mu_0)$ *at* the bifurcation point μ_0 is the crucial information needed. Our theorems determine precisely the local structure of $J^{-1}(0)$ when 0 is a bifurcation point.

b) Linearization Stability

Let $F : M \to N$ be a smooth map and let $x_0 \in M$ and $y_0 \in F(x_0)$. Let $T_{x_0}F : T_{x_0}M \to T_{y_0}N$ denote the tangent (derivative) of F. A vector $h \in \ker T_{x_0}F$ is called *integrable* if it is tangent to a C^1 curve $x(\lambda)$ in $F^{-1}(y_0)$. We call F *linearization stable* at x_0 if every $h \in \ker T_{x_0}F$ is integrable. This notion arose from perturbation theory in general relativity and was first introduced in Fisher and Marsden (1973). The work was motivated by Brill and Deser (1973). The connection with symmetries was made by Moncrief (1975a). If one seeks solutions x of $F(x) = y_0$ near to x_0, and if F is linearization stable at x_0, these solutions can be obtained in the form $x = x_0 + \lambda h + O(\lambda^2)$ for any $h \in \ker T_{x_0}F$.

If y_0 is a regular value of F, then F is linearization stable at x_0 by the implicit function theorem. Suppose, however, that y_0 is not a regular value of F. Then if $h \in \ker T_{x_0}F$ is integrable, it satisfies the necessary second order condition

$$\langle l, d_{x_0}^2 F \cdot (h, h) \rangle = 0$$

for any $l \in T_{y_0}N$ that annihilates the range of $T_{x_0}F$. (The second derivative on l is well defined since $\langle l, T_{x_0}F \cdot h \rangle = 0$.) This is obtained by differentiating $F(x(\lambda)) = y_0$ twice.

Theorems 3 and 5 imply that if $F = J$ is an Ad*-equivariant momentum mapping (associated with a group admitting a slice), and $J(x_0) = 0$, then the above second order condition is not only *necessary* for integrability, of $h \in \ker dJ(x_0)$, but is *sufficient* as well. This is basically another way of saying that the singularities in the set $J^{-1}(0)$ are homogeneous quadratic.

c) Reduction

Singularities can occur when one *reduces* a Hamiltonian system with symmetries. Suppose that (P, ω) is a symplectic manifold and that G acts on P and has an Ad*-equivariant momentum map $J : P \to \mathfrak{g}^*$. The first step in reduction (see Marsden and Weinstein, 1974) is to study the level surfaces of J. The proof of Theorem 1 shows that if $J(x_0) = \mu_0$, then x_0 is a regular point of J if and only if x_0 has no infinitesimal symmetries. In the regular case, $J^{-1}(x_0)$ is a manifold near x_0 and one can then proceed with reduction to get the symplectic manifold $P_{\mu_0} = J^{-1}(\mu_0)/G_{\mu_0}$. If, however, $\mathfrak{s}_{x_0} = \ker dJ(x_0)^* \neq \{0\}$ then singularities in $J^{-1}(\mu_0)$ will occur at x_0 (unless the cone C_{x_0} described in Theorem 3 happens to be a single point), and so the usual method of reduction does not apply. However, the set $\mathcal{N}_{x_0} \cap \mathscr{C} = \mathcal{N}_{x_0} \cap \mathscr{C}_\mathbb{P}$ $= N_{x_0} \cap S_{x_0} \cap \mathscr{C}$ is a manifold near x_0 (see Lemmas 6 and 7) with tangent space

$$T_{x_0}(\mathcal{N}_{x_0} \cap \mathscr{C}) = \ker dJ(x_0) \cap T_{x_0}\mathcal{N}_{x_0},$$

where $T_{x_0}\mathcal{N}_{x_0}$ is given by Lemma 1(b). In fact, it is a symplectic submanifold of P.

Lemma 18. *Assume $J(x_0) = 0$. Then $\mathcal{N}_{x_0} \cap \mathscr{C}$ is a symplectic submanifold of P.*

Proof. It suffices to show that the symplectic forms restricted to $T_{x_0}(\mathcal{N}_{x_0} \cap \mathscr{C})$ is nondegenerate. Let $u \in T_{x_0}(\mathcal{N}_{x_0} \cap \mathscr{C})$ $= \{u \in (\ker dJ(x_0) \cap \ker dJ(x_0) \circ \mathbb{J}) | \langle d^2J(x_0)(u, v), \xi \rangle = 0 \text{ for all } v \in T_{x_0}P \text{ and } \xi \in \mathfrak{s}_{x_0}\}$. We want to show that $\omega_{x_0}(u, v) = \langle\!\langle u, \mathbb{J}v \rangle\!\rangle = 0$ for all $v \in T_{x_0}(\mathcal{N}_{x_0} \cap \mathscr{C})$ implies that $u = 0$. To see this note that if $u \in T_{x_0}(\mathcal{N}_{x_0} \cap \mathscr{C})$ then $\mathbb{J}u$ lies in the same space. This follows from the $\mathbb{J}$-invariance of $d^2J(x_0)$ (see Lemma 4). Thus, putting $v = \mathbb{J}u$, we get $0 = \omega_{x_0}(u, \mathbb{J}u) = \langle\!\langle u, \mathbb{J}^2u \rangle\!\rangle = -\langle\!\langle u, u \rangle\!\rangle$. Thus $u = 0$. $\square$

If $H : P \to \mathbb{R}$ is a G-invariant Hamiltonian then the flow F_t of the corresponding Hamiltonian vector field X_H commutes with the action, so it leaves each submanifold N_{x_0} of points with the same symmetry type as x_0 invariant. It also leaves $N_{x_0} \cap \mathscr{C}$ invariant, since J is conserved. The dynamics on $N_{x_0} \cap \mathscr{C}$ can be reconstructed from the dynamics on $(N_{x_0} \cap \mathscr{C})/G \approx \mathcal{N}_{x_0} \cap \mathscr{C}$ as in Abraham and Marsden (1978, pp. 304–305). But the flow on $\mathcal{N}_{x_0} \cap \mathscr{C}$ is just the Hamiltonian flow of the Hamiltonian $H|(\mathcal{N}_{x_0} \cap \mathscr{C})$. Thus, while bifurcations can occur in the full space $\mathscr{C}$, reduction can still be done to obtain a Hamiltonian system on points with the same symmetry type as x_0. This procedure is useful, for example, when passing from the general Einstein equations to obtain a Hamiltonian structure for homogeneous cosmologies (cf. Jantzen, 1979).

d) Constraints in Field Theories

The reduction procedure or the topological program in mechanics naturally leads one to study the level sets $J^{-1}(\mu)$ of a momentum mapping. Here it plays the role of a conserved quantity in a mechanical system. However, in relativistic field theories, the zero set of a momentum map can often be identified with the constraint set in the sense of the Dirac theory of constraints. For relativity, the relevant group is the group of diffeomorphisms of spacetime and for gauge theories it is the group of bundle automorphisms (covering the identity if the theory is not coupled to gravity). In fact, in Gotay et al. (1980) it is shown that this

is the case for a wide variety of field theories (essentially any theory which is first class in the sense of Dirac)[3]. Thus, the results of this paper can be used to analyze the singularities that occur in spaces of solutions of relativistic field theories. How to do this for gauge theory and relativity will be outlined in the next section. It may be useful to attempt to resolve the singularities in the constraint sets by realizing them as projections of smooth Lagrangian submanifolds. It is not clear how to do this in detail, but the work of Kijowski and Tulczyjew (1979) may provide the correct context.

8. Examples

This section gives three applications of the general theory. Very likely there are several more of considerable interest. For example, bifurcations of the zero set of total angular momentum seem to play a role in the studies of Brown and Scriven (1980).

a) Angular Momentum of N particles

Let $P = T^*\mathbb{R}^{3N} \approx \mathbb{R}^{6N}$ be the phase space for N non-relativistic particles (allowing several particles to occupy the same position in $\mathbb{R}^3$). Consider the usual action of SO(3) on P. The associated momentum map is the total angular momentum J of the system. In terms of standard coordinates $(x_{(a)}^i, p_{(a)}^i)$, $a = 1, \ldots, N$, we have

$$J^i = \sum_{a,j,k} \varepsilon^{ijk} x_{(a)}^j p_{(a)}^k$$

or

$$J = \sum_a x_{(a)} \times p_{(a)}.$$

The level surfaces of the momentum map

$$J : T^*\mathbb{R}^{3N} \to \mathbb{R}^3 \approx \mathrm{so}(3)$$

will have singularities exactly at the points admitting nontrivial isotropy groups. The only point invariant under the full SO(3) action is clearly $x_{(a)} = 0$, $p_{(a)} = 0$. Other points having nontrivial isotropy groups have the form $x_{(a)} = \alpha_a \hat{n}$, $p_{(a)} = \beta_a \hat{n}$, where $\hat{n}$ is a unit vector in R^3 and α_a and β_a are constants (not all zero). Each such point is invariant under the circle action of rotations about the axis $\hat{n}$. The "degeneracy manifold" of points admitting such 1-dimensional isotropy groups is $2N + 2$ dimensional. Note that each such point has $J = 0$. We can expect that a cone of solutions of the equation $J = 0$ branches from each such symmetric point (for the case $N > 1$) since one can clearly have two or more particles each with non-vanishing individual angular momentum, but $J = 0$ by cancellation.

---

3   Relativistic fluids are covered if a suitable Hamiltonian formalism is used. Even in the "standard" formalism (see Hawking and Ellis, 1973), there are still singularities in the reduction process even though the equations themselves are linearization stable. The failing of the "standard" formalism is that it does not write the constraints as the zero set of a momentum mapping. This can be done with the right choice of variables; cf. Arms (1979b)

The case $N = 1$ is special, but also is interesting. Here

$$J^{-1}(0) = \{(x, p) \in \mathbb{R}^3 \times \mathbb{R}^3 \,|\, x \times p = 0\}$$

is a cone over a smooth compact 3-manifold M. In fact, it is not hard to see that M is $(S^2 \times S^1)/\sim$ where $\sim$ identifies simultaneous antipodal points in $S^2 \times S^1$.

[There is no bifurcation from non-zero points $(x, p) \in J^{-1}(0)$ since all such points have the *same* symmetry; here N_{x_0} coincides locally with $J^{-1}(0)$. For $N > 1$ this accident does not occur.]

b) Gauge Theory

Existence and uniqueness theory for the Cauchy problem shows that the structure of singularities in the solution space of the four dimensional Yang-Mills field equations on a fixed background spacetime is the same as that for the constraint equations. These constraint equations are well-known and may be described as follows (see, for example, Arms, 1979a): Let M be a fixed compact 3-manifold (a Cauchy surface in the fixed background spacetime). Let $\pi : B \to M$ be a principle G-bundle and let $\mathfrak{A}$ denote the space of $(W^{s,p}, s > 3/p + 1)$ connections on this bundle. Elements $A \in \mathfrak{A}$ represent vector potentials for gauge fields restricted to M. Let $P = T^*\mathfrak{A}$ be the basic symplectic space, elements of which are pairs (A, η); η represents the generalized electric field density. Assume $\mathfrak{g}$, the Lie algebra of G carries an adjoint-action invariant inner product $(\ ,\ )$, so $T_{(A,\eta)}(T^*\mathfrak{A})$ [elements of which are denoted (b, θ)] carries a preferred L_2 inner product $\langle\!\langle\ ,\ \rangle\!\rangle$. This, the canonical symplectic structure and the complex structure $\mathbb{J}(b, \theta) = (-\theta, b)$ [appropriately dualized by $(\ ,\ )$] are in the correct relationship (2).

The constraint equations are $J(A, \eta) = 0$ where $J(A, \eta) = d\eta + [A \wedge \eta]$ is the gauge covariant divergence of η using the connection A. In fact, J is the momentum map for the action of the group $\mathscr{G}$ of bundle automorphisms of B on P. This is the group G in the general theory; its Lie algebra is g, the $\mathfrak{g}$-valued functions on M. The dual g^* is the $\mathfrak{g}^*$ valued densities; thus $J : P \to g^*$. The adjoint operators $dJ(A, \eta)^*$ and $dJ(A, \eta)^+$ are elliptic and so one can construct a slice using $\ker(dJ(A, \eta) \circ \mathbb{J})$. In fact, all of the assumptions of Sect. 2 hold; the spaces here are infinite dimensional but ellipticity of $dJ(A, \eta)^*$ validates the technical points.

Moreover, J is quadratic. The quadratic term Q of Sect. 5 is

$$Q(b, \theta) = [b \wedge \theta] ,$$

where b and θ are perturbations of A and η, so $h = (b, \theta)$, and $[b \wedge \theta]$ is the bracket in $\mathfrak{g}$. (From this simple form gauge and $\mathbb{J}$-invariance can be verified directly.) Theorem 3 therefore directly applies. For gauge fields, infinitesimal symmetries of (A, η) are A-covariant constant $\mathfrak{g}$-valued functions on B that commute with η. The existence of such symmetries implies that the gauge field is reducible to a field with a smaller gauge group $H \subset G$, $H \neq G$. The space $\mathscr{N}_{(A,\eta)} \cap \mathscr{C}$ of Sect. 4 consists of solutions of the constraint equations which are reducible to the gauge group H; the rest of the solution set containing the conical singularities consists of solutions with a gauge group K intermediate between H and G, Theorem 4' then shows how conical singularities of specified symmetry type K fit together to produce the entire

conical singularity in the constraint set $\mathscr{C} = \{(A,\eta)|J(A,\eta)=0\}$. For more details, see Arms (1980).

For noncompact M, symmetries often are associated with a gauge group generating a momentum map representing a total energy-momentum tensor or charge. Such symmetries need not lie in the group generating the constraints, so need not indicate singularities in the space of solutions. (For example, in general relativity, Minkowski space is a regular point in the space of solutions of Einstein's equations.) For Yang-Mills fields however it is interesting to impose the constraint that the color charge q vanish for non-compact M, say $M = \mathbb{R}^3$. This is suggested by the classical limit of a quantized Yang-Mills theory with the property of "confinement".

Consider then a classical Yang-Mills theory on $M = \mathbb{R}^3$ with the constraint $q = 0$ added on, where

$$q = \int_M d\eta = - \int_M [A \wedge \eta].$$

[This procedure is to be done in the appropriate weighted Sobolev spaces to properly capture the asymptotic behavior; cf. Cantor (1979).] There is now a singularity in the space of all solutions at the trivial "vacuum" solution $A = 0$, $\eta = 0$; the constraint $q = 0$ leads to the second order condition

$$\int_M [b \wedge \theta] = 0$$

to be placed on first order perturbations, as in the compact case. Thus it seems that an appropriate free field approximation to the nonlinear theory is not the usual linearized Yang-Mills equations, but rather these equations supplemented by the above second order condition.

c) The Constraint Equations in General Relativity

We shall use the methods of Theorem 3 to study the conical singularities in the constraint equations of general relativity. Let M be a compact 3-manifold and $\mathscr{M}$ the space of $\left(W^{s,p}, s > \dfrac{3}{p} + 1\right)$ Riemannian metrics on M and let $P = T^*\mathscr{M}$ denote the "natural" cotangent bundle of $\mathscr{M}$; i.e. the fiber of $T^*\mathscr{M}$ over $g \in \mathscr{M}$ consists of all symmetric 2-contravariant tensor densities π (of class $W^{s-1,p}$). The constraint set of the vacuum Einstein equations on a 4 dimensional spacetime in which M is embedded as a compact hypersurface is well known to be the set

$$\mathscr{C} = \Phi^{-1}(0),$$

where $\Phi : T^*\mathscr{M} \to$ (Densities on M) × (One-form Densities on M) is defined by $\Phi(g,\pi) = (\mathscr{H}(g,\pi), \mathscr{J}(g,\pi))$ and $\mathscr{H}$ and $\mathscr{J}$ are given by

$$\mathscr{H}(g,\pi) = \{(\pi' \cdot \pi' - \tfrac{1}{2}(\operatorname{trace}\pi')^2) - R(g)\}\mu(g)$$

and

$$\mathscr{J}(g,\pi)_i = -2\pi_i{}^j{}_{|j}$$

Here $\pi = \pi' \otimes \mu(g)$, $\mu(g)$ is the volume form of g, and $R(g)$ is the scalar curvature of g.

In Fischer et al. (1979) it is shown that $\mathscr{C}$ has a conical singularity at (g_0, π_0) if (g_0, π_0) is the Cauchy data for a spacetime with one Killing field (either timelike or spacelike). Here *we show the same thing if there are k spacelike Killing fields.* We shall assume that trace $\pi' = \text{constant}$ on M (i.e. M is a hypersurface of constant mean curvature in the spacetime). Then, the aforementioned reference establishes the analogue of Eq. (8), namely that $\ker(D\Phi(g_0, \pi_0)^*)$ is k-dimensional and is spanned by elements of the form (O, X_a), $a = 1, \ldots, k$ where X_a is a vector field on M satisfying $L_{X_a}\pi_0 = 0$ and $L_{X_a}g_0 = 0$. The adjoint is taken relative to the L_2 metric on $T^*\mathcal{M}$ given by

$$\langle\!\langle (h_1, \omega_1), (h_2, \omega_2) \rangle\!\rangle = \int_M \{h_1 \cdot h_2 + \omega_1 \cdot \omega_2\} \mu(g),$$

where "$\cdot$" denotes contraction using g [(g, π) is the base point] and the natural pairing between (Densities) $\times$ (One-form Densities) and (Functions) $\times$ (Vector Fields). Thus $D\Phi(g, \pi)^*$: (Functions $\times$ Vector Fields) $\to T_{(g,\pi)}(T^*\mathcal{M})$, and one can compute this explicitly.

At this point we are not claiming that Φ is a momentum map for a group action. If we were dealing only with the momentum contraint $\mathscr{J}$, the results of Theorem 3 would apply directly since the momentum map for the action of the diffeomorphism group of M on $T^*\mathcal{M}$ is just the quadratic map $J : T^*\mathcal{M} \to$ (vector fields)*; $\langle J(g, \pi), X \rangle = \int_M X \cdot \mathscr{J}(g, \pi) = 2 \int_M (L_X g) \cdot \pi$. However the Hamiltonian constraint $\mathscr{H} = 0$ complicates this program.

Decomposition (7) is known to hold for Φ (see Moncrief, 1975b; Arms et al., 1975; Fischer et al., 1979)

$$T_{(g,\pi)}(T^*M) = \text{range}(-\mathbb{J}D\Phi(g, \pi)^*) \oplus \text{range}\, D\Phi(g, \pi)^*$$

$$\oplus [\ker(D\Phi(g, \pi) \circ \mathbb{J}) \cap \ker D\Phi(g, \pi)] \qquad (21)$$

since $\text{range}\,\mathbb{J}D\Phi(g_0, \pi_0)^* \subset \ker D\Phi(g_0, \pi_0)$ and $D\Phi(g_0, \pi_0)^*$ is elliptic. Here $\mathbb{J}_{(g,\pi)}(h, \omega) = (\omega^\flat, -h^\sharp)$, $\sharp$ and $\flat$ denote the raising and lowering of indices using the metric g.

Let

$$S_{(g_0, \pi_0)} = \{(g_0, \pi_0)\} + \text{a neighborhood of zero in } \ker D\Phi(g_0, \pi_0) \circ \mathbb{J}.$$

This makes sense since $\mathcal{M}$ is open in $S_2(M)$, the covariant symmetric two tensors on M, and so $T^*\mathcal{M}$ is open in the linear space $S_2(M) \times S_d^2(M)$, where $S_d^2(M)$ is the space of contravariant symmetric two tensor densities ($S_{(g_0, \pi_0)}$ plays the role of a slice for the four-dimensional diffeomorphism group of spacetime). The analogue of $\mathcal{N}_{x_0}$ (see Lemma 1) is

$$\mathcal{N}_{(g_0, \pi_0)} = \{(g, \pi) \in T^*\mathcal{M} \mid L_{X_a}g = L_{X_a}\pi = 0, \ a = 1, \ldots, k \quad \text{and} \quad (g, \pi) \in S_{(g_0, \pi_0)}\}.$$

Fischer et al. (1979) shows that $\mathcal{N}_{(g_0, \pi_0)}$ is a smooth manifold.

Now define the local diffeomorphism F of $T^*\mathcal{M}$ to $T^*\mathcal{M}$ near (g_0, π_0) as follows. First, let,

$$\Delta = D\Phi(g_0, \pi_0) \circ D\Phi(g_0, \pi_0)^*.$$

Since $D\Phi(g_0, \pi_0)^*$ is an elliptic operator, Δ is an isomorphism of range $D\Phi(g_0, \pi_0)$ to itself. Following what we did in Sect. 5, let $\mathbb{P}$ denote the orthogonal projection to range $D\Phi(g_0, \pi_0)$ and set $G = \Delta^{-1} \circ \mathbb{P}$. Write $h = (g, \pi) - (g_0, \pi_0)$ and let the remainder be given by

$$R(h) = \Phi(g, \pi) - D\Phi(g_0, \pi_0) \cdot h.$$

Next define F by the analogue of (15):

$$F(g, \pi) = (g, \pi) + D\Phi(g_0, \pi_0)^* \circ G \circ R(h).$$

Computations based on the results of Fischer et al. (1979) together with arguments like those in Sect. 5 establish the following:

1. $DF(g_0, \pi_0) = $ identity, so F is a local diffeomorphism.
2. F maps $S_{(g_0, \pi_0)}$ to itself.
3. If $\mathscr{C}_{\mathbb{P}} = \{(g, \pi) | \mathbb{P}\Phi(g, \pi) = 0\}$, which is a smooth manifold in a neighborhood of (g_0, π_0), then F maps $\mathscr{C}_{\mathbb{P}}$ in a $1-1$ way onto a neighborhood of (g_0, π_0) in (g_0, π_0) $+ \ker D\Phi(g_0, \pi_0)$.
4. F maps $\mathscr{C}_{\mathbb{P}} \cap S_{(g_0, \pi_0)}$ to $\{(g_0, \pi_0)\} + \ker D\Phi(g_0, \pi_0) \cap \ker(D\Phi(g_0, \pi_0) \circ \mathbb{J})$.

Now define the cone

$$C_{(g_0, \pi_0)} = \{(g_0, \pi_0)\} + [(\mathrm{Id} - \mathbb{P})R]^{-1}(0).$$

This is a cone because of the following crucial fact:

$$(\mathrm{Id} - \mathbb{P})R = (\mathrm{Id} - \mathbb{P})\Phi \text{ is quadratic in } (g, \pi).$$

This is because $(\mathrm{Id} - \mathbb{P})$ projects to the span of $\{(0, X_a) | a = 1, \ldots, k\}$ and so only the supermomentum component $\mathscr{J}$ is involved in $(\mathrm{Id} - \mathbb{P})\Phi$. Thus one can effect the last step in Sect. 5:

5. F maps $\mathscr{C} \cap S_{(g_0, \pi_0)}$ locally $1-1$ onto $C_{(g_0, \pi_0)}$.

The proof of Theorem 3 needs not only $(\mathrm{Id} - \mathbb{P})\Phi$ to be quadratic but also crucially uses $\mathbb{J}$ invariance and gauge invariance of $(\mathrm{Id} - \mathbb{P})D^2\Phi(g_0, \pi_0)$. The $\mathbb{J}$-invariance results from a direct computation using the expression for $\int X \cdot \mathscr{J}(g, \pi)$ given above. The gauge invariance also can be proven directly; see Fischer et al. (1979) for the general proof of gauge invariance of $D^2\Phi$. This reference also explains how to remove the gauge condition $S_{(g_0, \pi_0)}$.

Theorems 4 and 4' also carry over for this example. One of the crucial facts in this regard is Lemma 5, which can be interpreted as a covariance property. In particular, it entails that $\phi^* F(g, \pi) = F(\phi^* g, \phi^* \pi)$ for a diffeomorphism ϕ such that $\phi^* g_0 = g_0$, $\phi^* \pi_0 = \pi_0$. For this example these may be proved by a direct calculation.

These results describe, for example, the structure of the space solutions near Bianchi IX models, such as the Taub universe ($k = 4$), the Kasner solution ($k = 3$) or the Gowdy solutions ($k = 2$), cf. Jantzen (1979). In Arms et al. (1980) we shall generalize these arguments to allow a time-like Killing field as well as several spacelike Killing fields. The flat space time $T^3 \times \mathbb{R}$ is an especially interesting bifurcation point.

Glossary of Symbols

(P, ω)	symplectic manifold
$T_{x_0}P$	tangent space to P at $x_0 \in P$
$G, \mathfrak{g}$	Lie group, Lie algebra
$g \cdot x = \Phi_g(x)$	action of G on P
ξ_P	infinitesimal generator of the action on P corresponding to $\xi \in \mathfrak{g}$
$J : P \to \mathfrak{g}^*$	momentum mapping
$dJ(x_0) : T_{x_0}P \to \mathfrak{g}^*$	differential of J at x_0
$\mathbb{J}$	complex structure on P
S_{x_0}	slice for the G-action at x_0
I_{x_0}	isotropy group of x_0; $\{g \in G | gx_0 = x_0\}$
$\mathscr{I}_{x_0}, \mathfrak{s}_{x_0}$	identity component of I_{x_0}, its Lie algebra [Eq. (10)]
$\langle\!\langle \, , \, \rangle\!\rangle$	(weak) metric on P
$\langle \, , \, \rangle$	pairing between $\mathfrak{g}^*$ and $\mathfrak{g}$
$(\, , \,)_{x_0}$	inner product of $\mathfrak{g}^*$ depending on $x_0 \in P$
$dJ(x_0)^* : \mathfrak{g} \to T_{x_0}P$	adjoint of $dJ(x_0)$ relative to $\langle \, , \, \rangle$ and $\langle\!\langle \, , \, \rangle\!\rangle$
$dJ(x_0)^\dagger : \mathfrak{g}^* \to T_{x_0}P$	adjoint of $dJ(x_0)$ relative to $(\, , \,)$ and $\langle\!\langle \, , \, \rangle\!\rangle$
$T_{x_0}P = \text{Range}[\mathbb{J} \circ dJ(x_0)^*] \oplus \text{Range}[dJ(x_0)^*] \oplus [\ker(dJ(x_0) \circ \mathbb{J}) \cap \ker dJ(x_0)]$	
	Moncrief's decomposition
$\mathbb{P} : \mathfrak{g}^* \to \text{Range}\, dJ(x_0)$	orthogonal projection
$\mathscr{C} = J^{-1}(0)$	zero set of J (or constraint set)
$\mathscr{C}_{\mathbb{P}} = (\mathbb{P}J)^{-1}(0)$	zero set of $\mathbb{P} \circ J$
N_{x_0}	x's with the same orbit type as x_0
$\mathscr{N}_{x_0}$	$N_{x_0} \cap S_{x_0}$ (Lemma 1)
$\mathfrak{g}^*_{x_0}$	elements in $\mathfrak{g}^*$ with same symmetry as x_0 (Lemma 5)
f	$f = (\text{Id} - \mathbb{P}) \circ J : \mathscr{C}_{\mathbb{P}} \cap S_{x_0} \to \ker dJ(x_0)^\dagger$ (Lemma 8)
$\Delta = dJ(x_0) \circ dJ(x_0)^\dagger$	"elliptic" operator associated with J
$G = \Delta^{-1} \circ \mathbb{P}$	Greens' function for Δ
$F(x) = x + dJ(x_0)^\dagger \circ G \circ Q(h), h = x - x_0$	Kuranishi map (Lemma 9)
C_{x_0}	homogeneous cone associated with $d^2 J(x_0)$ (Theorem 3)
$\tilde{\mathbb{P}}$	orthogonal projection onto $\ker dJ(x_0)$ (Theorem 3)
$\mathfrak{h}, \mathscr{H}$	a Lie subalgebra of $\mathfrak{s}_{x_0}$, its Lie group
$\mathscr{A}_{\mathscr{H}}$	points with symmetry (at least) $\mathscr{H}$ (Theorem 4)
$\mathscr{B}_{\mathscr{H}}$	$\mathscr{A}_{\mathscr{H}} \cap S_{x_0}$

References

1. Abraham, R., Marsden, J.: Foundations of mechanics, second edition. Addison-Wesley, Reading 1978
2. Arms, J.: J. Math. Phys. NY **20**, 443–453 (1979a)
3. Arms, J.: Does matter break the link between symmetry and linearization stability? (preprint) (1979b)
4. Arms, J.: The structure of the solution set for the Yang-Mills equations (preprint) (1980)
5. Arms, J., Fisher, A., Marsden, J.: C.R. Acad. Sci. Paris **281**, 517–520 (1975)
6. Arms, J., Fisher, A., Marsden, J., Moncrief, V.: The structure of the space of solutions of Einstein's equations. II. Many killing fields (in preparation) (1980)
7. Atiyah, M.F., Hitchin, N.J., Singer, I.M.: Proc. R. Soc. London, Ser. A **362**, 425–461 (1978)
8. Brill, D., Deser, S.: Commun. Math. Phys. **32**, 291–304 (1973)
9. Bott, R.: Ann. Math. **60**, 248–261 (1954)
10. Brown, R.A., Scriven, L.E.: Proc. R. Soc. London (to appear) (1980)
11. Cantor, M.: Comp. Math. **38**, 3–35 (1979)
12. Ebin, D.: Proc. Symp. Pure Math. Am. Math. Soc. **XV**, 11–40 (1970)

13. Fischer, A., Marsden, J.: Bull. Am. Math. Soc. **79**, 995–1001 (1973)
14. Fischer, A., Marsden, J., Moncrief, V.: Ann. Inst. Henri Poincaré (to appear) (1980)
15. Gotay, M., Marsden, J., Sniatycki, J.: Lagrangian field theory and zero sets of momentum maps (in preparation) (1980)
16. Gotay, M., Nester, J., Hinds, G.: J. Math. Phys. NY **19**, 2388–2399 (1978)
17. Hawking, S., Ellis, G.: The large scale structure of spacetime. Cambridge: Cambridge University Press 1973
18. Hermann, R.: Differential geometry and the calculus of variations, New York: Academic Press 1968, second edition by Math. Sci. Press 1977
19. Jantzen, R.: Commun. Math. Phys. **64**, 211–232 (1979)
20. Kijowski, J., Tulczyjew, W.: A symplectic framework for field theories. Springer Lecture Notes in Physics, No. 107. Berlin, Heidelberg, New York: Springer 1979
21. Kuranishi, M.: New proof for the existence of locally complete families of complex structures. Proc. Conf. on Complex Analysis. A. Aeppli et al. (eds.). Berlin, Heidelberg, New York: Springer 1965
22. Marsden, J.: Bull. Am. Math. Soc. **84**, 1125–1148 (1978)
23. Marsden, J.: Geometric methods in mathematical physics. CBMS Conf. Series (to appear) (1980)
24. Marsden, J., Weinstein, A.: Rep. Math. Phys. **5**, 121–130 (1974)
25. Moncrief, V.: J. Math. Phys. NY **16**, 493–498; **17**, 1893–1902 (1975a)
26. Moncrief, V.: J. Math. Phys. NY **16**, 1556–1560 (1975b)
27. Moncrief, V.: Ann. Phys. **108**, 387–400 (1977)
28. Moncrief, V.: Phys. Rev. D **18**, 983–989 (1978)
29. Palais, R.: Mem. Am. Math. Soc. No. 22 (1957)
30. Segal, I.E.: Proc. Nat. Acad. Sci. USA **15**, 4638–4639 (1978)
31. Smale, S.: Inventiones Math. **10**, 305–331; **11**, 45–64 (1970)
32. Souriau, J.M.: Structure des systemes dynamique. Paris: Dunod 1970
33. Weinstein, A.: Lectures on symplectic manifolds. CBMS Conf. Series No. 29, AMS (1977)

Note added in proof. A very interesting example of a conical singularity in the angular momentum level surfaces of the spherical pendulum and its relation to action-angle variables is contained in the preprint of H. Duistermaat, "On global action angle coordinates".

2. Results on the symplectic structure on the solution space of the Yang-Mills equations at regular points, which reproduce the theorems of Ref. [24, 27] (but from the spacetime rather than dynamical point of views) are contained in Ref. [30] and in:

34. Garcia, P., Rendon, A.P.: Reducibility of the symplectic structure of minimal interactions. In: Lecture Notes in Mathematics, Vol. 676. Berlin, Heidelberg, New York: Springer 1978
35. Garcia, P.: Tangent structure of Yang-Mills equations and Hodge theory. In: Lecture Notes in Mathematics, Vol. 836. Berlin, Heidelberg, New York: Springer 1979