On Symplectic Reduction in Classical Mechanics

J. Butterfield[1]

All Souls College
Oxford OX1 4AL

Monday 5 December 2005; a Chapter of *The North Holland Handbook of Philosophy of Physics*

Abstract

This Chapter expounds the modern theory of symplectic reduction in finite-dimensional Hamiltonian mechanics. This theory generalizes the well-known connection between continuous symmetries and conserved quantities, i.e. Noether's theorem. It also illustrates one of mechanics' grand themes: exploiting a symmetry so as to reduce the number of variables needed to treat a problem. The exposition emphasises how the theory provides insights about the rotation group and the rigid body. The theory's device of quotienting a state space also casts light on philosophical issues about whether two apparently distinct but utterly indiscernible possibilities should be ruled to be one and the same. These issues are illustrated using "relationist" mechanics.

Keywords: symplectic reduction, symmetry, conserved quantities, Poisson manifolds, momentum maps, relationist mechanics.

Mottoes

The current vitality of mechanics, including the investigation of fundamental questions, is quite remarkable, given its long history and development. This vitality comes about through rich interactions with pure mathematics (from topology and geometry to group representation theory), and through new and exciting applications to areas like control theory. It is perhaps even more remarkable that absolutely fundamental points, such as a clear and unambiguous linking of Lie's work on the Lie-Poisson bracket on the dual of a Lie algebra ... with the most basic of examples in mechanics, such as the rigid body and the motion of ideal fluids, took nearly a century to complete.
    Marsden and Ratiu (1999, pp. 431-432).

In the ordinary theory of the rigid body, six different three-dimensional spaces $\mathbb{R}^3, \mathbb{R}^{3*}, \mathfrak{g}, \mathfrak{g}^*, TG_g, T^*G_g$ are identified.
    Arnold (1989, p. 324).

[1]email: jb56@cus.cam.ac.uk; jeremy.butterfield@all-souls.oxford.ac.uk

Contents

1 Introduction

1.1 Why classical mechanics?

All hail the rise of modern physics! Between 1890 and 1930, the quantum and relativity revolutions and the consolidation of statistical physics through the discovery of atoms, utterly transformed our understanding of nature; and had an enormous influence on philosophy; (e.g. Kragh 1999; Ryckman 2005). Accordingly, this Handbook concentrates on those three pillars of modern physics—quantum theories, spacetime theories and thermal physics. So some initial explanation of the inclusion of a Chapter on classical mechanics, indeed the classical mechanics of finite-dimensional systems, is in order.

The first point to make is that the various fields of classical physics, such as mechanics and optics, are wonderfully rich and deep, not only in their technicalities, but also in their implications for the philosophy and foundations of physics. From Newton's time onwards, classical mechanics and optics have engendered an enormous amount of philosophical reflection. As regards mechanics, the central philosophical topics are usually taken (and have traditionally been taken) to be space, time, determinism and the action-at-a-distance nature of Newtonian gravity. Despite their importance, I will not discuss these topics; but some other Chapters will do so (at least in part, and sometimes in connection with theories other than classical mechanics). I will instead focus on the theory of symplectic reduction, which develops the well-known connection between continuous symmetries and conserved quantities, summed up in Noether's "first theorem". I choose this focus partly by way of preparation for parts of some other Chapters; and partly because, as we will see in a moment, symplectic reduction plays a central role in the current renaissance of classical mechanics, and in its relation to quantum physics.

I said that classical physics engendered a lot of philosophical reflection. It is worth stressing two, mutually related, reasons for this: reasons which today's philosophical emphasis on the quantum and relativity revolutions tends to make us forget.

First: in the two centuries following Newton, these fields of classical physics were transformed out of all recognition, so that the framework for philosophical reflection about them also changed. Think of how in the nineteenth century, classical mechanics and optics gave rise to classical field theories, especially electromagnetism. And within this Chapter's specific field, the classical mechanics of finite-dimensional systems, think of how even its central theoretical principles were successively recast, in fundamental ways, by figures such Euler, Lagrange, Hamilton and Jacobi.

Second, various difficult problems beset the attempt to rigorously formulate classical mechanics and optics; some of which have considerable philosophical aspects. It is *not* true that once we set aside the familiar war-horse topics—space, time, determinism and action-at-a-distance—the world-picture of classical mechanics is straightforward: just "matter in motion". On the contrary. Even if we consider only finite-dimensional systems, we can ask, for example:

(i) For point-particles (material points): can they have different masses, and if so how? What happens when they collide? Indeed, for point-particles interacting only by Newtonian gravity, a collision involves infinite kinetic energy.

(ii) For extended bodies treated as finite-dimensional because rigid: what happens when they collide? Rigidity implies that forces, and displacements, are transmitted "infinitely fast" through the body. Surely that should not be taken literally? But if so, what justifies this idealization; and what are its scope and limits?

As to infinite-dimensional systems (elastic solids, fluids and fields), many parts of their theories remain active research areas, especially as regards rigorous formulations and results. For contemporary work on elastic solids, for example, cf. Marsden and Hughes (1982). As to fluids, the existence and uniqueness of rigorous solutions of the main governing equations, the Navier-Stokes equations, is still an open problem. This problem not only has an obvious bearing on determinism; it is regarded as scientifically significant enough that its solution would secure a million-dollar Clay Millennium prize.

These two reasons—the successive reformulations of classical mechanics, and its philosophical problems—are of course related. The monumental figures of classical mechanics recognized and debated the problems, and much of their technical work was aimed at solving them. As a result, there was a rich debate about the foundations of classical physics, in particular mechanics, for the two centuries after Newton's *Principia* (1687). A well-known example is Duhem's instrumentalist philosophy of science, which arose in large measure from his realization how hard it was to secure rigorous foundations at the microscopic level for classical mechanics. A similar example is Hilbert's being prompted by his contemporaries' continuing controversies about the foundations of mechanics, to choose as the sixth of his famous list of outstanding mathematical problems, the axiomatization of mechanics and probability; (but for some history of this list, cf. Grattan-Guinness (2000)). A third example, spanning both centuries, concerns variational principles: the various principles of least action formulated first by Maupertuis, then by Euler and later figures—first for finite classical mechanical systems, then for infinite ones—prompted much discussion of teleology. Indeed, this discussion ensnared the logical empiricists (Stöltzner 2003); it also bears on contemporary philosophy of modality (Butterfield 2004).

In the first half of the twentieth century, the quantum and relativity revolutions tended to distract physicists, and thereby philosophers, from these and similar problems. The excitement of developing the new theories, and of debating their implications for natural philosophy, made it understandable, even inevitable, that the foundational problems of classical mechanics were ignored.

Besides, this tendency was strengthened by the demands of pedagogy: the necessity of including the new theories in physics undergraduate degrees. By mid-century, the constraints of time on the physics curriculum had led many physics undergraduates' education in classical mechanics to finish with the elementary parts of analytical mechanics, especially of finite-dimensional systems: for example, with the material in Goldstein's well-known textbook (1950). Such a restriction is understandable, not least because: (i) the elementary theory of Lagrange's and Hamilton's equations requires knowledge of ordinary differential equations, and (ii) elementary Hamiltonian mechanics forms a springboard to learning elementary canonical quantization (as does Hamilton-Jacobi theory, from another perspective). Besides, as I mentioned: even this restricted body of theory provides plenty of material for philosophical analysis—witness my examples above, and the discussions of the great figures such Euler, Lagrange, Hamilton and Jacobi.

However, the second half of the twentieth century saw a renaissance in research in classical mechanics: hence my first motto. There are four obvious reasons for this: the first two "academic", and the second two "practical".

(i): Thanks partly to developments in mathematics in the decades after Hilbert's list of problems, the foundational questions were addressed afresh, as much by mathematicians and mathematically-minded engineers as by physicists. The most relevant developments lay in such fields as topology, differential geometry, measure theory and functional analysis. In this revival, the contributions of the Soviet school, always strong in mechanics and probability, were second to none. And relatedly:—

(ii): The quest to deepen the formulation of quantum theory, especially quantum field theory, prompted investigation of (a) the structure of classical mechanics and (b) quantization. For both (a) and (b), special interest attaches to the generally much harder case of infinite systems.

(iii): The coming of spaceflight, which spurred the development of celestial mechanics. And relatedly:—

(iv): The study of non-linear dynamics ("chaos theory"), which was spurred by the invention of computers.

With these diverse causes and aspects, this renaissance continues to flourish—and accordingly, I shall duck out of trying to further adumbrate it! I shall even duck out of trying to survey the philosophical questions that arise from the various formulations of mechanics from Newton to Jacobi and Poincaré. Suffice it to say here that to the various topics mentioned above, one could add, for example, the following two: the first broadly ontological, the second broadly epistemological.

(a): The analysis of notions such as mass and force (including how they change over time). For this topic, older books include Jammer (1957, 1961) and McMullin (1978); recent books include Boudri (2002), Jammer (2000), Lutzen (2005) and Slovik (2002); and Grattan-Guinness (2006) is a fine recent synopsis of the history, with many references.

(b): The analysis of what it is to have an explicit solution of a mechanical problem (including how the notion of explicit solution was gradually generalized). This topic is multi-faceted. It not only relates to the gradual generalization of the notion of function (a grand theme in the history of mathematics—well surveyed by Lutzen 2003), and to modern non-linear dynamics (cf. (iv) above). It also relates to the simplification of problems by exploiting a symmetry so as to reduce the number of variables one needs—and *this* is the core idea of symplectic reduction. I turn to introducing it.

1.2 Prospectus

The strategy of simplifying a mechanical problem by exploiting a symmetry so as to reduce the number of variables is one of classical mechanics' grand themes. It is theoretically deep, practically important, and recurrent in the history of the subject. The best-known general theorem about the strategy is undoubtedly Noether's theorem, which describes a correspondence between continuous symmetries and conserved quantities. There is both a Lagrangian and a Hamiltonian version of this theorem, though for historical reasons the name 'Noether's theorem' is more strongly attached to the Lagrangian version. However, we shall only need the Hamiltonian version of the theorem: it will be the "springboard" for our exposition of symplectic reduction.[2]

So I shall begin by briefly reviewing the Hamiltonian version in Section 2.1. For the moment, suffice it to make four comments (in ascending order of importance for what follows):

(i): Both versions are underpinned by the theorems in elementary Lagrangian and Hamiltonian mechanics about cyclic (ignorable) coordinates and their corresponding conserved momenta.[3]

(ii): In fact, the Hamiltonian version of the theorem is stronger. This reflects the fact that the canonical transformations form a "larger" group than the point transformations. A bit more precisely: though the point transformations $q \to q'$ on the configuration space Q induce canonical transformations on the phase space Γ of the qs and ps, $q \to q', p \to p'$, there are yet other canonical transformations which "mix" the qs and ps in ways that transformations induced by point transformations do not.

(iii): I shall limit our discussion to (a) time-independent Hamiltonians and (b) time-independent transformations. Agreed, analytical mechanics can be developed, in both Lagrangian and Hamiltonian frameworks, while allowing time-dependent dynamics and transformations. For example, in the Lagrangian framework, allowing velocity-

[2]For discussion of the Lagrangian version, cf. e.g. Brading and Castellani (this vol., ch. 13) or (restricted to finite-dimensional systems) Butterfield (2004a: Section 4.7). For an exposition of both versions that is complementary to this paper (and restricted to finite-dimensional systems), cf. Butterfield (2006). Brading and Castellani also bring out that, even apart from Noether's theorems in other branches of mathematics, there are other 'Noether's theorems' about symmetries in classical dynamics; so the present theorem is sometimes called Noether's "first theorem". Note also (though I shall not develop this point) that symplectic structure can be seen in the classical solution space of Lagrange's equations, so that symplectic reduction can be developed in the Lagrangian framework; cf. e.g. Marsden and Ratiu (1999: p. 10, Sections 7.2-7.5, and 13.5).

[3]Here we glimpse the long history of our subject: these theorems were of course clear to these subjects' founders. Indeed the strategy of exploiting a symmetry to reduce the number of variables occurs already in 1687, in Newton's solution of the Kepler problem; (or more generally, the problem of two bodies exerting equal and opposite forces along the line between them). The symmetries are translations and rotations, and the corresponding conserved quantities are the linear and angular momenta. In what follows, these symmetries and quantities will provide us with several examples.

dependent potentials and-or time-dependent constraints would prompt one to use what is often called the 'extended configuration space' $Q \times \mathbb{R}$. And in the Hamiltonian framework, time-dependence prompts one to use an 'extended phase space' $\Gamma \times \mathbb{R}$. Besides, from a philosophical viewpoint, it is important to consider time-dependent transformations: for they include boosts, which are central to the philosophical discussion of spacetime symmetry groups, and especially of relativity principles. But beware: rough-and-ready statements about symmetry, e.g. that the Hamiltonian must be invariant under a symmetry transformation, are liable to stumble on these transformations. To give the simplest example: the Hamiltonian of a free particle is just its kinetic energy, which can be made zero by transforming to the particle's rest frame; i.e. it is not invariant under boosts.

So a full treatment of symmetry in Hamiltonian mechanics, and thereby of symplectic reduction, needs to treat time-dependent transformations—and to beware! But I will set aside all these complications. Here it must suffice to assert, without any details, that the modern theory of symplectic reduction *does* cope with boosts; and more generally, with time-dependent dynamics and transformations.

(iv): As we shall see in detail, there are three main ways in which the theory of symplectic reduction generalizes Noether's theorem. As one might expect, these three ways are intimately related to one another.

(a): Noether's theorem is "one-dimensional" in the sense that for each symmetry (a vector field of a special kind on the phase space), it provides a conserved quantity, i.e. a real-valued function on the phase space, whose value stays constant over time. So in particular, different components of a conserved vector quantity, such as total linear momentum, are treated separately; (in this example, the corresponding vector fields generate translations in three different spatial directions). But in symplectic reduction, the notion of a *momentum map* provides a "unified" description of these different components.

(b): Given a symmetry, Noether's theorem enables us to confine our attention to the level surface of the conserved quantity, i.e. the sub-manifold of phase space on which the quantity takes its initial value: for the system's time-evolution is confined to that surface. In that sense, the number of variables we need to consider is reduced. But in symplectic reduction, we go further and form a *quotient space* from the phase space. That is, in the jargon of logic: we define on phase space an equivalence relation (not in general so simple as having a common value for a conserved quantity) and form the set of equivalence classes. In the jargon of group actions: we form the set of orbits. Passage to this quotient space can have various good technical, and even philosophical, motivations. And under good conditions, this set is itself a manifold with lower dimension.

(c): Hamiltonian mechanics, and so Noether's theorem, is usually formulated in terms of symplectic manifolds, in particular the cotangent bundle T^*Q of the configuration space Q. (Section 2.1 will give details.) But in symplectic reduction, we often need a (mild) generalization of the idea of a symplectic manifold, called a *Poisson manifold*, in which a bracket, with some of the properties of the Poisson bracket, is taken as the primitive notion. Besides, this is related to (b) in that we are often led to a Poisson manifold, and dynamics on it, by taking the quotient of a symplectic manifold (i.e. a phase space of the usual kind) by the action of a symmetry group.

As comment (iv) hints, symplectic reduction is a large subject. So there are several motivations for expounding it. As regards physics, many of the ideas and results can be developed for finite-dimensional classical systems (to which I will confine myself), but then generalized to infinite-dimensional systems. And in either the original or the generalized form, they underpin developments in quantum theories. So these ideas and results have formed part of the contemporary renaissance in classical mechanics; cf. (i) and (ii) at the end of Section 1.1.

As regards philosophy, symmetry is both a long-established focus for philosophical discussion, and a currently active one: cf. Brading and Castellani (2003). But philosophical discussion of symplectic reduction seems to have begun only recently, especially in some papers of Belot and Earman. This delay is presumably because the technical material is more sophisticated: indeed, the theory of symplectic reduction was cast in its current general form only in the 1970s. But as Belot and Earman emphasise, the philosophical benefits are worth the price of learning the technicalities. The most obvious issue is that symplectic reduction's device of quotienting a state space casts light on philosophical issues about whether two apparently distinct but utterly indiscernible possibilities should be ruled to be one and the same. In Section 2, I will follow Belot in illustrating this issue with "relationist" mechanics. Indeed, I have selected the topics for my exposition with an eye to giving philosophical readers the background needed for some of Belot's discussions. His papers (which I will cite in Section 2) make many judicious philosophical points, without burdening the reader with an exposition of technicalities: excellent stuff—but to fully appreciate the issues, one of course has to slog through the details.

Finally, in the context of this volume, symplectic reduction provides some background for the Chapters on the representation of time in mechanics (Belot, this vol., ch. 2), and on the relations between classical and quantum physics (Landsman, this vol., ch. 5, especially Sections 4.3-4.5 and 6.5; Dickson, this vol., ch. 4).

The plan of the Chapter is as follows. I first review Noether's theorem in Hamiltonian mechanics as usually formulated, in Section 2.1. Then I introduce the themes mentioned in (b) and (c) above, of quotienting a phase space, and Poisson manifolds (Section 2.2); and illustrate these themes with "relationist" mechanics (Section 2.3).

Thereafter, I expound the basics of symplectic reduction: (confining myself to finite-dimensional Hamiltonian mechanics). Section by Section, the plan will be as follows. Sections 3 and 4 review the modern geometry that will be needed. Section 3 is mostly about Frobenius' theorem, Lie algebras and Lie groups.[4] Section 4 expounds Lie group actions. It ends with the central idea of the co-adjoint representation of a Lie group G on the dual $\mathfrak{g}^*$ of its Lie algebra. This review enables us to better understand the

[4]Its first two Subsections also provide some pre-requisites for Malament (this vol.).

motivations for Poisson manifolds (5.1); and then to exhibit examples, and prove some main properties (Section 5.2 onwards). Section 6 applies this material to symmetry and conservation in mechanical systems. In particular, it expresses conserved quantities as momentum maps, and proves Noether's theorem for Hamiltonian mechanics on Poisson manifolds. Finally, in Section 7, we prove one of the several main theorems about symplectic reduction. It concerns the case where the natural configuration space for a system is itself a Lie group G: this occurs both for the rigid body and ideal fluids. In this case, quotienting the natural phase space (the cotangent bundle on G) gives a Poisson manifold that "is" the dual $\mathfrak{g}^*$ of G's Lie algebra.[5]

To sum up:— The overall effect of this exposition is, I hope, to illustrate this Chapter's mottoes: that classical mechanics is alive and kicking, not least through deepening our understanding of time-honoured systems such as the rigid body—whose analysis in traditional textbooks can be all too confusing!

2 Symplectic reduction: an overview

We begin by briefly reviewing Hamiltonian mechanics and Noether's theorem, in Section 2.1.[6] This prepares us for the idea of symplectic reduction, Section 2.2: which we then illustrate using "relationist" mechanics, Section 2.3.

2.1 Hamiltonian mechanics and Noether's theorem: a review

2.1.1 Symplectic manifolds; the cotangent bundle as a symplectic manifold

A *symplectic structure* or *symplectic form* on a manifold M is defined to be a differential 2-form ω on M that is closed (i.e. its exterior derivative $\mathbf{d}\omega$ vanishes) and is non-degenerate. That is: for any $x \in M$, and any two tangent vectors at x, $\sigma, \tau \in T_x$:

$$\mathbf{d}\omega = 0 \quad \text{and} \quad \forall\, \tau \neq 0,\ \exists \sigma : \quad \omega(\tau, \sigma) \neq 0 \ . \tag{2.1}$$

Such a pair (M, ω) is called a *symplectic manifold*. There is a rich theory of symplectic manifolds; but we shall only need a small fragment of it. (In particular, the fact that we mostly avoid the theory of canonical transformations means we will not need the theory of Lagrangian sub-manifolds.)

[5]In this endeavour, my sources are four books by masters of the subject: Abraham and Marsden (1978), Arnold (1989), Marsden and Ratiu (1999) and Olver (2000). But again, be warned: my selection is severe, as anyone acquainted with these or similar books will recognize.

[6]For more details about differential geometry, cf. Sections 3.1 and 3.2. For more details about the geometric formulation of mechanics, cf. Arnold (1989) or Marsden and Ratiu (1999); or Singer (2001) (more elementary than this exposition) or Abraham and Marsden (1978) (more advanced); or Butterfield (2006) (at the same level). Of many good textbooks of mechanics, I admire especially Desloge (1982) and Johns (2005).

First, it follows from the non-degeneracy of ω that M is even-dimensional. The reason lies in a theorem of linear algebra, which one then applies to the tangent space at each point of M. Namely, for any bilinear form $\omega : V \times V \to \mathbb{R}$: if ω is antisymmetric of rank $r \leq m \equiv \dim(V)$, then r is even. That is: $r = 2n$ for some integer n, and there is a basis $e_1, ..., e_i, ..., e_m$ of V for which ω has a simple expansion as wedge-products

$$\omega \;=\; \Sigma_{i=1}^{n}\, e^i \wedge e^{i+n} \;\; ; \tag{2.2}$$

equivalently, ω has the $m \times m$ matrix

$$\omega \;=\; \begin{pmatrix} \mathbf{0} & \mathbf{1} & \mathbf{0} \\ -\mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \;\; . \tag{2.3}$$

where $\mathbf{1}$ is the $n \times n$ identity matrix, and similarly for the zero matrices of various sizes. This *normal form* of antisymmetric bilinear forms is an analogue of the Gram-Schmidt theorem that an inner product space has an orthonormal basis, and is proved by an analogous argument.

So if an antisymmetric bilinear form is non-degenerate, then $r \equiv 2n = m$. That is: eq. 2.3 loses its bottom row and right column consisting of zero matrices, and reduces to the $2n \times 2n$ *symplectic matrix* ω given by

$$\omega := \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ -\mathbf{1} & \mathbf{0} \end{pmatrix} \;\; . \tag{2.4}$$

Second, the non-degeneracy of ω implies that at any $x \in M$, there is a basis-independent isomorphism $\omega^\flat$ from the tangent space T_x to its dual T_x^*. Namely: for any $x \in M$ and $\tau \in T_x$, the value of the 1-form $\omega^\flat(\tau) \in T_x^*$ is defined by

$$\omega^\flat(\tau)(\sigma) := \omega(\sigma, \tau) \quad \forall \sigma \in T_x \; . \tag{2.5}$$

This also means that a symplectic structure enables a covector field, i.e. a differential one-form, to determine a vector field. Thus for any function $H : M \to \mathbb{R}$, so that dH is a differential 1-form on M, the inverse of $\omega^\flat$ (which we might write as $\omega^\sharp$), carries dH to a vector field on M, written X_H. This is the key idea whereby in Hamiltonian mechanics, a scalar function H determines a dynamics; cf. Section 2.1.2.

So far, we have noted some implications of ω being non-degenerate. The other part of the definition of a symplectic form (for a manifold), viz. ω being closed, $\mathbf{d}\omega = 0$, is also important. We shall see in Section 2.1.3 that it implies that a vector field X on a symplectic manifold M preserves the symplectic form ω (i.e. in more physical jargon: generates (a one-parameter family of) canonical transformations) iff X is Hamiltonian in the sense that there is a scalar function f such that $X = X_f \equiv \omega^\sharp(df)$. Or in terms of the Poisson bracket, with $\cdot$ representing the argument place for a scalar function: $X(\cdot) = X_f(\cdot) \equiv \{\cdot, f\}$.

So much by way of introducing symplectic manifolds. I turn to showing that any cotangent bundle T^*Q is such a manifold. That is: it has, independently of a choice of coordinates or bases, a symplectic structure.

Given a manifold Q ($\dim(Q)=n$) which we think of as the system's configuration space, choose any local coordinate system q on Q , and the natural local coordinates q, p thereby induced on T^*Q. We define the 2-form

$$dp \wedge dq := dp_i \wedge dq^i := \Sigma_{i=1}^n dp_i \wedge dq^i \ . \tag{2.6}$$

In fact, eq. 2.6 defines the same 2-form, whatever choice we make of the chart q on Q. For $dp \wedge dq$ is the exterior derivative of a 1-form on T^*Q which is defined naturally (i.e. independently of coordinates or bases) from the derivative (also known as: tangent) map of the projection

$$\pi : (q,p) \in T^*Q \mapsto q \in Q. \tag{2.7}$$

Thus consider a tangent vector τ (not to Q, but) to the cotangent bundle T^*Q at a point $\eta = (q,p) \in T^*Q$, i.e. $q \in Q$ and $p \in T^*_q$. Let us write this as: $\tau \in T_\eta(T^*Q) \equiv T_{(q,p)}(T^*Q)$. The derivative map, $D\pi$ say, of the natural projection π applies to τ:

$$D\pi : \tau \in T_{(q,p)}(T^*Q) \mapsto (D\pi(\tau)) \in T_q \ \ . \tag{2.8}$$

Now define a 1-form θ_H on T^*Q by

$$\theta_H : \tau \in T_{(q,p)}(T^*Q) \mapsto p(D\pi(\tau)) \in \mathrm{I\!R} \ ; \tag{2.9}$$

where in this definition of θ_H, p is defined to be the second component of τ's base-point $(q,p) \in T^*Q$; i.e. $\tau \in T_{(q,p)}(T^*Q)$ and $p \in T^*_q$.

This 1-form is called the *canonical 1-form* on T^*Q. One now checks that in any natural local coordinates q, p, θ_H is given by

$$\theta_H = p_i dq^i. \tag{2.10}$$

Finally, we define a 2-form by taking the exterior derivative of θ_H:

$$\mathbf{d}(\theta_H) := \mathbf{d}(p_i dq^i) \equiv dp_i \wedge dq^i \ . \tag{2.11}$$

One checks that this 2-form is closed (since $\mathbf{d}^2 = 0$) and non-degenerate. So $(T^*Q, \mathbf{d}(\theta_H))$ is a symplectic manifold. Accordingly, $\mathbf{d}(\theta_H)$, or its negative $-\mathbf{d}(\theta_H)$, is called the *canonical symplectic form*, or *canonical 2-form*.

There is a theorem (Darboux's theorem) to the effect that locally, any symplectic manifold "looks like" a cotangent bundle: or in other words, a cotangent bundle is locally a "universal" example of symplectic structure. We will not go into details; but in Section 5.3.4, we will discuss the generalization of this theorem for Poisson manifolds. But first we review, in the next two Subsections, Hamilton's equations, and Noether's theorem.

2.1.2 Geometric formulations of Hamilton's equations

As we already emphasised, the main geometric idea behind Hamilton's equations is that a gradient, i.e. covector, field dH determines a vector field X_H. So to give a geometric formulation of Hamilton's equations at a point $x = (q, p)$ in a cotangent bundle T^*Q, let us write $\omega^\sharp$ for the (basis-independent) isomorphism from the cotangent space to the tangent space, $T^*_x \to T_x$, induced by $\omega := -\mathbf{d}(\theta_H) = dq^i \wedge dp_i$ (cf. eq. 2.5). Then Hamilton's equations may be written as:

$$\dot{x} = X_H(x) = \omega^\sharp(\mathbf{d}H(x)) = \omega^\sharp(dH(x)) \ . \tag{2.12}$$

There are various other formulations. Applying $\omega^\flat$, the inverse isomorphism $T_x \to T^*_x$, to both sides, we get

$$\omega^\flat X_H(x) = dH(x) \ . \tag{2.13}$$

In terms of the symplectic form ω at x, this is: for all vectors $\tau \in T_x$

$$\omega(X_H(x), \tau) = dH(x) \cdot \tau \ ; \tag{2.14}$$

or in terms of the contraction (also known as: interior product) $\mathbf{i}_X\alpha$ of a differential form α with a vector field X, with $\cdot$ marking the argument place of $\tau \in T_x$:

$$\mathbf{i}_{X_H}\omega := \omega(X_H(x), \cdot) = dH(x)(\cdot) \ . \tag{2.15}$$

More briefly, and now written for any function f, it is:

$$\mathbf{i}_{X_f}\omega = df \ . \tag{2.16}$$

Finally, recall the relation between the Poisson bracket and the directional derivative (or the Lie derivative $\mathcal{L}$) of a function: viz.

$$\mathcal{L}_{X_f} g = dg(X_f) = X_f(g) = \{g, f\} \ . \tag{2.17}$$

Combining this with eq. 2.16, we can state the relation between the symplectic form and Poisson bracket in the form:

$$\{g, f\} = dg(X_f) = \mathbf{i}_{X_f} dg = \mathbf{i}_{X_f}(\mathbf{i}_{X_g}\omega) = \omega(X_g, X_f) \ . \tag{2.18}$$

2.1.3 Noether's theorem

The core idea of Noether's theorem, in both the Lagrangian and Hamiltonian frameworks, is that to every continuous symmetry of the system there corresponds a conserved quantity (a first integral, a constant of the motion). The idea of a continuous symmetry is made precise along the following lines: a symmetry is a vector field on the state-space that (i) preserves the Lagrangian (respectively, Hamiltonian) and (ii) "respects" the structure of the state-space.

In the Hamiltonian framework, the heart of the proof is a "one-liner" based on the fact that the Poisson bracket is antisymmetric. Thus for any scalar functions f and H on a symplectic manifold (M, ω) (and so with a Poisson bracket given by eq. 2.18), we have that at any point $x \in M$

$$X_f(H)(x) \equiv \{H, f\}(x) = 0 \quad \text{iff} \quad 0 = \{f, H\}(x) \equiv X_H(f)(x) \ . \tag{2.19}$$

In words: around x, H is constant under the flow of the vector field X_f (i.e. under what the evolution would be if f was the Hamiltonian) iff f is constant under the flow X_H. Thinking of H as the physical Hamiltonian, so that X_H represents the real time-evolution (sometimes called: the dynamical flow), this means: around x, X_f preserves the Hamiltonian iff f is constant under time-evolution, i.e. f is a conserved quantity (a constant of the motion).

But we need to be careful about clause (ii) above: the idea that a vector field respects" the structure of the state-space. In the Hamiltonian framework, this is made precise as preserving the symplectic form. Thus we define a vector field X on a symplectic manifold (M, ω) to be *symplectic* (also known as: *canonical*) iff the Lie-derivative along X of the symplectic form vanishes, i.e. $\mathcal{L}_X \omega = 0$. (This definition is equivalent to X's generating (active) canonical transformations, and to its preserving the Poisson bracket. But I will not go into details about these equivalences: for they belong to the theory of canonical transformations, which, as mentioned, I will not need to develop.)

We also define a *Hamilton system* to be a triple (M, ω, H) where (M, ω) is a symplectic manifold and $H : M \to \mathbb{R}$, i.e. $M \in \mathcal{F}(M)$. And then we define a (continuous) *symmetry* of a Hamiltonian system to be a vector field X on M that:
(i) preserves the Hamiltonian function, $\mathcal{L}_X H = 0$; and
(ii) preserves the symplectic form, $\mathcal{L}_X \omega = 0$.

These definitions mean that to prove Noether's theorem from eq. 2.19, it will suffice to prove that a vector field X is symplectic iff it is locally of the form X_f. Such a vector field is called *locally Hamiltonian*. (And a vector field is called *Hamiltonian* if there is a global scalar $f : M \to \mathbb{R}$ such that $X = X_f$.) In fact, two results from the theory of differential forms, the Poincaré Lemma and Cartan's magic formula, make it easy to prove this; (for a vector field on any symplectic manifold (M, ω), i.e. (M, ω) does not need to be a cotangent bundle).

Again writing $\mathbf{d}$ for the exterior derivative, we recall that a k-form α is called:
(i): *exact* if there is a $(k-1)$-form β such that $\alpha = \mathbf{d}\beta$; (cf. the elementary definition of an exact differential);
(ii): *closed* if $\mathbf{d}\alpha = 0$.
The *Poincaré Lemma* states that every closed form is locally exact. To be precise: for any open set U of M, we define the vector space $\Omega^k(U)$ of k-form fields on U. Then the Poincaré Lemma states that if $\alpha \in \Omega^k(M)$ is closed, then at every $x \in M$ there is a neighbourhood U such that $\alpha \mid_U \in \Omega^k(U)$ is exact.

Cartan's magic formula is a useful formula (proved by straightforward calculation) relating the Lie derivative, contraction and the exterior derivative. It says that if X

is a vector field and α a k-form on a manifold M, then the Lie derivative of α with respect to X (i.e. along the flow of X) is

$$\mathcal{L}_X \alpha = \mathbf{d}\mathbf{i}_X \alpha + \mathbf{i}_X \mathbf{d}\alpha \ . \tag{2.20}$$

We now argue as follows. Since ω is closed, i.e. $\mathbf{d}\omega = 0$, Cartan's magic formula, eq. 2.20, applied to ω becomes

$$\mathcal{L}_X \omega \equiv \mathbf{d}\mathbf{i}_X \omega + \mathbf{i}_X \mathbf{d}\omega = \mathbf{d}\mathbf{i}_X \omega \ . \tag{2.21}$$

So for X to be symplectic is for $\mathbf{i}_X\omega$ to be closed. But by the Poincaré Lemma, if $\mathbf{i}_X\omega$ closed, it is locally exact. That is: there locally exists a scalar function $f : M \rightarrow \mathrm{I\!R}$ such that

$$\mathbf{i}_X \omega = df \ \text{ i.e. } \ X = X_f \ . \tag{2.22}$$

So for X to be symplectic is equivalent to X being locally Hamiltonian.

Thus we have

> **Noether's theorem for a Hamilton system** If X is a symmetry of a Hamiltonian system (M, ω, H), then locally $X = X_f$; so by the anti-symmetry of the Poisson bracket, eq. 2.19, f is a constant of the motion. And conversely: if $f : M \rightarrow \mathrm{I\!R}$ is a constant of the motion, then X_f is a symmetry.

We will see in Section 6.2 that most of this approach to Noether's theorem, in particular the "one-liner" appeal to the anti-symmetry of the Poisson bracket, eq. 2.19, carries over to the more general framework of Poisson manifolds. For the moment, we mention an example (which we will also return to).

For most Hamiltonian systems in euclidean space $\mathrm{I\!R}^3$, spatial translations and rotations are (continuous) symmetries. Let us consider in particular a system we will discuss in more detail in Section 2.3: N point-particles interacting by Newtonian gravity. The Hamiltonian is a sum of two terms, which are each individually invariant under translations and rotations:

(i) a kinetic energy term K; though I will not go into details, it is in fact defined by the euclidean metric of $\mathrm{I\!R}^3$, and is thereby invariant; and

(ii) a potential energy term V; it depends only on the particles' relative distances, and is thereby invariant.

The corresponding conserved quantities are the total linear and angular momentum.[7]

[7]By the way, this Hamiltonian is *not* invariant under boosts. But as I said in (iii) of Section 1.2, I restrict myself to time-independent transformations; the treatment of symmetries that "represent the relativity of motion" needs separate discussion.

2.2 The road ahead

In this Subsection, four comments will expand on the introductory comment (iv) of Section 1.2, and also give some information about the history of symplectic reduction and about some crucial examples.

(1): *Generalizing from Noether's theorem; Poisson manifolds*:—
Noether's theorem tells us that a continuous symmetry, i.e. a one-parameter group of symmetries, determines a first integral (i.e. a constant of the motion). So a larger group of symmetries, i.e. a group with several parameters, implies several first integrals. The phase flow is therefore confined to the intersection of the level surfaces of these integrals: an intersection which is in general a manifold. In other words: the simultaneous level manifold of these integrals is an invariant manifold of the phase flow.

It turns out that, in many useful cases, this manifold is also invariant under an appropriately chosen subgroup of the group of symmetries; and that the quotient space, i.e. the set of orbits under the action of this subgroup, is a manifold with a natural structure induced by the original Hamiltonian system that is sufficient to do mechanics in Hamiltonian style. The quotient space is therefore called the 'reduced phase space'.

But in some cases, this natural structure is not a symplectic form, but a (mild) generalization in which the the form is allowed to be degenerate; i.e. like eq. 2.3 rather than eq. 2.4. A manifold equipped with such a structure need not be a quotient manifold. It can instead be defined in terms of a generalization of the usual Poisson bracket, as defined in terms of the symplectic form by eq. 2.18.

The key idea is to postulate a bracket, acting on the scalar functions $F : M \to \mathbb{R}$ on *any* manifold M, and possessing four properties enjoyed by the usual Poisson bracket. One of the properties is anti-symmetry, emphasised in Section 2.1.3's proof of Noether's theorem. The other three are that the postulated bracket, again written $\{,\}$, is: to be bilinear; to obey the Jacobi identity for any real functions F, G, H on M, i.e.

$$\{\{F, H\}, G\} + \{\{G, F\}, H\} + \{\{H, G\}, F\} = 0 \ ; \tag{2.23}$$

and to obey Leibniz' rule for products, i.e.

$$\{F, H \cdot G\} = \{F, H\} \cdot G + H \cdot \{F, G\} \ . \tag{2.24}$$

We will see in Section 5 that such a bracket, again called 'Poisson bracket', provides a sufficient framework for mechanics in Hamiltonian style. In particular, it induces an anti-symmetric bilinear form that may be degenerate, as in eq. 2.3. A manifold M equipped with such a bracket is called a *Poisson manifold*.

The allowance of degeneracy means that a Poisson manifold can have *odd* dimension; while we saw in Section 2.1.1 that any symplectic manifold is even-dimensional. On the other hand, this generalized Hamiltonian mechanics will have clear connections with the usual formulation of Section 2.1. The main connection will be the result that any Poisson manifold M is a disjoint union of even-dimensional manifolds, on which M's degenerate antisymmetric bilinear form restricts to be non-degenerate.[8]

(2): *Historical roots*:—
The theory of symplectic reduction has deep historical roots in the work of classical mechanics' monumental figures. In part, this is no surprise. As mentioned in (i) of Section 1.2, cyclic coordinates underpin the role of symmetry in mechanics, and in particular Noether's theorem. And Newton's solution of the Kepler problem provides an example: witness textbooks' expositions of the transition to centre-of-mass coordinates, and of polar coordinates with the angle being cyclic (yielding angular momentum as the conserved quantity). So it is unsurprising that various results and ideas of symplectic reduction can be seen in the work of such masters as Euler, Lagrange, Hamilton, Jacobi, Lie and Poincaré; for example (as we will see), in Euler's theory of the rigid body.

But the history also holds a surprise. It turns out that Lie's epoch-making work on Lie groups already contained a detailed development of much of the general, modern theory.[9] The sad irony is that most of Lie's insights were not taken up—and were then repeatedly re-discovered. So this is yet another example (of which the history of mathematics has so many!) of the saying that he who does not learn from history is doomed to repeat it. The consolation is of course that it is often easier, and more fun, to re-discover something than to learn it from another...

Thus it was only from the mid-1960s that the theory, in essentially the form Lie had, was recovered and cast in the geometric language adopted by modern mechanics; namely, by contemporary masters such as Arnold, Kostant, Marsden, Meyer, Smale, Souriau and Weinstein; (cf. this Chapter's first motto). Happily, several of these modern authors are scholars of the history, and even their textbooks give some historical details: cf. Marsden and Ratiu (1999, pp. 336-8, 369-370, 430-432), and the notes to each Chapter of Olver (2000: especially p.427-428). (Hawkins (2000) is a full history of Lie groups from 1869 to 1926; for Lie, cf. especially its Sections 1.3, 2.5 and Chapter 3, especially 3.2.)

In any case, setting history aside: symplectic reduction has continued since the 1970s to be an active research area in contemporary mechanics, and allied fields such as symplectic geometry. So it has now taken its rightful place as a major part of the contemporary renaissance of classical mechanics: as shown by ...

(3): *Two examples: the rigid body and the ideal fluid*:—
Two examples illustrate vividly how symplectic reduction can give new physical understanding, even of time-honoured examples: the rigid body and the ideal fluid—as attested by this Chapter's mottoes. (Section 2.3 will develop a third example, more closely related to philosophy.)

As to the rigid body: we will see (especially in Section 5) that symplectic reduction considerably clarifies the elementary theory of the rigid body (Euler's equations, Euler

[8]Because of these clear connections, it is natural to still call the more general framework 'Hamiltonian'; as is usually done. But of course this is just a verbal matter.

[9]The main source is his (1890). Besides, Arnold (1989: 456) reports that the prototype example of a Poisson manifold, viz. the dual of a finite-dimensional Lie algebra, was already understood by Jacobi.

angles etc.): which, notoriously, can be all too confusing! For simplicity, I shall take the rigid body to be pivoted, so as to set aside translational motion. This will mean that the group of symmetries defining the quotienting procedure will be the rotation group. It will also mean that the rigid body's configuration space is given by the rotation group, since any configuration can be labelled by the rotation that obtains it from some reference-configuration. So in this application of symplectic reduction, the symmetry group (viz. the rotation group) will act on *itself* as the configuration space. This example will also give us our prototype example of a Poisson manifold.

As to the ideal fluid, i.e. a fluid that is incompressible and inviscid (with zero viscosity): this is of course an infinite-dimensional system, and so (as I announced in Section 1.2) outside the scope of this Chapter. So I will not go into any details, but just report the main idea.

The equations of motion of an ideal fluid, Euler's equations, are usually derived either by applying Newton's second law $\mathbf{F} = m\mathbf{a}$ to a small fluid element; or by a heuristic use of the Lagrangian or Hamiltonian approach (as in heuristic classical field theories). But in the mid-1960s, Arnold showed how the latter derivations could be understood in terms of a striking, even beautiful, analogy with the above treatment of the rigid body. Namely, the analogy shows that the configuration space of the fluid is an infinite-dimensional group; as follows. The configuration of an ideal fluid confined to some container occupying a volume $V \subset \mathbb{R}^3$ is an assignment to each spatial position $x \in V$ of an infinitesimal fluid element. Given such an assignment as a reference-configuration, any other configuration can be labelled by the volume-preserving diffeomorphism d from V to V that carries the reference-configuration to the given one, by dragging each fluid element along by d. So given a choice of reference-configuration, the fluid's configuration space is given by the infinite-dimensional group $\mathcal{D}$ of diffeomorphisms $d : V \rightarrow V$: just as the rotation group is the configuration space of a (pivoted) rigid body. $\mathcal{D}$ then forms the basis for rigorous Lagrangian and Hamiltonian theories of an ideal fluid.

These theories turn out to have considerable analogies with the Lagrangian and Hamiltonian theories of the rigid body, thanks to the fact that in both cases the symmetry group forms the configuration space. In particular, Euler's equations for ideal fluids are the analogues of Euler's equations for a rigid body. Besides, these rigorous theories of fluids (and symplectic reduction applied to them) are scientifically important: they have yielded various general theorems, and solved previously intractable problems. (For more details, cf. Abraham and Marsden (1978: Sections 4.4 and 4.6 for the rigid body, and 5.5.8 for the ideal fluid), Arnold (1989: Appendix 2:C to 2:F for the rigid body, and 2:G to 2:L for the ideal fluid), and Marsden and Ratiu (1999: Chapters 1.4 and 15 for the rigid body, and 1.5, p. 266, for the ideal fluid).)

(4): *Philosophical importance*:—
Symplectic reduction is also, I submit, philosophically important; in at least two ways. The first way is specific: it illustrates some methodological morals about how classical mechanics analyses problems. I develop this theme in (Butterfield 2005). The second

way is more general: the theory, or rather various applications of it, is directly relevant to disputes in the philosophy of space and time, and of mechanics. This relevance is recognized in contemporary philosophy of physics. So far as I know, the authors who develop these connections in most detail are Belot and Earman. They discuss symplectic reduction in connection with such topics as:

(i) the treatment of symmetries, including gauge symmetries;

(ii) the dispute between absolute and relationist conceptions of space and time; and

(iii) the interpretation of classical general relativity (a topic which connects (i) and (ii), and bears on heuristics for quantum gravity).

Thus Belot (1999, 2000, 2001, 2003, 2003a) and Earman (2003) discuss mainly (i) and-or (ii); Belot and Earman (2001) discusses (iii). For (i) and (ii), I also recommend Wallace (2003).

But these papers have a demanding pre-requisite: they invoke, but do not expound, the theory of symplectic reduction. They also discuss infinite-dimensional systems (especially classical electromagnetism and general relativity), without developing finite-dimensional examples like the rigid body. Indeed, there is, so far as I know, no article-length exposition of the theory which is not unduly forbidding for philosophers. So I aim to give such an exposition, to help readers of papers such as those cited.[10]

As an appetizer for this exposition, I will first (in Section 2.3) follow Belot in presenting the general features of a finite-dimensional symplectic reduction which has vivid philosophical connections, viz. to the absolute vs. relationist debate. This example concerns a system of point-particles in Euclidean space, either moving freely or interacting by a force such as Newtonian gravity. (The symmetries defining the quotienting procedure are given by the Euclidean group of translations and rotations.) For philosophers, this will be a good appetizer for symplectic reduction, since it sheds considerable light on relationism about space of the sort advocated by Leibniz and Mach.

2.3 Appetizer: Belot on relationist mechanics

2.3.1 Comparing two quotienting procedures

In several papers, Belot discusses how symplectic reduction bears on the absolute-vs.-relational debate about space. I shall pick out one main theme of his discussions: the comparison of a relational classical mechanical theory with what one gets by quotienting the orthodox absolutist (also called a 'substantivalist') classical mechanics, by

[10] As I said in Section 1.2, my material is drawn from the books by Abraham and Marsden, Arnold, Marsden and Ratiu, and Olver. More precisely, I will mostly draw on: Abraham and Marsden (1978: Sections 3.1-3.3, 4.1-4.3), Arnold (1989: Appendices 2, 5 and 14), Marsden and Ratiu (1999: Chapters 9-13) and Olver (2000: Chapter 6). And much of what follows—in spirit, and even in letter—is already in Lie (1890)! As a (non-philosophical) introduction to symplectic reduction, I also recommend Singer (2001). It is at a yet more elementary level than what follows; e.g. it omits Poisson manifolds and co-adjoint representations.

an appropriate symmetry group. His main contention—which I endorse—is that this comparison sheds considerable light on relationism: on both its motivation, and its advantages and disadvantages.[11]

Belot's overall idea is as follows. Where the relationist admits one possible configuration, as (roughly) a specification of all the distances (and thereby angles) between all the parts of matter, the absolutist (or substantivalist) sees an infinity of possibilities: one for each way the relationist's configuration (a *relative configuration*) can be embedded in the absolute space. This makes it natural to take the relationist to be envisaging a mechanics which is some sort of "quotient" of the absolutist's mechanics.

In particular, on the traditional conception of space as Euclidean (modelled by $\mathbb{R}^3$), each of the relationist's relative configurations corresponds to an equivalence class of absolutist configurations (i.e. embeddings of arrangements of matter into $\mathbb{R}^3$), with the members of the class related by spatial translations and rotations, i.e. elements of the Euclidean group. In the jargon of group actions, to be developed in Section 4: the Euclidean group acts on the set of all absolutist configurations, and a relative configuration corresponds to an orbit of this action. So it is natural to take the relationist to be envisaging a mechanics which quotients this action of the Euclidean group, to get a *relative configuration space.* A relationist mechanics, of Lagrangian or Hamiltonian type, is then to be built up on this space of relative configurations.

But as Belot emphasises, one can instead consider quotienting the absolutist's state-space—i.e. in a Hamiltonian framework, the phase space—rather than their configuration space. Indeed, this is exactly what one does in symplectic reduction. In particular, the Euclidean group's action on the absolutist's configuration space, Q say, can be lifted to give an action on the cotangent bundle T^*Q; which is accordingly called the 'cotangent lift'. One can then take the quotient, i.e. consider the orbits into which T^*Q is partitioned by the cotangent lift.

We thus have two kinds of theories to compare: (i) the relationist theories, built up from the relative configuration space; which for the sake of comparison with symplectic reduction we take to be Hamiltonian, rather than Lagrangian; (ii) theories obtained by quotienting "later", i.e. quotienting the absolutist's cotangent bundle.

I will now spell out this comparison. But I will not try to summarize Belot's more detailed conclusions, about what such a comparison reveals about the advantages and disadvantages of relationism. They are admirably subtle, and so defy summary: they can mainly be found at his (2000: p. 573-574, 582; 2001: Sections VIII to X). (Rovelli (this volume) also discusses relationism.)

As befits an appetizer, I will also (like Belot) concentrate on as simple a case as possible: a mechanics of N point-particles, which is to assume a Euclidean spatial geometry. Of course, the absolutist make this assumption by postulating a Euclidean space; but for the relationist, the assumption is encoded in constraints relating the

[11] The main references are Belot (1999, 2001, 2003: Sections 3.5, 5). Cf. also his (2000: Sections 4 to 5.3), (2003a: Section 6). Though I recommend all these papers, the closest template for what follows is (2001: Section VI et seq.).

various inter-particle distances. The main current example of a relationist mechanics of such a system is due to Barbour and Bertotti (1982), though they develop it in the Lagrangian framework; (to be precise, in terms of Jacobi's principle). Belot also discusses other relational theories, including field theories, i.e. theories of infinite systems; some of them also due to Barbour, and in a Lagrangian framework. But in this Section I only consider N point-particles.

Also, I will also not discuss boosts, though of course the relationist traditionally proposes to identify any two absolutist states of motion related by a boost. In terms of group actions, this means I will consider quotienting by an action of the euclidean group, but not the Galilei group. (Cf. how I set aside time-dependent transformations already in (iii) of Section 1.2.) I will also postpone to later Sections technical details, even when our previous discussion makes them accessible.

Finally, a warning to avoid later disappointment! The later Sections will not include a full analysis of the euclidean group's actions on configuration space and phase space, and their quotients. That would involve technicalities going beyond an appetizer. Instead (as mentioned at the end of Section 1.2), the material in later Sections is chosen so as to lead up to Section 7's theorem, the Lie-Poisson reduction theorem, about quotienting the phase space of a system whose configuration space is a Lie group. Further reasons for presenting the material for this theorem will be given in Section 5.1.

2.3.2 The spaces and group actions introduced

Let us begin by formulating the orthodox absolutist mechanics of N point-particles interacting by Newtonian gravity, together with the action of the Euclidean group.

Each point-particle occupies a point of $\mathbb{R}^3$, so that the configuration space Q is $\mathbb{R}^{3N}$: $\dim(Q) = 3N$. So the phase space for Hamiltonian mechanics will be the cotangent bundle $T^*Q \ni (q,p)$: $\dim(T^*Q) = 6N$.

The Hamiltonian is a sum of kinetic and potential terms, K and V. K depends only on the ps, and V only on the qs. In cartesian coordinates, with i now labelling particles $i = 1, ..., N$ rather than degrees of freedom, we have the familiar expressions:

$$H(q,p) = K(p) + V(q) \;\text{ with }\; K \;=\; \Sigma_i \frac{\mathbf{p}_i^2}{2m_i} \;\;,\;\; V(q) \;=\; G\,\Sigma_{i<j} \frac{m_i m_j}{\parallel \mathbf{q}_i - \mathbf{q}_j \parallel} \tag{2.25}$$

where m_i are the masses and G is the gravitational constant.[12,13]

[12]From the broader philosophical perspective, the most significant feature of eq. 2.25 is no doubt the fact that the potential is a sum of all the two-body potential energies for the configuration $q \in Q$: there are no many-body interactions.

[13]Incidental remark. In fact, the kinetic energy can be represented by a metric g on the configuration space. For Hamiltonian mechanics, this means that the kinetic energy scalar K on the cotangent bundle T^*Q can be defined by applying Q's metric g to the projections of the momenta p, where at each point $(q,p) \in T^*Q$ the projection is made with the preferred isomorphism $\omega^\sharp : T^*_q \to T_q$; (cf. eq.

The euclidean group E (aka: $E(3)$) is the group (under composition) of translations, rotations and reflections on $\mathbb{R}^3$. But since we will be interested in continuous symmetries, we will ignore reflections, and so consider the subgroup of orientation-preserving translations and rotations; i.e. the component of the group connected to the identity transformation (which I will also write as E). This is a Lie group, i.e. a group which is also a manifold, with the group operations smooth with respect to the manifold structure. Section 3 will give formal details. Here we just note that we need three real numbers to specify a translation ($\mathbf{x} = (x, y, z)$), and three to specify a rotation (two for an axis, and one for the angle through which to rotate); and accordingly, it is unsurprising that as a manifold, the dimension of E is 6: $\dim(E) = 6$.

E acts in the obvious sense on $\mathbb{R}^3$. For example, if $g \in E$ is translation by $\mathbf{x} \in \mathbb{R}^3$, g induces the map $\mathbf{q} \in \mathbb{R}^3 \mapsto \mathbf{q} + \mathbf{x}$. Similarly for a rotation induces: again, Section 3 will give a formal definition.

Now let E act in this way on each of the N factor spaces $\mathbb{R}^3$ of our system's configuration manifold $Q = \mathbb{R}^{3N}$. This defines an action Φ on Q: i.e. for all $g \in E$, there is a map $\Phi_g : Q \to Q$. For example, for $g =$ a translation by $\mathbf{x} \in \mathbb{R}^3$, we have

$$\Phi_g : (\mathbf{q}_j) = (\mathbf{q}_1, ..., \mathbf{q}_N) \in Q \mapsto (\mathbf{q}_1 + \mathbf{x}, ..., \mathbf{q}_N + \mathbf{x}) \in Q \ ; \tag{2.27}$$

and similarly for rotations. Since the potential function $V : Q \to \mathbb{R}$ of eq. 2.25 depends only on inter-particle distances, each map $\Phi_g : Q \to Q$ is a symmetry of the potential; i.e. we have $V(\Phi_g(q)) = V(q)$.

The action Φ (i.e. the assignment $g \in E \mapsto \Phi_g$) induces an action of E on $T^*Q = T^*\mathbb{R}^{3N}$, called the *cotangent lift* of Φ to T^*Q, and usually written as Φ^*; so that we have for each $g \in E$ a lifted map $\Phi^*_g : T^*Q \to T^*Q$. Again, the details can wait till later (Section 4). But the idea is that each map Φ_g on Q is smooth, and so maps curves to curves, and so vectors to vectors, and so covectors to covectors, and so on.

Unsurprisingly, each of the lifted maps $\Phi^*_g : T^*Q \to T^*Q$ leaves the potential V, now considered as a scalar on T^*Q, invariant: i.e. we have $V(\Phi^*_g(q,p)) = V(q,p) \equiv V(q)$. But furthermore, each of the lifted maps Φ^*_g is a symmetry of the Hamilton system, in our previous sense (Section 2.1.3). That is: Φ^*_g preserves the Hamiltonian (indeed the kinetic and potential terms are separately invariant); and it preserves the symplectic structure. This means the dynamics is invariant under the action of all $g \in G$: the dynamical histories of the system through (q,p) and through $\Phi^*_g(q,p)$ match exactly at each time. They are qualitatively indistinguishable: in contemporary metaphysical jargon, they are *duplicates*.

At this point, of course, we meet the absolute-vs.-relational debate about space. The absolutist asserts, and the relationist denies, that there being two such indistinguishable possibilities makes sense.[14] So the relationist, presented with the theory above, says

2.12). That is:—

$$K : (q,p) \in T^*Q \mapsto g_q(\omega^\sharp(p), \omega^\sharp(p)) \ . \tag{2.26}$$

[14] The *locus classicus* for this debate is of course the Leibniz-Clarke correspondence, though the

we should cut down the space of possibilities. As I said in Section 2.3.1, it is natural to make this precise in terms of quotienting the action of the euclidean group: a set of absolutist possibilities related one to another by elements of the euclidean group form an equivalence class (an orbit) which is to represent *one* relationist possibility.

But here we need to distinguish two different quotienting procedures. I will call them *Relationism* and *Reductionism* (with capital R's), since the former is close to both traditional and contemporary relationist proposals, and the latter is an example of the orthodox idea of symplectic reduction. As I said in Section 2.3.1, the main difference will be that:

(i): Relationism performs the quotient on E's action on the configuration space Q; the set of orbits form a *relative configuration space*, on which the relationist proposes to build a dynamics, whether Lagrangian or Hamiltonian—yielding in the latter case, a *relative phase space*; whereas

(ii): Reductionism performs the quotient on E's action on the usual phase space T^*Q, the set of orbits forming a *reduced phase space*.

Since our discussion adopts the Hamiltonian framework, it will not matter for what follows, that Relationism, as defined, can adopt the Lagrangian framework, while Reductionism is committed to the Hamiltonian one. What *will* matter is that (i) and (ii) make for phase spaces of different dimensions; the reduced phase space has six more dimensions than the relative phase space. The "dimension gap" is six.

We will see that four of the six variables that describe these dimensions are constants of the motion; the other two vary with time. And for certain choices of values of the constants of the motion (roughly: no rotation), the time-varying variables drop out, and the dynamics according to the Reductionist theory simplifies so as to coincide with that of the Relationist theory. In other words: if we impose no rotation, then the heterodox Relationist dynamics matches the conventional Reduced dynamics.

2.3.3 The Relationist procedure

The Relationist seeks a mechanics based on the *relative configuration space* (RCS). An element of the RCS is to be a pattern of inter-particle distances and angles that is geometrically possible, i.e. compatible with the N particles being embedded in $\mathbb{R}^3$. So, roughly speaking, an element of the RCS is a euclidean configuration, modulo isometries; and the RCS will be the set of orbits $\mathbb{R}^{3N}/E$.

Even before giving a more precise statement, we can state the "punchline" about dimensions, as follows. Since $\dim(E) = 6$, quotienting by E subtracts six dimensions:

protagonists' argumentation is of course sometimes theological. Clarke the absolutist maintains that there are many possible arrangements of bits of matter in space consistent with a specification of all relative distances, saying 'if [the mere will of God] could in no case act without a pre-determining cause ... this would tend to take away all power of choosing, and to introduce fatality.' Leibniz claims there is only one such arrangement: 'those two states ... would not at all differ from one another. Their difference therefore is only to be found in our chimerical supposition of the reality of space in itself.'

that is, the dimension of the RCS will be $3N$-6.

But we need to be more precise about the RCS. For the orbits and quotient spaces to be manifolds, and for dimensions to add or subtract in this simple way, we need to excise two classes of "special" points from $\mathbb{R}^{3N}$, before we quotient. (But I postpone till Section 4 the technical rationale for these excisions.)

Let $\delta_Q \subset \mathbb{R}^{3N}$ be the set of configurations which are symmetric: i.e. each is fixed by some element of E (other than the identity element!). Any configuration in which all the point-particles are collinear provides an example: the configuration is fixed by any rotation about the line as axis. Let Δ_Q be the set of collision configurations; i.e. configurations in which two or more particles are coincident in the usual configuration space $\mathbb{R}^{3N}$. (The Q subscripts will later serve as a reminder that these sets are sets of configurations.) δ_Q and Δ_Q are both of measure zero in $\mathbb{R}^{3N}$. Excise both of them, and call the resulting space, which is again of dimension $3N$: $Q := \mathbb{R}^{3N} - (\delta_Q \cup \Delta_Q)$.

δ_Q and Δ_Q are each closed under the action of E. That is, each is a union of orbits: a euclidean transformation of a symmetric (collision) configuration is also symmetric (collision). So E acts on Q. Now quotient Q by E. Q/E is the Relationist's RCS. Since $\dim(E) = 6$, we have: $\dim(Q/E) = 3N$-6.

These $3N-6$ variables encode all of a (relative) configuration's particle-pair relative distances, $r_{ij} \in \mathbb{R}$ (with i, j labelling particles). Note that there are $N(N-1)/2$ such relative distances; and for $N > 4$, this is greater than $3N - 6$: (for $N >> 4$, it is much greater). So the relative distances, though physically intuitive, give an *over-complete* set of coordinates on Q/E. (So they cannot be freely chosen: there are *constraints* between them.)

So the Relationist seeks a mechanics that uses this RCS. Newton's second law being second-order in time means that she will also need quantities like velocities (in a Lagrangian framework) or like momenta (in a Hamiltonian framework). For the former, she will naturally consider the $N(N-1)/2$ relative velocities $\dot{r}_{ij} := \frac{d}{dt} r_{ij}$; and for the latter, the corresponding momenta $p_{ij} := \frac{\partial L}{\partial \dot{r}_{ij}}$. Again, she must beware of constraints. The tangent and cotangent bundles built on her RCS Q/E will each have dimension $2(3N - 6) = 6N - 12$. So again, for $N > 4$, the number $N(N-1)/2$ of relative velocities $\dot{r}_{ij}$, or of relative momenta p_{ij}, is greater than the number of degrees of freedom concerned; and for $N >> 4$, it is much greater. So again, the relative velocities or relative momenta are over-complete: there are constraints.

On the other hand, if the Relationist uses only these relative quantities, r_{ij} and either $\dot{r}_{ij}$ or p_{ij} (or "equivalent" coordinates on $T(Q/E)$ or $T^*(Q/E)$ that are not over-complete), she faces a traditional problem—whatever the other details of her theory. At least, she faces a problem if she hopes for a deterministic theory which is empirically equivalent to the orthodox absolutist theory. I will follow tradition and state the problem in terms of relative velocities rather than momenta.

The problem concerns rotation; (and herein lies the strength of Newton's and Clarke's position in the debate against Leibniz). For according to the absolutist theory two systems of point-particles could match with respect to all relative distances and relative velocities, and yet have different future evolutions; so that a theory allowing the same possibilities as the absolutist one, yet using only these relative quantities (or "equivalent" variables) would have to be indeterministic.

The simplest example is an analogue for point-particles of Newton's two globes thought-experiment. Thus the systems could each consist of just two point-particles with zero relative velocity. One system could be non-rotating, so that the point-particles fall towards each other under gravity; while the other system could be rotating about an axis normal to the line between the particles, and rotating at just such a rate as to balance the attractive force of gravity.

The Relationist has traditionally replied that they do *not* hope for a theory empirically equivalent to the absolutist one. Rather, they envisage a mechanics in which, of the two systems mentioned, only the non-rotating evolution is possible: more generally, a mechanics in which the universe as a whole must have zero angular momentum. Originally, in authors like Leibniz and Mach, this reply was a promissory note. But modern Relationist theories such as Barbour and Bertotti's (1982) have made good the promise; and they have been extended well beyond point-particles interacting by Newtonian gravity. Besides, since the universe seems in fact to be non-rotating, these theories can even claim to be empirically adequate, at least as regards this principal difference from absolutist theories.[15]

But it is not my brief to go into these theories' details, except by way of comparison with a quotiented version of the absolutist theory: cf. Section 2.3.4.

2.3.4 The Reductionist procedure

The Reductionist's main idea is to quotient only after passing to the orthodox phase space for N point-particles, i.e. the cotangent bundle $T^*\mathbb{R}^{3N}$ of $\mathbb{R}^{3N}$. So the idea is to consider $(T^*\mathbb{R}^{3N})/E$, i.e. the quotient of $T^*\mathbb{R}^{3N}$ by the cotangent-lifted action Φ^* of the euclidean group E.

More precisely, we again proceed by first excising special points that would give technical trouble. But now the points to be excised are in the cotangent bundle $T^*\mathbb{R}^{3N}$, not in $\mathbb{R}^{3N}$. So let $\delta \subset T^*\mathbb{R}^{3N}$ be the set of phase space states whose configurations are symmetric (in the sense of Section 2.3.3's δ_Q). Let $\Delta \subset T^*\mathbb{R}^{3N}$ be the set of collision points; i.e. states in which two or more particles are coincident in the configuration space $\mathbb{R}^{3N}$. Both δ and Δ are of measure zero. Excise both of them, and call the resulting phase space, which is again of dimension $6N$: $M := T^*\mathbb{R}^{3N} - (\delta \cup \Delta)$.

δ and Δ are each closed under the cotangent-lifted action of E on $T^*\mathbb{R}^{3N}$. That is, each is a union of orbits: the cotangent lift of a euclidean transformation acting on a phase space state with a symmetric (collision) configuration yields a state which also

[15] An advocate of the absolutist theory might say that it is odd to make what seems a contingent feature of the universe, non-rotation, a principle of mechanics; and the Relationist might reply that their view has the merit of predicting that the universe does not rotate! I fear there are no clear criteria for settling this methodological dispute; anyway, I will not pursue it.

has a symmetric (collision) configuration. So E acts on M. Now quotient M by E, getting $\bar{M} := M/E$. This is called *reduced phase space.* We have: $\dim(\bar{M}) = \dim(M)$ - $\dim(E) = 6N - 6$.

As emphasised at the end of Section 2.3.2, $\bar{M}$ has six more dimensions than the corresponding Relationist phase space (whether the velocity phase space (tangent bundle) or the momentum phase space (cotangent bundle)). The dimension of *those* phase spaces is $2(3N - 6) = 6N - 12$. Indeed, we can better understand both the reduced phase space $\bar{M}$ and Relationist phase spaces by considering this "dimension gap". There are two extended comments to make.

(1): *Obtaining the Relationist phase space*:—
We can obtain the Relationist momentum phase space from our original phase space M. Thus let M_0 be the subspace of M in which the system has total linear momentum and total angular momentum both equal to zero. Since these are constants of the motion, M_0 is dynamically closed and so supports a Hamiltonian dynamics given just by restriction of the original dynamics. With linear and angular momentum each contributing three real numbers, $\dim(M_0) = \dim(M)$ - $6 = 6N - 6$. Furthermore, M_0 is closed under (is a union of orbits under) the cotangent-lifted action of E. So let us quotient M_0 by this action of E, and write $\bar{M}_0 := M_0/E$. Then $\dim(\bar{M}_0) = 6N - 6 - 6 = 6N - 12$.

Now recall that this is the dimension of the phase space of the envisaged Relationist theory built on the RCS Q/E. And indeed, as one would hope: $\bar{M}_0$ is the Hamiltonian version of Barbour and Bertotti's 1982 Relational theory; (recall that they work in a Lagrangian framework).

That is: $\bar{M}_0$ is a symplectic manifold, and points in $\bar{M}_0$ are parametrized by all the particle-to-particle relative distances and relative velocities. There is a deterministic dynamics which matches that of the original absolutist theory, once the original dynamics is projected down to Section 2.3.3's relative configuration space Q/E.

In short: the vanishing total linear and angular momenta mean that an initial state comprising only relative quantities *is* sufficient to determine all future relative quantities.

(2): *Decomposing the Reductionist reduced phase space*:—
Let us return to the reduced phase space $\bar{M}$. The first point to make is that since the Hamiltonian H on M, or indeed on $T^*\mathbb{R}^{3N}$, is invariant under the cotangent-lifted action of E, the usual dynamics on M projects down to $\bar{M} = M/E$. That is: the reduced phase space dynamics captures all the E-invariant features of the usual dynamics.

In fact, $\bar{M}$ is a *Poisson manifold.* So it is our first example of the more general framework for Hamiltonian mechanics announced in (1) of Section 2.2. Again, I postpone technical detail till later (especially Sections 5.1 and 5.2.4). But the idea is that a Poisson manifold has a *degenerate* antisymmetric bilinear map, which implies that the manifold is a disjoint union of symplectic manifolds. Each symplectic manifold is called a *leaf* of the Poisson manifold. The leaves' symplectic structures "mesh" with one another; and within each leaf there is a conventional Hamiltonian dynamics.

Even without a precise definition of a Poisson manifold, we can describe how M is decomposed into symplectic manifolds, each with a Hamiltonian dynamics. Recall that we have: $\dim(\bar{M}) = \dim(M)$ - $\dim(E) = 6N - 6$. This breaks down as:

$$6N - 6 = (6N - 12) + 3 + 3 = 2(3N - 6) + 3 + 3 =: \alpha + \beta + \gamma \tag{2.28}$$

where the right hand side defines α, β, γ respectively as $2(3N - 6), 3$ and 3. In terms of $\bar{M}$, this means the following.

(i): α corresponds to (1)'s $\bar{M}_0$, i.e. to $T^*(Q/E)$. As discussed, $3N - 6$ variables encode all the particle-pair relative distances; and the other $3N - 6$ variables encode all the particle-pair relative momenta.

The six extra variables additional to these $6N$-12 relative quantities consist of: four constants of the motion, and two other variables which are dynamical, i.e. change in time.

(ii): β stands for three of the four constants of the motion: viz. the three variables that encode the total linear momentum of the system, i.e. the momentum of the centre of mass. These constants of the motion are "just parameters" in the sense that: (a) not only does specifying a value for all three of them fix a surface, i.e. a $(6N - 9)$-dimensional hypersurface in $\bar{M}$, on which there is a Hamiltonian dynamics; also (b) this Hamiltonian and symplectic structure is independent of the values we specify.[16]

(iii): γ stands for the three variables that encode the total angular momentum of the system. One of these is a fourth constant of the motion, viz. the magnitude L of the total angular momentum. The other two time-varying quantities fix a point on a sphere (2-sphere) of radius L, encoding the direction of the angular momentum of the system *in a frame rotating with it.* The situation is as in the elementary theory of the rigid body: though the total angular momentum relative to coordinates fixed in space is a constant of the motion (three constant real numbers), the total angular momentum relative to the body is constant only in magnitude (one real number L), not in direction. This will be clearer in Section 5 onwards, when we describe the Poisson manifold structure in the theory of the rigid body. For the moment, there are two main comments to make about the N particle system:—

(a): If we specify L, in addition to the momentum of the centre of mass of the system, we get a $(6N - 10)$-dimensional hypersurface in $\bar{M}$, on which (as in (ii)) there is a Hamiltonian dynamics. So we can think of $\bar{M}$ as consisting of the four real-parameter family of these hypersurfaces, with each point of each hypersurface being equipped with a sphere of radius L; (subject to a qualification in (b) below).

Note that here 'each point being equipped' does not mean that the sphere gives

[16] As mentioned at the end of Section 2.3.1, the relationist traditionally proposes to identify absolutist states of motion that differ just by the value of the total momentum. And indeed, the proposal can be implemented by considering an action of the Galilean group on the absolutist phase space M, and identifying points related by Galilean boosts. For discussion and references, cf. Belot (2000: Section 5.3).

the extra dimensions that would constitute $\bar{M}$ as a fibre bundle; (there would be two dimensions lacking). Rather: in the point's representation by $6N - 10$ real numbers, two of the numbers can be taken to represent a point on a sphere.

(b): But unlike the situation for β in (ii) above, the Hamiltonian dynamics on such a hypersurface *depends* on the value of L. In particular, if $L = 0$ the sphere representing the body angular momentum is degenerate: it is of radius zero, and the other two time-varying quantities drop out. A point in the hypersurface is represented by $6N - 12$ real numbers; i.e. the hypersurface is $6N - 12$-dimensional.

Now recall from Section 2.3.3 or (1) above that $6N - 12$ is the dimension of the phase space of the envisaged Relationist theory built on the RCS Q/E. And indeed, just as one would hope: the hypersurface with $L = 0$ and also with vanishing linear momentum, with its dynamics, *is* the symplectic manifold and dynamics that is the Hamiltonian version of Barbour and Bertotti's 1982 Relational theory of N point-particles. In terms of (1)'s notation, this hypersurface is $\bar{M}_0$.

We can sum up this comparison as follows. On this hypersurface $\bar{M}_0$, the dynamics in the reduced phase space coincides with the dynamics one obtains for the relative variables, if one arbitrarily embeds their initial values in the usual absolutist phase space $T^*\mathbb{R}^{3N}$, subject to the constraint that the total angular and linear momenta vanish, and then reads off (just by projection) their evolution from the usual evolution in $T^*\mathbb{R}^{3N}$.

2.3.5 Comparing the Relationist and Reductionist procedures

In comparing the Relationist and Reductionist procedures, I shall just make just two extended comments, and refer to Belot for further discussion. The gist of both comments is that Reductionism suffices: Relationism is not needed. The first is a commonplace point; the second is due to Belot.

2.3.5.A Reductionism allows for rotation The first comment reiterates the Reductionist's ability, and the Relationist's inability, to endorse Newton's globes (or bucket) thought-experiment. The Reductionist can work in either

(i) the $(6N - 6)$-dimensional phase space $\bar{M} = M/E$; or

(ii) the $(6N - 9)$-dimensional hypersurface got from (i) by specifying the centre of mass' linear momentum; or

(iii) the $(6N - 10)$-dimensional hypersurface got from (ii) by also specifying a non-zero value of L.

In all three cases, the Reductionist can describe rotation in a way that the Relationist with their $(6N - 12)$-dimensional space cannot. For she has to hand the three extra non-relative variables (L and two others) that describe the rotation of the system as a whole. (Incidentally: that they describe the system as a whole is suggested by there being just three of them, whatever the value of N.) In particular, she can distinguish states of rotation and non-rotation ($L = 0$), in the sense of endorsing the distinctions advocated by the globes and bucket thought-experiments.

The Reductionist can also satisfy a traditional motivation for relationism, which concerns general philosophy, rather than the theory of motion. It is especially associated with Leibniz: namely, our theory (or our metaphysics) should not admit distinct but utterly indiscernible possibilities. One might well ask why we should endorse this "principle of the identity of indiscernibles" for possibilities rather than objects. For Leibniz himself, the answer lies (as Belot's (2001) brings out) in his principle of sufficient reason, and ultimately in theology.

But in any case the Reductionist can satisfy the requirement. Agreed, the usual absolutist theory, cast in $T^*\mathbb{R}^{3N}$ (or if you prefer, $M = T^*\mathbb{R}^{3N} - (\delta \cup \Delta)$) has nine variables that describe (i) the position of the centre of mass, (ii) the orientation of the system about its centre of mass, and (iii) the system's total linear momentum: i.e. three variables, a vector in $\mathbb{R}^3$, for each of (i)-(iii). So the usual absolutist theory has a nine-dimensional "profligacy" of distinct but indiscernible possibilities. But as we have seen, the Reductionist quotients by the action of the euclidean group E, and so works in $\bar{M} = M/E$: which removes the profligacy about (i) and (ii).

As to (iii), I agree that for all I have said, a job remains to be done. The foliation of $\bar{M}$ by a three real-parameter family of $(6N-9)$-dimensional hypersurfaces, labelled by the system's total linear momentum, codifies the profligacy—but does not eliminate it. But as I mentioned above (cf. footnote 16), the Reductionist can in fact quotient further, by considering the action of Galilean boosts and identifying phase space points that differ by a boost; i.e. defining orbits transverse to these hypersurfaces.

2.3.5.B Analogous reductions in other theories I close my philosophers' appetizer for symplectic reduction by summarizing some general remarks of Belot's (2001: Sections VIII-IX); cf. also his (2003a, Sections 12, 13). They are about how our discussion of relational mechanics is typical of many cases; and how symplectic reduction can be physically important. I label them (1)-(3).

(1): *A general contrast: when to quotient*:—
The example of N point-particles interacting by Newtonian gravity is typical of a large class of cases (infinite-dimensional, as well as finite-dimensional). There is a configuration space Q, acted on by a continuous group G of symmetries, which lifts to the cotangent bundle T^*Q, with the cotangent lift leaving invariant the Hamiltonian, and so the dynamics. So we can quotient T^*Q by G to give a reduced theory. (There is a Lagrangian analogue; but as above, we set it aside.) But there is also some motivation for quotienting G's action on Q, irrespective of how we then go on the construct dynamics. Let us adopt 'relationism' as a mnemonic label for whatever motivates quotienting the configuration space. Then with suitable technical conditions assumed (recall our excision of δ and Δ), we will have:

(i): for the reduced Hamiltonian theory: $\dim((T^*Q)/G) = 2 \dim Q$ - $\dim G$;

(ii): for the relationist theory, in a Lagrangian or Hamiltonian framework:
$\dim(T(Q/G)) = \dim(T^*(Q/G)) = 2(\dim Q$ - $\dim G)$

So we have in the reduced theory, dim G variables that do not occur in the relationist theory: let us call them 'non-relational variables'.

(2): *The non-relational variables*:—
Typically, these non-relational variables represent global, i.e. collective, properties of the system. That is unsurprising since the number, dim G, of these variables is independent of the number of degrees of freedom of the system (dim Q, or 2dim Q if you count rate of change degrees of freedom separately).

Some of these variables are conserved quantities, which arise (by Noether's theorem) from the symmetries. Furthermore, there can be specific values of the conserved quantities, like the vanishing angular momentum of Section 2.3.4, for which the reduced theory collapses into the relationist theory. That is, not only are the relevant state spaces of equal dimension; but also their dynamics agree.

(3): *The reduced theory*:—
Typically, the topology and geometry of the reduced phase space $(T^*Q)/G$, and the Hamiltonian function on it, $\bar{H} : (T^*Q)/G \rightarrow \mathbb{R}$ say, are more complex than the corresponding features of the unreduced theory on T^*Q. In particular, the reduced Hamiltonian $\bar{H}$ typically has potential energy terms corresponding to forces that are absent from the unreduced theory. But this should not be taken as necessarily a defect, for two reasons.

First, there are famous cases in which the reduced theory has a distinctive motivation. One example is Hertz' programme in mechanics, viz. to "explain away" the apparent forces of our macroscopic experience (e.g. gravity) as arising from reduction of a theory that has suitable symmetries. (The programme envisaged cyclic variables for microscopic degrees of freedom that were unknown to us; cf. Lutzen (1995, 2005).) Another famous example is the Kaluza-Klein treatment of the force exerted on a charged particle by the electromagnetic field. That is: the familiar Lorentz force-law describing a charged particle's motion in four spacetime dimensions can be shown to arise by symplectic reduction from a theory postulating a spacetime with a fifth (tiny and closed) spatial dimension, in which the particle undergoes straight-line motion. Remarkably, the relevant conserved quantity, viz. momentum along the fifth dimension, can be identified with electric charge; so that the theory can claim to explain the conservation of electric charge. (This example generalizes to other fields: for details and references, cf. Marsden and Ratiu (1999, Section 7.6).)

Second, the reduced theory need not be so complicated as to be impossible to work with. Indeed, these two examples prove this point, since in them the reduced theory is entirely tractable: for it is the familiar theory—that one might resist abandoning for the sake of the postulated unreduced theory.[17] Besides, Belot describes how, even when the reduced theory seems complicated (and not just because it is unfamiliar!), the general theory of symplectic reduction, as developed over the last forty years, has

[17]And here one should resist being prejudiced because of familiarity. Why *not* have Newtonian gravity arise from a microscopic cyclic degree of freedom? Why *not* have the Lorentz force law arise from geodesic motion in a five-dimensional spacetime with the fifth dimension wrapped up, so that conservation of charge is explained, in Noether's theorem fashion, by a symmetry?

shown that one can often "do physics" in the reduced phase space: and that, as in the Kaluza-Klein example, the physics in the reduced phase space can be heuristically, as well as interpretatively, valuable.

3 Some geometric tools

So much by way of an appetizer. The rest of the Chapter, comprising this Section and the next four, is the five-course banquet! This Section expounds some modern differential geometry, especially about Lie algebras and Lie groups. Section 4 takes up actions by Lie groups. Then Section 5 describes Poisson manifolds as a generalized framework for Hamiltonian mechanics. As I mentioned in (2) of Section 2.2, Lie himself developed this framework; so in effect, he knew everything in these two Sections—so it is a true (though painful!) pun to say that these three Sections give us the "Lie of the land". In any case, these two Sections will prepare us for Section 6's description of symmetry and conservation in terms of momentum maps. Finally, Section 7 will present one of the main theorems about symplectic reduction. It concerns the case where the natural configuration space for a system is itself a Lie group G; (cf. (3) of Section 2.2). Quotienting the natural phase space (the cotangent bundle on G) will give a Poisson manifold that "is" the dual of G's Lie algebra.

In this Section, I first sketch some notions of differential geometry, and fix notation (Section 3.1). Then I introduce Lie algebras and Lie brackets of vector fields (Section 3.2). Though most of this Section (indeed this Chapter!) is about differential rather than integral notions, I will later need Frobenius' theorem, which I present in Section 3.3. Then I give some basic information about Lie groups and their Lie algebras (Section 3.4).

3.1 Vector fields on manifolds

3.1.1 Manifolds, vectors, curves and derivatives

By way of fixing ideas and notation, I begin by giving details about some ideas in differential geometry (some already used in Section 2.1), and introducing some new notation for them.

A manifold M will be finite-dimensional, except for obvious and explicit exceptions such as the infinite-dimensional group of diffeomorphisms of a (as usual: finite-dimensional!) manifold. I will not be concerned about the degree of differentiability in the definition of a manifold, or of any associated geometric objects: 'smooth' can be taken throughout what follows to mean C^∞. I will often not be concerned with global, as against local, structures and results; (though the reduction results we are driving towards *are* global in nature). For example, I will not be concerned about whether curves are inextendible, or flows are complete.

I shall in general write a vector at a point $x \in M$ as X; or in terms of local coordinates x^i, as $X = X^i \frac{\partial}{\partial x^i}$ (summation convention). *From now on*, I shall write the tangent space at a point $x \in M$ as $T_x M$ (rather than just T_x), thus explicitly indicating the manifold M to which it is tangent. As before, I write the tangent bundle, consisting of the "meshing collection" of these tangent spaces, as TM. Similarly, I write a 1-form (covector) at a point $x \in M$ as α; and so the cotangent space at $x \in M$ as $T^*_x M$; and as before, the cotangent bundle as T^*M.

A smooth map $f : M \to N$ between manifolds M and N (maybe $N = M$) maps smooth curves to smooth curves, and so tangent vectors to tangent vectors; and so on for 1-forms and higher tensors. It is convenient to write Tf, called the *derivative* or *tangent* of f (also written as f_* or *df* or Df), for the induced map on the tangent bundle.

In more detail: let us take a curve c in M to be a smooth map from an interval $I \subset \mathbb{R}$ to M, and a tangent vector at $x \in M$, $X \in T_x M$, to be an equivalence class $[c]_x$ of curves through x. (The equivalence relation is that the curves be tangent at x, with respect to every local chart at x; but I omit the details of this.) Then we define $Tf : TM \to TN$ (also written $f_* : TM \to TN$) by

$$f_*([c]_x) \equiv Tf([c]_x) := [f \circ c]_{f(x)}, \text{ for all } x \in M. \tag{3.1}$$

We sometimes write $T_x f$ for the restriction of Tf to just the tangent space $T_x M$ at x; i.e.

$$T_x f : [c]_x \in T_x M \mapsto [f \circ c]_{f(x)} \in T_{f(x)} N. \tag{3.2}$$

In Section 3.1.2.B, we will discuss how one can instead define tangent vectors to be differential operators on the set of all scalar functions defined in some neighbourhood of the point in question, rather than equivalence classes of curves. One can then define the tangent map $f_* \equiv Tf$ in a way provably equivalent to that above.

3.1.2 Vector fields, integral curves and flows

We will be especially concerned with vector fields defined on M, i.e. $X : x \in M \mapsto X(x) \in T_x M$, or on a subset $U \subset M$. So suppose that X is a vector field on M and $f : M \to N$ is a smooth map, so that $T_x f : T_x M \to T_{f(x)} N$.

3.1.2.A Push-forwards and pullbacks It is important to note $(T_x f)(X(x))$ does *not* in general define a vector field on N. For $f(M)$ may not be all of N, so that for $y \in (N - \text{ran}(f))$ $(T_x f)(X(x))$ assigns no element of $T_y N$. And f may not be injective, so that we could have $x, x' \in M$ and $f(x) = f(x')$ with $(T_x f)(X(x)) \neq (T_{x'} f)(X(x'))$. Thus we say that vector fields do not *push forward*.

On the other hand, suppose that $f : M \to N$ is a *diffeomorphism* onto N: that is, the smooth map f is a bijection, and its inverse f^{-1} is also smooth. Then for any vector field X on M, $Tf(X)$ is a vector field on N. So in this case, the vector field does push forward. Accordingly, $Tf(X)$ is called the *push-forward* of X; it is often written as $f_*(X)$. So for any $x \in M$, the pushed forward vector field at the image point $f(x)$ is given by

$$(f_*(X))(f(x)) := T_x f \cdot X(x) \ . \tag{3.3}$$

(Note the previous use of the asterisk-subscript for the derivative of f, in eq. 3.1.)

This prompts three more general comments.

(1): More generally: we say that two vector fields, X on M and Y on N, are *f-related* on M (respectively: on $S \subset M$) if $(Tf)(X) = Y$ at all $x \in M$ (respectively: $x \in S$).

(2): We can generalize the idea that a diffeomorphism implies that a vector field can be pushed forward, in two ways. First, the diffeomorphism need only be defined locally, on some neighbourhood of the point $x \in M$ of interest. Second, a diffeomorphism establishes a one-one correspondence, not just between vector fields defined on its domain and codomain, but also between all differential geometric objects defined on its domain and codomain: in particular, 1-form fields, and higher rank tensors.

(3): (This continues comment (2).) Though vector fields do not in general push forward, 1-form fields *do* in general *pull back*. This is written with an asterisk-superscript. That is: for any smooth $f : M \to N$, not necessarily a diffeomorphism (even locally), and any 1-form field (differential 1-form) α on N, we define the pullback $f^*(\alpha)$ to be the 1-form on M whose action, for each $x \in M$, and each $X \in T_x M$, is given by:

$$(f^*(\alpha))(X) := \alpha \mid_{f(x)} (Tf(X)) \ . \tag{3.4}$$

Similarly, of course if the map f is defined only locally on a subset of M: a 1-form defined on the range of f pulls back to a 1-form on the domain of f.

3.1.2.B The correspondence between vector fields and flows The leading idea about vector fields is that, for any manifold, the theorems on the local existence, uniqueness and differentiability of solutions of systems of ordinary differential equations (e.g. Arnold (1973: 48-49, 77-78, 249-250), Olver (2000: Prop 1.29)) secure a one-one correspondence between four notions:

(i): Vector fields X on a subset $U \subset M$, on which they are non-zero; $X : x \in U \mapsto X(x) \in T_x M, X(x) \neq 0$;

(ii): Non-zero directional derivatives at each point $x \in U$, in the direction of the vector $X(x)$. In terms of coordinates $\mathbf{x} = x^1, ..., x^n$, these are first-order linear differential operators $X^1(\mathbf{x})\frac{\partial}{\partial x^1} + \ldots + X^n(\mathbf{x})\frac{\partial}{\partial x^n}$, with $X^i(\mathbf{x})$ the i-component in this coordinate system of the vector $X(x)$. Such an operator is often introduced abstractly as a *derivation*: a map on the set of smooth real-valued functions defined on a neighbourhood of x, that is linear and obeys the Leibniz rule.

(iii): Integral curves (aka: solution curves) of the fields X in U; i.e. smooth maps $\phi : I \to M$ from a real open interval $I \subset \mathrm{I\!R}$ to U, with $0 \in I$, $\phi(0) = x \in U$, and whose tangent vector at each $\phi(\tau), \tau \in I$ is $X(\phi(\tau))$.

(iv): Flows X^τ mapping, for each field X and each $x \in U$, some appropriate subset of U to another: $X^\tau : U \to M$. This flow is guaranteed to exist only in some neighbourhood of a given point x, and for τ in some neighbourhood of $0 \in \mathbb{R}$; but this will be enough for us. Such a flow is a one-parameter subgroup of the "infinite-dimensional group" of all local diffeomorphisms.

I spell out this correspondence in a bit more detail:— In local coordinates $x^1, ..., x^n$, any smooth curve $\phi : I \to M$ is given by n smooth functions $\phi(\tau) = (\phi^1(\tau), ..., \phi^n(\tau))$, and the tangent vector to ϕ at $\phi(\tau) \in M$ is

$$\dot{\phi}(\tau) = \dot{\phi}^1(\tau)\frac{\partial}{\partial x^1} + \ldots + \dot{\phi}^n(\tau)\frac{\partial}{\partial x^n}. \tag{3.5}$$

So for ϕ to be an integral curve of X requires that for all $i = 1, ..., n$ and all $\tau \in I$

$$\dot{\phi}^i(\tau) = X^i(\tau). \tag{3.6}$$

The local existence and uniqueness, for a given vector field X and $x \in M$, of the integral curve $\phi_{X,x}$ through x (with $\phi(0) = x$) then ensures that the flow, written either as X^τ or as $\phi_X(\tau)$

$$X^\tau : x \in M \mapsto X^\tau(x) \equiv \phi_{X,x}(\tau) \in M \quad , \tag{3.7}$$

is (at least locally) well-defined. The flow is a one-parameter group of transformations of M, and X is called its *infinitesimal generator*.

The exponential notation

$$\exp(\tau X)(x) := X^\tau(x) \equiv \phi_{X,x}(\tau) \tag{3.8}$$

is suggestive. For example, the group operation in the flow, i.e.

$$X^{\tau+\sigma}(x) = X^\tau(X^\sigma(x)) \quad , \tag{3.9}$$

is written in the suggestive notation

$$\exp((\tau + \sigma)X)(x) = \exp(\tau X)(\exp(\sigma X)(x)) \quad . \tag{3.10}$$

So computing the flow for a given X (i.e. solving a system of n first-order differential equations!) is called *exponentiation* of the vector field X.

Remark:— The above correspondence can be related to our discussion of diffeomorphisms and pushing forward vector fields. In particular: if two vector fields, X on M and Y on N, are *f-related* by $f : M \to N$, so that $(Tf)(X(x)) = Y(f(x))$, then f induces a map from integral curves of X to integral curves of Y. We can express this in terms of exponentiation of X and $Y = (Tf)(X)$:

$$f(\exp(\tau X)x) = \exp(\tau(Tf)(X))(f(x)). \tag{3.11}$$

Remark:— I emphasise that the above correspondence between (i), (ii), (iii) and (iv) is not true at a *single* point. More precisely:

(a): On the one hand: the correspondence between (i) and (ii) holds at a point; and also holds for zero vectors. That is: a single vector $X \in T_xM$ corresponds to a directional derivative operator (derivation) at x; and $X = 0$ corresponds to the zero derivative operator mapping all local scalars to 0. (Indeed, as I mentioned: vectors are often defined as such operators/derivations). But:

(b): On the other hand: the correspondence between (i) and (iii), or between (i) and (iv), requires a neighbourhood. For a single vector $X \in T_xM$ corresponds to a whole class of curves (and so: of flows) through x, not to a single curve. Namely, it corresponds to all the curves (flows) with X as their tangent vector.

However, we shall see (starting in Section 3.4) that for a manifold with suitable extra structure, a single vector *does* determine a curve. (And we will again talk of exponentiation.)

We need to generalize one aspect of the above correspondence (i)-(iv), namely the (i)-(ii) correspondence between vectors and directional derivatives. This generalization is the Lie derivative.

3.1.3 The Lie derivative

Some previous Sections have briefly used the Lie derivative. Since we will use it a lot in the sequel, we now introduce it more thoroughly.

We have seen that given a vector field X on a manifold M, a point $x \in M$, and any scalar function f defined on a neighbourhood of x, there is a naturally defined rate of change of f along X at x: the directional derivative $X(x)(f)$.

Now we will define the Lie derivative along X as an operator $\mathcal{L}_X$ that defines a rate of change along X: not only for locally defined functions (for which the definition will agree with our previous notion, i.e. we will have $\mathcal{L}_X(f) = X(f)$); but also for vector fields and differential 1-forms.[18] We proceed in three stages.

(1): We first define the Lie derivative as an operator on scalar functions, in terms of the vector field X on M. We define the *Lie derivative* along the field X (aka: the derivative in the direction of X), $\mathcal{L}_X$, as the operator on scalar functions $f : M \to \mathrm{I\!R}$ defined by:

$$\mathcal{L}_X : f \mapsto \mathcal{L}_X f : M \to \mathrm{I\!R} \text{ with } \forall x \in M : \quad (\mathcal{L}_X f)(x) := \frac{d}{d\tau} |_{\tau=0}\, f(X^\tau(x)) \;\equiv\; X(x)(f). \tag{3.12}$$

Though this definition assumes that both X and f are defined globally, i.e. on all of M, it can of course be restricted to a neighbourhood. Thus defined, $\mathcal{L}_X$ is linear and obeys the Leibniz rule, i.e.

$$\mathcal{L}_X(fg) = f\mathcal{L}_X(g) + g\mathcal{L}_X(f) \;\; ; \tag{3.13}$$

[18]Indeed, the definition can be extended to all higher rank tensors. But I will not develop those details, since—apart from Section 2.1.3's mention of the Lie derivative of the symplectic form $\mathcal{L}_X\omega$ (viz. the requirement that if X is a symmetry, $\mathcal{L}_X\omega = 0$)—we shall not need them.

In coordinates $\mathbf{x} = x^1, ..., x^n$, $\mathcal{L}_X f$ is given by

$$\mathcal{L}_X f = X^1(\mathbf{x}) \frac{\partial f}{\partial x^1} + \ldots + X^n(\mathbf{x}) \frac{\partial f}{\partial x^n}, \tag{3.14}$$

with $X^i(\mathbf{x})$ the i-component of the vector $X(x)$. Eq. 3.14 means that despite eq. 3.12's mention of the flow X^τ, the Lie derivative of a scalar agrees with our previous notion of directional derivative: that is, for all f, $\mathcal{L}_X(f) = X(f)$.

(2): In (1), the vector field X determined the operator $\mathcal{L}_X$: in terms of Section 3.1.2.B's correspondence, we moved from (i) to (ii). But we can conversely define a vector field in terms of its Lie derivative; and in Section 3.2.2's discussion of the Lie bracket, we shall do exactly this.

In a bit more detail:— We note that the set $\mathcal{F}(M)$ of all scalar fields on M, $f : M \to \mathbb{R}$ forms an (infinite-dimensional) real vector space under pointwise addition. So also does the set $\mathcal{X}(M)$ of all vector fields on M, $X : x \in M \mapsto X(x) \in T_x M$. Furthermore, $\mathcal{X}(M)$ is isomorphic as a real vector space, and as an module over the scalar fields, to the collection of operators $\mathcal{L}_X$. The isomorphism is given by the map $\theta : X \mapsto \mathcal{L}_X$ defined in (1).

(3): We now extend the definition of $\mathcal{L}_X$ so as to define it on vector fields Y and 1-forms α. We can temporarily use θ as notation for *either* a vector field Y *or* a differential 1-form α. Given a vector field X and flow $X^\tau \equiv \phi_X(\tau)$, we need to compare θ at the point $x \in M$ with θ at the nearby point $X^\tau(x) \equiv \phi_{X,x}(\tau)$, in the limit as τ tends to zero. But the value of θ at $X^\tau(x)$ is in the tangent space, or cotangent space, at $X^\tau(x)$: $T_{X^\tau(x)}M$ or $T^*_{X^\tau(x)}M$. So to make the comparison, we need to somehow transport back this value to $T_x M$ or $T^*_x M$.

Fortunately, the vector field X provides a natural way to define such a transport. For the vector field Y, we use the differential (i.e. push-forward) of the inverse flow, to "get back" from $X^\tau(x)$ to x. Using $\phi^*(\tau)$ for this "pullback" of $\phi_{X,x}(\tau)$, we define

$$\phi^*(\tau) := T(\exp(-\tau X)) \equiv d\exp(-\tau X) : T_{X^\tau(x)}M \equiv T_{\exp(\tau X)(x)}M \to T_x M \quad . \tag{3.15}$$

For the 1-form α, we define the transport by the pullback, already defined by eq. 3.4:

$$\phi^*(\tau) := (\exp(-\tau X))^* : T^*_{X^\tau(x)}M \equiv T^*_{\exp(\tau X)(x)}M \to T^*_x M \quad . \tag{3.16}$$

With these definitions of $\phi^*(\tau)$, we now define the *Lie derivative* $\mathcal{L}_X \theta$, where θ is a vector field Y *or* a differential 1-form α, as the vector field or differential 1-form respectively, with value at x given by

$$\lim_{\tau \to 0} \frac{\phi^*(\tau)(\theta\mid_{X^\tau(x)}) - \theta\mid_x}{\tau} \quad = \quad \frac{d}{d\tau}\mid_{\tau=0} \; \phi^*(\tau)(\theta\mid_{X^\tau(x)}) \quad . \tag{3.17}$$

Finally, an incidental result to illustrate this Chapter's "story so far". It connects Noether's theorem, from Section 2.1.3, to this Section's details about the Lie derivative, and to the theorem stating the local existence and uniqueness of solutions of ordinary differential equations (cf. the start of Section 3.1.2.B). This latter theorem implies that on any manifold any vector field X can be "straightened out", in the sense that around any point at which X is non-zero, there is a local coordinate system in which X has all but one component vanish and the last component equal to 1. Using this theorem, it is straightforward to show that on any even-dimensional manifold *any* vector field X is locally Hamiltonian, with respect to *some* symplectic form, around a point where X is non-zero. One just defines the symplectic form by Lie-dragging from a surface transverse to X's integral curves.

3.2 Lie algebras and brackets

I now introduce Lie algebras and the Lie bracket of two vector fields.

3.2.1 Lie algebras

A *Lie algebra* is a vector space V equipped with a bilinear anti-symmetric operation, usually denoted by square brackets (and called 'bracket' or 'commutator'), $[,] : V \times V \to V$, that satisfies the Jacobi identity, i.e.

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0 \ . \tag{3.18}$$

3.2.1.A Examples; rotations introduced Here are three examples.

(i): $n \times n$ matrices equipped with the usual commutator, i.e. $[X,Y] := XY - YX$. (So the matrix multiplication "contributes" to the bracket, but not to the underlying vector space structure.)

(ii): 3×3 anti-symmetric matrices, equipped with the usual commutator.

(iii): $\mathbb{R}^3$ equipped with vector multiplication. In fact, example (iii) is essentially the same as example (ii); and this example will recur in what follows, in connection with rotations and the rigid body. (We will also see that example (ii) is in a sense more fundamental.)

To explain this, we first recall that every anti-symmetric operator A on a three-dimensional oriented euclidean space is the operator of vector multiplication by a fixed vector, ω say. That is: for all $\mathbf{q}, A\mathbf{q} = [\omega, \mathbf{q}] \equiv \omega \wedge \mathbf{q}$. (Proof: the anti-symmetric operators on $\mathbb{R}^3$ for a 3-dimensional vector space, since an anti-symmetric 3×3 matrix has three independent components. Vector multiplication by a vector ω is a linear and anti-symmetric operator; varying ω we get a subspace of the space of all anti-symmetric operators on $\mathbb{R}^3$; but this subspace has dimension 3; so it coincides with the space of all anti-symmetric operators.)

With this result in hand, the following three points are all readily verified.

(1): The matrix representation of A in cartesian coordinates is then

$$A = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \tag{3.19}$$

We can write

$$A \leftrightarrow \omega \text{ or } A_{ij} = -\epsilon_{ijk}\omega_k \text{ or } \omega_i = -\frac{1}{2}\epsilon_{ijk}A_{jk}. \tag{3.20}$$

(2): The plane Π of vectors perpendicular to ω is an invariant subspace for A, i.e. $A(\Pi) = \Pi$. And ω is an eigenvector for A with eigenvalue 0. This suggest a familiar elementary interpretation, which will be confirmed later (Section 3.4): viz. that any 3×3 anti-symmetric matrix A represents a infinitesimal rotation, and ω represents instantaneous angular velocity. That is, we will have, for all $\mathbf{q} \in \mathrm{I\!R}^3$: $\dot{\mathbf{q}} = A\mathbf{q} = [\omega, \mathbf{q}]$.

(3): The commutator of any two 3×3 anti-symmetric matrices A, B, i.e. $[A, B] := AB - BA$, corresponds by eq. 3.20 to vector multiplication of the axes of rotation. That is: writing eq. 3.20's bijection from vectors to matrices as $\Theta : \omega \mapsto A =: \Theta(\omega)$, we have for vectors $\mathbf{q}, \mathbf{r}, \mathbf{s}$

$$(\Theta(\mathbf{q})\Theta(\mathbf{r}) - \Theta(\mathbf{r})\Theta(\mathbf{q}))\mathbf{s} = \Theta(\mathbf{q})[\mathbf{r}, \mathbf{s}] - \Theta(\mathbf{r})[\mathbf{q}, \mathbf{s}] \tag{3.21}$$

$$= [\mathbf{q}, [\mathbf{r}, \mathbf{s}]] - [\mathbf{r}, [\mathbf{q}, \mathbf{s}]] \tag{3.22}$$

$$= [[\mathbf{q}, \mathbf{r}], \mathbf{s}] = \Theta([\mathbf{q}, \mathbf{r}]) \cdot \mathbf{s}. \tag{3.23}$$

where the [,] represents vector multiplication, i.e. $[\mathbf{q}, \mathbf{r}] \equiv \mathbf{q} \wedge \mathbf{r}$.

Eq. 3.23 means that Θ gives a Lie algebra isomorphism; and so our example (iii) is essentially the same as example (ii).

Besides, we can already glimpse why example (ii) is in a sense more fundamental. For this correspondence between anti-symmetric operators (or matrices) and vectors, eq. 3.20, is specific to three dimensions. In n dimensions, the number of independent components of an anti-symmetric matrix is $n(n-1)/2$: only for $n = 3$ is this equal to n. Yet we will see later (Section 3.4.4) that rotations on euclidean space $\mathrm{I\!R}^n$ of any dimension n are generated, in a precise sense, by the Lie algebra of $n \times n$ anti-symmetric matrices. So only for $n = 3$ is there a corresponding representation of rotations by vectors in $\mathrm{I\!R}^n$.

In the next two Subsections, we shall see other examples of Lie algebras: whose vectors are vector fields (Section 3.2.2), or tangent vectors at the identity element of a Lie group (Section 3.4). The first example will be an infinite-dimensional Lie algebra; the second finite-dimensional (since we will only consider finite-dimensional Lie groups). Besides, the above examples (i) and (ii) (equivalently: (i) and (iii)) will recur: each will be the vector space of tangent vectors at the identity element of a Lie group.

3.2.1.B Structure constants A finite-dimensional Lie algebra is characterized, relative to a basis, by a set of numbers, called *structure constants* that specify the bracket operation. Thus if $\{v_1, ..., v_n\}$ is a basis of a Lie algebra V, we define the structure constants $c^k_{ij}, (i, j, k = 1, ..., n)$ by expanding, in terms of this basis, the bracket of any two basis elements

$$[v_i, v_j] = \Sigma_k c^k_{ij} v_k \quad . \tag{3.24}$$

The bilinearity of the bracket implies that eq. 3.24 determines the bracket of all pairs of vectors $v, w \in V$. And the bracket's obeying anti-symmetry and the Jacobi identity implies that, for any basis, the structure constants obey

$$c^k_{ij} = -c^k_{ji} \quad ; \quad \Sigma_k (c^k_{ij} c^m_{kl} + c^k_{li} c^m_{kj} + c^k_{jl} c^m_{ki}) = 0 \tag{3.25}$$

Conversely, any set of constants c^k_{ij} obeying eq. 3.25 are the structure constants of an n-dimensional Lie algebra.

3.2.2 The Lie bracket of two vector fields

Given two vector fields X, Y on a manifold M, the corresponding flows do not in general commute: $X^t Y^s \neq Y^s X^t$. The non-commutativity is measured by the commutator of the Lie derivatives of X and of Y, i.e. $\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X$. (Cf. eq. 3.12 and 3.17 for a definition of the Lie derivative.) Here, 'measured' can be made precise by considering Taylor expansions; but I shall not go into detail about this.

What matters for us is that this commutator, which is at first glance seems to be a second-order operator, is in fact a first-order operator. This is verified by calculating in a coordinate system, and seeing that the second derivatives occur twice with opposite signs:

$$(\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X) f = \Sigma_i \; X^i \frac{\partial}{\partial x^i} \left(\Sigma_j Y^j \frac{\partial f}{\partial x^j} \right) \; - \; \Sigma_j \; Y^j \frac{\partial}{\partial x^j} \left(\Sigma_i X^i \frac{\partial f}{\partial x^i} \right) \tag{3.26}$$

$$= ... = \Sigma_{i,j} \; \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^i \frac{\partial X^j}{\partial x^i} \right) \frac{\partial f}{\partial x^j}. \tag{3.27}$$

So $\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X$ corresponds to a vector field: (recall (2) of Section 3.1.3, about defining a vector field from its Lie derivative). We call this field Z the *Lie bracket* (also known as: Poisson bracket, commutator, and Jacobi-Lie bracket!) of the fields X and Y, and write it as $[X, Y]$. It is also written as $\mathcal{L}_X Y$ and called the *Lie derivative of* Y *with respect to* X. (Beware: some books use an opposite sign convention.)

Thus $Z \equiv [X, Y] \equiv \mathcal{L}_X Y$ is defined to be the vector field such that

$$\mathcal{L}_Z \equiv \mathcal{L}_{[X,Y]} = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X \quad . \tag{3.28}$$

It follows that $Z \equiv [X, Y]$'s components in a coordinate system are given by eq. 3.27. This formula can be remembered by writing it (with summation convention, i.e. omitting the Σ) as

$$\left[X^i \frac{\partial}{\partial x^i}, Y^j \frac{\partial}{\partial x^J} \right] = X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \frac{\partial}{\partial x^i} \tag{3.29}$$

Another way to write eq. 3.27 is as:

$$[X, Y]^j = \ (X \cdot \nabla)Y^j \ - \ (Y \cdot \nabla)X^j \ ; \tag{3.30}$$

or without coordinates, writing $\mathbf{D}$ for the derivative map given by the Jacobian matrix, as

$$[X, Y] = \mathbf{D}Y \cdot X - \mathbf{D}X \cdot Y. \tag{3.31}$$

Again, the vector field $Z \equiv [X, Y]$ measures the non-commutation of the flows X^t and Y^s: in particular, these flows commute iff $[X, Y] = 0$.

We will need three results about the Lie bracket. They concern, respectively, the relation to Lie algebras, to Poisson brackets, and to Frobenius' theorem.

(1): The Lie bracket is obviously a bilinear and anti-symmetric operation on the (infinite-dimensional) vector space $\mathcal{X}(M)$ of all vector fields on M: $[,] : \mathcal{X}(M) \times \mathcal{X}(M) \to \mathcal{X}(M)$. One readily checks that it satisfied the Jacobi identity. (Expand $\mathcal{L}_{[[X,Y],Z]} = \mathcal{L}_{[X,Y]}\mathcal{L}_Z - \mathcal{L}_Z\mathcal{L}_{[X,Y]}$ etc.) So: $\mathcal{X}(M)$ is an (infinite-dimensional) Lie algebra.

(2): Returning to Hamiltonian mechanics (Section 2.1): there is a simple and fundamental relation between the Lie bracket and the Poisson bracket, via the notion of Hamiltonian vector fields (Section 2.1.3).

Namely: the Hamiltonian vector field of the Poisson bracket of two scalar functions f, g on the symplectic manifold M is, upto a sign, the Lie bracket of the Hamiltonian vector fields, X_f and X_g, of f and g:

$$X_{\{f,g\}} = -[X_f, X_g] = [X_g, X_f]. \tag{3.32}$$

Proof: apply the rhs to an arbitrary scalar $h : M \to \mathbb{R}$. One easily obtains $X_{\{f,g\}}(h)$, by using:

(i) the definition of a Hamiltonian vector field;

(ii) the Lie derivative of a function equals its elementary directional derivative eq. 3.12; and

(iii) the Poisson bracket is antisymmetric and obeys the Jacobi identity.

This result means that the Hamiltonian vector fields on a symplectic manifold M, equipped with the Poisson bracket, form an (infinite-dimensional) Lie subalgebra of the Lie algebra $\mathcal{X}(M)$ of all vector fields on the symplectic manifold M. Later, it will be important that this result extends from symplectic manifolds to Poisson manifolds; (details in Section 5.2.2).

(3): For Frobenius' theorem (Section 3.3), we need to relate the Lie bracket to Section 3.1.2's idea of vector fields being f-related by a map $f : M \to N$ between manifolds M and N. In short: if two pairs of vector fields are f-related, so is their Lie bracket. More explicitly: if X, Y are vector fields on M, and $f : M \to N$ is a map such that $(Tf)(X), (Tf)(Y)$ are well-defined vector fields on N, then Tf commutes with the Lie bracket:

$$(Tf)[X, Y] = [(Tf)X, (Tf)Y] \ . \tag{3.33}$$

3.3 Submanifolds and Frobenius' theorem

This Subsection differs from the preceding ones in three ways. First, it emphasises integral, rather than differential, notions.

Second: Section 3.1.2.B have emphasised that the integral curves of a vector field correspond to integrating a system of ordinary differential equations. Since such curves are one-dimensional submanifolds of the given manifold, our present topic, viz. higher-dimensional submanifolds, naturally suggests partial differential equations. For their integration involves finding, given an assignment to each point x of a manifold M of a subspace S_x (with dimension greater than one) of the tangent space T_xM, an integral surface, i.e. a submanifold S of M whose tangent space at each of its points is S_x.[19]

However, we will *not* be concerned with partial differential equations. For us, submanifolds of dimension higher than one arise when the span S_x of the tangent vectors at x to a set of *vector* fields fit together to form a submanifold. Thus Frobenius' theorem states, roughly speaking, that a finite set of vector fields is integrable in this sense iff the vector fields are in involution. That is: iff their pairwise Lie brackets are expandable in terms of the fields; i.e. the vector fields form a Lie subalgebra of the entire Lie algebra of vector fields. We will not need to prove this theorem. But we need to state it and use it—in particular, for the foliation of Poisson manifolds.

Third: a warning is in order. The intuitive idea of a subset $S \subset M$ that is a smooth manifold "in its own right" can be made precise in different ways. So there are subtleties about the definition of 'submanifold', and terminology varies between expositions—in a way it does not for the material in previous Sections. I will adopt what seems to be a widespread, if not majority, terminology.[20]

3.3.1 Submanifolds

The fundamental definition is:

Given a manifold M ($\dim(M)=n$), a *submanifold* of M of dimension k is a subset $N \subset M$ such that for every $y \in N$ there is an admissible local chart (i.e. a chart in

[19]Beware: there is no analogue for partial differential equations of the local existence and uniqueness theorem for ordinary differential equations. Even a field of two-dimensional planes in three-dimensional space is in general not integrable, e.g. the field of planes given by the equation $dz = ydx$. So integrable fields of planes, or other tangent subspaces on a manifold, are an exception; and accordingly, the integration theory for partial differential equations is less unified, and more complicated, than that for ordinary differential equations.

[20]My treatment is based on Marsden and Ratiu (1999, p. 124-127, 140) for Section 3.3.1, and Olver (2000, p. 38-40) for Section 3.3.2. As to varying terminology: Olver (2000, p. 9) defines 'submanifold' to be what we will call an immersed submanifold; (which latter, for us, does *not* have to be a submanifold, since the immersion need not be an embedding). Bishop and Goldberg (1980, p. 40-41) provide a similar example. For a detailed introduction to the different notions of submanifold, cf. Darling (1994, Chapters 3 and 5). Note that I will also omit some details, in particular about Frobenius' theorem providing regular immersions.

M's maximal atlas) (U, ϕ) with $y \in U$ and with the *submanifold property*, viz.

$$\text{(SM).}\ \phi : U \to \mathbb{R}^k \times \mathbb{R}^{n-k} \ \text{ and } \ \phi(U \cap N) = \phi(U) \cap (\mathbb{R}^k \times \{\mathbf{0}\}). \tag{3.34}$$

The set N becomes a manifold, generated by the atlas of all charts of the form $(U \cap N, \phi \mid (U \cap N))$, where (U, ϕ) is a chart of M having the submanifold property. (This makes the topology of N the relative topology.)

We need to take note of two ways in which submanifolds can be specified in terms of smooth functions between manifolds.

(1): A submanifold can be specified as the set on which a smooth function $f : M \to P$ between manifolds takes a certain value. In effect, this will be a generalization of eq. 3.34's requirement that $n - k$ coordinate-components of a chart ϕ take the value zero. This will involve the idea that the tangent map Tf is surjective, in which case f will be called a *submersion*. We will need this approach for quotients of actions of Lie groups.

(2): A submanifold can be specified parametrically, as the set of values of a local parametrization: i.e. as the range of a smooth function f with M as *codomain*. This will involve the idea that the tangent map Tf is injective, in which case f will be called an *immersion*. We will need this approach for Frobenius' theorem.

(1): *Submersions*:—
If $f : M \to P$ is a smooth map between manifolds, a point $x \in M$ is called a *regular point* if the tangent map $T_x f$ is surjective; otherwise x is a *critical point* of f. If $C \subset M$ is the set of critical points of M, we say $f(C)$ is the set of *critical values* of f, and $P - f(C)$ is the set of *regular values* of f. So if $p \in P$ is a regular value of f, then at every $x \in M$ with $f(x) = p$, $T_x f$ is surjective.

The *submersion theorem* states that if $p \in P$ is a regular value of f, then:

(i): $f^{-1}(p)$ is a submanifold of M of dimension $\dim(M)$ - $\dim(P)$; and

(ii): the tangent space of this submanifold at any point $x \in f^{-1}(p)$ is the kernel of f's tangent map:

$$T_x(f^{-1}(p)) = \ker T_x f \ . \tag{3.35}$$

If $T_x f$ is surjective for every $x \in M$, f is called a *submersion*.

(2): *Immersions*:—
A smooth map between manifolds $f : M \to P$ is called an *immersion* if $T_x f$ is injective at every $x \in M$. The *immersion theorem* states that $T_x f$ is injective iff there is a neighbourhood U of x in M such that $f(U)$ is a submanifold of P and $f\mid_U : U \to f(U)$ is a diffeomorphism.

NB: This does not say that $f(M)$ is a submanifold of P. For f may not be injective (so that $f(M)$ has self-intersections). And even if f is injective, f can fail to be a homeomorphism between M and $f(M)$, equipped with the relative topology induced from P. A standard simple example is an injection of an open interval of $\mathbb{R}$ into an "almost-closed" figure-of-eight in $\mathbb{R}^2$.

Nevertheless, when $f : M \to P$ is an immersion, and is also injective, we call $f(M)$

an *injectively immersed submanifold* (or shorter: an *immersed submanifold*): though $f(M)$ might not be a submanifold.

We also define an *embedding* to be an immersion that is also a homeomorphism (and so injective) between M and $f(M)$ (where the latter has the relative topology induced from P). If f is an embedding, $f(M)$ *is* a submanifold of N and f is a diffeomorphism $f : M \to f(M)$.

In fact, Frobenius' theorem will provide injectively immersed submanifolds that need not be embedded, and so need not be submanifolds. (They must also obey another condition, called 'regularity', that I will not go into.)

3.3.2 The theorem

We saw at the end of Section 3.2.2 that if two pairs of vector fields are f-related, so is their Lie bracket: cf. eq. 3.33. This result immediately yields a necessary condition for two vector fields to be tangent to an embedded submanifold: namely

If X_1, X_2 are vector fields on M that are tangent to an embedded submanifold S (i.e. at each $x \in S$, $X_i(x) \in T_xS < T_xM$), then their Lie bracket $[X_1, X_2]$ is also tangent to S.

This follows by considering the diffeomorphism $f : \tilde{S} \to S$ that gives an embedding of S in M. One then uses the fact that Tf commutes with the Lie bracket, eq. 3.33. That is: the Lie bracket of the f-related vector fields $\tilde{X}_1, \tilde{X}_2$ on $\tilde{S}$, which is of course tangent to $\tilde{S}$, is carried by Tf to the Lie bracket $[X_1, X_2]$ of X_1 and X_2. So $[X_1, X_2]$ is tangent to S.

The idea of Frobenius' theorem will be that this necessary condition of two vector fields being tangent to a submanifold is also sufficient. To be more precise, we need the following definitions.

A *distribution* D on a manifold M is a subset of the tangent bundle TM such that at each $x \in M$, $D_x := D \cap T_xM$ is a vector space. The dimension of D_x is the *rank* of D at x. If the rank of D is constant on M, we say the distribution is *regular*.

A distribution is *smooth* if for every $x \in M$, and every $X_0 \in D_x$, there is a neighbourhood $U \subset M$ of x, and a smooth vector field X on U such that (i) $X(x) = X_0$, (ii) for all $y \in U$, $X(y) \in D_y$. Such a vector field X is called a *local section* of D. Example: a set of r vector fields, $X_1, ..., X_r$ each defined on M, together define a smooth distribution of rank at most r.

A distribution is *involutive* if for any pair X_1, X_2 of local sections, the Lie bracket $[X_1, X_2](y) \in D_y$ in the two sections' common domain of definition.

We similarly say that a set of r smooth vector fields, $X_1, ..., X_r$, on a manifold M is *in involution* if everywhere in M they span their Lie brackets. That is: there are smooth real functions $h^k_{ij} : M \to R, i, j, k = 1, ..., r$ such that at each $x \in M$

$$[X_i, X_j](x) = \Sigma_k \, h^k_{ij}(x) X_k(x). \tag{3.36}$$

(Beware: involution is used in a different sense in connection with Liouville's theorem, viz. a set of real functions on phase space is said to be in involution when all their pairwise Poisson brackets vanish.)

A distribution D on M is *integrable* if for each $x \in M$ there is a local submanifold $N(x)$ of M whose tangent bundle equals the restriction of D to $N(x)$. If D is integrable, the various $N(x)$ can be extended to get, through each $x \in M$, a unique maximal connected set whose tangent space at each of its elements y is D_y. Such a set is called a (maximal) *integral manifold*.

NB: In general, each integral manifold is injectively immersed in M, but not embedded in it; and so, by the discussion in (2) of Section 3.3.1, an integral manifold might not be a submanifold of M. But (like most treatments), I shall ignore this point, and talk of them as submanifolds, *integral submanifolds*.

If the rank of D is constant on M, all the integral submanifolds have a common dimension: the rank of D. But in general the rank of D varies across M, and so does the dimension of the integral submanifolds.

We similarly say that a set of r vector fields, $X_1, ..., X_r$, is *integrable*; viz. if through every $x \in M$ there passes a local submanifold $N(x)$ of M whose tangent space at each of its points is spanned by $X_1, ..., X_r$. (Again: we allow that at some x, $X_1(x), ..., X_r(x)$ may be linearly dependent, so that the dimension of the submanifolds varies.)

We say (both for distributions and sets of vector fields) that the collection of integral manifolds is a *foliation* of M, and its elements are *leaves*. Again: if the dimension of the leaves is constant on M, we say the foliation is *regular*.

With these definitions in hand, we can now state Frobenius' theorem: both in its usual form, which concerns the case of constant rank, i.e. regular distributions and vector fields that are everywhere linearly independent; and in a generalized form. The usual form is:

> **Frobenius' theorem (usual form)** A smooth regular distribution is integrable iff it is involutive.
> Or in terms of vector fields: a set of r smooth vector fields, $X_1, ..., X_r$, on a manifold M, that are everywhere linearly independent, is integrable iff it is in involution.

The generalization comes in two stages. The first stage concerns varying rank, but assumes a finite set of vector fields. It is straightforward: this very same statement holds. That is: a set of r smooth vector fields, $X_1, ..., X_r$, on a manifold M (perhaps not everywhere linearly independent) is integrable iff it is in involution.

But for the foliation of Poisson manifolds (Section 5.3.3), we need to consider an infinite set of vector fields, perhaps with varying rank; and for such a set, this statement fails. Fortunately, there is a useful generalization; as follows.

Let $\mathcal{X}$ be a set of vector fields on a manifold M, that forms a vector space. So in the above discussion of r vector fields, $\mathcal{X}$ can be taken as all the linear combinations $\Sigma_{i=1}^{r}\ f_i(x)X_i(x), x \in M$, where the f_i are arbitrary smooth functions $f : M \to \mathbb{R}$.

Such an $\mathcal{X}$ is called *finitely generated.*

For any $\mathcal{X}$ forming a vector space, we say (as before) that $\mathcal{X}$ is *in involution* if $[X, Y] \in \mathcal{X}$ whenever $X, Y \in \mathcal{X}$. Let $\mathcal{X}_x$ be the subspace of T_xM spanned by the $X(x)$ for all $X \in \mathcal{X}$. As before, we define: an *integral manifold* of $\mathcal{X}$ is a submanifold $N \subset M$ such that for all $y \in N$, $T_yN = \mathcal{X}_y$; and $\mathcal{X}$ is called *integrable* iff through each $x \in M$ there passes an integral manifold.

As before: if $\mathcal{X}$ is integrable, it is in involution. But the converse fails. A further condition is needed, as follows.

We say that $\mathcal{X}$ is *rank-invariant* if for any vector field $X \in \mathcal{X}$, the dimension of the subspace $\mathcal{X}_{\exp(\tau X)(x)}$ along the flow generated by X is a constant, independent of τ. (But it can depend on the point x.)

Since the integral curve $\exp(\tau X)(x)$ through x should be contained in any integral submanifold, rank-invariance is certainly a necessary condition of integrability. (It also follows from $\mathcal{X}$ being finitely generated.) In fact we have:

> **Frobenius' theorem (generalized form)** A system $\mathcal{X}$ of vector fields on M is integrable iff it is rank-invariant and in involution.

The idea of the proof is to directly construct the integral submanifolds. The submanifold through x is obtained as

$$N = \{\exp(X_1)\exp(X_2)....\exp(X_p)(x) : p \geq 1, X_i \in \mathcal{X}\}. \tag{3.37}$$

The rank-invariance secures that for any $y \in N$, $\mathcal{X}_y$ has dimension $\dim(N)$.

3.4 Lie groups, and their Lie algebras

I introduce Lie groups and their Lie algebras. By the last two Subsections (Sections 3.4.3 and 3.4.4), we will have enough theory to compute efficiently the Lie algebra of a fundamentally important Lie group, the rotation group.

3.4.1 Lie groups and matrix Lie groups

A *Lie group* is a group G which is also a manifold, and for which the product and inverse operations $G \times G \to G$ and $G \to G$ are smooth.

Examples:—-

(i): $\mathbb{R}^n$ under addition.

(ii): The group of linear isomorphisms of $\mathbb{R}^n$ to $\mathbb{R}^n$, denoted $GL(n, \mathbb{R})$ and called the *general linear group*; represented by the real invertible $n \times n$ matrices. This is an open subset of $\mathbb{R}^{n^2}$, and so a manifold of dimension n^2; and the formulas for the product and inverse of matrices are smooth in the matrix components.

(iii) The group of rotations about the origin of $\mathbb{R}^3$, represented by 3×3 orthogonal matrices of determinant 1; denoted $SO(3)$, where S stands for 'special' (i.e. determinant 1), and O for 'orthogonal'.

In fact, all three examples can be regarded as Lie groups of matrices, with matrix multiplication as the operation. In example (i), consider the isomorphism θ between $\mathbb{R}^n$ under addition and $(n+1) \times (n+1)$ matrices with diagonal entries all equal to 1, other rightmost column entries equal to the given vector in $\mathbb{R}^n$, and all other entries zero. Thus consider, for the case $n = 3$:

$$\theta : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{3.38}$$

This suggests that we define a *matrix Lie group* to be any set of invertible real matrices, under matrix multiplication, that is closed under multiplication, inversion and taking of limits. That a matrix Lie group is a Lie group will then follow from $GL(n, \mathbb{R})$ being a Lie group, and the theorem below (in Section 3.4.3) that any closed subgroup of a Lie group is itself a Lie group.

For matrix Lie groups, some of the theory below simplifies. For example, the definition of exponentiation of an element of the group's Lie algebra reduces to exponentiation of a matrix. But we will develop some of the general theory, since (as always!) it is enlightening and powerful.

3.4.2 The Lie algebra of a Lie group

The main result in this Subsection is that for any Lie group G, the tangent space T_eG at the identity $e \in G$ has a natural Lie algebra structure that is induced by certain natural vector fields on G; as follows.

3.4.2.A Left-invariant vector fields define the Lie algebra :
Let G be a Lie group. Each $g \in G$ defines a diffeomorphism of G onto itself by *left translation*, and similarly by *right translation*:

$$L_g : h \in G \mapsto gh \in G \ \ ; \ \ R_g : h \in G \mapsto hg \in G. \tag{3.39}$$

Remark: In Section 4 we will describe this in the language of group actions, saying that in eq. 3.39 *G acts on itself* by left and right translation.

Now consider the induced maps on the tangent spaces, i.e. the tangent (aka: derivative) maps; cf. eq.s 3.1, 3.2. They are $(L_g)_* =: L_{g*}, (R_g)_* =: R_{g*}$ where for each $h \in G$:

$$L_{g*} : T_hG \to T_{gh}G \quad \text{and} \quad R_{g*} : T_hG \to T_{hg}G. \tag{3.40}$$

In particular: the derivative $(R_g)_*$ at $e \in G$ maps T_eG to T_gG. This implies that every vector $\xi \in T_eG$ defines a vector field on G: its value at any $g \in G$ is the image

$(R_g)_*\xi$ of ξ under $(R_g)_*$. Such a vector field is called a *right-invariant* vector field: it is uniquely defined by (applying the derivative of right translation to) its value at the identity $e \in G$.

In more detail, and now defining *left-invariant* vector fields:—
A vector field X on G is called *left-invariant* if for every $g \in G$, $(L_g)_*X = X$. More explicitly, let us write $T_h L_g$ for the tangent or derivative of L_g at h, i.e. for $L_{g*} : T_h G \to T_{gh} G$. Then left-invariance requires that

$$(T_h L_g)X(h) = X(gh) \text{ for every } g \text{ and } h \in G. \tag{3.41}$$

Thus every vector $\xi \in T_e G$ defines a left-invariant vector field, written X_ξ, on G: X_ξ's value at any $g \in G$ is the image $(L_g)_*\xi$ of ξ under $(L_g)_*$. In other words: $X_\xi(g) := (T_e L_g)\xi$.

Not only is a left-invariant vector field uniquely defined by its value at the identity $e \in G$. Also, the set $\mathcal{X}_L(G)$ of left-invariant vector fields on G is isomorphic as a vector space to the tangent space $T_e G$ at the identity e. For the linear maps α, β defined by

$$\alpha : X \in \mathcal{X}_L(G) \mapsto X(e) \in T_e G \ \ ; \ \text{ and } \ \beta : \xi \in T_e G \mapsto \{g \mapsto X_\xi(g) := (T_e L_g)\xi\} \in \mathcal{X}_L(G) \tag{3.42}$$

compose to give the identity maps:

$$\beta \circ \alpha = id_{\mathcal{X}_L(G)} \ \ ; \ \ \alpha \circ \beta = id_{T_e G}. \tag{3.43}$$

$\mathcal{X}_L(G)$ is a Lie subalgebra of the Lie algebra of all vector fields on G, because it is closed under the Lie bracket. That is: the Lie bracket of left-invariant vector fields X and Y is itself left-invariant, since one can check that for every $g \in G$ we have (with L meaning 'left' not 'Lie'!)

$$L_{g*}[X, Y] = [L_{g*}X, L_{g*}Y] = [X, Y]. \tag{3.44}$$

If we now define a bracket on $T_e G$ by

$$[\xi, \eta] := [X_\xi, X_\eta](e) \tag{3.45}$$

then $T_e G$ becomes a Lie algebra. It is called *the Lie algebra of* G, written $\mathfrak{g}$ (or, to avoid ambiguity about which Lie group is in question: $\mathfrak{g}(G)$). It follows from eq. 3.44 that

$$[X_\xi, X_\eta] = X_{[\xi,\eta]} \ \ ; \tag{3.46}$$

that is to say, the maps α, β are Lie algebra isomorphisms.

This result, that $T_e G$ has a natural Lie algebra structure, is *very* important. For, as we shall see in the rest of Section 3.4: the structure of a Lie group is very largely determined by the structure of this Lie algebra. Accordingly, as we shall see in Sections 4 and 5 et seq.: this Lie algebra underpins most of the constructions made with the Lie group, e.g. in Lie group actions. Thus Olver writes that this result 'is the cornerstone of Lie group theory ... almost the entire range of applications of Lie groups to differential equations ultimately rests on this one construction!' (Olver 2000: 42).

Before turning in the next Subsection to examples, and the topic of subgroups and subalgebras, I end with four results, (1)-(4), which will be needed later; and a remark.

3.4.2.B Four results :
(1): Lie group structure determines Lie algebra structure in the following sense. If G, H are Lie groups, and $f : G \to H$ is a smooth homomorphism, then the derivative of f at the identity $T_e f : \mathfrak{g}(G) \to \mathfrak{g}(H)$ is a Lie algebra homomorphism. In particular, for all $\xi, \eta \in \mathfrak{g}(G)$, $(T_e f)[\xi, \eta] = [T_e f(\xi), T_e f(\eta)]$. (Cf. eq. 3.33.)

(2): *Exponentiation again; a correspondence between left-invariant vector fields and one-dimensional subgroups*:
Recall from Section 3.1, especially eq. 3.8, that each vector field X on the manifold G determines an integral curve ϕ_X in G passing through the identity e (with $\phi_X(0) = e$). We now write the points in (the image of) this curve as g_τ (X and e being understood):

$$\exp(\tau X)(e) \equiv X^\tau(e) \equiv \phi_{X,e}(\tau) =: g_\tau . \tag{3.47}$$

It is straightforward to show that if X is left-invariant, this (image of a) curve is a *one-parameter subgroup* of G: i.e. not just as eq. 3.7 et seq., a one-parameter subgroup of the group of diffeomorphisms of the manifold G. In fact:

$$g_{\tau+\sigma} = g_\tau g_\sigma \quad g_0 = e \quad g_\tau^{-1} = g_{-\tau} \ . \tag{3.48}$$

Besides, the group is defined for all $\tau \in \mathbb{R}$; and is isomorphic to either $\mathbb{R}$ or the circle group $SO(2)$. Conversely, any connected one-parameter subgroup of G is generated by a left-invariant vector field in this way.

Accordingly, we define exponentiation of elements ξ of $\mathfrak{g}$ by reference to the isomorphisms eq. 3.42 and 3.43. It is also convenient to define this as a map taking values in G. Thus for $\xi \in \mathfrak{g}$ and its corresponding left-invariant vector field X_ξ that takes as value at $g \in G$, $X_\xi(g) := (T_e L_g)(\xi)$, we write the integral curve of X_ξ that passes through e (with value e for argument $\tau = 0$) as

$$\phi_\xi : \tau \in \mathbb{R} \mapsto \exp(\tau X_\xi)(e) \in G \ . \tag{3.49}$$

Then we define the *exponential map* of $\mathfrak{g}$ into G to be the map

$$\exp : \xi \in \mathfrak{g} \mapsto \phi_\xi(1) \in G \ . \tag{3.50}$$

Using the linearity of β as defined by eq. 3.42, these two equations, eq. 3.49 and 3.50, are related very simply:

$$\exp(\tau\xi) := \phi_{\tau\xi}(1) := \exp(1.X_{\tau\xi})(e) = \exp(\tau X_\xi) \ . \tag{3.51}$$

We write $\exp_G$ rather than exp when the context could suggest a Lie group other than G.

The map exp is a local diffeomorphism of a neighbourhood of $0 \in \mathfrak{g}$ to a neighbourhood of $e \in G$; but not in general a global diffeomorphism onto G. In modern terms, this result follows by applying the inverse function theorem to the discussion above. (It also represents an interesting example of the history of subject; cf. Hawkins (2000: 82-83) for Lie's version of this result, without explicit mention of its local nature.)

The map exp also has the basic property, adding to result (1) above, that ...

(3): *Homomorphisms respect exponentiation*:
If $f : G \to H$ is a smooth homomorphism of Lie groups, then for all $\xi \in \mathfrak{g}$,

$$f(\exp_G \xi) = \exp_H((T_e f)(\xi)). \tag{3.52}$$

(4): *Right-invariant vector fields as an alternative approach*:
We have followed the usual practice of defining $\mathfrak{g}$ in terms of left-invariant vector fields. One can instead use right-invariant vector fields. This produces some changes in signs, and in whether certain defined operations respect or reverse the order of two elements used in their definition. I will not go into many details about this. But some will be needed when we consider:

(i): Lie group actions, and especially their infinitesimal generators (Section 4.4 and 4.5);

(ii): reduction on the cotangent bundle of a Lie group—as occurs in the theory of the rigid body (Section 6.5 and 7.3.3).

For the moment we just note two basic results, (A) and (B); postponing others to Section 4.4 et seq..

(A): Corresponding to the vector space isomorphism between $\mathfrak{g}$ and the left-invariant vector fields, as in eq. 3.42. viz.

$$\xi \in T_e G \mapsto X_\xi \in \mathcal{X}_L(G) \;\text{ with }\; X_\xi(g) := (T_e L_g)\xi \;, \tag{3.53}$$

there is a vector space isomorphism to the set of right-invariant vector fields

$$\xi \in T_e G \mapsto Y_\xi \in \mathcal{X}_R(G) \;\text{ with }\; Y_\xi(g) := (T_e R_g)\xi \;. \tag{3.54}$$

Besides, the Lie bracket of right-invariant vector fields is itself right-invariant. So corresponding to our previous definition, eq. 3.45, of a Lie bracket on $T_e G$, and its corollary eq. 3.46, i.e. $[X_\xi, X_\eta] = X_{[\xi,\eta]}$, that makes $T_e G \cong \mathcal{X}_L(G)$ a Lie algebra isomorphism: we can also define a Lie bracket on $T_e G$ by

$$[\xi, \eta]_R := [Y_\xi, Y_\eta](e) \;, \tag{3.55}$$

and get a Lie algebra isomorphism $T_e G \cong \mathcal{X}_R(G)$.

(B): But the two Lie brackets, eq. 3.45 and 3.55, on $T_e G$ are different. In fact one can show that:

(i): X_ξ and Y_ξ are related by

$$I_* X_\xi = -Y_\xi \tag{3.56}$$

where $I : G \to G$ is the inversion map $I(g) := g^{-1}$, and I_* is the push-forward on vector fields induced by I, cf. eq. 3.3, i.e.

$$(I_* X_\xi)(g) := (TI \circ X_\xi \circ I^{-1})(g) \;. \tag{3.57}$$

Besides, since I is a diffeomorphism, eq. 3.56 makes I_* a vector space isomorphism.

(ii): It follows from eq. 3.56 that

$$[X_\xi, X_\eta](e) = -[Y_\xi, Y_\eta](e) \ \ ; \ \ \text{so} \ \ [\xi, \eta] = -[\xi, \eta]_R \ . \tag{3.58}$$

Finally, a remark about physics. In applications to physics, G is usually the group of symmetries of a physical system, and so a vector field on G is the infinitesimal generator of a one-parameter group of symmetries. For mechanics, we saw this repeatedly in Section 2, especially as regards the group of translations and rotations about the origin, in physical space $\mathbb{R}^3$. This Subsection's isomorphism between the Lie algebra $\mathfrak{g}$ and left-invariant vector fields on G means that we can think of $\mathfrak{g}$ also as consisting of infinitesimal symmetries of the system. (The $\xi \in \mathfrak{g}$ are also called generators of the group G.)

3.4.3 Examples, subgroups and subalgebras

I begin with the first two of Section 3.4.1's three examples. That will prompt a little more theory, which will enable us to deal efficiently in the next Subsection with the third example, viz. the rotation group.

(1): *Examples*:—

(i): $G := \mathbb{R}^n$ under addition. G is abelian so that left and right translation coincide. The invariant vector fields are just the constant vector fields, so that $\mathcal{X}_L(G) \equiv \mathcal{X}_R(G) \cong \mathbb{R}^n$. So the tangent space at the identity T_eG, i.e. the Lie algebra $\mathfrak{g}$, is itself $\mathbb{R}^n$. The bracket structure is wholly degenerate: for all invariant vector fields X, Y, $[X, Y] = 0$; and for all $\xi, \eta \in \mathfrak{g}$, $[\xi, \eta] = 0$.

(ii): $G := GL(n, \mathbb{R})$, the general linear group. Since G is open in $End(\mathbb{R}^n, \mathbb{R}^n)$, the vector space of all linear maps on $\mathbb{R}^n$ ('*End*' for 'endomorphism'), G's Lie algebra, as a vector space, is $End(\mathbb{R}^n, \mathbb{R}^n)$; (cf. example (i)). To compute what the Lie bracket is, we first note that any $\xi \in End(\mathbb{R}^n, \mathbb{R}^n)$ defines a corresponding vector field on $GL(n, \mathbb{R})$ by

$$X_\xi : A \in GL(n, \mathbb{R}) \mapsto A\xi \in End(\mathbb{R}^n, \mathbb{R}^n) \ . \tag{3.59}$$

Besides, X_ξ is left-invariant, since for every $B \in GL(n, \mathbb{R})$, the left translation

$$L_B : A \in GL(n, \mathbb{R}) \mapsto BA \in GL(n, \mathbb{R}) \tag{3.60}$$

is linear, and so

$$X_\xi(L_B A) = BA\xi = T_A L_B X_\xi(A) \ . \tag{3.61}$$

Applying now eq. 3.31 at the identity $I \in GL(n, \mathbb{R})$ to the definition of the bracket in the Lie algebra, eq. 3.45, we have:

$$[\xi, \eta] := [X_\xi, X_\eta](I) = \mathbf{D}X_\eta(I) \cdot X_\xi(I) - \mathbf{D}X_\xi(I) \cdot X_\eta(I). \tag{3.62}$$

But $X_\eta A = A\eta$ is linear in A, so $\mathbf{D}X_\eta(I) \cdot B = B\eta$. This means that

$$\mathbf{D}X_\eta(I) \cdot X_\xi(I) = \xi\eta \ ; \tag{3.63}$$

and similarly

$$\mathbf{D}X_\xi(I) \cdot X_\eta(I) = \eta\xi. \tag{3.64}$$

So the Lie algebra $End(\mathbb{R}^n, \mathbb{R}^n)$ has the usual matrix commutator as its bracket: $[\xi, \eta] = \xi\eta - \eta\xi$. This Lie algebra is often written $\mathfrak{gl}(n, \mathbb{R})$.

Let us apply to this example, result (2) from Section 3.4.2.B. In short, the result said that left-invariant vector fields correspond (by exponentiation through $e \in G$) to connected one-parameter subgroups of G. To find the one-parameter subgroup $\exp(\tau X_\xi)(e)$ of $GL(n, \mathbb{R})$, we take the matrix entries $x_{ij}, (i, j = 1, ..., n)$ as the n^2 coordinates on $GL(n, \mathbb{R})$, so that the tangent space at the identity matrix I is the set of vectors

$$\Sigma_{ij} \;\; \xi_{ij} \, \frac{\partial}{\partial x_{ij}} \, |_I \tag{3.65}$$

with $\xi = (\xi_{ij})$ an arbitrary matrix. For given ξ, $\exp(\tau X_\xi)e$ is found by integrating the n^2 ordinary differential equations

$$\frac{dx_{ij}}{d\tau} = \Sigma_k \xi_{ik} x_{kj} \;\; ; \;\; x_{ij}(0) = \delta_{ij}. \tag{3.66}$$

The solution is just the matrix exponential:

$$X(\tau) = \exp(\tau\xi). \tag{3.67}$$

More generally, let us return to Section 3.4.1's idea of a matrix Lie group. For a matrix Lie group G, the definition of its Lie algebra can be given as:

$$\mathfrak{g} = \{ \text{ the set of matrices } \xi = \phi'(0) : \; \phi \text{ a differentiable map} : \mathbb{R} \to G, \phi(0) = e_G \}. \tag{3.68}$$

The deduction of the structure of the Lie algebra then proceeds straightforwardly. In particular, we get the result that the one-parameter subgroup generated by $\xi \in \mathfrak{g}$ is given by matrix exponentials, as in eq. 3.67: the group is $\{\exp(\tau\xi) : \tau \in \mathbb{R}\}$.

This result will help us compute our third example: finding the Lie algebra of the rotation group. But for that example, it is worth first developing a little the result (2) from Section 3.4.2.B: i.e. the correspondence between left-invariant vector fields and connected one-parameter subgroups of G.

(2): *More theory*:—

First, a warning remark. We will later need to take notice of the fact that a subgroup, even a one-parameter subgroup, of a Lie group G need not be a submanifold of G. Here we recall Section 3.3.1's definitions of immersion and embedding. Accordingly, we now define a subgroup H of a Lie group G to be a *Lie subgroup* of G if the inclusion map $i : H \to G$ is an injective immersion.

Just as we saw in Section 3.3.1 that not every injective immersion is an embedding, so also there are examples of Lie subgroups that are not submanifolds. Example: the torus T^2 can be made into a Lie group in a natural way (exercise: do this!); the one-parameter subgroups on the torus T^2 that wind densely on the torus are Lie subgroups that are not submanifolds. (For more details about this example, cf. Arnold (1973: 160-167) or Arnold (1989: 72-74) or Butterfield (2004a: Section 2.1.3.B).)

But it turns out that being closed is a sufficient, and necessary, further condition. That is:

> If H is a closed subgroup of a Lie group G, then H is a submanifold of G and in particular a Lie subgroup. And conversely, if H is a Lie subgroup that is also a submanifold, then H is closed.

Result (2) from Section 3.4.2.B, i.e. the correspondence between one-dimensional subgroups of G and one-dimensional subspaces (and so subalgebras) of $\mathfrak{g}$, generalizes to higher-dimensional subgroups and subalgebras. That is to say:

> If $H \subset G$ is a Lie subgroup of G, then its Lie algebra $\mathfrak{h} := \mathfrak{g}(H)$ is a subalgebra of $\mathfrak{g} \equiv \mathfrak{g}(G)$. In fact
>
> $$\mathfrak{h} = \{\xi \in \mathfrak{g} : \exp(\tau X_\xi)(e) \in H \text{ , for all } \tau \in \mathrm{I\!R}\}. \tag{3.69}$$
>
> And conversely, if $\mathfrak{h}$ is any m-dimensional subalgebra of $\mathfrak{g}$, then there is a unique connected m-dimensional Lie subgroup H of G with Lie algebra $\mathfrak{h}$.

The proof of the first two statements uses result (1) of Section 3.4.2.B. For the third, i.e. converse, statement, the main idea is that $\mathfrak{h}$ defines m vector fields on G that are linearly independent and in involution, so that one can apply Frobenius' theorem to infer an integral submanifold. One then has to prove that H is a Lie subgroup: Olver (2000: Theorem 1.51) and Marsden and Ratiu (1999: 279-280) give details and references. (Historical note: to see that this result, sometimes called Lie's 'third fundamental theorem', is close to what Lie himself called the main theorem of his theory of groups, cf. Hawkins (2000: 83).)

This general correspondence between Lie subgroups and Lie subalgebras prompts the question whether every finite-dimensional Lie algebra $\mathfrak{g}$ is the Lie algebra of a Lie group. The answer is Yes. Besides, the question reduces to the case of a matrix Lie group (i.e. a Lie subgroup of $GL(n, \mathrm{I\!R})$), in the sense that: every finite-dimensional Lie algebra $\mathfrak{g}$ is isomorphic to a subalgebra of $\mathfrak{gl}(n, \mathrm{I\!R})$, for some n. But be warned: this does not imply (and it is not true) that every Lie group is realizable as a matrix Lie group, i.e. that every Lie group is isomorphic to a Lie subgroup of $GL(n, \mathrm{I\!R})$.

This general correspondence also simplifies greatly the computation of the Lie algebras of Lie groups, for example $H := SO(3)$, that are Lie subgroups of $GL(n, \mathrm{I\!R})$. We only need to combine it with example (ii) above, that $\mathfrak{gl}(n, \mathrm{I\!R})$ is $End(\mathrm{I\!R}^n, \mathrm{I\!R}^n)$ with the usual matrix commutator as its bracket: $[\xi, \eta] = \xi\eta - \eta\xi$.

Thus we infer that the Lie algebra of $SO(3)$, written $\mathfrak{so}(3)$, is a subalgebra of $End(\mathrm{I\!R}^n, \mathrm{I\!R}^n)$ with the matrix commutator as bracket. Besides, we can identify $\mathfrak{so}(3)$ by looking at all the one-dimensional subgroups of G contained in it. Combining eq. 3.67 and 3.69, we have

$$\mathfrak{so}(3) = \{\xi \in \mathfrak{gl}(n, \mathrm{I\!R}) : \text{the matrix exponential } \exp(\tau\xi) \in SO(3), \ \forall\tau \in \mathrm{I\!R}\}. \tag{3.70}$$

With this result in hand, we can now compute $\mathfrak{so}(3)$.

3.4.4 The Lie algebra of the rotation group

Our first aim is to calculate the Lie algebra $\mathfrak{so}(3)$ (also written: $so(3)$) of $H := SO(3)$, the rotation group. This will lead us back to Section 3.2.1.A's correspondence between anti-symmetric matrices and vectors in $\mathbb{R}^3$.

$SO(3)$ is represented by 3×3 orthogonal matrices of determinant 1. So the requirement in eq. 3.70 becomes, now writing e, not exp:

$$(e^{\tau\xi})(e^{\tau\xi})^T = I \text{ and } \det(e^{\tau\xi}) = 1. \tag{3.71}$$

Differentiating the first equation with respect to τ and setting $\tau = 0$ yields

$$\xi + \xi^T = 0. \tag{3.72}$$

So ξ must be anti-symmetric, i.e. represented by an anti-symmetric matrix. Conversely, for any such anti-symmetric matrix ξ, we can show that $\det(e^{\tau\xi}) = 1$. So, indeed:

$$\mathfrak{so}(3) = \{3 \times 3 \text{ antisymmetric matrices}\}. \tag{3.73}$$

Notice that the argument is independent of choosing $n = 3$. It similarly computes $\mathfrak{so}(n)$ for any integer n:

$$\mathfrak{so}(n) = \{n \times n \text{ antisymmetric matrices}\}. \tag{3.74}$$

Thus the rotations on euclidean space $\mathbb{R}^n$ of any dimension n are generated by the Lie algebra of $n \times n$ anti-symmetric matrices.

This justifies our assertion at the end of Section 3.2.1.A that the rotation group in three dimensions is special in being representable by vectors in the space on which it acts, i.e. $\mathbb{R}^3$. For as we have just seen, in general the infinitesimal generators of rotations are anti-symmetric matrices, which in n dimensions have $n(n-1)/2$ independent components. But only for $n = 3$ does this equal n.

Remark: An informal computation of $\mathfrak{so}(3)$, based on the idea that higher-order terms in $e^{\tau\xi}$ can be neglected (cf. the physical idea that ξ represents an infinitesimal rotation), goes as follows.

For $(I + \tau\xi)$ to be a rotation requires that

$$(I + \tau\xi)(I + \tau\xi)^T = I \text{ and } \det(I + \xi\tau) = 1. \tag{3.75}$$

Dropping higher-order terms, the first equation yields

$$I + \tau(\xi + \xi^T) = I \quad \text{i.e. } \xi + \xi^T = 0. \tag{3.76}$$

Besides, the second equation in eq. 3.75 yields no further constraint, since for any anti-symmetric matrix ξ written as (cf. eq. 3.19)

$$\xi = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix}, \tag{3.77}$$

we immediately compute that $\det(I + \xi\tau) = 1 + \tau^2(\xi_1^2 + \xi_2^2 + \xi_3^2)$. So, dropping higher-order terms, $\det(I + \xi\tau) = 1$. In short, we again conclude that

$$\mathfrak{so}(3) = \{3 \times 3 \text{ antisymmetric matrices}\}. \tag{3.78}$$

For later use (e.g. Sections 4.4 and 4.5.1), we note that the three matrices

$$A^x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A^y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A^z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{3.79}$$

span $\mathfrak{so}(3)$, and generate the one-parameter subgroups

$$R^x_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad R^y_\theta = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \quad R^z_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.80}$$

representing anticlockwise rotation around the respective coordinate axes in the physical space $\mathbb{R}^3$.

Having computed $\mathfrak{so}(3)$ to consist of antisymmetric matrices, we can use Section 3.2.1.A's correspondence between these and vectors in $\mathbb{R}^3$ so as to realize $\mathfrak{so}(3)$ as vectors with the Lie bracket as vector multiplication. With these realizations in hand, we can readily obtain several further results about rotations. We will not need any. But a good example, which uses eq. 3.20's isomorphism Θ from vectors $\omega \in \mathbb{R}^3$ to matrices $A \in \mathfrak{so}(3)$, is as follows:—

$\exp(\tau\ \Theta(\omega))$ is a rotation about the axis ω by the angle $\tau \parallel \omega \parallel$.

We can now begin to see the point of this Chapter's second motto (from Arnold), that the elementary theory of the rigid body confuses six conceptually different three-dimensional spaces. For our discussion has already distinguished three of the six spaces which Arnold lists (in a different notation). Namely, we have just distinguished:

(i) $\mathbb{R}^3$, especially when taken as physical space; from (ii) $\mathfrak{so}(3) \equiv T_e(SO(3))$, the generators of rotations; though they are isomorphic as Lie algebras, by eq. 3.20's bijection Θ from vectors $\omega \in \mathbb{R}^3$ to matrices $A \in \mathfrak{so}(3)$;

(ii) $\mathfrak{so}(3) \equiv T_e(SO(3))$ from its isomorphic copy under the derivative of left translation by g (i.e. under $(L_g)_*$), viz. $T_g(SO(3))$: cf. eq. 3.41. (In the motto, Arnold writes g for $\mathfrak{so}(3)$ and G for $SO(3)$.)

In Section 5.2.4 we will grasp (even without developing the theory of the rigid body!) the rest of the motto. That is, we will see why Arnold also mentions the three corresponding dual spaces, $\mathbb{R}^{3*}, \mathfrak{so}(3)^*$ and $T_g^*(SO(3))$. But we can already say more about the two tangent spaces $\mathfrak{so}(3) \equiv T_e(SO(3))$ and $T_g(SO(3))$, in connection with the idea that for a pivoted rigid body, the configuration space can be taken as $SO(3)$; (cf. (3) of Section 2.2). We will show that there are two isomorphisms from $T_g(SO(3))$ to $T_e(SO(3))$ that are natural, not only in the mathematical sense of being basis-independent but also in the sense of having a physical interpretation. Namely, they represent the computation of the angular velocity from the Lagrangian generalized velocity, i.e. $\dot{q}$. In effect, one isomorphism computes the angular velocity's components with respect to an orthonormal frame fixed in space (called *spatial coordinates*); and the other computes it with respect to a frame fixed in the rigid body (*body coordinates*). In fact, these isomorphisms are the derivatives of right and left translation, respectively; (cf. eq. 3.39 and 3.40).

So suppose a pivoted rigid body has a right-handed orthonormal frame $\{a, b, c\}$ fixed in it. We can think of the three unit vectors a, b, c as column vectors in $\mathbb{R}^3$. Arranging them in a 3×3 matrix $g := (a\ b\ c) \in GL(3, \mathbb{R})$, we get a matrix that maps the unit x-vector e_1 to a, the unit y-vector e_2 to b, etc. That is: g maps the standard frame e_1, e_2, e_3 to a, b, c, and g is an orthogonal matrix: $g \in S0(3) = \{g \in GL(3, \mathbb{R}) \mid \tilde{g}g = I\ \}$. Thus g represents the configuration of the body, and the configuration space is $SO(3)$.

By differentiating the condition $\tilde{g}g = I$, we deduce that the tangent space at a specific g $T_g(SO(3))$, i.e. the space of velocities $\dot{g}$, is the 3-dimensional vector subspace of $GL(3, \mathbb{R})$:

$$T_g(SO(3)) = \{\dot{g} \in GL(3, \mathbb{R}) \mid \ \dot{\tilde{g}}g + \tilde{g}\dot{g} = 0\ \ \} \tag{3.81}$$

Now recall examples (ii) and (iii) of Section 3.2.1.A. We saw there that though the angular velocity of the body is usually taken to be the vector ω such that, with our "body-vectors" a, b, c,

$$\dot{a} = \omega \wedge a, \quad \dot{b} = \omega \wedge b, \quad \dot{c} = \omega \wedge c \ : \tag{3.82}$$

we can instead encode the angular velocity by the antisymmetric matrix $A := \Theta(\omega) \in \mathfrak{g} \equiv T_e(SO(3))$. As we saw, eq. 3.82 then becomes

$$\dot{a} = \Theta(\omega)a, \quad \dot{b} = \Theta(\omega)b, \quad \dot{c} = \Theta(\omega)c \ : \tag{3.83}$$

or equivalently the matrix equation for the configuration $g = (a\ b\ c)$,

$$\dot{g} \equiv (\dot{a}\ \dot{b}\ \dot{c}) = \Theta(\omega)g \ \ ; \ \text{ i.e. } \ \Theta(\omega) = \dot{g}g^{-1} \ . \tag{3.84}$$

Thus we see that the map from $T_g(SO(3))$ to $\mathfrak{g} = T_e(SO(3))$

$$\dot{g} \in T_g(SO(3)) \mapsto \dot{g}g^{-1} \equiv \dot{g}\tilde{g} \in \mathfrak{g} \tag{3.85}$$

maps the generalized velocity $\dot{g}$ to the angular velocity $\Theta(\omega)$. This is the angular velocity represented in the usual elementary way, with respect to coordinates fixed in space. One immediately checks that it is an isomorphism (exercise!).

On the other hand, let us consider $\Theta(\omega)$ as a linear transformation $\Theta(\omega) : \mathbb{R}^3 \to \mathbb{R}^3$, and express it in the *body* coordinates a, b, c. This gives $g^{-1}\Theta(\omega)g \equiv g^{-1}\dot{g}$. Thus the map

$$\dot{g} \in T_g(SO(3)) \mapsto g^{-1}\dot{g} \equiv \tilde{g}\dot{g} \in \mathfrak{g} \tag{3.86}$$

maps the generalized velocity $\dot{g}$ to the angular velocity expressed in body coordinates. It also is clearly an isomorphism.

Summing up: we have two natural isomorphisms that compute the angular velocity, in spatial and body coordinates respectively, from the generalized velocity $\dot{g}$.

Incidentally, one can verify directly that the images $\dot{g}\tilde{g}$ and $\tilde{g}\dot{g}$ of the isomorphisms eq. 3.85 and 3.86 lie in $\mathfrak{g}$, i.e. are antisymmetric matrices. Thus with $\cdot$ for the elementary dot-product, we have:

$$g^{-1}\dot{g} \equiv \tilde{g}\dot{g} = \begin{pmatrix} \tilde{a} \\ \tilde{b} \\ \tilde{c} \end{pmatrix} (\dot{a}\ \dot{b}\ \dot{c}) = \begin{pmatrix} 0 & a \cdot \dot{b} & a \cdot \dot{c} \\ b \cdot \dot{a} & 0 & b \cdot \dot{c} \\ c \cdot \dot{a} & c \cdot \dot{b} & 0 \end{pmatrix} . \tag{3.87}$$

This is an antisymmetric matrix, since differentiating $a \cdot b = b \cdot c = a \cdot c = 0$ with respect to time gives $a \cdot \dot{b} + \dot{a} \cdot b = 0$ etc. Finally, we deduce that $\dot{g}\tilde{g}$ is antisymmetric from the facts that $\dot{g}\tilde{g} = g(g^{-1}\dot{g})g^{-1}$ and antisymmetry is preserved by conjugation by g.

We end this Subsection with two incidental remarks; (they will not be used in what follows).

(1): In Section 2.1.1, we could have specialized the discussion from a symplectic manifold to a symplectic vector space, i.e. a (real, finite-dimensional) vector space equipped with a non-degenerate anti-symmetric bilinear form $\omega : Z \times Z \to \mathbb{R}$. It follows that Z is of even dimension. The question then arises which linear maps $A : Z \to Z$ preserve the normal form of ω given by eq. 2.4. It is straightforward to show that this is equivalent to A preserving the form of Hamilton's equations (for any Hamiltonian); so that these maps A are called *symplectic* (or *canonical*, or *Poisson*). The set of all such maps form a Lie group, the *symplectic group*, written $\mathrm{Sp}(Z, \omega)$. But since this Chapter will not need the theory of canonical transformations, I leave the study of $\mathrm{Sp}(Z, \omega)$'s structure as an exercise! (For details, cf. e.g. Abraham and Marsden (1978: 167-174), Marsden and Ratiu (1999: 69-72, 293-299).)

(2): Finally, a glimpse of the infinite-dimensional manifolds that this Chapter has foresworn. Consider the infinite-dimensional Lie group $Diff(M)$ of all diffeomorphisms on M. An element of its Lie algebra, i.e. a vector $A \in T_e(Diff(M))$, is a vector field, or equivalently a flow, on M. Besides, the Lie bracket in this Lie algebra $T_e(Diff(M))$, as defined by eq. 3.45 turns out to be the usual Lie bracket of the vector fields on M, as defined in Section 3.2.2.

4 Actions of Lie groups

We turn to actions of Lie groups on manifolds. The notions, results and examples in this Section will be crucial from Section 5.4 onwards. Fortunately, the foregoing provides several examples of the notions and results we need. Section 4.1 will give basic material, including the crucial notion of cotangent lifts. Sections 4.2 and 4.3 describe conditions for orbits and quotient spaces to be manifolds. Section 4.4 describes actions infinitesimally, i.e. in terms of their infinitesimal generators. Section 4.5 presents two important representations of a Lie group, its adjoint and co-adjoint representations, on its Lie algebra $\mathfrak{g}$ and on the dual $\mathfrak{g}^*$ respectively. Finally, Section 4.6 gathers some threads concerning our central, recurring example, viz. the rotation group.

4.1 Basic definitions and examples

A *left action* of a Lie group G on a manifold M is a smooth map $\Phi : G \times M \to M$ such that:

(i): $\Phi(e, x) = x$ for all $x \in M$

(ii): $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$ for all $g, h \in G$ and all $x \in M$.

We sometimes write $g \cdot x$ for $\Phi(g, x)$.

Similarly, a *right action* of a Lie group G on a manifold M is a smooth map $\Psi : M \times G \to M$ satisfying (i) $\Psi(x, e) = x$ and (ii) $\Psi(\Psi(x, g), h) = \Psi(x, gh)$. We sometimes write $x \cdot g$ for $\Psi(x, g)$.

It is convenient to also have a subscript notation. For every $g \in G$, we define

$$\Phi_g : M \to M \quad : \quad x \mapsto \Phi(g, x). \tag{4.1}$$

In this notation, (i) becomes $\Phi_e = id_M$ and (ii) becomes $\Phi_{gh} = \Phi_g \circ \Phi_h$. For right actions, (ii) becomes $\Psi_{gh} = \Psi_h \circ \Psi_g$.

One immediately verifies that any left action Φ of G on a manifold M, $g \mapsto \Phi_g : M \to M$, defines a right action Ψ by

$$g \mapsto \Psi_g := \Phi_{g^{-1}} : M \to M \;\; ; \;\; \text{i.e.} \;\; \Psi : (x, g) \in M \times G \mapsto \Phi(g^{-1}, x) \in M \; . \tag{4.2}$$

(Use the fact that in G, $(gh)^{-1} = h^{-1}g^{-1}$.) Similarly, a right action defines a left action, by taking the inverse in G. We will occasionally make use of this left-right "flip".

The definition of left action is equivalent to saying that the map $g \mapsto \Phi_g$ is a homomorphism of G into Diff(M), the group of diffeomorphisms of M. In the special case where M is a Banach space V and each $\Phi_g : V \to V$ is a continuous linear transformation, the action of G on V is called a *representation* of G on V.

The *orbit* of $x \in M$ (under the action Φ) is the set

$$\text{Orb}(x) = \{\Phi_g(x) : g \in G\} \subset M. \tag{4.3}$$

The action is called *transitive* if there is just one orbit, i.e. for all $x, y \in M$, there is a $g \in G$ such that $g \cdot x = y$. It is called *effective* (or *faithful*) if $\Phi_g = \text{id}_M$ implies $g = e$, i.e. if $g \mapsto \Phi_g$ is one-to-one. It is called *free* if it has no fixed points for any $g \neq e$: that is, $\Phi_g(x) = x$ implies $g = e$. In other words, it is free if for each $x \in M$, $g \mapsto \Phi_g(x)$ is one-to-one. (So: every free action is faithful.)

4.1.A Examples; cotangent lifts We begin with geometric examples; and then return to mechanics, giving first some general theory, followed by some examples.

(1): *Geometric examples*:—

(i): $SO(3)$ acts on $\mathbb{R}^3$ by $(A, x) \mapsto Ax$. The action is faithful. But it is neither free (each rotation fixes the points on its axis) nor transitive (the orbits are the spheres centred at the origin).

(ii): $GL(n, \mathbb{R})$ acts on $\mathbb{R}^n$ by $(A, x) \mapsto Ax$. The action is faithful, not free, and "almost transitive": the zero subspace $\{\mathbf{0}\}$ is an orbit, and so is $\mathbb{R}^n - \{\mathbf{0}\}$.

(iii): Suppose X is a vector field on M which is *complete* in the sense that the flow $\phi_X(\tau)$ of eq. 3.7 is defined for all $\tau \in \mathbb{R}$. Then this flow defines an action of $\mathbb{R}$ on M.

We turn to two examples which will be central, and recurring, in our discussion of symplectic reduction.

(iv): Left translation by each $g \in G$, $L_g : h \in G \mapsto gh \in G$ (cf. eq. 3.39), defines a left action of G on itself. Since G is a group, it is transitive and free (and so faithful). Similarly, right translation, $g \mapsto R_g$ with $R_g : h \in G \mapsto hg \in G$, defines a *right* action. And $g \mapsto R_{g^{-1}}$ defines a *left* action; cf. eq. 4.2.

One readily proves that left translation lifts to the tangent bundle TG as a left action. That is: one verifies by the chain rule that

$$\Phi_g : TG \to TG \quad : \quad v \equiv v_h \in T_hG \ \mapsto \ (T_hL_g)(v) \in T_{gh}G \tag{4.4}$$

defines a left action on TG. Similarly, right translation lifts to a right action on TG. But our interest in Hamiltonian mechanics of course makes us more interested in *cotangent* lifts. See (2) below for the general definitions, and example (viii) in (3) below for the cotangent lift of left translation.

(v): G acts on itself by *conjugation* (inner automorphism): $g \mapsto K_g := R_{g^{-1}} \circ L_g$. That is: $K_g : h \in G \mapsto ghg^{-1} \in G$. Each K_g is an isomorphism of G. The orbits are conjugacy classes. Section 4.5 will introduce two "differentiated versions" of action by conjugation, viz. the adjoint and co-adjoint actions, which will be important in symplectic reduction.

(2): *Hamiltonian symmetries and cotangent lifts*:—
We turn to Hamiltonian mechanics. Following the discussion in Section 2.1.3, we say: given a Hamilton system (M, ω, H) with (M, ω) a symplectic manifold and $H : M \to \mathbb{R}$, a *Hamiltonian group of symmetries* is a Lie group G acting on M such that each $\Phi_g : M \to M$ preserves both ω and H. Then the simplest possible examples are spatial translations and-or rotations acting on the free particle. The details of these examples, (vi) and (vii) below, will be clearer if we first develop some general theory.

This theory will illustrate the interaction between the left-right contrast for actions, and the tangent-cotangent contrast for bundles. Besides, both the general theory and the examples' details will carry over straightforwardly, i.e. component by component, to the case of N particles interacting by Newtonian gravity, discussed in Section 2.3.2: the action defined on a single particle is just repeated for each of the N particles.

So we will take $M := (\mathbb{R}^3) \times (\mathbb{R}^3)^*, \omega := dq^i \wedge dp^i, H := p^2/2m$. In the first place, both translations (by $\mathbf{x} \in \mathbb{R}^3$) and rotations (by $A \in SO(3)$) act on the configuration space $Q = \mathbb{R}^3$. We have actions of $\mathbb{R}^3$ and $SO(3)$ on $\mathbb{R}^3$ by

$$\Phi_{\mathbf{x}}(\mathbf{q}) = \mathbf{q} + \mathbf{x} \quad ; \quad \Phi_A(\mathbf{q}) = A\mathbf{q} \; . \tag{4.5}$$

But these actions *lift* to the cotangent bundle $T^*Q = (\mathbb{R}^3) \times (\mathbb{R}^3)^* \cong \mathbb{R}^6$; (as mentioned in Section 2.3.2). The lift of these actions is defined using a result that does not use the notion of an action. Namely:

> Any diffeomorphism $f : Q_1 \rightarrow Q_2$ induces a *cotangent lift* $T^*f : T^*Q_2 \rightarrow T^*Q_1$ (i.e. in the opposite direction) which is symplectic, i.e. maps the canonical one-form, and so symplectic form, on T^*Q_2 to that of T^*Q_1.

To define the lift of an action, it is worth going into detail about the definition of T^*f. (But I will not prove the result just stated; for details, cf. Marsden and Ratiu (1999: Section 6.3).)

The idea is that T^*f is to be the "pointwise adjoint" of the tangent map $Tf : TQ_1 \rightarrow TQ_2$ (eq. 3.1). That is: we define T^*f in terms of the contraction of its value, for an arbitrary argument $\alpha \in T^*_{q_2}Q_2$, with an arbitrary tangent vector $v \in T_{f^{-1}(q_2)}Q_1$. (Here it will be harmless to (follow many presentations and) conflate a point in T^*Q_2, i.e. strictly speaking a pair $(q_2, \alpha), q_2 \in Q_2, \alpha \in T^*_{q_2}Q_2$, with its form α. And similarly it will be harmless to conflate a point (q_1, v) in TQ_1 with its vector $v \in T_{q_1}Q_1$.)

We recall that any finite-dimensional vector space is naturally, i.e. basis-independently, isomorphic to its double dual: $(V^*)^* \cong V$; and we will use angle brackets $< ; >$ for the natural pairing between V and V^*. So we define $T^*f : T^*Q_2 \rightarrow T^*Q_1$ by requiring:

$$< \; (T^*f)(\alpha); v \; > \; := \; < \; \alpha; (Tf)(v) \; > , \quad \forall \; \alpha \in T^*_{q_2}Q_2, \; v \in T_{f^{-1}(q_2)}Q_1 \quad . \tag{4.6}$$

NB: Because T^*f "goes in the opposite direction", the composition of lift with function-composition involves a reversal of the order. That is: if $Q_1 = Q_2 \equiv Q$ and f, g are two diffeomorphisms of Q, then

$$T^*(f \circ g) = T^*g \circ T^*f. \tag{4.7}$$

With this definition of T^*f, a left action Φ of G on the manifold Q induces for each $g \in G$ the cotangent lift of $\Phi_g : Q \rightarrow Q$. That is: we have the map

$$T^*\Phi_g \equiv T^*(\Phi_g) : T^*Q \rightarrow T^*Q, \text{ with } \alpha \in T^*_qQ \mapsto (T^*\Phi_g)(\alpha) \in T^*_{g^{-1}\, \cdot \, q}Q \; . \tag{4.8}$$

Now consider the map assigning to each $g \in G$, $T^*\Phi_g$:

$$g \in G \mapsto T^*\Phi_g : T^*Q \rightarrow T^*Q \; . \tag{4.9}$$

To check that this is indeed an action of G on T^*Q, we first check that since $\Phi_e = id_Q$, $T\Phi_e : TQ \rightarrow TQ$ is id_{TQ} and $T^*(\Phi_e)$ is id_{T^*Q}. But beware: eq. 4.7 yields

$$T^*\Phi_{gh} = T^*(\Phi_g \circ \Phi_h) = T^*\Phi_h \circ T^*\Phi_g \; , \tag{4.10}$$

so that eq. 4.9 defines a *right* action.

But here we recall that any left action defines a right action by using the inverse; cf. eq. 4.2. Combining this with the idea of the cotangent lift of an action on Q, we get:

The left action Φ on Q defines, not only the right action eq. 4.9 on T^*Q, but also a *left* action on T^*Q, viz. by

$$g \in G \mapsto \Psi_g := T^*(\Phi_{g^{-1}}) : T^*Q \to T^*Q \ . \tag{4.11}$$

For since $(gh)^{-1} = h^{-1}g^{-1}$,

$$\Psi_{gh} \equiv T^*(\Phi_{(gh)^{-1}}) = T^*(\Phi_{h^{-1}g^{-1}}) = T^*(\Phi_{h^{-1}} \circ \Phi_{g^{-1}}) = T^*\Phi_{g^{-1}} \circ T^*\Phi_{h^{-1}} \equiv \Psi_g \circ \Psi_h \ . \tag{4.12}$$

In short, the two reversals of order cancel out. This sort of left-right flip will recur in some important contexts in the following, in particular in Sections 6.5 and 7.

(3): *Mechanical examples*:—
So much by way of generalities. Now we apply them to translations and rotations of a free particle, to rotations of a pivoted rigid body, and to N point-particles.

(vi): Let the translation group $G = (\mathbb{R}^3, +)$ act on the free particle's configuration space $Q = \mathbb{R}^3$ by

$$\Phi_{\mathbf{x}}(\mathbf{q}) = \mathbf{q} + \mathbf{x} \ . \tag{4.13}$$

Since G is abelian, the distinction between left and right actions of G collapses. (And if we identify G with Q, this is left=right translation by $\mathbb{R}^3$ on itself, i.e. example (iv) again: and so transitive and free.) But of course the lifted actions we have defined, "with g" and "with g^{-1}", eq. 4.9 and 4.11 respectively, remain distinct actions.

Then, writing $\alpha = (\mathbf{q}, \mathbf{p}) \in T^*_{\mathbf{q}}Q$, and using the fact that $T\Phi_{\mathbf{x}}(\mathbf{q} - \mathbf{x}, \dot{\mathbf{q}}) = (\mathbf{q}, \dot{\mathbf{q}})$, we see that eq. 4.6 implies that: first,

$$T^*(\Phi_{\mathbf{x}})(\mathbf{q}, \mathbf{p}) \in T^*_{\mathbf{q}-\mathbf{x}}Q \ ; \tag{4.14}$$

and second, that for all $\dot{\mathbf{q}} \in T_{\mathbf{q}-\mathbf{x}}Q$,

$$< T^*(\Phi_{\mathbf{x}})(\mathbf{q}, \mathbf{p}); (\mathbf{q} - \mathbf{x}, \dot{\mathbf{q}}) \ > \ = \ < \ (\mathbf{q}, \mathbf{p}); (\mathbf{q}, \dot{\mathbf{q}}) > \ \equiv \ \mathbf{p}(\dot{\mathbf{q}}) \ . \tag{4.15}$$

For eq. 4.15 to hold for all $\dot{\mathbf{q}} \in T_{\mathbf{q}-\mathbf{x}}Q$ requires that $T^*(\Phi_{\mathbf{x}})(\mathbf{q}, \mathbf{p})$ does not affect $\mathbf{p}$, i.e.

$$T^*(\Phi_{\mathbf{x}})(\mathbf{q}, \mathbf{p}) = (\mathbf{q} - \mathbf{x}, \mathbf{p}) \ . \tag{4.16}$$

So this is the lifted action "with g", corresponding to eq. 4.9. Similarly, the lifted action "with g^{-1}", corresponding to eq. 4.11, is: $\Psi_{\mathbf{x}}(\mathbf{q}, \mathbf{p}) := T^*(\Phi_{-\mathbf{x}})(\mathbf{q}, \mathbf{p}) = (\mathbf{q} + \mathbf{x}, \mathbf{p})$.

One readily checks that these lifted actions preserve both $\omega = dq^i \wedge dp^i$ (an exercise in manipulating the exterior derivative) and $H := p^2/2m$. So we have a Hamiltonian symmetry group. The action is not transitive: the orbits are labelled by their values of $\mathbf{p} \in (\mathbb{R}^3)^*$. But it is free.

(vii): Let $SO(3)$ act on the left on $Q = \mathbb{R}^3$ by

$$\Phi_A(\mathbf{q}) = A\mathbf{q} \ . \tag{4.17}$$

(This is example (i) again.) Let us lift this action "with g", i.e. eq. 4.9, so as to get a right action on T^*Q.

As in example (vi), we write $\alpha = (\mathbf{q}, \mathbf{p}) \in T^*_{\mathbf{q}}Q$. Using the fact that $T\Phi_A(\mathbf{q}, \dot{\mathbf{q}}) = (A\mathbf{q}, A\dot{\mathbf{q}})$, eq. 4.6 then implies that: first,

$$T^*(\Phi_A)(\mathbf{q}, \mathbf{p}) \in T^*_{A^{-1}\mathbf{q}}Q \ ; \tag{4.18}$$

and second, that for all $\dot{\mathbf{q}} \in T_{A^{-1}\mathbf{q}}Q$,

$$< T^*(\Phi_A)(\mathbf{q}, \mathbf{p}); (A^{-1}\mathbf{q}, \dot{\mathbf{q}}) \ > \ = \ < \ (\mathbf{q}, \mathbf{p}); (\mathbf{q}, A\dot{\mathbf{q}}) > \ \equiv \ \mathbf{p}(A\dot{\mathbf{q}}) \ \equiv \ p_i A^i_j \dot{q}^j \ . \tag{4.19}$$

For eq. 4.19 to hold for all $\dot{\mathbf{q}} \in T_{A^{-1}\mathbf{q}}Q$ requires that

$$T^*(\Phi_A)(\mathbf{q}, \mathbf{p}) = (A^{-1}\mathbf{q}, \mathbf{p}A) \ , \tag{4.20}$$

where $\mathbf{p}A$ is a row-vector. Or if one thinks of the $\mathbf{p}$ components as a column vector, it requires:

$$T^*(\Phi_A)(\mathbf{q}, \mathbf{p}) = (A^{-1}\mathbf{q}, \tilde{A}\mathbf{p}) = (A^{-1}\mathbf{q}, A^{-1}\mathbf{p}) \ , \tag{4.21}$$

where $\tilde{}$ represents the transpose of a matrix, and the last equation holds because A is an orthogonal matrix.

So this is the lifted action "with g", corresponding to eq. 4.9. Similarly, the lifted action "with g^{-1}", corresponding to eq. 4.11, is: $\Psi_A(\mathbf{q}, \mathbf{p}) := T^*(\Phi_{A^{-1}})(\mathbf{q}, \mathbf{p}) = (A\mathbf{q}, A\mathbf{p})$.

Again, one readily checks that these lifted actions preserve both $\omega = dq^i \wedge dp^i$ (another exercise in manipulating the exterior derivative!) and $H := p^2/2m$. So $SO(3)$ is a Hamiltonian symmetry group.

Like the original action of $SO(3)$ on Q, these actions are faithful. But they are not transitive: the orbits are labelled by the radii of two spheres centred at the origins of $\mathbb{R}^3$ and $(\mathbb{R}^3)^*$. And they are not free: suppose $\mathbf{q}$ and $\mathbf{p}$ are parallel and on the axis of rotation of A.

(viii): Now we consider the pivoted rigid body. But unlike examples (vi) and (vii), we will consider only kinematics, not dynamics: even for a free body. That is, we will say nothing about the definitions of, and invariance of, ω and H; for details of these, cf. e.g. Abraham and Marsden (1978: Sections 4.4 and 4.6) and the other references given in (3) of Section 2.2. We will in any case consider the dynamics of this example in more general terms (using momentum maps) in Sections 6.5.3 and 7.

We recall from the discussion at the end of Section 3.4.4 that the configuration space of the pivoted rigid body is $SO(3) =: G$. We also saw there that the space and body representations of the angular velocity $v = \dot{g} \in T_gG$ are given by right and left translation. Thus eq. 3.85 and 3.86 give:

$$v^S \equiv \dot{g}^S := T_g R_{g^{-1}}(\dot{g}) \quad \text{and} \quad v^B \equiv \dot{g}^B := T_g L_{g^{-1}}(\dot{g}) \ . \tag{4.22}$$

But we are now concerned with the *cotangent* lift of left (or right) translation. So let $SO(3)$ act on itself by left translation: $\Phi_g h \equiv L_g h = gh$. Let us lift this action "with g", i.e. eq. 4.9, to get a right action on T^*G. So let $\alpha \in T^*_h G$ and $(TL_g)(h, \dot{h}) = (gh, g\dot{h})$. Then eq. 4.6 implies that: first

$$(T^* L_g)(\alpha) \in T^*_{g^{-1}h} G \ , \tag{4.23}$$

and second that for all $v \in T_{g^{-1}h}G$

$$< \ T^*(L_g)(\alpha); v \ > = < \ \alpha; gv \ > \ . \tag{4.24}$$

In other words, on analogy with eq. 4.16 and 4.20: for eq. 4.24 to hold for all $v \in T_{g^{-1}h}G$ requires that with $gv \in T_h G$:

$$T^*(L_g)(\alpha) : v \in T_{g^{-1}h}G \mapsto \alpha(gv) \ . \tag{4.25}$$

Similarly, the lifted action "with g^{-1}" corresponding to eq. 4.11, i.e. the left action on T^*G, is

$$< \ T^*(L_{g^{-1}})(\alpha); v \ > = < \ \alpha; g^{-1}v \ > , \forall \alpha \in T^*_h G, v \in T_{gh}G \tag{4.26}$$

We will continue this example in Section 4.6, after developing more of the theory of Lie group actions.

Finally, let us sketch another mechanical example: the case of N particles with configuration space $Q := \mathbb{R}^{3N}$ interacting by Newtonian gravity—discussed in Section 2.3.2. This will combine and generalize examples (vi) and (vii); and lead on to the next Sections' discussions of orbits and quotients.

(ix): As I mentioned above (before eq. 4.5), the cotangent-lifted actions of translations and rotations on a single particle carry over straightforwardly to the case of N particles: the action defined on a single particle is just repeated, component by component, for each of the N particles to give an action on $T^*Q \cong \mathbb{R}^{3N} \times (\mathbb{R}^{3N})^*$.

Furthermore, the groups of translations and rotations are subgroups of a single group, the Euclidean group E. I shall not define E exactly. Here, let it suffice to say that:

(a): E's component-wise action on the configuration space $Q := \mathbb{R}^{3N}$ has a cotangent lift, which is of course also component by component.

(b): E's cotangent-lifted action is not transitive, nor free; but it is faithful.

(c): If we take as the Hamiltonian function the H of eq. 2.25, describing the particles as interacting by Newtonian gravity, then E is a Hamiltonian symmetry group. In fact, the kinetic and potential energies are separately invariant, essentially because the particles' interaction depends only on the inter-particle distances, not on their positions or orientations; cf. the discussion in Section 2.3.2.

A final comment about example (ix), which points towards the following Sections:—

Recall that in Sections 2.3.3 and 2.3.4, we used this example as a springboard to discussing Relationist and Reductionist procedures, which quotiented the configuration space or phase space. But in order for the quotient spaces (and orbits) to be manifolds, and in particular for dimensions to add or subtract in a simple way, we needed to excise two classes of "special" points, before quotienting. These were: the class of symmetric configurations or states (i.e. those fixed by some element of E), and the class of collision configurations or states. For the quotienting of phase space advocated by Reductionism, the classes of states were $\delta \subset T^*\mathbb{R}^{3N}$ and $\Delta \subset T^*\mathbb{R}^{3N}$; (cf. Section 2.3.4 for definitions.)

With examples (vi) to (ix) in hand, we can now see that:

(a): δ and Δ are each closed under the cotangent-lifted action of E on $T^*\mathbb{R}^{3N}$; i.e., each is a union of orbits. So E acts on $M := T^*\mathbb{R}^{3N} - (\delta \cup \Delta)$.

(b): E acts freely on M.

We will see in the sequel (especially in Sections 4.3.B and 5.5) that an action being free is one half (one conjunct) of an important sufficient condition for orbits and quotient spaces to be manifolds. The other conjunct will be the notion of an action being *proper*: which we will define in Section 4.3.

4.2 Quotient structures from group actions

In finite dimensions, any orbit $\mathrm{Orb}(x)$ is an immersed submanifold of M. This can be proved directly (Abraham and Marsden (1978: Ex. 1.6F(b), p. 51, and 4.1.22 p. 265)). But for our purposes, this is best seen as a corollary of some conditions under which quotient structures are manifolds; as follows.

The relation, $x \cong y$ if there is a $g \in G$ such that $g \cdot x = y$, is an equivalence relation, with the orbits as equivalence classes. We denote the quotient space, i.e. the set of orbits, by M/G (sometimes called the *orbit space*). We write the canonical projection as

$$\pi : M \to M/G, \quad x \mapsto \mathrm{Orb}(x) \; ; \tag{4.27}$$

and we give M/G the quotient topology by defining $U \subset M/G$ to be open iff $\pi^{-1}(U)$ is open in M.

Simple examples (e.g. (ii) of Section 4.1.A) show that this quotient topology need not be Hausdorff. However, it is easy to show that if the set

$$R := \{(x, \Phi_g x) \in M \times M : \; (g, x) \in G \times M\} \tag{4.28}$$

is a closed subset of $M \times M$, then the quotient topology on M/G is Hausdorff.

But to ensure that M/G has a manifold structure, further conditions are required. The main one (and a much harder theorem) is:

> R is a closed submanifold of $M \times M$ iff M/G is a manifold with $\pi : M \to M/G$ a submersion.

This theorem has two Corollaries which are important for us.

(1): A map $h : M/G \to N$, from the manifold M/G, for which $\pi : M \to M/G$ is a submersion, to the manifold N, is smooth iff $h \circ \pi : M \to N$ is smooth.

This corollary has a useful implication, called *passage to the quotients*, about the notion of *equivariance*—which will be important in symplectic reduction.

A smooth map $f : M \to N$ is called *equivariant* if it respects the action of a Lie group G on the manifolds. That is: Let G act on M and N by $\Phi_g : M \to M$ and $\Psi_g : N \to N$ respectively. $f : M \to N$ is called equivariant with respect to these actions if for all $g \in G$

$$f \circ \Phi_g = \Psi_g \circ f. \tag{4.29}$$

That is, f is equivariant iff for all g, the following diagram commutes:

$$\begin{array}{ccc} M & \xrightarrow{f} & N \\ \uparrow{\scriptstyle \Phi_g} & & \uparrow{\scriptstyle \Psi_g} \\ M & \xrightarrow{f} & N \end{array} \tag{4.30}$$

Equivariance immediately implies that f naturally induces a map, $\hat{f}$ say, on the quotients. That is: the map

$$\hat{f} : \mathrm{Orb}(x) \in M/G \mapsto \mathrm{Orb}(f(x)) \in N/G \tag{4.31}$$

is well-defined, i.e. independent of the chosen representative x for the orbit.

Applying the corollary we have: If $f : M \to N$ is equivariant, and the quotients M/G and M/N are manifolds with the canonical projections both submersions, then f being smooth implies that $\hat{f}$ is smooth. This is called *passage to the quotients*.

(2): Let H be a closed subgroup of the Lie group G. (By (2) of Section 3.4.3, this is equivalent to H being a subgroup that is a submanifold of G.) Let H act on G by left translation: $(h, g) \in H \times G \mapsto hg \in G$, so that the orbits are the right cosets Hg. Then G/H is a manifold and $\pi : G \to G/H$ is a submersion.

4.3 Proper actions

By adding to the Section 4.2's main theorem (i.e., R is a closed submanifold of $M \times M$ iff M/G is a manifold with $\pi : M \to M/G$ a submersion), the notion of a proper action we can give useful sufficient conditions for:

(A): orbits to be submanifolds;

(B): M/G to be a manifold.

An action $\Phi : G \times M \to M$ is called *proper* if the map

$$\tilde{\Phi} : (g, x) \in G \times M \mapsto (x, \Phi(g, x)) \in M \times M \tag{4.32}$$

is proper. By this we mean that if $\{x_n\}$ is a convergent sequence in M, and $\{\Phi_{g_n}(x_n)\}$ is a convergent sequence in M, then $\{g_n\}$ has a convergent subsequence in G. In finite dimensions, this means that compact sets have compact inverse images; i.e. if $K \subset M \times M$ is compact, then $\tilde{\Phi}^{-1}(K)$ is compact.

If G is compact, this condition is automatically satisfied. Also, the action of a group on itself by left (or by right) translation (Example (iv) of Section 4.1.A) is always proper. Furthermore, the cotangent lift of left (or right) translation ((2) and Example (viii) of Section 4.1.A) is always proper. We shall not prove this, but it will be important in the sequel.

4.3.A Isotropy groups; orbits as manifolds For $x \in M$ the *isotropy* (or *stabilizer* or *symmetry*) group of Φ at x is

$$G_x := \{g \in G : \Phi_g(x) \equiv \Phi(g,x) = x\} \subset G. \tag{4.33}$$

(So an action is free iff for all $x \in M$, $G_x = \{e\}$.)

So if we define

$$\Phi^x : G \to M \quad : \quad \Phi^x(g) := \Phi(g,x) \tag{4.34}$$

we have: $G_x = (\Phi^x)^{-1}(x)$. (The notation Φ^x is a "cousin" of the notation Φ_g defined in eq. 4.1.)

So since Φ^x is continuous, G_x is a closed subgroup of G. So, by the result in (2) of Section 3.4.3 (i.e. the result before eq. 3.69), G_x is a submanifold (as well as Lie subgroup) of G. And if the action is proper, G_x is compact.

Furthermore, the fact that for all $h \in G_x$ we have $\Phi^x(gh) = \Phi_g \circ \Phi_h(x) = \Phi_g(x)$, implies that Φ^x naturally induces a map

$$\tilde{\Phi}^x : [g] = gG_x \in G/G_x \mapsto \Phi_g x \in \mathrm{Orb}(x) \subset M \; . \tag{4.35}$$

That is, this map is well-defined. $\tilde{\Phi}^x$ is injective because if $\Phi_g x = \Phi_h x$ then $g^{-1}h \in G_x$, so that $gG_x = hG_x$.

It follows from Section 4.2's main theorem (i.e., R is a closed submanifold of $M \times M$ iff M/G is a manifold with $\pi : M \to M/G$ a submersion) that:

(a): If $\Phi : G \times M \to M$ is an action and $x \in M$, then $\tilde{\Phi}^x$ defined by eq. 4.35 is an injective immersion.

Here we recall from Section 3.3.1 that injective immersions need not be embeddings. But:—

(b): If also Φ is proper, the orbit $\mathrm{Orb}(x)$ is a closed submanifold of M and $\tilde{\Phi}^x$ is a diffeomorphism. In other words: the manifold structure of $\mathrm{Orb}(x)$ is given by the bijective map $[g] \in G/G_x \mapsto g \cdot x \in \mathrm{Orb}(x)$ being a diffeomorphism.

Examples:—

(We use the numbering of corresponding examples in Section 4.1.A):—

(i): $G = SO(3)$ acts on $M = \mathbb{R}^3$ by $(A,x) \mapsto Ax$. Since $\mathrm{Orb}(x)$ is a sphere centred at the origin of radius $\parallel x \parallel$, $M/G \cong \mathbb{R}^+$: which is not a manifold. But results (a) and (b) are illustrated: the isotropy group G_x at x is the group of rotations with x on the axis; the action is proper (for G is compact); the orbit $\mathrm{Orb}(x)$ is a closed manifold of M; and the isotropy group's cosets $[g] \in G/G_x$ are mapped diffeomorphically by $\tilde{\Phi}^x$ to points on the sphere $\mathrm{Orb}(x)$.

(iii): Let X be the constant vector field ∂_x on $M = \mathbb{R}^3$. X is complete. The action of $\mathbb{R}$ on M has as orbit through the point $\mathbf{x} = (x, y, z) \in \mathbb{R}^3$, the line $y =$ constant, $z =$ constant. The action is free, and therefore faithful and the isotropy groups are trivial. So $G/G_x \equiv G$. The action is proper. Again results (a) and (b) are illustrated: the orbits $\mathrm{Orb}(\mathbf{x})$ are closed submanifolds of M, viz. copies of the real line $\mathbb{R} = G \equiv G/G_x$ that are diffeomorphic to $\mathbb{R}$ by $\tilde{\Phi}^{\mathbf{x}}$.

4.3.B A sufficient condition for the orbit space M/G to be a manifold With result (b) from the end of Section 4.3.A,, we can prove that:

> If $\Phi : G \times M \to M$ is a proper free action, then the orbit space M/G is a manifold with $\pi : M \to M/G$ a submersion.

Examples: (again using the numbering in Section 4.1.A):—

(i): $G = SO(3)$ acts on $M = \mathbb{R}^3$ by $(A, x) \mapsto Ax$. Since $\mathrm{Orb}(x)$ is a sphere centred at the origin of radius $\parallel x \parallel$, $M/G \cong \mathbb{R}^+$: which is not a manifold, and indeed the action is not free.

(iii): Let X be the constant vector field ∂_x on $M = \mathbb{R}^3$. X is complete, and the action of $\mathbb{R}$ on M has as orbits the lines $y =$ constant, $z =$ constant. The action is faithful, free and proper, so that the orbit space M/G is a manifold: $M/G \cong \mathbb{R}^2$.

(iv): Left (or right) translation is obviously a free action of a group G on itself, and we noted above that it is proper. But since it is transitive, the orbit space G/G is the trivial 0-dimensional manifold (the singleton set of G).

(viii): The cotangent lift of left (or right) translation by $SO(3)$, or more generally, by a Lie group G. This action is proper (noted after eq. 4.32), and obviously free.

(ix): The Euclidean group E acts freely on $M := T^*\mathbb{R}^{3N} - (\delta \cup \Delta)$. This action is also proper: a (harder!) exercise for the reader.

4.4 Infinitesimal generators of actions

We now connect this Subsection's topic, group actions, with the Lie algebra of the Lie group concerned, i.e. with the topic of Section 3.4, especially 3.4.2.

Let $\Phi : G \times M \to M$ be a (left) action by the Lie group G on a manifold M. Then each $\xi \in \mathfrak{g}$ defines an action of $\mathbb{R}$ on M, which we write as Φ^ξ, in the following way.

We can think either in terms of exponentiation of ξ's corresponding left-invariant vector field X_ξ (cf. eq. 3.8 and 3.47); or in terms of of exponentiating ξ itself (cf. eq. 3.50 and 3.51):

$$\Phi^\xi : \mathbb{R} \times M \to M \quad : \quad \Phi^\xi(\tau, x) := \Phi(\exp(\tau X_\xi), x) \equiv \Phi(\exp(\tau\xi), x). \tag{4.36}$$

That is, in terms of our subscript notation for the original action Φ (cf. eq. 4.1): $\Phi_{\exp(\tau X_\xi)} \equiv \Phi_{\exp(\tau\xi)} : M \to M$ is a flow on M.

That the flow is complete, i.e. that an action of all of $\mathbb{R}$ is defined, follows from (2) *Exponentiation again* of Section 3.4.2, especially after eq. 3.48. Cf. also example (iii) of Section 4.1.

We say that the corresponding vector field on M, written ξ_M, i.e. the vector field defined at $x \in M$ by

$$\xi_M(x) := \frac{d}{d\tau} |_{\tau=0} \Phi_{\exp(\tau X_\xi)}(x) \equiv \frac{d}{d\tau} |_{\tau=0} \Phi_{\exp(\tau\xi)}(x) \qquad (4.37)$$

is the *infinitesimal generator* of the action corresponding to ξ.

In terms of the map Φ^x defined in eq. 4.34, we have that for all $\xi \in \mathfrak{g}$

$$\xi_M(x) = (T_e\Phi^x)(\xi) \; . \qquad (4.38)$$

So NB: the words 'infinitesimal generator' are used in different, though related, ways. In Remark (2) at the end of Section 3.4.2, a vector field on the group G, or an element $\xi \in \mathfrak{g}$, was called an 'infinitesimal generator'. Here the infinitesimal generator is a vector field on the action-space M. Similarly, beware the notation: ξ_M is a vector field on M, while X_ξ is a vector field on G.

As an example, we again take the rotation group $SO(3)$ acting on $\mathbb{R}^3$: $(A, \mathbf{x}) \in SO(3) \times \mathbb{R} \mapsto A\mathbf{x}$. One readily checks that with $\omega \in \mathbb{R}^3$, so that $\Theta(\omega) \in \mathfrak{so}(3)$, the infinitesimal generator of the action corresponding to $\xi \equiv \Theta(\omega)$ is the vector field on $\mathbb{R}^3$

$$\xi_{\mathbb{R}^3}(\mathbf{x}) \equiv (\Theta(\omega))_{\mathbb{R}^3}(\mathbf{x}) = \omega \wedge \mathbf{x} \; . \qquad (4.39)$$

In particular, the vector field on $\mathbb{R}^3$ representing infinitesimal anti-clockwise rotation about the x-axis is $e_1 := y\partial_z - z\partial_y$ (cf. eq. 3.79). Similarly, the infinitesimal generators of the action of rotating about the y axis and about the z-axis are, respectively: $e_2 := z\partial_x - x\partial_z$ and $e_3 := x\partial_y - y\partial_x$. The Lie brackets are given by:

$$[e_1, e_2] = -e_3 \quad [e_3, e_1] = -e_2 \quad [e_2, e_3] = -e_1. \qquad (4.40)$$

The minus signs here are a general feature of the transition $\xi \in \mathfrak{g} \mapsto \xi_M \in \mathcal{X}(M)$; cf. result (4) below.

As another example, we take the infinitesimal generator of left and right translation on the group G. (We will need this example for our theorems about symplectic reduction; cf. Sections 6.5.3, 7.2 and 7.3.3.) NB: There will be a "left-right flip" here, which continues the discussion in (4) of Section 3.4.2.B, comparing using left-invariant vs. right-invariant vector fields to define the Lie algebra of a Lie group.

For left translation $\Phi(g, h) \equiv L_g h := gh$, we have for all $\xi \in \mathfrak{g}$:

$$\Phi^\xi(\tau, h) = (\exp\tau\xi)h = R_h(\exp\tau\xi) \; ; \qquad (4.41)$$

so that the infinitesimal generator is

$$\xi_G(g) = (T_e R_g)\xi \; . \qquad (4.42)$$

So ξ_G is a *right*-invariant vector field; and unless G is abelian, it is *not* equal to the left-invariant vector field $g \mapsto X_\xi(g) := (T_e L_g)\xi$; cf. eq. 3.40 and 3.42.

Similarly, for right translation (which is a *right* action, cf. (1) (iv) in Section 4.1.A), the infinitesimal generator is the left-invariant vector field

$$g \mapsto X_\xi(g) := (T_e L_g)\xi \ . \tag{4.43}$$

Three straightforward results connect the notion of an infinitesimal generator with previous ideas. I will not give proofs, but will present them in the order of the previous ideas.

(1): Recall the correspondence between Lie subgroups and Lie subalgebras, at the end of Section 3.4.3; eq. 3.69. This implies that the Lie algebra of the isotropy group $G_x, x \in M$ (called the *isotropy algebra*), is

$$\mathfrak{g}_x = \{\xi \in \mathfrak{g} : \xi_M(x) = 0\} \ . \tag{4.44}$$

(2): Infinitesimal generators ξ_M give a differential version of the notion of equivariance, discussed in (1) of Section 4.2: a version called *infinitesimal equivariance.*

In eq. 4.29, we set $g = \exp(\tau\xi)$ and differentiate with respect to τ at $\tau = 0$. This gives $Tf \circ \xi_M = \xi_N \circ f$. That is: ξ_M and ξ_N are f-related. In terms of the pullback f^* of f, we have: $f^*\xi_N = \xi_M$.

(3): Suppose the action Φ is proper, so that by result (b) at the end of Section 4.3.A: the orbit $\mathrm{Orb}(x)$ of any point $x \in M$ is a (closed) submanifold of M. Then the tangent space to $\mathrm{Orb}(x)$ at a point y in $\mathrm{Orb}(x)$ is

$$T\mathrm{Orb}(x)_y = \{\xi_M(y) : \xi \in \mathfrak{g}\} \ . \tag{4.45}$$

Finally, there is a fourth result relating infinitesimal generators ξ_M to previous ideas; as follows. (But it is less straightforward than the previous (1)-(3): its proof requires the notion of the adjoint representation, described in the next Section.)

(4): The infinitesimal generator map $\xi \mapsto \xi_M$ establishes a Lie algebra *anti*-homomorphism between $\mathfrak{g}$ and the Lie algebra $\mathcal{X}_M$ of all vector fields on M. (Contrast the Lie algebra isomorphism between $\mathfrak{g}$ and the set $\mathcal{X}_L(G)$ of left-invariant vector fields on the group G; Section 3.4.2 especially eq. 3.42.) That is:

$$(a\xi + b\eta)_M = a\xi_M + b\eta_M \ \ ; \ \ [\xi_M, \eta_M] = -[\xi, \eta]_M \ \ \forall \xi, \eta \in \mathfrak{g}, \text{ and } a, b \in \mathbb{R}. \tag{4.46}$$

Incidentally, returning to (4) of Section 3.4.2.B, which considered defining the Lie algebra of a Lie group in terms of right-invariant vector fields, instead of left-invariant vector fields: had we done so, the corresponding map $\xi \mapsto \xi_M$ would have been a Lie algebra homomorphism.

4.5 The adjoint and co-adjoint representations

A leading idea of later Sections (especially Sections 5.4, 6.4 and 7) will be that there is a natural symplectic structure in the orbits of a certain natural representation of any

Lie group: namely a representation of the group on the dual of its own Lie algebra, called the *co-adjoint representation*. Here we introduce this representation. But we lead up to it by first describing the *adjoint representation* of a Lie group on its own Lie algebra. Even apart from symplectic structure (and so applications in mechanics), both representations illustrate the ideas of previous Subsections. I will again use $SO(3)$ and $\mathfrak{so}(3)$ as examples.

4.5.1 The adjoint representation

We proceed in four stages. We first define the representation, then discuss infinitesimal generators, then discuss matrix Lie groups, and finally discuss the rotation group.

(1): *The representation defined*:—
Let G be a Lie group and $\mathfrak{g}$ its Lie algebra, i.e. the tangent space to the group at the identity $e \in G$, equipped with the commutator bracket operation $[,]$.

Recall (e.g. from the beginning of Section 3.4.2) that G acts on itself by left and right translation: each $g \in G$ defines diffeomorphisms of G onto itself by

$$L_g : h \in G \mapsto gh \in G \;\; ; \;\; R_g : h \in G \mapsto hg \in G. \tag{4.47}$$

The induced maps of the tangent spaces are, for each $h \in G$:

$$L_{g*} : TG_h \to TG_{gh} \quad \text{and} \quad R_{g*} : TG_h \to TG_{hg}. \tag{4.48}$$

The diffeomorphism $K_g := R_{g^{-1}} \circ L_g$ (i.e. conjugation by $g, K_g : h \mapsto ghg^{-1}$) is an inner automorphism of G. (Cf. example (v) at the end of Section 4.1.) Its derivative at the identity $e \in G$ is a linear map from the Lie algebra $\mathfrak{g}$ to itself, which is denoted:

$$Ad_g := (R_{g^{-1}} \circ L_g)_{*e} : \mathfrak{g} \to \mathfrak{g}. \tag{4.49}$$

So letting g vary through G, the map $Ad : g \mapsto Ad_g$ assigns to each g a member of $\mathrm{End}(\mathfrak{g})$, the space of linear maps on (endomorphisms of) $\mathfrak{g}$. The chain rule implies that $Ad_{gh} = Ad_g Ad_h$. So

$$Ad : g \mapsto Ad_g \tag{4.50}$$

is a left action, a representation, of G on $\mathfrak{g}$: $G \times \mathfrak{g} \to \mathfrak{g}$. It is called the *adjoint representation*.

Three useful results about Ad follow from our results (1) and (3) in Section 3.4.2.B (cf. eq. 3.52: *Homomorphisms respect exponentiation*):

[1]: If $\xi \in \mathfrak{g}$ generates the one-parameter subgroup $H = \{\exp(\tau X_\xi) : \tau \in \mathbb{R}\}$, then $Ad_g(\xi)$ generates the conjugate subgroup $K_g(H) = gHg^{-1}$.

$$\exp(Ad_g(\xi)) = K_g(\exp \xi) := g(\exp \xi)g^{-1}. \tag{4.51}$$

Incidentally, eq. 4.51 has a many-parameter generalization. Let H and H' be two connected r-dimensional Lie subgroups of the Lie group G, with corresponding Lie subalgebras $\mathfrak{h}$ and $\mathfrak{h}'$ of the Lie algebra $\mathfrak{g} = \mathfrak{g}(G)$. Then H and H' are conjugate subgroups, $H' = gHg^{-1}$, iff $\mathfrak{h}$ and $\mathfrak{h}'$ are corresponding conjugate subalgebras, i.e. $\mathfrak{h}' = Ad_g(\mathfrak{h})$.

[2]: Eq. 4.51 also implies another result which will be needed for a crucial result about symplectic reduction, in Section 6.5.2. (The many-parameter generalization just mentioned will not be needed.) It relates Ad to the pullback of an arbitrary action Φ.

Thus let Φ be a left action of G on M. Then for every $g \in G$ and $\xi \in \mathfrak{g}$

$$(Ad_g\xi)_M = \Phi^*_{g^{-1}}\xi_M \ , \tag{4.52}$$

where Φ^* indicates pullback of the vector field. For we have:

$$\begin{aligned}
(Ad_g\xi)_M(x) &:= \frac{d}{d\tau}\ |_{\tau=0}\ \Phi(\exp(\tau Ad_g\xi), x) && (4.53)\\
&= \frac{d}{d\tau}\ |_{\tau=0}\ \Phi(g(\exp\tau\xi)g^{-1}, x) \quad \text{by eq. 4.51} && (4.54)\\
&= \frac{d}{d\tau}\ |_{\tau=0}\ (\Phi_g \circ \Phi_{\exp\tau\xi} \circ \Phi_{g^{-1}}(x)) && (4.55)\\
&= T_{\Phi_{g^{-1}}(x)}\Phi_g(\xi_M(\Phi_{g^{-1}}(x))) \quad \text{by the chain rule and eq. 4.37} && (4.56)\\
&= \left(\Phi^*_{g^{-1}}\xi_M\right)(x) \quad \text{by the definition of pullback.} && (4.57)
\end{aligned}$$

Not only is this result needed later. Also, incidentally: it is the main part of the proof of result (4) at the end of Section 4.4, that $\xi \mapsto \xi_M$ is a Lie algebra anti-homomorphism.

[3]: Ad_g is an algebra homomorphism, i.e.

$$Ad_g[\xi, \eta] = [Ad_g\xi, Ad_g\eta] \ \ , \ \ \xi, \eta \in \mathfrak{g}. \tag{4.58}$$

(2): *Infinitesimal generators: the map ad*:—
The map Ad is differentiable. Its derivative at $e \in G$ is a linear map from the Lie algebra $\mathfrak{g}$ to the space of linear maps on $\mathfrak{g}$. This map is called ad, and its value for argument $\xi \in \mathfrak{g}$ is written ad_ξ. That is:

$$ad := Ad_{*e} : \mathfrak{g} \to \mathrm{End}\mathfrak{g} \quad ; \quad ad_\xi = \frac{d}{d\tau}\ |_{\tau=0}\ Ad_{\exp(\tau\xi)} \tag{4.59}$$

where $\exp(\tau\xi)$ is the one-parameter subgroup with tangent vector ξ at the identity. But if we apply the definition eq. 4.37 of the infinitesimal generator of an action, to the adjoint action Ad, we get that for each $\xi \in \mathfrak{g}$, the generator $\xi_\mathfrak{g}$, i.e. a vector field on $\mathfrak{g}$, is

$$\xi_\mathfrak{g} : \eta \in \mathfrak{g} \mapsto \xi_\mathfrak{g}(\eta) \in \mathfrak{g} \quad \text{with} \quad \xi_\mathfrak{g}(\eta) := \frac{d}{d\tau}\ |_{\tau=0}\ Ad_{\exp(\tau\xi)}(\eta). \tag{4.60}$$

Comparing eq. 4.59, we see that ad_ξ is just the infinitesimal generator $\xi_\mathfrak{g}$ of the adjoint action corresponding to ξ:

$$ad_\xi = \xi_\mathfrak{g} \ . \tag{4.61}$$

We now compute the infinitesimal generators of the adjoint action. It will be crucial to later developments (especially Section 5.4) that these are given by the Lie bracket in $\mathfrak{g}$.

We begin by considering the function $Ad_{\exp(\tau\xi)}(\eta)$ to be differentiated. By eq. 4.49, we have

$$\begin{aligned} Ad_{\exp(\tau\xi)}(\eta) &= T_e(R_{\exp(-\tau\xi)} \circ L_{\exp(\tau\xi)})(\eta) \\ &= (T_{\exp(\tau\xi)}(R_{\exp(-\tau\xi)}) \circ T_e L_{\exp(\tau\xi)})(\eta) \\ &= (T_{\exp(\tau\xi)}(R_{\exp(-\tau\xi)}) \cdot X_\eta(\exp(\tau\xi)) \end{aligned} \tag{4.62}$$

where the second line follows by the chain rule, and the third by definition of left-invariant vector field. Writing the flow of X_ξ as $\phi_\tau(g) = g \exp \tau\xi = R_{\exp(\tau\xi)} g$, and applying the definition of the Lie derivative (eq. 3.17), we then have

$$\begin{aligned} \xi_{\mathfrak{g}}(\eta) := \frac{d}{d\tau} \mid_{\tau=0} Ad_{\exp(\tau\xi)}(\eta) = \frac{d}{d\tau} \left[T_{\phi_\tau(e)} \phi_\tau^{-1} \cdot X_\eta(\phi_\tau(e)) \right] \mid_{\tau=0} \\ = [X_\xi, X_\eta](e) = [\xi, \eta]. \end{aligned} \tag{4.63}$$

where the final equation is the definition eq. 3.45 of the Lie bracket in the Lie algebra.

So for the adjoint action, the infinitesimal generator corresponding to ξ is taking the Lie bracket: $\eta \mapsto [\xi, \eta]$. To sum up: eq. 4.59 and 4.60 now become

$$ad = Ad_{*e} : \mathfrak{g} \to \mathrm{End}\mathfrak{g} \quad ; \quad ad_\xi = \frac{d}{d\tau} \mid_{\tau=0} Ad_{\exp(\tau\xi)} = \xi_{\mathfrak{g}} : \eta \in \mathfrak{g} \mapsto [\xi, \eta] \in \mathfrak{g}. \tag{4.64}$$

(3): *Example: matrix Lie groups*:—
In the case where $G \subset GL(n, \mathrm{I\!R})$ is a matrix Lie group with Lie algebra $\mathfrak{g} \subset \mathfrak{gl}(n)$, these results are easy to verify. Writing $n \times n$ matrices as $A, B \in G$, conjugation is $K_A(B) = ABA^{-1}$, and the adjoint map Ad is also given by conjugation

$$Ad_A(X) = AXA^{-1}, \quad A \in G, X \in \mathfrak{g}. \tag{4.65}$$

So with $A(\tau) = \exp(\tau X)$, so that $A(0) = I$ and $A'(0) = X$, we have with $Y \in \mathfrak{g}$

$$\begin{aligned} \frac{d}{d\tau} \mid_{\tau=0} Ad_{\exp \tau X} Y = \frac{d}{d\tau} \mid_{\tau=0} \left[A(\tau) Y A(\tau)^{-1} \right] \\ = A'(0) Y A^{-1}(0) + A(0) Y A^{-1'}(0). \end{aligned} \tag{4.66}$$

But differentiating $A(\tau)A^{-1}(\tau) = I$ yields

$$\frac{d}{d\tau}(A^{-1}(\tau)) = -A^{-1}(\tau)A'(\tau)A^{-1}(\tau), \quad \text{and so } \ A^{-1'}(0) = -A'(0) = -X \tag{4.67}$$

so that indeed we have

$$\frac{d}{d\tau} \mid_{\tau=0} Ad_{\exp \tau X} Y = XY - YX = [X, Y]. \tag{4.68}$$

(4): *Example: the rotation group:—*
It is worth giving details for the case of $G = SO(3)$, $\mathfrak{g} = \mathfrak{so}(3)$. We saw in Section 3.4.4 (eq. 3.79) that the three matrices

$$A^x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A^y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A^z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{4.69}$$

span $\mathfrak{so}(3)$, and generate the one-parameter subgroups

$$R^x_\theta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \quad R^y_\theta = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \quad R^z_\theta = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4.70}$$

representing anticlockwise rotation around the respective coordinate axes in the physical space $\mathbb{R}^3$. To calculate the adjoint action of R^x_θ on the generator A^y, we differentiate the product $R^x_\theta R^y_\tau R^x_{-\theta}$ with respect to τ and set $\tau = 0$. That is, we find

$$Ad_{R^x_\theta}(A^y) = R^x_\theta(A^y)R^x_\theta = \begin{pmatrix} 0 & -\sin\theta & \cos\theta \\ \sin\theta & 0 & 0 \\ -\cos\theta & 0 & 0 \end{pmatrix} = \cos\theta \cdot A^y + \sin\theta \cdot A^z. \tag{4.71}$$

We similarly find

$$Ad_{R^x_\theta}(A^x) = A^x, \;\; Ad_{R^x_\theta}(A^z) = -\sin\theta \cdot A^y + \cos\theta \cdot A^z. \tag{4.72}$$

So the adjoint action of the subgroup R^x_θ representing rotations around the x-axis of physical space is given by rotations around the A^x-axis in the Lie algebra space $\mathfrak{so}(3)$. Similarly for the other subgroups representing rotations around the y or z-axis. And so for any rotation matrix $R \in SO(3)$, relative to given axes x, y, z for $\mathbb{R}^3$, its adjoint map Ad_R acting on $\mathfrak{so}(3) \cong \mathbb{R}^3$ has the same matrix representation relative to the induced basis $\{A^x, A^y, A^z\}$ of $\mathfrak{so}(3)$. (NB: This agreement between $SO(3)$'s adjoint representation and its natural physical interpretation is special to $SO(3)$: it does not hold for other matrix Lie groups.)

Finally, the infinitesimal generators of the adjoint action are given by differentiation. For example, using eq. 4.71, we find that

$$ad_{A^x}(A^y) := \frac{d}{d\theta} \mid_{\theta=0} Ad_{R^x_\theta} A^y = A^z \;\; ; \tag{4.73}$$

which agrees with the commutator: $A^z = [A^x, A^y]$.

4.5.2 The co-adjoint representation

Again we proceed in stages. We first define the representation, then discuss infinitesimal generators, and then take the rotation group as an example.

(1): *The representation defined*:—
We recall that a linear map $A : V \to W$ induces (basis-independently) a *transpose* (dual) map, written A^* (or $\tilde{A}$ or A^T), $A^* : W^* \to V^*$ on the dual spaces, $V^* := \{\alpha : V \to \mathrm{I\!R} \mid \alpha \text{ linear } \}$ and similarly for W^*; by

$$\forall \alpha \in W^*, \forall v \in V : \quad A^*(\alpha)(v) \; \equiv < A^*(\alpha) \; ; \; v > := \; \alpha(A(v)) \; \equiv \; (\alpha \circ A)(v) \; . \quad (4.74)$$

So any representation, $\mathcal{R}$ say, of a group G on a vector space V, $\mathcal{R} : G \to \mathrm{End}(V)$, induces a representation $\mathcal{R}^*$ of G on the dual space V^*, by taking the transpose. We shall call $\mathcal{R}^*$ the *dual* or *transpose* of $\mathcal{R}$; it is also sometimes called a 'contragredient representation'. That is: for $\mathcal{R}(g) : V \to V$, we define $\mathcal{R}^*(g) : V^* \to V^*$ by

$$\mathcal{R}^*(g) : \alpha \in V^* \;\mapsto\; \mathcal{R}^*(g)(\alpha) := \alpha(\mathcal{R}(g)) \in V^* \; . \quad (4.75)$$

Thus the adjoint representation of G on $\mathfrak{g}$ induces a *co-adjoint representation* of G on the dual $\mathfrak{g}^*$ of its Lie algebra $\mathfrak{g}$, i.e. on the cotangent space to the group G at the identity, $\mathfrak{g}^* = T^*_e G$. The co-adjoint representation will play a central role in symplectic reduction (starting in Section 5.4).

So let $Ad^*_g : \mathfrak{g}^* \to \mathfrak{g}^*$ be the dual (aka: transpose) of Ad_g, defined by

$$\forall \alpha \in \mathfrak{g}^*, \xi \in \mathfrak{g} : \; < Ad^*_g \alpha; \xi > := < \alpha; Ad_g \xi > . \quad (4.76)$$

Since $Ad : g \mapsto Ad_g$ is a left action ($Ad_{gh} = Ad_g Ad_h$), the assignment $g \mapsto Ad^*_g$ is a right action. So to define a left action, we use the inverse g^{-1}; cf. eq. 4.2 and 4.11. Namely, we define the left action

$$(g, \alpha) \in G \times \mathfrak{g}^* \mapsto Ad^*_{g^{-1}} \alpha \in \mathfrak{g}^* \; ; \quad (4.77)$$

called the *co-adjoint action* of G on $\mathfrak{g}^*$. And the corresponding *co-adjoint representation* of G on $\mathfrak{g}^*$ is denoted by

$$Ad^* : G \to \mathrm{End}(\mathfrak{g}^*), \quad Ad^*_{g^{-1}} = (T_e(R_g \circ L_{g^{-1}}))^* \; . \quad (4.78)$$

(2): *The map ad^*; infinitesimal generators*:—
The map Ad^* is differentiable. Its derivative at $e \in G$ is a linear map from the Lie algebra $\mathfrak{g}$ to the space of linear maps on $\mathfrak{g}^*$. This map is called ad^*, and its value for argument $\xi \in \mathfrak{g}$ is written ad^*_ξ. Thus ad^*_ξ is an endomorphism of $\mathfrak{g}^*$, and we have

$$ad^* = Ad^*_{*e} : \xi \in \mathfrak{g} \to ad^*_\xi \in \mathrm{End}\mathfrak{g}^* \; . \quad (4.79)$$

Now recall our deduction from eq. 4.59 and 4.60 that $ad_\xi = \xi_{\mathfrak{g}}$, i.e. eq. 4.61. In the same way we here deduce an equality to the infinitesimal generator of the co-adjoint action:

$$ad^*_\xi = \xi_{\mathfrak{g}^*} \; . \quad (4.80)$$

In fact, ad^*_ξ is, *modulo* a minus sign, the *adjoint* of ad_ξ, in the usual sense of the natural pairing of a vector space with its dual: as we now show. (So the notation ad^* is justified, *modulo* a minus sign.)

Let us compute for this action, the value of the infinitesimal generator $\xi_{\mathfrak{g}^*}$ (a vector field on $\mathfrak{g}^*$, induced by $\xi \in \mathfrak{g}$) at the point $\alpha \in \mathfrak{g}^*$. That is, we will compute the value $\xi_{\mathfrak{g}^*}(\alpha)$. As usual, we identify the tangent space $(T\mathfrak{g}^*)_\alpha$ in which this value lives, with $\mathfrak{g}^*$ itself; and similarly for $\mathfrak{g}$. So, with $\xi_{\mathfrak{g}^*}$ acting on $\eta \in \mathfrak{g}$, we compute:

$$< ad^*_\xi(\alpha); \eta > \; \equiv \; < \xi_{\mathfrak{g}^*}(\alpha); \eta > \; = \; \left\langle \frac{d}{d\tau} |_{\tau=0} Ad^*_{\exp -\tau\xi}(\alpha); \eta \right\rangle \tag{4.81}$$

$$= \; \frac{d}{d\tau} |_{\tau=0} \left\langle Ad^*_{\exp -\tau\xi}(\alpha); \eta \right\rangle \; = \; \frac{d}{d\tau} |_{\tau=0} \langle \alpha; Ad_{\exp -\tau\xi}\eta \rangle \tag{4.82}$$

$$= \; \left\langle \alpha; \frac{d}{d\tau} |_{\tau=0} Ad_{\exp -\tau\xi}\eta \right\rangle \; = \; < \alpha; -[\xi, \eta] > \; = \; - < \alpha; ad_\xi(\eta) > . \tag{4.83}$$

So ad^*_ξ, defined as the derivative of Ad^* is, up to a sign, the adjoint of ad_ξ.

(3): *Example: the rotation group*:—
Let us now write the elementary vector product in $\mathbb{R}^3$ as $\wedge$, and identify $\mathfrak{so}(3) \cong (\mathbb{R}^3, \wedge)$ and $\mathfrak{so}(3)^* \cong \mathbb{R}^{3^*}$. And let us have the natural pairing given by the elementary euclidean inner product $\cdot$. Then the result just obtained (now with $\bullet$ marking the argument-place)

$$< \xi_{\mathfrak{g}^*}(\alpha); \bullet >= - < \alpha; [\xi, \bullet] > \tag{4.84}$$

becomes for $\alpha \in \mathfrak{so}(3)^*$ and $\xi \in so(3)$

$$\xi_{\mathfrak{so}(3)^*}(\alpha) \cdot \bullet = -\alpha \cdot (\xi \wedge \bullet) \; . \tag{4.85}$$

So for $\eta \in \mathfrak{so}(3)$, we have

$$< \xi_{\mathfrak{so}(3)^*}(\alpha); \eta > = \xi_{\mathfrak{so}(3)^*}(\alpha) \cdot \eta = -\alpha \cdot (\xi \wedge \eta) = -(\alpha \wedge \xi) \cdot \eta = - < \alpha \wedge \xi; \eta > . \tag{4.86}$$

In short:

$$\xi_{\mathfrak{so}(3)^*}(\alpha) = -\alpha \wedge \xi = \xi \wedge \alpha. \tag{4.87}$$

Now since $SO(3)$ is compact, we know that the co-adjoint action is proper; so $\mathrm{Orb}(\alpha)$ is a closed submanifold of $\mathfrak{so}(3)^*$, and eq. 4.45 of Section 4.4 applies. So if we fix α, and let ξ vary through $\mathfrak{so}(3) \cong \mathbb{R}^3$, we get all of the tangent space $T_\alpha\mathrm{Orb}(\alpha)$ to the orbit passing through α. Then eq. 4.87 implies that the tangent space is the plane normal to α, and passing through α's end-point. Letting α vary through $\mathfrak{so}(3)^*$, we conclude that the *co-adjoint orbits are the spheres centred on the origin.*

In the following Sections, we will see that the orbits of the co-adjoint representation of *any* Lie group G have a natural symplectic structure. So the orbits are always even-dimensional; and by considering all Lie groups and all possible orbits, we can get a series of examples of symplectic manifolds.

Besides, this fact will play a central role in our generalized formulation of Hamiltonian mechanics, and in symplectic reduction. And we will (mercifully!) get a good understanding of that role, already in Section 5.1. To prepare for that, it is worth gathering some threads about our recurrent example, $SO(3)$; and generalizing them to other Lie groups ...

4.6 Kinematics on Lie groups

To summarize some aspects of this Section, and to make our later discussion of reduction clearer, it is worth collecting and generalizing some of our results about $SO(3)$ and the description it provides of the rigid body. More precisely, we will now combine:

(i): the description of space and body coordinates in terms of left and right translation, at the end of Section 3.4.4;

(ii): the cotangent lift of translation (example (viii) of Section 4.1.A);

(iii): the adjoint and co-adjoint representations of $SO(3)$ (as in (4) of Section 4.5.1, and (3) of Section 4.5.2.

We will also generalize: namely, we will consider (i) to (iii) for an arbitrary Lie group G, not just for $SO(3)$. (The point of doing so will become clear in (3) of Section 5.1.) This will occur already in Section 4.6.1. Then in Section 4.6.2, we will show how this material yields natural diffeomorphisms

$$TG \to G \times \mathfrak{g} \quad \text{and} \quad T^*G \to G \times \mathfrak{g}^* \ ; \tag{4.88}$$

(so if $\dim G = n$, then all four manifolds are $2n$-dimensional). We will also see that by applying Section 4.2's notion of equivariance, we can "pass to the quotients", and get from eq. 4.88, the natural diffeomorphisms

$$TG/G \to \mathfrak{g} \quad \text{and} \quad T^*G/G \to \mathfrak{g}^* \ ; \tag{4.89}$$

where the quotients on the left hand sides (the domains) is by the action of left translation; (to be precise: by the action of its derivative for TG, and its cotangent lift for T^*G).

4.6.1 Space and body coordinates generalized to G

So let a (finite-dimensional) Lie group G act on itself by left and right translation, L_g and R_g. For any $g \in G$, we define

$$\lambda_g : T_gG \to \mathfrak{g} \ \text{by} \ \ v \in T_gG \mapsto (T_eL_g)^{-1}(v) \equiv (T_gL_{g^{-1}})(v) \in \mathfrak{g} \ . \tag{4.90}$$

We similarly define

$$\rho_g : v \in T_gG \mapsto (T_eR_g)^{-1}(v) \equiv (T_gR_{g^{-1}})(v) \in \mathfrak{g} \ . \tag{4.91}$$

On analogy with the case of the pivoted rigid body (cf. eq. 3.85 and 3.86, or eq. 4.22), we say that λ_g represents $v \in T_gG$ in *body coordinates*, and ρ_g represents v in *space coordinates*. We also speak of body and space *representations*. The transition from body to space coordinates is then an isomorphism of $\mathfrak{g}$; viz. by eq. 4.49

$$\forall \xi \in \mathfrak{g}, \;\; (\rho_g \circ \lambda_g^{-1})(\xi) = \rho_g(T_e L_g(\xi)) \equiv Ad_g \xi \; . \tag{4.92}$$

So we can combine the S and B superscript notation of eq. 4.22 with Section 4.5.1's notion of the adjoint representation, and write

$$v^S = Ad_g v^B \; . \tag{4.93}$$

In a similar way, the cotangent lifts of left and right translation provide isomorphisms between the dual spaces $T_g^*G, g \in G$ and $\mathfrak{g}^*$. Thus for any $g \in G$, we define

$$\bar{\lambda}_g : T_g^*G \to \mathfrak{g}^* \;\text{ by }\; \alpha \in T_g^*G \mapsto \alpha \circ T_e L_g \equiv (T_e L_g)^*(\alpha) \equiv (T_e^* L_g)(\alpha) \in \mathfrak{g}^* \; ; \tag{4.94}$$

and similarly

$$\bar{\rho}_g : \alpha \in T_g^*G \mapsto \alpha \circ T_e R_g \equiv (T_e^* R_g)(\alpha) \in \mathfrak{g}^* \; . \tag{4.95}$$

And we again use the S and B superscript notation of eq. 4.22, and define for $\alpha \in T_g^*G$

$$\alpha^S := (T_e^* R_g)(\alpha) \equiv \bar{\rho}_g(\alpha) \quad \text{and} \quad \alpha^B := (T_e^* L_g)(\alpha) \equiv \bar{\lambda}_g(\alpha) \; , \tag{4.96}$$

which are called the *space* (or 'spatial') and *body* representations, respectively, of α. The transition from body to space representations is now an isomorphism of $\mathfrak{g}^*$; viz.

$$\forall \alpha \in \mathfrak{g}^*, \;\; (\bar{\rho}_g \circ \bar{\lambda}_g^{-1})(\alpha) = Ad^*_{g^{-1}}(\alpha) \; , \quad \text{i.e.} \quad \alpha^S = Ad^*_{g^{-1}}(\alpha^B) \; . \tag{4.97}$$

4.6.2 Passage to the quotients

For later purposes, we need to develop the details of how the element $g \in G$ "carries along throughout" in eq. 4.90 to 4.97. More precisely, we have two isomorphisms:

$$TG \cong G \times \mathfrak{g} \quad \text{and} \quad T^*G \cong G \times \mathfrak{g}^* \; . \tag{4.98}$$

These are isomorphisms of vector bundles; but we shall not develop the language of fibre bundles. What matters for us is that once we exhibit these isomorphisms, we will see that we have equivariant maps relating two group actions, in the sense of eq. 4.29 and 4.30. And this will mean that we can pass to the quotients to infer that TG/G is diffeomorphic to $\mathfrak{g}$, and correspondingly that T^*G/G is diffeomorphic to $\mathfrak{g}^*$.

This last diffeomorphism will form the first part of Section 7's main theorem, the Lie-Poisson reduction theorem, which says that T^*G/G and $\mathfrak{g}^*$ are isomorphic as Poisson manifolds. In Section 5 onwards, we will develop the notion of a Poisson manifold, and the significance of this isomorphism for the reduction of mechanical problems.

I should note here that there is a parallel story about the first diffeomorphism, i.e.

about TG/G being diffeomorphic to $\mathfrak{g}$. It forms the first part of another reduction theorem, which is the Lagrangian analogue of Section 7's Lie-Poisson theorem. But since this Chapter has adopted the Hamiltonian approach, I will not go into details. They can be found in Marsden and Ratiu (1999: Sections 1.2, 13.5, 13.6), under the title 'Euler-Poincaré reduction'.

Thus corresponding to eq. 4.90, we define the isomorphism

$$\lambda : TG \to G \times \mathfrak{g} \quad \text{by} \quad \lambda(v) := (g, (T_e L_g)^{-1}(v)) \equiv (g, (T_g L_{g^{-1}})(v)) \tag{4.99}$$

with $v \in T_g G$, i.e. $g = \pi_G(v)$ and $\pi_G : TG \to G$ the canonical projection. (As mentioned concerning eq. 4.6, it is harmless to (follow many presentations and) conflate a point in TG, i.e. strictly speaking a pair $(g, v), g \in G, v \in T_g G$, with its vector v.) And corresponding to eq. 4.91, we define the isomorphism

$$\rho : TG \to G \times \mathfrak{g} \quad \text{by} \quad \rho(v) := (g, (T_e R_g)^{-1}(v)) \equiv (g, (T_g R_{g^{-1}})(v)) \; . \tag{4.100}$$

The transition from body to space representations given by eq. 4.92 now implies

$$(\rho \circ \lambda^{-1})(g, \xi) = \rho(g, T_e L_g(\xi)) = (g, (T_e R_g)^{-1} \circ T_e L_g(\xi)) = (g, Ad_g \xi). \tag{4.101}$$

In a similar way, the cotangent bundle T^*G is isomorphic in two ways to $G \times \mathfrak{g}^*$: namely by

$$\bar{\lambda}(\alpha) := (g, \alpha \circ T_e L_g) \equiv (g, (T_e^* L_g)\alpha) \in G \times \mathfrak{g}^* \; , \tag{4.102}$$

and by

$$\bar{\rho}(\alpha) := (g, \alpha \circ T_e R_g) \equiv (g, (T_e^* R_g)\alpha) \in G \times \mathfrak{g}^* \tag{4.103}$$

where $\alpha \in T_g^* G$, i.e. $g = \pi_G^*(\alpha)$ with $\pi_G^* : T^*G \to G$ the canonical projection. (Again, we harmlessly conflate a point (g, α) in T^*G with its form $\alpha \in T_g^* G$.)

Let us now compute in the body representation, the actions of: (i) the (derivative of the) left translation map, TL_g, and (ii) the corresponding cotangent lift T^*L_g. This will show that λ and $\bar{\lambda}$ are equivariant maps for certain group actions.

(i): We compute:

$$(\lambda \circ TL_g \circ \lambda^{-1})(h, \xi) = (\lambda \circ TL_g)(h, TL_h(\xi)) = \lambda(gh, (TL_g \circ TL_h)(\xi)) \tag{4.104}$$

$$= (gh, ((TL_{gh})^{-1} \circ TL_{gh})(\xi)) = (gh, \xi). \tag{4.105}$$

So in the body representation, left translation does not act on the vector component. (That is intuitive, in that the vector ξ is "attached to the body" and so should not vary relative to coordinates fixed in it.) Eq. 4.105 means that λ is an equivariant map relating left translation TL_g on TG to the G-action on $G \times \mathfrak{g}$ given just by left translation on the first component:

$$\Phi_g((h, \xi)) \equiv g \cdot (h, \xi) := (gh, \xi) \; . \tag{4.106}$$

Equivariance means that λ induces a map $\hat{\lambda}$ on the quotients. That is: as in eq. 4.31, the map

$$\hat{\lambda} : TG/G \to (G \times \mathfrak{g})/G \tag{4.107}$$

defined as mapping, for any g, the orbit of any $v \in T_gG$ to the orbit of $\lambda(v)$, i.e.

$$\hat{\lambda} : \mathrm{Orb}(v) \equiv \{u \in TG \mid T_gL_h(v) = u, \text{ some } h \in G\} \mapsto \mathrm{Orb}(\lambda(v)) \tag{4.108}$$
$$\equiv \{(hg, (T_eL_g)^{-1}(v)) \mid \text{some } h \in G\} \tag{4.109}$$

is well-defined, i.e. independent of the chosen representative v of the orbit.

Besides, since the canonical projections, $v \in TG \mapsto \mathrm{Orb}(v) \in TG/G$ and $(g, \xi) \mapsto \mathrm{Orb}((g, \xi)) \in (G \times \mathfrak{g})/G$, are submersions, we can apply result (1) of Section 4.2 and conclude that $\hat{\lambda}$ is smooth.

Finally, we notice that since the action of left translation is transitive, we can identify each orbit of the Φ of eq. 4.106 with its right component $\xi \in \mathfrak{g}$; and so we can identify the set of orbits $(G \times \mathfrak{g})/G$ with $\mathfrak{g}$.

To sum up: we have shown that TG/G and $(G \times \mathfrak{g})/G$, i.e. in effect $\mathfrak{g}$, are diffeomorphic:

$$\hat{\lambda} : TG/G \;\to\; (G \times \mathfrak{g})/G \equiv \mathfrak{g} \; . \tag{4.110}$$

(ii): The results for the cotangent bundle are similar to those in (i). On analogy with eq. 4.105, the action of the cotangent lift of left translation T^*L_g is given in body representation by applying eq. 4.102 to get

$$(\bar{\lambda} \circ (T^*L_g) \circ \bar{\lambda}^{-1})(h, \alpha) = (g^{-1}h, \alpha) \; ; \tag{4.111}$$

or equivalently, now taking the cotangent lift of left translation to define a left action (cf. eq. 4.11),

$$(\bar{\lambda} \circ (T^*L_{g^{-1}}) \circ \bar{\lambda}^{-1})(h, \alpha) = (gh, \alpha) \; . \tag{4.112}$$

So in body representation, left translation does not act on the covector component; (again, an intuitive result in so far as α is "attached to the body"). So eq. 4.112 means that $\bar{\lambda}$ is an equivariant map relating the cotangent lifted left action of left translation on T^*G to the G-action on $G \times \mathfrak{g}^*$ given just by left translation on the first component:

$$\Phi_g((h, \alpha)) \equiv g \cdot (h, \alpha) := (gh, \alpha) \; . \tag{4.113}$$

So, on analogy with eq. 4.107 and 4.109, we can pass to the quotients, defining a map

$$\hat{\bar{\lambda}} : T^*G/G \to (G \times \mathfrak{g}^*)/G \tag{4.114}$$

by requiring that for $\alpha \in T^*_gG$, so that $T^*L_{h^{-1}}\alpha \in T^*_{hg}G$:

$$\hat{\bar{\lambda}} : \mathrm{Orb}(\alpha) \equiv \{\beta \in T^*G \mid \beta = T^*L_{h^{-1}}(\alpha), \text{ some } h \in G\} \mapsto \tag{4.115}$$
$$\mathrm{Orb}(\bar{\lambda}(\alpha)) \equiv \{(hg, (T^*_eL_g)(\alpha)) \mid \text{some } h \in G\} \equiv \{(h, (T^*_eL_g)\alpha) \mid \text{some } h \in G\} \tag{4.116}$$

And finally, we identify the set of orbits $(G \times \mathfrak{g}^*)/G$ with $\mathfrak{g}^*$, so that we conclude that T^*G/G and $\mathfrak{g}^*$ are diffeomorphic. That is, we think of the diffeomorphism $\hat{\bar{\lambda}}$ as mapping T^*G/G to $\mathfrak{g}^*$:

$$\hat{\bar{\lambda}} : \mathrm{Orb}(\alpha) \equiv \{\beta \in T^*G \mid \beta = T^*L_{h^{-1}}(\alpha), \text{ some } h \in G\} \in T^*G/G \mapsto (T^*_e L_g)(\alpha) \in \mathfrak{g}^* . \tag{4.117}$$

As I said above, this diffeomorphism is the crucial first part of Section 7's main reduction theorem. But we will see its role there, already in (3) of Section 5.1.

Finally, a result which will *not* be needed later. To calculate the derivatives and cotangent lifts of left translation in *space* representation, we replace λ and $\bar{\lambda}$ by ρ and $\bar{\rho}$ as defined by eq. 4.100 and 4.103. We get as the analogues of eq. 4.105 and 4.111 respectively:

$$(\rho \circ TL_g \circ \rho^{-1})(h, \xi) = (gh, Ad_g(\xi)) \; , \tag{4.118}$$

and

$$(\bar{\rho} \circ T^*L_g \circ \bar{\rho}^{-1})(h, \alpha) = (g^{-1}h, Ad^*_g(\alpha)) \; . \tag{4.119}$$

Though these results are not needed later, they are also analogues of some later results, eq. 6.89 and 6.90, which we will need. (Note that, in accordance with the discussion between eq. 4.76 and 4.77, eq. 4.119 involves right actions.)

5 Poisson manifolds

5.1 Preamble: three reasons for Poisson manifolds

Now that we are equipped with Sections 3 and 4's toolbox of modern geometry, we can develop, in this Section and the two to follow, the theory of symplectic reduction. This Section develops the general theory of Poisson manifolds, as a framework for a generalized Hamiltonian mechanics. Its main results concern the foliation, and quotienting, of Poisson manifolds. Then Section 6 returns us to symmetries and conserved quantities: topics which are familiar from Section 2.1.3, but which Section 6 will discuss in the generalized framework using the notion of a momentum map. Finally, in Section 7 all the pieces of our jigsaw puzzle will come together, in our symplectic reduction theorem.

We already glimpsed in (1) of Section 2.2 the idea of a Poisson manifold as a generalization of a symplectic manifold, that provides the appropriate framework for a generalized Hamiltonian mechanics. It is a manifold equipped with a bracket, called a 'Poisson bracket', that has essentially the same formal defining properties as in symplectic geometry except that it can be "degenerate". In particular, the dimension m of a Poisson manifold M can be even or odd. As we will see, Hamiltonian mechanics can be set up on Poisson manifolds, in a natural generalization of the usual formalism: there are m first-order ordinary differential equations for the time evolution of local coordinates $x^1, ..., x^m$, and the time-derivative of any dynamical variable (scalar function on the Poisson manifold M) is given by its Poisson bracket with the Hamiltonian.

Besides, this generalization reduces to the usual formalism in the following sense. Any Poisson manifold M is foliated into symplectic manifolds, and any Hamiltonian mechanics of our generalized kind defined on M restricts on each symplectic leaf to a conventional Hamiltonian mechanics using the induced symplectic form.

This last point, the invariance of the symplectic leaves under the dynamics, prompts the question 'why bother with the Poisson manifold, since the dynamics can be written down on each leaf?'. There are three reasons. I will just mention the first; the rest of Section 5 will develop the second; and the two subsequent Sections will develop the third.

(1): *Parameters and stability*:—
The first two reasons concern the fact that for many problems in Hamiltonian mechanics, it is natural to consider an odd-dimensional state-space. One principal way this happens is if the system is characterized by some odd number, say s (maybe $s = 1$), of parameters that are constant in time. Then even though for a fixed value of the parameter(s), there is a Hamiltonian mechanics on a symplectic manifold, of dimension $2n$ say, it is useful to envisage the $2n + s$ dimensional space in order to keep track of how the behaviour of systems depends on the parameters. For example, this is very useful for analysing stability, especially if one can somehow control the value of the parameters. Stability theory (and related fields such as bifurcation theory) are crucially important, and vast, topics—which I will not go into.[21]

(2): *Odd-dimensional spaces: the rigid body again*:—
Secondly, even in the absence of such controllable parameters, there are mechanical systems whose description leads naturally to an odd-dimensional state-space. The paradigm elementary example is the rigid body pivoted at a point (mentioned in (3) of Section 2.2). An elementary analysis, repeated in every textbook, leads to a description of the body by the three components of the angular momentum (relative to body coordinates, i.e. coordinates fixed in the body): these components evolve according to the three first-order Euler equations.

This situation prompts two foundational questions; (which of course most textbooks ignore!). First, we note that a configuration of the body is given by three real numbers: viz. to specify the rotation required to rotate the body into the given configuration, from a fiducial configuration. So a conventional Hamiltonian description of the rigid body would use six first-order equations. (Indeed, similarly for a Lagrangian description, if we treat the three $\dot{q}$s as variables.) So how is the description by Euler's equations related to a six-dimensional Hamiltonian (or indeed Lagrangian) description?

Second, can the description by the Euler equations be somehow regarded as itself Hamiltonian, or Lagrangian?

This Chapter will not pursue these questions about the rigid body; for details, cf.

[21]Except to note a broad philosophical point. These parameters illustrate the modal or counterfactual involvements of mechanics. The s dimensions of the state-space, and the mathematical constructions built on them, show how rich and structured these involvement are. For a detailed discussion of the modal involvements of mechanics, cf. Butterfield (2004).

the references at the end of (3) of Section 2.2. For us, the important point is that the theory of symplectic reduction shows that the answer to the second question is Yes. Indeed, a "resounding Yes". For we will see very soon (in Section 5.2.4.A) that the three-dimensional space of the components, in body coordinates, of the angular momentum is our prototype example of a Poisson manifold; and the evolution by Euler's equations is the Hamiltonian mechanics on each symplectic leaf of this manifold. In short: in our generalized framework, Euler's equations are *already* in Hamiltonian form.

Furthermore, this Poisson manifold is already familiar: it is $\mathfrak{so}(3)^*$, the dual of the Lie algebra of the rotation group. Here we connect with several previous discussions (and this Chapter's second motto).

First: we connect with the discussion of rotation in Relationist and Reductionist mechanics (Sections 2.3.3 to 2.3.5). In particular, cf. comment (iii) about γ, the three variables encoding the total angular momentum of the system, at the end of Section 2.3.4. (So as regards (1)'s idea of labelling the symplectic leaves by parameters constant in time: in this example, it is the magnitude L of the total body angular momentum which is the parameter.)

Second: we connect with Section 3.4.4's discussion of $\mathfrak{so}(3)$, with Section 4.5.2's discussion of the co-adjoint representation on $\mathfrak{so}(3)^*$, and with Section 4.6's discussion of kinematics on an arbitrary Lie group. As regards the rigid body, the main physical idea is that the action of $SO(3)$ on itself by left translation is interpreted in terms of the coordinate transformation, i.e. rotation, between the space and body coordinate systems.

But setting aside the rigid body: recall that in Section 4.5.2 we saw that for $\mathfrak{so}(3)^*$, the co-adjoint orbits are the spheres centred on the origin. I also announced that they have a natural symplectic structure—and that this was true for the orbits of the co-adjoint representation of *any* Lie group. Now that we have the notion of a Poisson manifold, we can say a bit more, though of course the proofs are yet to come:—

> For any Lie group G, the dual of its Lie algebra $\mathfrak{g}^*$ is a Poisson manifold; and G has on $\mathfrak{g}^*$ a co-adjoint representation, whose orbits are the symplectic leaves of $\mathfrak{g}^*$ as a Poisson manifold.

In particular, we remark that the theory of the rigid body just sketched is independent of the dimension of physical space being three: it carries over to $\mathfrak{so}(n)^*$ for any n. So we can readily do the Hamiltonian mechanics of the rigid body in arbitrary dimensions. That sounds somewhat academic! But it leads to a more general point, which is obviously of vast practical importance.

In engineering we often need to analyse or design bodies consisting of two or more rigid bodies jointed together, e.g. at a universal joint. Often the configuration space of such a jointed body can be given by a sequence of rotations (in particular about the joints) and-or translations from a fiducial configuration; so that we can take an appropriate Lie group G as the body's configuration space. If so, we can try to mimic our strategy for the rigid body, i.e. to apply the result just announced. And indeed, for such bodies, the action of left translation, and so the adjoint and co-adjoint representations of G on $\mathfrak{g}$ and $\mathfrak{g}^*$, can often be physically significant.

But leaving engineering aside, let us sum up this second reason for Poisson manifolds as follows. For some mechanical systems the natural state-space for a Hamiltonian mechanics is a Poisson manifold. And in the paradigm case of the rigid body, there is a striking interpretation of the Poisson manifold's leaves as the orbits of the co-adjoint representation of the rotation group $SO(3)$.

(3): *Reduction*:—
My first two reasons have not mentioned reduction. But unsurprisingly, they have several connections with the notion. Here I shall state just one main connection, which links Section 4.6's kinematics on Lie groups to our main reduction theorem: this will be my third motivation for studying Poisson manifolds.

In short, the connection is that:—

(i): For various systems, the configuration space is naturally taken to be a Lie group G; (as we have just illustrated with the rigid body).

(ii): So it is natural to set up an orthodox Hamiltonian mechanics of the system on the cotangent bundle T^*G. But (as in the Reductionist procedure of Section 2.3.4) it is also natural to quotient by the lift to the cotangent bundle of G's action on itself by left translation.

(iii): When we do this, the resulting reduced phase space T^*G/G is a Poisson manifold. Indeed it is an isomorphic copy of $\mathfrak{g}^*$. That is, we have an isomorphism of Poisson manifolds: $\mathfrak{g}^* \cong T^*G/G$. This is the Lie-Poisson reduction theorem.

I shall give a bit more detail about each of (i)-(iii).

(i): For various systems, any configuration can be obtained by acting with an element of the Lie group G on some reference configuration which can itself be labelled by an element of G, say the identity $e \in G$. So we take the Lie group G to be the configuration space. As mentioned in (3) of Section 2.2, there is even an infinite-dimensional example of this: the ideal fluid.

(ii): So T^*G is the conventional Hamiltonian phase space of the system. But G acts on itself by left translation. We can then consider the quotient of T^*G by the cotangent lift of left translation. Intuitively, this is a matter of "rubbing out" the way that T^*G encodes (i)'s choice of reference configuration. By passing to the quotients as in Section 4.6, we infer that T^*G/G is a manifold. But of course it is in general *not* even-dimensional. For its dimension is $\frac{1}{2}\text{dim}(T^*G) \equiv \text{dim}(G)$. So consider any odd-dimensional G: for example, our old friend, the three-dimensional rotation group $SO(3)$.

(iii): But T^*G/G *is* always a Poisson manifold. And it is always isomorphic as a Poisson manifold to $\mathfrak{g}^*$, with its symplectic leaves being the co-adjoint orbits of $\mathfrak{g}^*$: $\mathfrak{g}^* \cong T^*G/G$.

I end this third reason for studying Poisson manifolds with two remarks about examples.

The first remark echoes the end of Section 4.5.2, where I said that by considering all possible Lie groups and all the orbits of their co-adjoint representations, we get a series of examples of symplectic manifolds. We can now put this together with the notion of a Poisson manifold, and with the comment at the end of Section 3.4.3, that every (finite-dimensional) Lie algebra is the Lie algebra of a Lie group. In short: we get a series of examples of Poisson manifolds, in either of two equivalent ways: from the dual $\mathfrak{g}^*$ of any (finite-dimensional) Lie algebra $\mathfrak{g}$; or from the quotient T^*G/G of the cotangent lift of left translation. In either case, the example is the co-adjoint representation.

The second remark is that there are yet other examples of Poisson manifolds and reductions. Indeed, we noted one in Section 2.3.4: viz. the Reductionist's reduced phase space $\bar{M} := M/E$, obtained by quotienting the phase space $M := T^*\mathrm{I\!R}^{3N} - (\delta \cup \Delta)$ by the (cotangent lift) of the action of the euclidean group E on $\mathrm{I\!R}^{3N}$. But I shall not go into further details about this example; (for which cf. the Belot papers listed in Section 2.3.1, and references therein). Here it suffices to note that this example is not of the above form: $\mathrm{I\!R}^{3N}$ is not E, and the action of E on $\mathrm{I\!R}^{3N}$ is not left translation. This of course echoes my remarks at the end of Section 1.2 that the theory of symplectic reduction is too large and intricate for this Chapter to be more than an "appetizer".

So much by way of motivating Poisson manifolds. The rest of this Section will cover reasons (1) and (2); but reason (3), about reduction, is postponed to Sections 6 and 7. We give some basics about Poisson manifolds, largely in coordinate-dependent language, in Section 5.2. In Section 5.3, we move to a more coordinate-independent language and show that Poisson manifolds are foliated into symplectic manifolds. In Section 5.4, we show that the leaves of the foliation of a finite-dimensional Lie algebra $\mathfrak{g}^*$ are the orbits of the co-adjoint representation of G on $\mathfrak{g}^*$. Finally in Section 5.5, we prove a general theorem about quotienting a Poisson manifold by the action of Lie group, which will be important for Section 7's main theorem.

5.2 Basics

In Sections 5.2.1 to 5.2.3, we develop some basic definitions and results about Poisson manifolds. This leads up to Section 5.2.4, where we see that the dual of any finite-dimensional Lie algebra has a natural (i.e. basis-independent) Poisson manifold structure. Throughout, there will be some obvious echoes of previous discussions of anti-symmetric forms, Poisson brackets, Hamiltonian vector fields and Lie brackets (Sections 2.1 and 3.2). But I will for the most part *not* articulate these echoes.

5.2.1 Poisson brackets

A manifold M is called a *Poisson manifold* if it is equipped with a *Poisson bracket* (also known as: *Poisson structure*). A *Poisson bracket* is an assignment to each pair

of smooth real-valued functions $F, H : M \to \mathbb{R}$, of another such function, denoted by $\{F, H\}$, subject to the following four conditions:—
(a) Bilinearity:

$$\{aF + bG, H\} = a\{F, H\} + b\{G, H\} \; ; \; \{F, aG + bH\} = a\{F, G\} + b\{F, H\} \;\; \forall a, b \in \mathbb{R}. \tag{5.1}$$

(b) Anti-symmetry:

$$\{F, H\} = -\{H, F\} \; . \tag{5.2}$$

(c) Jacobi identity:

$$\{\{F, H\}, G\} + \{\{G, F\}, H\} + \{\{H, G\}, F\} = 0 \; . \tag{5.3}$$

(d) Leibniz' rule:

$$\{F, H \cdot G\} = \{F, H\} \cdot G + H \cdot \{F, G\} \; . \tag{5.4}$$

In other words: M is a Poisson manifold iff both: (i) the set $\mathcal{F}(M)$ of smooth scalar functions on M, equipped with the bracket $\{,\}$, is a Lie algebra; and (ii) the bracket $\{,\}$ is a derivation in each factor.

Any symplectic manifold is a Poisson manifold. The Poisson bracket is defined by the manifold's symplectic form; cf. eq. 2.18.

"Canonical" Example:—
Let $M = \mathbb{R}^m, m = 2n+l$, with standard coordinates $(q, p, z) = (q^1, ..., q^n, p^1, ..., p^n, z^1, ..., z^l)$. Define the Poisson bracket of any two functions $F(q, p, z)$, $H(q, p, z)$ by

$$\{F, H\} := \Sigma_i^n \left(\frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p^i} - \frac{\partial F}{\partial p^i} \frac{\partial H}{\partial q^i} \right). \tag{5.5}$$

Thus this bracket ignores the z coordinates; and if l were equal to zero, it would be the standard Poisson bracket for $\mathbb{R}^{2n}$ as a symplectic manifold. We can immediately deduce the Poisson brackets for the coordinate functions. Those for the qs and ps are as for the usual symplectic case:

$$\{q^i, q^j\} = 0 \quad \{p^i, p^j\} = 0 \quad \{q^i, p^j\} = \delta_{ij}. \tag{5.6}$$

On the other hand, all those involving the zs vanish:

$$\{q^i, z^j\} = \{p^i, z^j\} = \{z^i, z^j\} \equiv 0. \tag{5.7}$$

Besides, any function F depending only on the z's, $F \equiv F(z)$ will have vanishing Poisson brackets with all functions $H : \{F, H\} = 0$.

This example seems special in that M is foliated into $2n$-dimensional symplectic manifolds, each labelled by l constant values of the zs. But Section 5.3.4 will give a generalization for Poisson manifolds of Darboux's theorem (mentioned at the end of

Section 2.1.1): a generalization saying, roughly speaking, that every Poisson manifold "looks locally like this".

For any Poisson manifold, we say that a function $F : M \to \mathbb{R}$ is *distinguished* or *Casimir* if its Poisson bracket with all smooth functions $H : M \to \mathbb{R}$ vanishes identically: $\{F, H\} = 0$.

5.2.2 Hamiltonian vector fields

Given a smooth function $H : M \to \mathbb{R}$, consider the map on smooth functions: $F \mapsto \{F, H\}$. The fact that the Poisson bracket is bilinear and obeys Leibniz's rule implies that this map $F \mapsto \{F, H\}$ is a derivation on the space of smooth functions, and so determines a vector field on M; (cf. (ii) of Section 3.1.2.B). We call this vector field the *Hamiltonian vector field* associated with (also known as: generated by) H, and denote it by X_H.

But independently of the Poisson structure, the action of any vector field X_H on a smooth function F, $X_H(F)$, also equals $L_{X_H}(F) \equiv dF(X_H)$; (cf. eq. 3.12). So we have for all smooth F

$$L_{X_H}(F) \equiv dF(X_H) \equiv X_H(F) = \{F, H\} \ . \tag{5.8}$$

The equations describing the flow of X_H are called *Hamilton's equations*, for the choice of H as "Hamiltonian".

In the previous example with $M = \mathbb{R}^{2n+l}$, we have

$$X_H = \Sigma_i^n \left(\frac{\partial H}{\partial p^i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p^i} \right) , \tag{5.9}$$

and the flow is given by the ordinary differential equations

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p^i} \quad \frac{dp^i}{dt} = -\frac{\partial H}{\partial q^i} \quad \frac{dz^j}{dt} = 0. \quad i = 1, ..., n; \;\; j = 1, ..., l. \tag{5.10}$$

Again, the zs, and any function $F(z)$ solely of them, are distinguished and have a vanishing Hamiltonian vector field. On the other hand, the coordinate functions q^i and p^i generate the Hamiltonian vector fields $-\frac{\partial}{\partial p^i}$ and $\frac{\partial}{\partial q^i}$ respectively.

Two further remarks about eq. 5.8:—

(1): It follows that a function H is distinguished (i.e. has vanishing Poisson brackets with all functions) iff its Hamiltonian vector field X_H vanishes everywhere. And since the Poisson bracket is antisymmetric, this is so iff H is constant along the flow of all Hamiltonian vector fields.

(2): This equation is the beginning of the theory of constants of the motion (first integrals), and of Noether's theorem, for Poisson manifolds; just as the corresponding equation was the beginning for the symplectic case. This will be developed in Section 6.

Poisson brackets and Lie brackets:—
With the definition eq. 5.8 in hand, we can readily establish our first important connection between Poisson manifolds and Section 3's Lie structures. Namely: result (2) at the end of Section 3.2.2, eq. 3.32, is also valid for Poisson manifolds.

That is: the Hamiltonian vector field of the Poisson bracket of scalars F, H on a Poisson manifold M is, upto a sign, the Lie bracket of the Hamiltonian vector fields, X_F and X_H, of F and H:

$$X_{\{F,H\}} = -[X_F, X_H] = [X_H, X_F] \ . \tag{5.11}$$

The proof is exactly as for eq. 3.32.

So the Hamiltonian vector fields, with the Poisson bracket, form a Lie subalgebra of the Lie algebra $\mathcal{X}_M$ of all vector fields on the Poisson manifold M. This result will be important in Section 5.3.3's proof that every Poisson manifold is a disjoint union of symplectic manifolds.

5.2.3 Structure functions

We show that to compute the Poisson bracket of any two functions given in some local coordinates $\mathbf{x} = x^1, ..., x^m$, it suffices to know the Poisson brackets of the coordinates. For any function $H : M \to \mathbb{R}$, let the components of its Hamiltonian vector field in the coordinate system $\mathbf{x}$ be written as $h^i(x)$. So $X_H = \Sigma_i^m\ h^i(x)\frac{\partial}{\partial x^i}$. Then for any other function F, we have

$$\{F, H\} = X_H(F) = \Sigma_i^m\ h^i(x)\frac{\partial F}{\partial x^i}. \tag{5.12}$$

Taking x^i as the function F, we get: $\{x^i, H\} = X_H(x^i) = h^i(x)$. So eq. 5.12 becomes

$$\{F, H\} = \Sigma_i^m\ \{x^i, H\}\frac{\partial F}{\partial x^i}. \tag{5.13}$$

If we now put x^i for H and H for F in eq. 5.13, we get

$$\{x^i, H\} = -\{H, x^i\} = -X_{x^i}(H) = -\Sigma_j^m\ \{x^j, x^i\}\frac{\partial H}{\partial x^j}.. \tag{5.14}$$

Combining eq.s 5.13 and 5.14, we get the basic formula for the Poisson bracket of any two functions in terms of the Poisson bracket of local coordinates:

$$\{F, H\} = \Sigma_i^m\ \Sigma_j^m\ \{x^i, x^j\}\frac{\partial F}{\partial x^i}\frac{\partial H}{\partial x^j}. \tag{5.15}$$

We assemble these basic brackets, which we call the *structure functions* of the Poisson manifold,

$$J^{ij}(x) := \{x^i, x^j\} \quad i, j = 1, ..., m \tag{5.16}$$

into a $m \times m$ anti-symmetric matrix of functions, $J(x)$, called the *structure matrix* of M. More precisely, it is the structure matrix for M relative to our coordinate system

$\mathbf{x}$. Of course, the transformation of J under a coordinate change $x'^i := x'^i(x^1, ..., x^m)$ is determined by setting $F := x'^i, H := x'^j$ in the basic formula eq. 5.15.

Then, writing ∇H for the (column) gradient vector of H, eq. 5.15 becomes

$$\{F, H\} = \nabla F \cdot J\nabla H. \tag{5.17}$$

For example, the canonical bracket on $\mathbb{R}^{2n+l}$, eq.5.5, written in the (q, p, z) coordinates, has the simple form

$$J = \begin{pmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.18}$$

where I is the $n \times n$ identity matrix.

We can write the Hamiltonian vector field, and the Hamilton's equations, associated with the function H in terms of J. Since

$$\{x^i, H\} = \Sigma_j^m \{x^i, x^j\} \frac{\partial H}{\partial x^j} \tag{5.19}$$

we get:

$$X_H = \Sigma_i^m \left(\Sigma_j^m J^{ij}(x) \frac{\partial H}{\partial x^j} \frac{\partial}{\partial x^i} \right), \tag{5.20}$$

or in matrix notation: $X_H = (J\nabla H) \cdot \partial_x$. Similarly, Hamilton's equations

$$\frac{dx^i}{dt} = \{x^i, H\} \tag{5.21}$$

get the matrix form

$$\frac{dx}{dt} = J(x)\nabla H(x) \ \ ; \quad \text{i.e.} \ \ \frac{dx^i}{dt} = \Sigma_j^m J^{ij}(x) \frac{\partial H}{\partial x^j} \ . \tag{5.22}$$

To summarize how we have generalized from the usual form of Hamilton's equations: compare eq. 5.22, 5.17 and 5.18 respectively with eq. 2.12, 2.18 and 2.3.

Note that not every $m \times m$ anti-symmetric matrix of functions on an m-dimensional manifold (or even: on an open subset of $\mathbb{R}^m$) is the structure matrix of a Poisson manifold: for the Jacobi identity constrains the functions. In fact it is readily shown that the Jacobi identity corresponds to the following m^3 partial differential equations governing the $J^{ij}(x)$, which are in general non-linear. Writing as usual ∂_l for $\partial/\partial x^l$:

$$\Sigma_{l=1}^m \left(J^{il}\partial_l J^{jk} + J^{kl}\partial_l J^{ij} + J^{jl}\partial_l J^{ki} \right) = 0 \quad i, j, k, = 1, ..., m; \forall x \in M. \tag{5.23}$$

In particular, any constant anti-symmetric matrix J defines a Poisson structure.

5.2.4 The Poisson structure on $\mathfrak{g}^*$

We can now show that any m-dimensional Lie algebra $\mathfrak{g}$ defines a Poisson structure, often called the *Lie-Poisson bracket*, on any m-dimensional vector space V. We proceed in two stages.

(1): We first present the definition in a way that seems to depend on a choice of bases, both in $\mathfrak{g}$ (where the definition makes a choice of structure constants) and in the space V.

(2): Then we will see that choosing V to be $\mathfrak{g}^*$, the definition is in fact basis-independent.

This Poisson structure on $\mathfrak{g}^*$ will be of central importance from now on. As Marsden and Ratiu write: 'Besides the Poisson structure on a symplectic manifold, the Lie-Poisson bracket on $\mathfrak{g}^*$, the dual of a Lie algebra, is perhaps the most fundamental example of a Poisson structure' (1999: 415). Here we return to our motivating discussion of Poisson manifolds, especially reasons (2) and (3) of Section 5.1: which concerned the rigid body and reduction, respectively. Indeed, we will see already in the Example at the end of this Subsection (Section 5.2.4.A) how the Lie-Poisson bracket on the special case $\mathfrak{g}^* := \mathfrak{so}(3)^*$ clarifies the theory of the rigid body. And we will see in Sections 7.2 and 7.3.3 how for any $\mathfrak{g}$, the Lie-Poisson bracket on $\mathfrak{g}^*$ is induced by reduction, from the canonical Poisson (viz. symplectic) structure on the cotangent bundle T^*G. This will be our reduction theorem, that $T^*G/G \cong \mathfrak{g}^*$.

After (2), we will see that the Lie-Poisson bracket on $\mathfrak{g}^*$ implies that Hamilton's equations on $\mathfrak{g}^*$ can be expressed using ad^*: a form that will be needed later. This will be (3) below. Then we will turn in Section 5.2.4.A to the example $\mathfrak{g}^* := \mathfrak{so}(3)^*$.

(1): *A Poisson bracket on any vector space V:—*
Take a basis, say $e_1, ..., e_m$, in $\mathfrak{g}$, and so structure constants c^k_{ij} (cf. eq. 3.24). Consider the space V as a manifold, and coordinatize it by taking a basis, $\epsilon_1, ..., \epsilon_m$ say, determining coordinates $x^1, ..., x^m$. We now define the Poisson bracket (in this case, often called the *Lie-Poisson bracket*) between two smooth functions $F, H : V \to \mathbb{R}$ to be

$$\{F, H\} := \Sigma^m_{i,j,k=1} \; c^k_{ij} x^k \frac{\partial F}{\partial x^i} \frac{\partial H}{\partial x^j}. \tag{5.24}$$

This takes the form of eq. 5.15, with linear structure functions $J^{ij}(x) = \Sigma^m_k \; c^k_{ij} x^k$. One easily checks that anti-symmetry, and the Jacobi identity, for the structure constants, eq. 3.25, implies that these J^{ij} are anti-symmetric and obey their Jacobi identity eq. 5.23. So eq. 5.24 defines a Poisson bracket on V.

In particular, the associated Hamiltonian equations, eq.s 5.21 and 5.22, take the form

$$\frac{dx^i}{dt} = \Sigma^m_{j,k=1} \; c^k_{ij} x^k \frac{\partial H}{\partial x^j}. \tag{5.25}$$

(2): *The Lie-Poisson bracket on $\mathfrak{g}^*$:—*
To give a basis-independent characterization of the Lie-Poisson bracket, we first recall

that:

(i): the gradient $\nabla F(x)$ of $F : V \rightarrow \mathrm{I\!R}$ at any point $x \in V$ is in the dual space V^* of (continuous) linear functionals on V;:

(ii): any finite-dimensional vector space is canonically, i.e. basis-independently, isomorphic to its double dual: $(V^*)^* \cong V$.

Then writing $< ; >$ for the natural pairing between V and V^*, we have, for any $y \in V$

$$< \nabla F(x); y >:= \lim_{\tau\rightarrow 0} \frac{F(x+\tau y) - F(x)}{\tau} . \tag{5.26}$$

Now let us take V in our definition of the Lie-Poisson bracket to be $\mathfrak{g}^*$. So we will show that $\mathfrak{g}$ makes $\mathfrak{g}^*$ a Poisson manifold, in a basis-independent way. And let the basis $\epsilon_1, ..., \epsilon_m$ be dual to the basis $e_1, ..., e_m$ of $\mathfrak{g}$. If $F : \mathfrak{g}^* \rightarrow \mathrm{I\!R}$ is any smooth function, its gradient $\nabla F(x)$ at any point $x \in \mathfrak{g}^*$ is an element of $(\mathfrak{g}^*)^* \cong \mathfrak{g}$. One now checks that the Lie-Poisson bracket defined by eq. 5.24 has the basis-independent expression

$$\{F, H\}(x) \ = < x; [\nabla F(x), \nabla H(x)] > \ , \ \ x \in \mathfrak{g}^* \tag{5.27}$$

where $[,]$ is the ordinary Lie bracket on the Lie algebra $\mathfrak{g}$ itself.

(3): *Hamilton's equations on* $\mathfrak{g}^*$:—
We can also give a basis-independent expression of the Hamilton's equations eq. 5.25: viz. by expressing the Lie bracket in eq. 5.27 in terms of *ad*, as indicated by eq. 4.64.

Thus let $F \in \mathcal{F}(\mathfrak{g}^*)$ be an arbitrary smooth scalar function on $\mathfrak{g}^*$. By the chain rule

$$\frac{dF}{dt} = \mathbf{D}F(x) \cdot \dot{x} =< \dot{x}; \nabla F(x) > \ . \tag{5.28}$$

But applying eq.s 4.64 and 4.83 to eq. 5.27 implies:

$$\{F, H\}(x) =< x; [\nabla F(x), \nabla H(x)] >= - < x; ad_{\nabla H(x)}(\nabla F(x)) >=< ad^*_{\nabla H(x)}(x); \nabla F(x) > \ . \tag{5.29}$$

Since F is arbitrary and the pairing is non-degenerate, we deduce that Hamilton's equations take the form

$$\frac{dx}{dt} = ad^*_{\nabla H(x)}(x) \ . \tag{5.30}$$

7.2.4.A Example: $\mathfrak{so}(3)$ and $\mathfrak{so}(3)^*$ As an example of the dual of a Lie algebra as a Poisson manifold, let us consider again our standard example $\mathfrak{so}(3)^*$. We will thereby make good our promise in (2) of Section 5.1, to show that Euler's equations for a rigid body are *already* in Hamiltonian form—in our generalized sense. We will also see why in the Chapter's second motto, Arnold mentions the three dual spaces, $\mathrm{I\!R}^{3*}, \mathfrak{so}(3)^*$ and $T^*(SO(3))_g$; (cf. the discussion at the end of Section 3.4.4).

The Lie algebra $\mathfrak{so}(3)$ of $SO(3)$ has a basis e_1, e_2, e_3 representing infinitesimal rotations around the x-, y- and z-axes of $\mathrm{I\!R}^3$. As we have seen, we can think of these basis elements: as vectors in $\mathrm{I\!R}^3$ with $[,]$ as elementary vector multiplication; or as anti-symmetric matrices with $[,]$ as the matrix commutator; or as left-invariant vector fields on $SO(3)$ with $[,]$ as the vector field commutator (i.e. Lie bracket).

Let $\epsilon_1, \epsilon_2, \epsilon_3$ be a dual basis for $\mathfrak{so}(3)^*$, with $x = x^1\epsilon_1 + x^2\epsilon_2 + x^3\epsilon_3$ a typical point therein. If $F : \mathfrak{so}(3)^* \to \mathbb{R}$, its gradient at x is the vector

$$\nabla F = \frac{\partial F}{\partial x^1} e_1 + \frac{\partial F}{\partial x^2} e_2 + \frac{\partial F}{\partial x^3} e_3 \quad \in \mathfrak{so}(3). \tag{5.31}$$

Then eq. 5.27 tells us that, if we write $\mathfrak{so}(3)$ as $\mathbb{R}^3$ with $\times$ for elementary vector multiplication, the Lie-Poisson bracket on $\mathfrak{so}(3)^*$ is

$$\{F, H\}(x) = x^1 \left(\frac{\partial F}{\partial x^3}\frac{\partial H}{\partial x^2} - \frac{\partial F}{\partial x^2}\frac{\partial H}{\partial x^3} \right) + ... + x^3 \left(\frac{\partial F}{\partial x^2}\frac{\partial H}{\partial x^1} - \frac{\partial F}{\partial x^1}\frac{\partial H}{\partial x^2} \right) \tag{5.32}$$

$$= -x \cdot (\nabla F \times \nabla H). \tag{5.33}$$

So the structure matrix $J(x)$ is

$$J(x) = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix}, \quad x \in \mathfrak{so}(3)^*. \tag{5.34}$$

Hamilton's equations corresponding to the Hamiltonian function $H(x)$ are therefore

$$\frac{dx}{dt} = x \times \nabla H(x)\ . \tag{5.35}$$

Now consider the Hamiltonian representing the kinetic energy of a free pivoted rigid body

$$H(x) = \frac{1}{2}\left(\frac{(x^1)^2}{I_1} + \frac{(x^2)^2}{I_2} + \frac{(x^3)^2}{I_2} \right), \tag{5.36}$$

in which the I_i are the moments of inertia about the three coordinate axes, and the x^i are the corresponding components of the *body* angular momentum. For this Hamiltonian, Hamilton's equations eq. 5.35 become

$$\frac{dx^1}{dt} = \frac{I_2 - I_3}{I_2 I_3} x^2 x^3 \ , \quad \frac{dx^2}{dt} = \frac{I_3 - I_1}{I_3 I_1} x^3 x^1 \ , \quad \frac{dx^3}{dt} = \frac{I_1 - I_2}{I_1 I_2} x^1 x^2 \ , \tag{5.37}$$

Indeed, these are the Euler equations for a free pivoted rigid body. I shall not go into details about the rigid body. I only note that:

(i): In the elementary theory of such a body, the magnitude L of the angular momentum is conserved, and eq. 5.37 describes the motion of the x^i on a sphere of radius L centred at the origin.

(ii): In Section 5.4, we will return to seeing these spheres as the orbits of the co-adjoint representation of $SO(3)$ on $\mathfrak{so}(3)^*$ (cf. Section 4.5.2).

(iii): Let us sum up this theme by saying, with Marsden and Ratiu (1999, p.11) that here we see: 'a simple and beautiful Hamiltonian structure for the rigid body equations'.

5.3 The symplectic foliation of Poisson manifolds

We first reformulate some ideas of Section 5.2 in more coordinate-independent language, starting with Section 5.2.3's idea of the structure matrix $J(x)$ (Section 5.3.1). Then we discuss canonical transformations on a Poisson manifold (Section 5.3.2). This will lead up to showing that any Poisson manifold is foliated by symplectic leaves (Section 5.3.3). Finally, we state a generalization of Darboux's theorem; and again take $\mathfrak{so}(3)$ as an example (Section 5.3.4).

5.3.1 The Poisson structure and its rank

We now pass from the structure matrix J, eq. 5.16, to a coordinate-independent object, the *Poisson structure* (also known as: co-symplectic structure), written B. Whereas J multiplied naive gradient vectors, as in eq. 5.17 and 5.22, B is to map the 1-form dH into its Hamiltonian vector field; as follows.

At each point x in a Poisson manifold M, there is a unique linear map B_x, which we will also write as B

$$\mathsf{B} \equiv \mathsf{B}_x : T^*_x M \rightarrow T_x M \tag{5.38}$$

such that

$$\mathsf{B}_x(dH(x)) = X_H(x). \tag{5.39}$$

For the requirement eq. 5.39 implies, by eq. 5.20, that for each $j = 1, ..., m$

$$\mathsf{B}_x(dx^j) = \Sigma_i J^{ij}(x) \frac{\partial}{\partial x^i} \mid_x \tag{5.40}$$

Since the differentials dx^i span $T^*_x M$, this fixes B_x, by linearity. B_x's action on any one-form $\alpha = \Sigma a_j dx^j$ is:

$$\mathsf{B}_x(\alpha) = \Sigma_{i,j} J^{ij}(x) a_j \frac{\partial}{\partial x^i} \mid_x \tag{5.41}$$

so that B_x is essentially matrix multiplication by $J(x)$. Here, compare again eq. 5.21 and 5.22.

Here we recall that any linear map between (real finite-dimensional) vector spaces, $B : V \rightarrow W^*$, has an associated bilinear form $B^\sharp$ on $V \times W^{**} \cong V \times W$ given by

$$B^\sharp(v, w) \ := \ < B(v) \ ; \ w > \ . \tag{5.42}$$

Accordingly, some authors introduce the Poisson structure as a bilinear form $\mathsf{B}^\sharp_x : T^*_x M \times T^*_x M \rightarrow \mathrm{I\!R}$, often called the *Poisson tensor*. Thus eq. 5.42 gives, for $\alpha, \beta \in T^*_x M$

$$\mathsf{B}^\sharp_x(\alpha, \beta) \ := < \mathsf{B}(\alpha), \beta > \ . \tag{5.43}$$

$\mathsf{B}^\sharp_x$ is antisymmetric, since the matrix $J(x)$ is. So, if we now let x vary over M, we can sum up in the traditional terminology of tensor analysis: $\mathsf{B}^\sharp$ is an antisymmetric contravariant two-tensor field.

Example:— Consider our first example, $M = \mathbb{R}^{2n+l}$ with the "usual bracket" eq. 5.5, from the start of Section 5.2.1. For any one-form

$$\alpha = \Sigma_{i=1}^{n}(a_i dq^i + b_i dp^i) + \Sigma_{j=1}^{l} c_j dz^j \tag{5.44}$$

we have

$$\mathsf{B}(\alpha) = \Sigma_{i=1}^{n}\left(b_i \frac{\partial}{\partial q^i} - a_i \frac{\partial}{\partial p^i}\right) . \tag{5.45}$$

In this example the form of B is the same from point to point. In particular, the kernel of B has everywhere the same dimension, viz. l, the number of distinguished coordinates.

We now define the *rank* at x of a Poisson manifold M to be the rank of its Poisson structure B at x, i.e. the dimension of the range of B_x. This range is also the span of all the Hamiltonian vector fields on M at x:

$$\text{ran}(\mathsf{B}_x) := \{X \in T_x M : \; X = \mathsf{B}_x(\alpha), \text{ some } \; \alpha \in T_x^* M\} \; = \; \{X_H(x) : \; H : M \to \mathbb{R} \text{ smooth } \} . \tag{5.46}$$

So the rank of M at x is also equal to the dimension of B_x's domain, i.e. $\dim(T_x^* M) = \dim(M)$, minus the dimension of the kernel, $\dim(\mathsf{B}_x)$.

Since in local coordinates, B_x is given by multiplication by the structure matrix $J(x)$, the rank of M at x is the rank (the same in any coordinates) of the matrix $J(x)$. That $J(x)$ is anti-symmetric implies that the rank of M is *even*: cf. again the normal form of antisymmetric bilinear forms, eq. 2.2 and 2.3.

The manifold M being symplectic corresponds, of course, to the rank of B being everywhere maximal, i.e. equal to $\dim(M)$.

In this case, the kernel of B is trivial, and any distinguished function H is constant on M. For H is distinguished iff $X_H = 0$; and if the rank is maximal, then $dH = 0$, so that H is constant.

Besides, each of the Poisson structure and symplectic form on M determine the other. In particular, the Poisson tensor $\mathsf{B}^\sharp$ of eq. 5.43 is, up to a sign, the "contravariant cousin" of M's symplectic form ω. For recall: (i) the relation between a symplectic manifold's Poisson bracket and its form, eq. 2.18, viz.

$$\{F, H\} = dF(X_H) = \omega(X_F, X_H) \; ; \tag{5.47}$$

and (ii) eq. 5.8 for Hamiltonian vector fields on a Poisson manifold, viz.

$$X_H(F) = \{F, H\} \; . \tag{5.48}$$

Applying these equations yields, if we start from eq. 5.43 and eq. 5.39:

$$\mathsf{B}^\sharp(dH, dF) \; := < B(dH), dF >= dF(X_H) = X_H(F) = \{F, H\} = \omega(X_F, X_H) \; . \tag{5.49}$$

We have also seen examples where the Poisson structure B is of non-maximal rank:

(i): In our opening "canonical" example, the Poisson bracket eq. 5.5 on $M =$

$\mathbb{R}^{2n+l}$ has rank $2n$ everywhere.

(ii): In the Lie-Poisson structure on $\mathfrak{so}(3)^*$, the rank varies across the manifold: it is 2 everywhere, except at the origin $x = 0$ where it is 0. (Cf. the rank of the matrix J in eq. 5.34.)

5.3.2 Poisson maps

Already at the beginning of our development of Poisson manifolds, we saw that a scalar function $H : M \to \mathbb{R}$ defines equations of motion, with H as "Hamiltonian", for all other functions $F : M \to \mathbb{R}$, of the familiar Poisson bracket type:

$$\dot{F} = \{F, H\}. \tag{5.50}$$

(Cf. Section 5.2.2, especially the remarks around eq. 5.8.) We now develop the generalization for Poisson manifolds of some related notions and results.

We say that a smooth map $f : M_1 \to M_2$ between Poisson manifolds $(M_1, \{,\}_1)$ and $(M_2, \{,\}_2)$ is *Poisson* or *canonical* iff it preserves the Poisson bracket. To be precise: we first need the idea of the *pullback* of a function; cf. Section 3.1.2.A. In this context, the pullback f^* of a function $F : M_2 \to \mathbb{R}$ is given by

$$f^*F := F \circ f; \quad \text{i.e. } f^*F : x \in M_1 \mapsto F(f(x)) \in \mathbb{R}. \tag{5.51}$$

Then we say that $f : M_1 \to M_2$ is *Poisson* iff for all smooth functions $F, G : M_2 \to \mathbb{R}$ $(F, G \in \mathcal{F}(M_2))$

$$f^*\{F, G\}_2 = \{f^*F, f^*G\}_1 \ ; \tag{5.52}$$

where by the definition eq. 5.51, the lhs $\equiv \{F, G\}_2 \circ f$, and the rhs $\equiv \{F \circ f, G \circ g\}_1$.

We note the special case where $M_1 = M_2 =: M$ and M is symplectic; i.e. the Poisson bracket is of maximal rank, and so defines a symplectic form on M, as in eq. 5.49. In this case, we return to the equivalence in Section **??**'s usual formulation of Hamiltonian mechanics, between preserving the Poisson bracket and preserving the symplectic form. That is: a map $f : M \to M$ on a symplectic manifold M is Poisson iff it is symplectic.

Besides, we already have for symplectic manifolds an infinitesimal version of the idea of a Poisson or symplectic map: viz. the idea of a locally Hamiltonian vector field; cf. Section 2.1.3. Similarly for Poisson manifolds, we will need the corresponding infinitesimal version of a Poisson map; but not till Section 6.1.1.

One can show (using in particular the Jacobi identity) that the flows of a Hamiltonian vector field are Poisson. (Here of course, $(M_1, \{,\}_1) = (M_2, \{,\}_2)$.) That is: if ϕ_τ is the flow of X_H (i.e. $\phi_\tau = \exp(\tau X_H)$), then

$$\phi_\tau^*\{F, G\} = \{\phi_\tau^*F, \phi_\tau^*G\} \quad \text{i.e. } \{F, G\} \circ \phi_\tau = \{F \circ \phi_\tau, G \circ \phi_\tau\} \ . \tag{5.53}$$

Similarly, one can readily show the equivalent proposition, that along the flow of a Hamiltonian vector field the Lie derivative of the Poisson tensor $\mathsf{B}^\sharp$ vanishes. That is:

for any smooth function $H : M \to \mathbb{R}$, we have:

$$\mathcal{L}_{X_H} \mathsf{B}^\sharp = 0 \; . \tag{5.54}$$

Since preserving the Poisson bracket implies in particular preserving its rank, it follows from eq. 5.53 (or from eq. 5.54) that:

If X_H is a Hamiltonian vector field on a Poisson manifold M, then for any $\tau \in \mathbb{R}$ and $x \in M$, the rank of M at $\exp(\tau X_H)(x)$ is the same as the rank at x. In other words: Hamiltonian vector fields are rank-invariant in the sense used in the general form of Frobenius' theorem (Section 3.3.2).

This result will be important for the foliation theorem for Poisson manifolds.

We will also need the result (also readily shown) that Poisson maps push Hamiltonian flows forward to Hamiltonian flows. More precisely: let $f : M_1 \to M_2$ be a Poisson map; so that at each $x \in M_1$, we have the derivative map on the tangent space, $Tf : (TM_1)_x \to (TM_2)_{f(x)}$. And let $H : M_2 \to \mathbb{R}$ be a smooth function. If ϕ_τ is the flow of X_H and ψ_τ is the flow (on M_1) of $X_{H \circ f}$, then:

$$\phi_\tau \circ f = f \circ \psi_\tau \;\text{ and }\; Tf \circ X_{H\circ f} = X_H \circ f \;\; . \tag{5.55}$$

In particular, this square commutes:

$$\begin{array}{ccc} M_1 & \xrightarrow{f} & M_2 \\ \uparrow{\scriptstyle \psi_\tau} & & \uparrow{\scriptstyle \phi_\tau} \\ M_1 & \xrightarrow{f} & M_2 \end{array} \tag{5.56}$$

5.3.3 Poisson submanifolds: the foliation theorem

To state the foliation theorem for Poisson manifolds, we need the idea of a *Poisson immersion*, which leads to the closely related idea of a *Poisson submanifold.* In effect, these ideas combine the idea of a Poisson map with the ideas about injective immersions in (2) of Section 3.3.1. We recall from that discussion that for an injective immersion, $f : N \to M$, the range $f(N)$ is not necessarily a submanifold of M: but $f(N)$ is nevertheless called an 'injectively immersed submanifold' of M. (But as mentioned in Section 3.3.2, many treatments ignore this point: they in effect assume that an injective immersion f is also an embedding, i.e. a homeomorphism between N and $f(N)$, so that $f(N)$ is indeed a submanifold of M and f is a diffeomorphism.)

An injective immersion $f : N \to M$, with M a Poisson manifold, is called a *Poisson immersion* if any Hamiltonian vector field defined on an open subset of M containing $f(N)$ is in the range of the derivative map of f at $y \in N$, i.e. $\mathrm{ran}(T_y f)$, at all points $f(y)$ for $y \in N$.

Being a Poisson immersion is equivalent to the following rather technical condition.

Characterization of Poisson immersions An injective immersion $f : N \to M$, with M a Poisson manifold, is a Poisson immersion iff:

if $F, G : V \subset N \to \mathrm{I\!R}$, where V is open in N, and if $\bar{F}, \bar{G} : U \to \mathrm{I\!R}$ are extensions of $F \circ f^{-1}, G \circ f^{-1} : f(V) \to \mathrm{I\!R}$ to an open neighbourhood U of $f(V)$ in M, then $\{\bar{F}, \bar{G}\} \mid_{f(V)}$ is well-defined and independent of the extensions.

The main point of this equivalence is that it ensures that if $f : N \to M$ is a Poisson immersion, then N has a Poisson structure, and $f : N \to M$ is a Poisson map. It is worth seeing how this comes about—by proving the equivalence.

Proof: Let $f : N \to M$ be a Poisson immersion, and let $F, G : V \subset N \to \mathrm{I\!R}$ and let $\bar{F}, \bar{G} : U \supset f(V) \to \mathrm{I\!R}$ be extensions of $F \circ f^{-1}, G \circ f^{-1} : f(V) \to \mathrm{I\!R}$. Then for $y \in V$, there is a unique vector $v \in TN_y$ such that

$$X_{\bar{G}}(f(y)) = (T_y f)(v) \ . \tag{5.57}$$

So evaluating the Poisson bracket of $\bar{F}$ and $\bar{G}$ at $f(y)$ yields, by eq. 5.8,

$$\{\bar{F}, \bar{G}\}(f(y)) = d\bar{F}(f(y)) \cdot X_{\bar{G}}(f(y)) = d\bar{F}(f(y)) \cdot (T_y f)(v) = d(\bar{F} \circ f)(y) \cdot v \equiv dF(y) \cdot v \ . \tag{5.58}$$

So $\{\bar{F}, \bar{G}\}(f(y))$ is independent of the extension $\bar{F}$ of $F \circ f^{-1}$. Since the Poisson bracket is antisymmetric, it is also independent of the extension $\bar{G}$ of $G \circ f^{-1}$. So we can define a Poisson structure on N by defining for any y in an open $V \subset N$

$$\{F, G\}_N(y) := \{\bar{F}, \bar{G}\}_M(f(y)) \ . \tag{5.59}$$

This makes $f : N \to M$ a Poisson map, since for any $\bar{F}, \bar{G}$ on M and any $y \in N$, we have that

$$[f^*\{\bar{F}, \bar{G}\}_M](y) \equiv [\{\bar{F}, \bar{G}\}_M \circ f](y) = \{F, G\}_N(y) \equiv \{f^*\bar{F}, f^*\bar{G}\}_N(y) \ ; \tag{5.60}$$

where the middle equality uses eq. 5.59.

For the converse implication, assume that eq. 5.58 holds, and let $H : U \to \mathrm{I\!R}$ be a Hamiltonian defined on an open subset U of M that intersects $f(N)$. Then as we have just seen, N is a Poisson manifold and $f : N \to M$ is a Poisson map. Because f is Poisson, it pushes $X_{H \circ f}$ to X_H. That is: eq. 5.55 implies that if $y \in N$ is such that $f(y) \in U$, then

$$X_H(f(y)) = (T_y f)(X_{H \circ f}(y)) \ . \tag{5.61}$$

So $X_H(f(y))$ is in the range of $T_y f$; so $f : N \to M$ is a Poisson immersion. QED.

Now suppose that the inclusion $id : N \to M$ is a Poisson immersion. Then we call N a *Poisson submanifold* of M. We emphasise, in line with the warning we recalled from (2) of Section 3.3.1, that N need not be a submanifold of M; but it is nevertheless called an 'injectively immersed submanifold' of M.

From the definition of a Poisson immersion, it follows that any Hamiltonian vector field must be tangent to a Poisson submanifold. In other words: writing $\mathcal{X}$ for the system of Hamiltonian vector fields on M, and $\mathcal{X} \mid_x$ for their values at $x \in M$, we have: if N is a Poisson submanifold of M, and $x \in N$, $\mathcal{X} \mid_x \subset TN_x$.

For the special case where M is a symplectic manifold, we have $\mathcal{X} \mid_x = T_x M$, and the only Poisson submanifolds of M are its open sets.

Finally, we define the following equivalence relation on a Poisson manifold M. Two points $x_1, x_2 \in M$ are *on the same symplectic leaf* if there is a piecewise smooth curve in M joining them, each segment of which is an integral curve of a locally defined Hamiltonian vector field. An equivalence class of this equivalence relation is a *symplectic leaf*.

We can now state and prove that Poisson manifolds are foliated.

7.3.3.A Foliation theorem for Poisson manifolds The result is:—

> A Poisson manifold M is the disjoint union of its symplectic leaves. Each symplectic leaf is an injectively immersed Poisson submanifold, and the induced Poisson structure on the leaf is symplectic. The leaf through the point x, N_x say, has dimension equal to the rank of the Poisson structure at x; and the tangent space to the leaf at x equals
>
> $$TN_x = \mathrm{ran}(\mathsf{B}_x) := \{X \in T_x M \ : \ X = \mathsf{B}_x(\alpha), \text{ some } \alpha \in T_x^* M\} \quad (5.62)$$
> $$= \{X_H(x) : H \in \mathcal{F}(U), \ U \text{ open in } M \ \} \quad (5.63)$$

Proof: We apply the general form of Frobenius' theorem (Section 3.3.2) to the system $\mathcal{X}$ of Hamiltonian vector fields on M. We know from eq. 5.11 (Section 5.2.2) that $\mathcal{X}$ is involutive, and from eq. 5.53 above that it is rank-invariant. So by Frobenius' theorem, $\mathcal{X}$ is integrable. The integral submanifolds are by definition given by the rhs of eq. 5.63. QED.

One also readily shows that:

(i): One can evaluate the Poisson bracket of $F, G : M \to \mathbb{R}$ at $x \in M$ by restricting F and G to the symplectic leaf N_x through x, and evaluating the Poisson bracket that is defined by the symplectic form on the leaf N_x; (i.e. the Poisson bracket defined in eq. 2.18).

(ii): A distinguished function is constant on any symplectic leaf N_x of M.

We end with two remarks. The first is a mathematical warning; the second concerns physical interpretation.

(1): Recall our warning that symplectic leaves need not be submanifolds. This also means that all the distinguished functions being constants does *not* imply that the Poisson structure is non-degenerate. Indeed, one can readily construct an example in which the symplectic leaves are not manifolds, all distinguished functions are constants, and the Poisson structure is degenerate. Namely, one adapts an example mentioned before, in Section 3.4.3: the flows on the torus T^2 that wind densely around it. (For more details about this example, cf. Arnold (1973: 160-167) or Arnold (1989: 72-74) or Butterfield (2004a: Section 2.1.3.B); for how to adapt it, cf. Marsden and Ratiu

(1999: 347).

(2): As we have seen, any integral curve of any Hamiltonian vector field X_H is confined to one of the symplectic leaves. So if we are interested only in the behaviour of a single solution through a point $x \in M$, we can restrict our attention to the symplectic leaf N_x through x: for the solution will always remain in N_x. But as stressed in Section 5.1, there are at least three good reasons not to ignore the more general Poisson structure!

5.3.4 Darboux's theorem

At the end of Section 2.1.1, we mentioned Darboux's theorem: it said that any symplectic manifold "looks locally like" a cotangent bundle. The generalization for Poisson manifolds says that any Poisson manifold "looks locally like" our canonical example on $\mathbb{R}^m, m = 2n + l$, given at the start of Section 5.2.1. More precisely, we have:

> Let M be an m-dimensional Poisson manifold, and let $x \in M$ be a point with an open neighbourhood $U \subset M$ throughout which the rank is a constant $2n \leq m$. Then defining $l := m - 2n$, there is a possibly smaller neighbourhood $U' \subset U$ of x, on which there exist local coordinates $(q, p, z) = (q^1, ..., q^n, p^1, ..., p^n, z^1, ..., z^l)$, for which the Poisson bracket takes the form
>
> $$\{F, H\} := \Sigma_i^n \left(\frac{\partial F}{\partial q^i} \frac{\partial H}{\partial p^i} - \frac{\partial F}{\partial p^i} \frac{\partial H}{\partial q^i} \right) . \tag{5.64}$$
>
> (So the Poisson brackets for the coordinate functions take the now-familiar form given by eq. 5.6 and 5.7.) The symplectic leaves of M intersect the coordinate chart in the slices $\{z^1 = c_1, \dots, z^l = c_l\}$ given by constant values of the distinguished coordinates z.

We shall not give the proof. Suffice it to say that:

(i): Like Darboux's theorem for symplectic manifolds: it proceeds by induction on the "half-rank" n; and it begins by taking any function F as the "momentum" p^1 and constructing the canonically conjugate coordinate q^1 such that $\{q^1, p^1\} = 1$.

(ii): The induction step invokes a version of Frobenius' theorem in which the fact that the rank $2n$ is constant throughout U secures a coordinate system in which the $2n$-dimensional integral manifolds are given by slices defined by constant values of the remaining l coordinates. The Poisson structure then secures that these remaining coordinates are distinguished.

7.3.4.A Example: $\mathfrak{so}(3)^*$ yet again We illustrate (1) the foliation theorem and (2) Darboux's theorem, with $\mathfrak{so}(3)^*$; whose Lie-Poisson structure we described in Section 5.2.4.A.

(1): At $x \in \mathfrak{so}(3)^*$, the subspace $\mathcal{X} \mid_x := \{X_H(x) : H \in \mathcal{F}(U), \;\; U \text{ open in } \; M \; \}$ of values of locally Hamiltonian vector fields is spanned by $e_1 := y\partial_z - z\partial_y$ representing infinitesimal rotation about the x-axis (cf. eq. **??**); $e_2 := z\partial_x - x\partial_z$ for rotation about the y-axis; and $e_3 := x\partial_y - y\partial_x$ for rotation about the z-axis. If $x \neq 0$, these vectors span a two-dimensional subspace of $T\mathfrak{so}(3)^*_x$: viz. the tangent plane to the sphere $S_{|x|}$ of radius $\mid x \mid$ centred at the origin. So the foliation theorem implies that $\mathfrak{so}(3)^*$'s symplectic leaves are these spheres; and the origin.

We can compute the Poisson bracket of $F, G : S_{|x|} \rightarrow \mathrm{I\!R}$ by extending F and G to a neighbourhood of $S_{|x|}$; cf. eq. 5.59. That is: we can consider extensions $\bar{F}, \bar{G} : U \supset S_{|x|} \rightarrow \mathrm{I\!R}$, and calculate the Poisson bracket in $\mathfrak{so}(3)^*$, whose Poisson structure we already computed in eq. 5.33.

Adopting spherical polar coordinates with $r =\mid x \mid$, i.e. $x^1 = r\cos\theta\sin\phi, x^2 = r\sin\theta\sin\phi, x^3 = r\cos\phi$, we can define $\bar{F}, \bar{G}$ merely by $\bar{F}(r,\theta,\phi) := F(\theta,\phi), \bar{G}(r,\theta,\phi) := G(\theta,\phi)$; so that the partial derivatives with respect to the spherical angles θ, ϕ are equal, i.e. $\bar{F}_\theta = F_\theta, \bar{F}_\phi = F_\phi, \bar{G}_\theta = G_\theta, \bar{G}_\phi = G_\phi$.

Besides, eq. 5.15 implies that we need only calculate the Poisson bracket in $\mathfrak{so}(3)^*$ of the spherical angles θ and ϕ. So eq. 5.33 gives

$$\{\theta, \phi\} = -x \cdot (\nabla\theta \times \nabla\phi) = \frac{-1}{r\sin\phi} \; ; \tag{5.65}$$

and eq. 5.59 and 5.15 give

$$\{F, G\} = \{\bar{F}, \bar{G}\} = \frac{-1}{r\sin\phi}(F_\theta G_\phi - F_\phi G_\theta) \; . \tag{5.66}$$

(2): $z := x^3$ defines the Hamiltonian vector field $X_z = x^2\partial_{x^1} - x^1\partial_{x^2}$ that generates clockwise rotation about the $z \equiv x^3$-axis. So away from the origin the polar angle $\theta := \arctan(x^2/x^1)$ has a Poisson bracket with z equal to: $\{\theta, z\} = X_z(\theta) = -1$. Exprssing $F, H : \mathfrak{so}(3)^* \rightarrow \mathrm{I\!R}$ in terms of the coordinates z, θ and $r :=\mid x \mid$, we find that the Lie-Poisson bracket is: $\{F, H\} = F_z H_\theta - F_\theta H_z$. So (z, θ, r) are canonical coordinates.

5.4 The symplectic structure of the co-adjoint representation

Section 5.2.4 described how the dual $\mathfrak{g}^*$ of a finite-dimensional Lie algebra of a Lie group G has the structure of a Poisson manifold. In this case, the foliation established in the previous Subsection has an especially neat interpretation. Namely: the leaves are the orbits of the co-adjoint representation of G on $\mathfrak{g}^*$.

This symplectic structure in the co-adjoint representation sums up themes from Sections 4.5 (especially 4.5.2), and 5.2.4 and 5.3. In particular, it connects two properties of the Lie bracket in $\mathfrak{g}$, which we have already seen: viz.

(i): The Lie bracket in $\mathfrak{g}$ gives the infinitesimal generators of the adjoint action; cf. eq. 4.64.

(ii): The Lie bracket in $\mathfrak{g}$ defines (in a basis-independent way) a Lie-Poisson bracket on $\mathfrak{g}^*$, thus making $\mathfrak{g}^*$ a Poisson manifold. (Cf. the definition in eq. 5.24, shown to be basis-independent by eq. 5.27.)

In fact, there is a wealth of instructive results and examples about the structure of the co-adjoint representation: we will only scratch the surface—as in other Sections! We will give a proof, under a simplifying assumption, of one main result; and then make a few remarks about other results.

The result is:

The orbits of the co-adjoint representation are $\mathfrak{g}^*$'s leaves
Let G be a Lie group, with its co-adjoint representation Ad^* on $\mathfrak{g}^*$. That is, recalling eq. 4.78, we have:

$$Ad^* : G \rightarrow \mathrm{End}(\mathfrak{g}^*), \quad Ad^*_{g^{-1}} = (T_e(R_g \circ L_{g^{-1}}))^* \ . \tag{5.67}$$

The orbits of this representation are the symplectic leaves of $\mathfrak{g}^*$, taken as equipped with its natural Poisson structure, i.e. the Lie-Poisson bracket eq. 5.27.

Proof:— We shall prove this under the simplifying assumption that the co-adjoint action of G on $\mathfrak{g}^*$ is proper. (We recall from the definition of proper actions, eq. 4.32, that for any compact Lie group, such as $SO(3)$, this condition is automatically satisfied.) Then we know from result (3) and eq. 4.45, at the end of Section 4.4, that this implies that the co-adjoint orbit $\mathrm{Orb}(\alpha)$ of any $\alpha \in \mathfrak{g}^*$ is a closed submanifold of $\mathfrak{g}^*$, and that the tangent space to $\mathrm{Orb}(\alpha)$ at a point $\beta \in \mathrm{Orb}(\alpha)$ is

$$T\mathrm{Orb}(\alpha)_\beta \ = \ \{ \ \xi_{\mathfrak{g}^*}(\beta) \ : \ \xi \in \mathfrak{g} \ \} \ . \tag{5.68}$$

We will see shortly how this assumption implies that $\mathfrak{g}^*$'s symplectic leaves are submanifolds.[22]

We now argue as follows. For $\xi \in \mathfrak{g}$, consider the scalar function on $\mathfrak{g}^*$, $K_\xi : \alpha \in \mathfrak{g}^* \mapsto K_\xi(\alpha) \ := < \alpha; \xi > \in \mathbb{R}$; and its Hamiltonian vector field X_{K_ξ}. At each $\alpha \in \mathfrak{g}^*$, the gradient $\nabla K_\xi(\alpha) \equiv dK_\xi(\alpha)$, considered as an element of $(T^*\mathfrak{g}^*)_\alpha \cong \mathfrak{g}$, is just ξ itself. Now we will compute $X_{K_\xi}(F)(\alpha)$ for any $F : \mathfrak{g}^* \rightarrow \mathbb{R}$ and any $\alpha \in \mathfrak{g}^*$, using in order:

(i): the intrinsic definition of the Lie-Poisson bracket on $\mathfrak{g}^*$, eq. 5.27;

(ii): the fact that the infinitesimal generator of the adjoint action is the Lie bracket in $\mathfrak{g}$, eq. 4.64;

(iii): the fact that the derivative ad^* of the co-adjoint action Ad^* is, up to a sign, the adjoint of ad_ξ; eq. 4.83.

Thus we get, for all $F : \mathfrak{g}^* \rightarrow \mathbb{R}$ and $\alpha \in \mathfrak{g}^*$:

$$X_{K_\xi}(F)(\alpha) \ \equiv \ \{F, K_\xi\}(\alpha) \ = \ < \ \alpha \ ; \ [\nabla F(\alpha), \nabla K_\xi(\alpha)] \ > \tag{5.69}$$

$$= \ < \ \alpha \ ; \ [\nabla F(\alpha), \xi] \ > \ = \ - \ < \ \alpha \ ; \ [\xi, \nabla F(\alpha)] \ > \tag{5.70}$$

$$= \ - \ < \ \alpha \ ; \ ad_\xi(\nabla F(\alpha)) \ > \tag{5.71}$$

$$= \ < \ ad^*_\xi(\alpha) \ ; \ \nabla F(\alpha) \ > . \tag{5.72}$$

[22] To verify that our condition is indeed simplifying—i.e. that in general the co-adjoint orbits in $\mathfrak{g}^*$ are not submanifolds—consider the example in Marsden and Ratiu (1999: 14.1.(f), p. 449); taken from Kirillov (1976: 293).

But on the other hand, the vector field X_{K_ξ} is uniquely determined by its action on all such functions F at all $\alpha \in \mathfrak{g}^*$:

$$X_{K_\xi}(F)(\alpha) \ \equiv \ < \ X_{K_\xi}(\alpha) \, ; \nabla F(\alpha) \ > \ . \tag{5.73}$$

So we conclude that at each $\alpha \in \mathfrak{g}^*$:

$$X_{K_\xi} = ad^*_\xi \ . \tag{5.74}$$

But the subspace $\mathcal{X} \mid_\alpha$ of values at α of Hamiltonian vector fields is spanned by the $X_{K_\xi}(\alpha)$, with ξ varying through $\mathfrak{g}$. And as ξ varies through $\mathfrak{g}$, $ad^*_\xi(\alpha)$ is the tangent space $T\mathrm{Orb}(\alpha)_\alpha$ to the co-adjoint orbit $\mathrm{Orb}(\alpha)$ of G through α. So

$$\mathcal{X} \mid_\alpha = T\mathrm{Orb}(\alpha)_\alpha \ . \tag{5.75}$$

So the integral submanifolds of the system $\mathcal{X}$ of Hamiltonian vector fields, which are the symplectic leaves of $\mathfrak{g}^*$ by Section 5.3.3.A's foliation theorem, are the co-adjoint orbits. QED.

For the illustration of this theorem by our standard example, $\mathfrak{so}(3)^*$, cf. our previous discussions of it: in Section 4.5.2 for its co-adjoint structure; in Section 5.2.4.A for its Lie-Poisson structure; and in Section 5.3.4.A for its symplectic leaf structure.

We end this Subsection by stating two other results. They are not needed later, but they are enticing hints of how rich is the theory of co-adjoint orbits.

(1): For each $g \in G$, the co-adjoint map $Ad^*_g : \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ is a Poisson map that preserves the symplectic leaves of $\mathfrak{g}^*$.

(2): A close cousin of the theorem just proven is that the Lie bracket on $\mathfrak{g}$ defines (via its definition of the Lie-Poisson bracket on $\mathfrak{g}^*$, eq. 5.27) a symplectic form, i.e. a non-degenerate closed two-form, on each co-adjoint orbit, by:

$$\omega(\alpha)(ad^*_\xi(\alpha), ad^*_\eta(\alpha)) \ := \ < \ \alpha \, ; [\xi, \eta]_{\mathfrak{g}} \ > \ , \ \ \forall \alpha \in \mathfrak{g}^*, \ \forall \xi, \eta \in \mathfrak{g} \ . \tag{5.76}$$

This theorem is proven in detail (without our simplifying assumption that G's action is proper) by Marsden and Ratiu (1999: Thm 14.3.1, pp. 453-456); and much more briefly by Arnold (1989: 321, 376-377, 457); and rather differently (even without using the notion of a Poisson manifold!) in Abraham and Marsden (1978: 302-303).

5.5 Quotients of Poisson manifolds

We now end Section 5 with the simplest general theorem about quotienting a Lie group action on a Poisson manifold, so as to get a quotient space (set of orbits) that is itself a Poisson manifold. So this theorem combines themes from Sections 4—in particular, the idea from Section 4.3.B that for a free and proper group action, the orbits and quotient space are manifolds—with material about Poisson manifolds from Section 5.2. (The material in Sections 5.3 and 5.4 will not be needed.) This theorem will be important in Section 7. We call this result the

Poisson reduction theorem: Suppose the Lie group G acts on Poisson manifold M is such a way that each $\Phi_g : M \to M$ is a Poisson map. Suppose also that the quotient space M/G is a manifold and the projection $\pi : M \to M/G$ is a smooth submersion (say because G's action on M is free and proper, cf. Section 4.3.B). Then there is a unique Poisson structure on M/G such that π is a Poisson map. The Poisson bracket on M/G is called the *reduced* Poisson bracket.

Proof: Let us first assume that M/G is a Poisson manifold and that π is a Poisson map; and show uniqueness. We first note that for any $f : M/G \to \mathbb{R}$, the function $\bar{f} := f \circ \pi : M \to \mathbb{R}$ is obviously the unique G-invariant function on M that projects by π to f. That is: if $[x] \equiv \mathrm{Orb}(x) \equiv G \cdot x$ is the orbit of $x \in M$, then $\bar{f}$ assigns the same value $f([x])$ to all elements of the orbit $[x]$. Besides, in terms of pullbacks (eq. 5.51), $\bar{f} = \pi^* f$.

Then the condition that π be Poisson, eq. 5.52, is that for any two smooth scalars $f, h : M/G \to \mathbb{R}$, we have an equation of smooth scalars on M:

$$\{f, h\}_{M/G} \circ \pi = \{f \circ \pi, h \circ \pi\}_M = \{\bar{f}, \bar{h}\}_M \tag{5.77}$$

where the subscripts indicate on which space the Poisson bracket is defined. Since π is surjective, eq. 5.77 determines the value $\{f, h\}_{M/G}$ uniquely.

But eq. 5.77 also *defines* $\{f, h\}_{M/G}$ as a Poisson bracket; in two stages. (1): The facts that Φ_g is Poisson, and $\bar{f}$ and $\bar{h}$ are constant on orbits imply that

$$\{\bar{f}, \bar{h}\}(g \cdot x) \;=\; (\{\bar{f}, \bar{h}\} \circ \Phi_g)(x) \;=\; \{\bar{f} \circ \Phi_g, \bar{h} \circ \Phi_g\}(x) \;=\; \{\bar{f}, \bar{h}\}(x). \tag{5.78}$$

That is: $\{\bar{f}, \bar{h}\}$ is also constant on orbits, and so defines $\{f, h\}$ uniquely.

(2): We show that $\{f, h\}$, as thus defined, is a Poisson structure on M/G, by checking that the required properties, such as the Jacobi identity, follow from the Poisson structure $\{\ ,\ \}_M$ on M. QED.

This theorem is a "prototype" for material to come. We spell this out in two brief remarks, which look forward to the following two Sections.

(1): *Other theorems*:— This theorem is one of many that yield new Poisson manifolds and symplectic manifolds from old ones by quotienting. In particular, as we will see in detail in Section 7, this theorem is exemplified by the case where $M = T^*G$ (so here M is symplectic, since it is a cotangent bundle), and G acts on itself by left translations, and so acts on T^*G by a cotangent lift. In this case, we will have $M/G \cong \mathfrak{g}^*$; and the reduced Poisson bracket just defined, by eq. 5.77, will be the Lie-Poisson bracket we have already met in Section 5.2.4.

(2): *Reduction of dynamics*:— Using this theorem, we can already fill out a little what is involved in reduced dynamics; which we only glimpsed in our introductory discussions, in Section 2.3 and 5.1. We can make two basic points, as follows.

(A): If H is a G-invariant Hamiltonian function on M, it defines a corresponding function h on M/G by $H = h \circ \pi$. The fact that Poisson maps push Hamiltonian flows forward to Hamiltonian flows (eq. 5.55) implies, since π is Poisson, that π transforms X_H on M to X_h on M/G. That is:

$$T\pi \circ X_H = X_h \circ \pi \quad ; \tag{5.79}$$

i.e. X_H and X_h are π-related. Accordingly, we say that the Hamiltonian system X_H on M *reduces* to that on M/G.

(B): We shall see in Section 6.2 that G-invariance of H is associated with a family of conserved quantities (constants of the motion, first integrals), viz. a constant of the motion $J(\xi) : M \to \mathbb{R}$ for each $\xi \in \mathfrak{g}$. Here, J being conserved means $\{J, H\} = 0$; just as in our discussion of Noether's theorem in ordinary Hamiltonian mechanics (Section 2.1.3). Besides, if J is also G-invariant, then the corresponding function j on M/G is conserved by X_h since

$$\{j, h\} \circ \pi = \{J, H\} = 0 \ \text{ implies } \ \{j, h\} = 0 \ . \tag{5.80}$$

6 Symmetry and conservation revisited: momentum maps

We now develop the topics of symmetry and conserved quantities (and so Noether's theorem) in the context of Poisson manifolds. At the centre of these topics lies the idea of a *momentum map* of a Lie group action on a Poisson manifold; which we introduce in Section 6.1. This is the modern geometric generalization of a conserved quantity, such as linear or angular momentum for the Euclidean group—hence the name. Formally, it will be a map $\mathbf{J}$ from the Poisson manifold M to the dual $\mathfrak{g}^*$ of the Lie algebra of the symmetry group G. Since its values lie in a vector space, it has components. So our description of conserved quantities will no longer be "one-dimensional", i.e. focussed on a single vector field in the state space, as it was in Sections **??** and **??**. The map $\mathbf{J}$ will be associated with a linear map J from $\mathfrak{g}$ to $\mathcal{F}(M)$, the scalar functions on the manifold M. That is: for each $\xi \in \mathfrak{g}$, $J(\xi)$ will be a conserved quantity if the Hamiltonian H is invariant under the infinitesimal generator ξ_M, i.e. if $\xi_M(H) = 0$.

The conservation of momentum maps will be expressed by the Poisson manifold version of Noether's theorem (Section 6.2), and illustrated by the familiar examples of linear and angular momentum (Section 6.3). Then we discuss the equivariance of momentum maps, with respect to the co-adjoint representation of G on $\mathfrak{g}^*$; Section 6.4. Finally in Section 6.5, we discuss the crucial special case of momentum maps on cotangent bundles, again with examples.

6.1 Canonical actions and momentum maps

We first apply the definition of Poisson maps (from Section 5.3.2) to group actions (Section 6.1.1). This will lead to the idea of the momentum map (Section 6.1.2).

6.1.1 Canonical actions and infinitesimal generators

Let G be a Lie group acting on a Poisson manifold M by a smooth left action $\Phi : G \times M \to M$; so that as usual we write $\Phi_g : x \in M \mapsto \Phi_g(x) := g \cdot x \in M$. As in the definition of a Poisson map (eq. 5.52), we say the action is *canonical* if

$$\Phi_g^*\{F_1, F_2\} = \{\Phi_g^* F_1, \Phi_g^* F_2\} \tag{6.1}$$

for any $F_1, F_2 \in \mathcal{F}(M)$ and any $g \in G$. If M is symplectic with symplectic form ω, then the action is canonical iff it is symplectic, i.e. $\Phi_g^*\omega = \omega$ for all $g \in G$.

We will be especially interested in the infinitesimal version of this notion; and so with infinitesimal generators of actions. We recall from eq. 4.37 that the infinitesimal generator of the action corresponding to a Lie algebra element $\xi \in \mathfrak{g}$ is the vector field ξ_M on M obtained by differentiating the action with respect to g at the identity in the direction ξ:

$$\xi_M(x) = \frac{d}{d\tau}[\exp(\tau\xi) \cdot x] \mid_{\tau=0} . \tag{6.2}$$

So we differentiate eq. 6.1 with respect to g in the direction ξ, to give:

$$\xi_M(\{F_1, F_2\}) = \{\xi_M(F_1), F_2\} + \{F_1, \xi_M(F_2)\} . \tag{6.3}$$

Such a vector field ξ_M is called an *infinitesimal Poisson automorphism*.

Side-remark:— We will shortly see that it is the universal quantification over $g \in G$ in eq. 6.1, and correspondingly in eq. 6.3 and 6.5 below, that means our description of conserved quantities is no longer focussed on a single vector field; and in particular, that a momentum map representing a conserved quantity has components.

In the symplectic case, differentiating $\Phi_g^*\omega = \omega$ implies that the Lie derivative $\mathcal{L}_{\xi_M}\omega$ of ω with respect to ξ vanishes: $\mathcal{L}_{\xi_M}\omega = 0$. We saw in Section 2.1.3 that this is equivalent to ξ_M being locally Hamiltonian, i.e. there being a local scalar $J : U \subset M \to \mathbb{R}$ such that $\xi_M = X_J$. This was how Section 2.1.3 vindicated eq. 2.19's "one-liner" approach to Noether's theorem: because the vector field X_f is locally Hamiltonian, it preserves the symplectic structure, i.e. Lie-derives the symplectic form $\mathcal{L}_{X_f}\omega = 0$—as a symmetry should.

We also saw in result (2) at the end of Section 3.2.2 that the "meshing", up to a sign, of the Poisson bracket on scalars with the Lie bracket on vector fields implied that the locally Hamiltonian vector fields form a Lie subalgebra of the Lie algebra $\mathcal{X}(M)$ of all vector fields.

Turning to the context of Poisson manifolds, we need to note two points. The first is a similarity with the symplectic case; the second is a contrast.

(1): One readily checks, just by applying eq. 6.3, that the infinitesimal Poisson automorphisms are closed under the Lie bracket. So we write the Lie algebra of these vector fields as $\mathcal{P}(M)$: $\mathcal{P}(M) \subset \mathcal{X}(M)$.

(2): On the other hand, Section 2.1.3's equivalence between a vector field being locally Hamiltonian and preserving the geometric structure of the state-space breaks down.

Agreed, the first implies the second: a locally Hamiltonian vector field preserves the Poisson bracket. We noted this already in Section 5.3.2. The differential statement was that such a field X_H Lie-derives the Poisson tensor: $\mathcal{L}_{X_H}\mathsf{B}^\sharp = 0$ (eq. 5.54). The finite statement was that the flows of such a field are Poisson maps: $\phi^*_\tau\{F, G\} = \{\phi^*_\tau F, \phi^*_\tau G\}$ (eq. 5.53).

But the converse implication fails: an infinitesimal Poisson automorphism on a Poisson manifold need *not* be locally Hamiltonian. For example, make $\mathbb{R}^2$ a Poisson manifold by defining the Poisson structure

$$\{F, H\} = x\left(\frac{\partial F}{\partial x}\frac{\partial H}{\partial y} - \frac{\partial H}{\partial x}\frac{\partial F}{\partial y}\right) \; ; \tag{6.4}$$

then the vector field $X = \partial/\partial y$ in a neighbourhood of a point on the y-axis is a non-Hamiltonian infinitesimal Poisson automorphism.

This point will affect the formulation of Noether's theorem for Poisson manifolds, in Section 6.2.

Nevertheless, we shall from now on be interested in cases where for all ξ, ξ_M is globally Hamiltonian. This means there is a map $J : \mathfrak{g} \rightarrow \mathcal{F}(M)$ such that

$$X_{J(\xi)} = \xi_M \tag{6.5}$$

for all $\xi \in \mathfrak{g}$. There are three points we need to note about this condition.

(1): Since the right hand side of eq. 6.5 is linear in ξ, we can require such a J to be a linear map. For given any J obeying eq. 6.5, we can take a basis $e_1, \ldots, e_m$ of $\mathfrak{g}$ and define a new linear $\bar{J}$ by setting, for any $\xi = \xi^i e_i$, $\bar{J}(\xi) := \xi^i J(e_i)$.

(2): Eq. 6.5 does not determine $J(\xi)$. For by the linearity of the map $\mathsf{B} : dJ(\xi) \mapsto X_{J(\xi)}$, we can add to such a $J(\xi)$ any distinguished function, i.e. an $F : M \rightarrow \mathbb{R}$ such that $X_F = 0$. That is: $X_{J(\xi)+F} \equiv X_{J(\xi)}$. (Of course, in the symplectic case, the only distinguished functions are constants.)

(3): It is worth expressing eq. 6.5 in terms of Poisson brackets. Recalling that for any $F, H \in \mathcal{F}(M)$, we have $X_H(F) = \{F, H\}$, this equation becomes

$$\{F, J(\xi)\} = \xi_M(F) \;\; , \forall F \in \mathcal{F}(M), \;\; \forall \xi \in \mathfrak{g} \; . \tag{6.6}$$

We will also need the following result:

$$X_{J([\xi,\eta])} = X_{\{J(\xi),J(\eta)\}_M} \; . \tag{6.7}$$

To prove this, we just apply two previous results, each giving a Lie algebra anti-homomorphism.

(i): Result (4) at the end of Section 4.4: for any left action of Lie group G on any manifold M, the map $\xi \mapsto \xi_M$ is a Lie algebra anti-homomorphism between $\mathfrak{g}$ and the Lie algebra $\mathcal{X}_M$ of all vector fields on M:

$$(a\xi + b\eta)_M = a\xi_M + b\eta_M \;\; ; \;\; [\xi_M, \eta_M] = -[\xi, \eta]_M \;\; \forall \xi, \eta \in \mathfrak{g}, \text{ and } a, b \in \mathbb{R}. \tag{6.8}$$

(ii): The "meshing" up to a sign, just as in the symplectic case, of the Poisson bracket on scalars with the Lie bracket on vector fields, as in eq. 5.11 at the end of Section 5.2.2:

$$X_{\{F,H\}} = -[X_F, X_H] = [X_H, X_F] \ . \tag{6.9}$$

So for a Poisson manifold M, the map $F \in \mathcal{F}(M) \mapsto X_F \in \mathcal{X}(M)$ is a Lie algebra anti-homomorphism.

Applying (i) and (ii), we deduce eq. 6.7 by:

$$X_{J([\xi,\eta])} = [\xi, \eta]_M = -[\xi_M, \eta_M] = -[X_{J(\xi)}, X_{J(\eta)}] = X_{\{J(\xi),J(\eta)\}_M} \ . \tag{6.10}$$

6.1.2 Momentum maps introduced

So suppose that there is a canonical left action of G on a Poisson manifold M. And suppose there is a linear map $J : \mathfrak{g} \to \mathcal{F}(M)$ such that

$$X_{J(\xi)} = \xi_M \tag{6.11}$$

for all $\xi \in \mathfrak{g}$.

The two requirements—that the action be infinitesimally canonical (i.e. each $\xi_M \in \mathcal{P}(M)$) and that each ξ_M be globally Hamiltonian—can be expressed as requiring that there be a $J : \mathfrak{g} \to \mathcal{F}(M)$ such that there is a commutative diagram. Namely, the map $\xi \in \mathfrak{g} \mapsto \xi_M \in \mathcal{P}(M)$ is to equal the composed map:

$$\mathfrak{g} \xrightarrow{J} \mathcal{F}(M) \xrightarrow{F \mapsto X_F} \mathcal{P}(M) \ . \tag{6.12}$$

Then the map $\mathbf{J} : M \to \mathfrak{g}^*$ defined by

$$< \mathbf{J}(x) \, ; \, \xi > \ := \ J(\xi)(x) \tag{6.13}$$

for all $\xi \in \mathfrak{g}$ and $x \in M$, is called the *momentum map* of the action.

Another way to state this definition is as follows. Any smooth function $\mathbf{J} : M \to \mathfrak{g}^*$ defines at each $\xi \in \mathfrak{g}$ a scalar $J(\xi) : x \in M \mapsto (\mathbf{J}(x))(\xi) \in \mathbb{R}$. By taking $J(\xi)$ as a Hamiltonian function, one defines a Hamiltonian vector field $X_{J(\xi)}$. But since G acts on M, each $\xi \in \mathfrak{g}$ defines a vector field on M, viz. ξ_M. So we say that $\mathbf{J}$ is a *momentum map* for the action if for each $\xi \in \mathfrak{g}$, these two vector fields are identical: $X_{J(\xi)} = \xi_M$.

Three further remarks by way of illustrating this definition:—

(1): *An isomorphism*:— One readily checks that eq. 6.13 defines an isomorphism between the space of smooth maps $\mathbf{J}$ from M to $\mathfrak{g}^*$, and the space of linear maps J from $\mathfrak{g}$ to scalar functions $\mathcal{F}(M)$. We can take J to define $\mathbf{J}$ by saying that at each $x \in M$, $\mathbf{J}(x) : \xi \in \mathfrak{g} \mapsto \mathbf{J}(x)(\xi) \in \mathbb{R}$ is to be given by the composed map

$$\mathfrak{g} \xrightarrow{J} \mathcal{F}(M) \xrightarrow{|_x} \mathbb{R} \ , \tag{6.14}$$

where $|_x$ means evaluation at $x \in M$. Or we can take $\mathbf{J}$ to define J by saying that at each $\xi \in \mathfrak{g}$, $J(\xi) : x \in M \mapsto J(\xi)(x) \in \mathbb{R}$ is to be given by the composed map

$$M \xrightarrow{\mathbf{J}} \mathfrak{g}^* \xrightarrow{|_\xi} \mathbb{R} \ , \tag{6.15}$$

where $|_\xi$ means evaluation at $\xi \in \mathfrak{g}$.

(2): *Differential equations for the momentum map*:— Using Hamilton's equations, we can readily express the definition of momentum map as a set of differential equations. Recall that on a Poisson manifold, Hamilton's equations are determined by eq. 5.39, which was that at each $x \in M$

$$\mathsf{B}_x(dH(x)) = X_H(x) \ \ ; \tag{6.16}$$

or in local coordinates $x^i, i = 1, \ldots, m \equiv \dim(M)$, with $J^{ij}(x) \equiv \{x^i, x^j\}$ the structure matrix,

$$\mathsf{B}_x(\frac{\partial H}{\partial x^j} dx^j) = \Sigma_{i,j} J^{ij}(x) \frac{\partial H}{\partial x^j} \frac{\partial}{\partial x^i} \ |_x \ ; \tag{6.17}$$

(cf. eq. 5.41). So in local coordinates, Hamilton's equations are given by eq. 5.22, which was:

$$\frac{dx^i}{dt} = \Sigma^m_j \ J^{ij}(x) \frac{\partial H}{\partial x^j} \ . \tag{6.18}$$

So the condition for a momentum map $X_{J(\xi)} = \xi_M$ is that for all $\xi \in \mathfrak{g}$ and all $x \in M$

$$\mathsf{B}_x(d(J(\xi))(x)) = \xi_M(x) \ \ . \tag{6.19}$$

In coordinates, this is the requirement that for all $i = 1, \ldots, m$

$$\Sigma^m_j \ J^{ij}(x) \frac{\partial J(\xi)}{\partial x^j} \ = \ (\xi_M)^i(x) \ , \tag{6.20}$$

where—apologies!—the two Js on the left hand side have very different meanings.

In the symplectic case, $\dim(M) \equiv m = 2n$ and we have Hamilton's equations as eq. 2.15, viz.

$$\mathbf{i}_{X_H}\omega := \omega(X_H, \cdot) = dH(\cdot) \ \ . \tag{6.21}$$

So the condition for a momentum map is that for all ξ

$$\omega(\xi_M, \cdot) = d(J(\xi))(\cdot) \ \ . \tag{6.22}$$

In Hamiltonian mechanics, it is common to write the $2n$ local coordinates q, p as ξ, i.e. to write

$$\xi^\alpha := q^\alpha, \ \ \alpha = 1, ..., n \ \ \ ; \ \ \ \xi^\alpha := p_{\alpha - n}, \ \ \alpha = n+1, ..., 2n \ \ . \tag{6.23}$$

So in order to express eq. 6.22 in local coordinates, let us temporarily write η for the arbitrary element of $\mathfrak{g}$. Then writing $\eta_M = (\eta_M)^\alpha \frac{\partial}{\partial \xi^\alpha}$ and $\omega_{\alpha\beta} := \omega(\frac{\partial}{\partial \xi^\alpha}, \frac{\partial}{\partial \xi^\beta})$, eq. 6.22 becomes

$$\omega_{\alpha\beta}(\eta_M)^\alpha = \frac{\partial J(\eta)}{\partial \xi^\beta} \ \ . \tag{6.24}$$

(3): *Components: an example*:— As discussed after eq. 6.3, we think of the collection of functions $J(\xi)$, as ξ varies through $\mathfrak{g}$, as the *components* of $\mathbf{J}$.

To take our standard example: the angular momentum of a particle in Euclidean space, in a state $x = (\mathbf{q}, \mathbf{p})$ is $\mathbf{J}(x) := \mathbf{q} \wedge \mathbf{p}$. Identifying $\mathfrak{so}(3)^*$ with $\mathbb{R}^3$ so that the natural pairing is given by the dot product (cf. (3) at the end of Section 4.5.2), we get that the component of $\mathbf{J}(x)$ around the axis $\xi \in \mathbb{R}^3$ is $< \mathbf{J}(x) \; ; \; \xi >= \xi \cdot (\mathbf{q} \wedge \mathbf{p})$. The Hamiltonian vector field determined by this Hamiltonian function $x = (\mathbf{q}, \mathbf{p}) \mapsto \xi \cdot (\mathbf{q} \wedge \mathbf{p})$ is of course the infinitesimal generator of rotations about the ξ-axis. In Section 6.3, we will see more examples of momentum maps.

6.2 Conservation of momentum maps: Noether's theorem

In ordinary Hamiltonian mechanics, we saw that Noether's theorem had a simple expression as a "one-liner" based on the antisymmetry of the Poisson bracket: namely, in eq. 2.19, which was that for any scalar functions F, H

$$X_F(H) = \{H, F\} = 0 \quad \text{iff} \quad 0 = \{F, H\} = X_H(F) \ . \tag{6.25}$$

In words: the Hamiltonian H is constant under the flow induced by F iff F is a constant of the motion under the dynamical flow X_H.

More precisely, Section 2.1.3 vindicated this one-liner as expressing Noether's theorem. For the one-liner respected the requirement that a symmetry should preserve the symplectic form (equivalently, the Poisson bracket), and not just (as in the left hand side of eq. 6.25) the Hamiltonian function H; for, by Cartan's magic formula, a vector field's preserving the symplectic form was equivalent to its being locally Hamiltonian.

For Poisson manifolds, the equivalence corresponding to this last statement fails. That is, as we noted in (2) of Section 6.1.1: an infinitesimal Poisson automorphism need not be locally Hamiltonian.

Nevertheless, most of the "one-liner" approach to Noether's theorem carries over to the framework of Poisson manifolds. In effect, we just restrict discussion to cases where the relevant Hamiltonian vector fields exist: recall our saying after (2) of Section 6.1.1 that we would concentrate on cases where all the ξ_M are globally Hamiltonian.

Thus, it is straightforward to show that for a Poisson manifold M, just as for symplectic manifolds: if $F, H \in \mathcal{F}(M)$, H is constant along the integral curves of X_F iff $\{H, F\} = 0$ iff F is constant along the integral curves of X_H. (We could have proved this already in Section 5.2.2; but postponed it till now, when it will be used.)

With this result as a lemma, one immediately gets

> **Noether's theorem for Poisson manifolds** Suppose that G acts canonically on a Poisson manifold M and has a momentum map $\mathbf{J} : M \to \mathfrak{g}^*$; and that H is invariant under ξ_M for all $\xi \in \mathfrak{g}$, i.e. $\{H, J(\xi)\} = \xi_M(H) = 0$, $\forall \xi \in \mathfrak{g}$; (cf. eq. 6.6). Then $\mathbf{J}$ is a constant of the motion determined by

H. That is:

$$\mathbf{J} \circ \phi_\tau = \mathbf{J} \tag{6.26}$$

where ϕ_τ is the flow of X_H.

Proof: By the lemma, the fact that $\{H, J(\xi)\} = \xi_M(H) = 0$ implies that $J(\xi)$ is constant along the flow of X_H. So by the definition of momentum map, eq. 6.13, the corresponding $\mathfrak{g}^*$-valued map $\mathbf{J}$ is also a constant of the motion. QED.

It follows immediately that H itself, and any distinguished function, is a constant of the motion. Besides, as remarked in (2) at the end of Section 6.1.1: a constant of the motion $J(\xi)$ is determined only up to an arbitrary choice of a distinguished function. Indeed, though this Chapter has set aside (ever since (iii) of Section 1.2) time-dependent functions: if one considers them, then there is here an arbitrary choice of a time-dependent distinguished function.

6.3 Examples

We give two familiar examples; and then, as a glimpse of the general power of the theory, two abstract examples (which will *not* be needed later on).

(1): *Total linear momentum of N particles* :—
In (3) at the end of Section 4.1.A, we showed that the *left* cotangent lift of the action of the translation group $\mathbb{R}^3$ on $Q = \mathbb{R}^{3N}$ to $M = T^*\mathbb{R}^{3N}$, i.e. the left action corresponding to eq. 4.11, is

$$\Psi_{\mathbf{x}}(\mathbf{q}_i, \mathbf{p}^i) := T^*(\Phi_{-\mathbf{x}})(\mathbf{q}_i, \mathbf{p}^i) = (\mathbf{q}_i + \mathbf{x}, \mathbf{p}^i) \ , \ \ i = 1, ..., N \ . \tag{6.27}$$

(Here we combine the discussions of examples (vi) and (ix) in Section 4.1.A.)

To find the momentum map, we: (a) compute the infinitesimal generator ξ_M for an arbitrary element ξ of $\mathfrak{g} = \mathbb{R}^3$; and then (b) solve eq. 6.22, or in coordinates eq. 6.24.

(a): We differentiate eq. 6.27 with respect to $\mathbf{x}$ in the direction ξ, getting

$$\xi_M(\mathbf{q}_i, \mathbf{p}^i) = (\xi, ..., \xi, \mathbf{0}, ..., \mathbf{0}) \ . \tag{6.28}$$

(b): Any function $J(\xi)$ has Hamiltonian vector field

$$X_{J(\xi)}(\mathbf{q}_i, \mathbf{p}^i) = \left(\frac{\partial J(\xi)}{\partial \mathbf{p}^i}, -\frac{\partial J(\xi)}{\partial \mathbf{q}_i} \right) \ ; \tag{6.29}$$

so that the desired $J(\xi)$ with $X_{J(\xi)} = \xi_M$ solves

$$\frac{\partial J(\xi)}{\partial \mathbf{p}^i} = \xi \quad \text{and} \quad \frac{\partial J(\xi)}{\partial \mathbf{q}_i} = \mathbf{0} \ , \ 1 \leq i \leq N \ . \tag{6.30}$$

Choosing constants so that J is linear, the solution is

$$J(\xi)(\mathbf{q}_i, \mathbf{p}^i) = \left(\Sigma_{i=1}^N \mathbf{p}^i \right) \cdot \xi \ , \quad \text{i.e.} \quad \mathbf{J}(\mathbf{q}_i, \mathbf{p}^i) = \Sigma_{i=1}^N \mathbf{p}^i \ ; \tag{6.31}$$

i.e. the familiar total linear momentum.

(2): *Angular momentum of a single particle* :—
$SO(3)$ acts on $Q = \mathbb{R}^3$ by $\Phi_A(\mathbf{q}) = A\mathbf{q}$. So the tangent (derivative) map is

$$T_{\mathbf{q}}\Phi_A : (\mathbf{q}, \mathbf{v}) \in T\mathbb{R}^3_{\mathbf{q}} \mapsto (A\mathbf{q}, A\mathbf{v}) \in T\mathbb{R}^3_{A\mathbf{q}} \ . \tag{6.32}$$

As we saw in example (vii) of Section 4.1.A, the *left* cotangent lift of the action to $M = T^*\mathbb{R}^3$ (the lifted action "with g^{-1}", corresponding to eq. 4.11) is:

$$T^*_{A\mathbf{q}}(\Phi_{A^{-1}})(\mathbf{q}, \mathbf{p}) = (A\mathbf{q}, A\mathbf{p}) \ . \tag{6.33}$$

To find the momentum map, we proceed in two stages, (a) and (b), as in example (1).

(a): We differentiate eq. 6.33 with respect to A in the direction $\xi = \Theta(\omega) \in \mathfrak{so}(3)$, where $\omega \in \mathbb{R}^3$ and Θ is as in eq. 3.20 and 3.23. We get

$$\xi_M(\mathbf{q}, \mathbf{p}) = (\xi\mathbf{q}, \xi\mathbf{p}) = (\omega \wedge \mathbf{q}, \omega \wedge \mathbf{v}) \ . \tag{6.34}$$

(b): So the desired $J(\xi)$ is the solution linear in ξ to the Hamilton's equations

$$\frac{\partial J(\xi)}{\partial \mathbf{p}} = \xi\mathbf{q} \quad \text{and} \quad \frac{\partial J(\xi)}{\partial \mathbf{q}} = -\xi\mathbf{p} \ . \tag{6.35}$$

So a solution is given by

$$J(\xi)(\mathbf{q}, \mathbf{p}) = (\xi\mathbf{q}) \cdot \mathbf{p} = (\omega \wedge \mathbf{q}) \cdot \mathbf{p} = (\mathbf{q} \wedge \mathbf{p}) \cdot \omega \ , \tag{6.36}$$

so that

$$\mathbf{J}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \wedge \mathbf{p} \ , \tag{6.37}$$

i.e. the familiar angular momentum.

(3): *Dual of a Lie algebra homomorphism* :—
We begin by stating a *Lemma*, which we will not prove; for details cf. Marsden and Ratiu (1999: 10.7.2, p. 372). Namely: let G, H be Lie groups and let $\alpha : \mathfrak{g} \to \mathfrak{h}$ be a linear map between their Lie algebras. Then α is a Lie algebra homomorphism iff its dual $\alpha^* : \mathfrak{h}^* \to \mathfrak{g}^*$ is a (linear) Poisson map (where $\mathfrak{h}^*, \mathfrak{g}^*$ are equipped with their natural Lie-Poisson brackets as in Section 5.2.4).

Now let G, H be Lie groups, let $A : H \to G$ be a Lie group homomorphism, and let $\alpha : \mathfrak{h} \to \mathfrak{g}$ be the induced Lie algebra homomorphism; so that by the Lemma, $\alpha^* : \mathfrak{g}^* \to \mathfrak{h}^*$ is a Poisson map. We will prove that α^* is also a momentum map for the action of H on $\mathfrak{g}^*$ given by, with $h \in H, x \in \mathfrak{g}^*$:

$$\Phi(h, x) \equiv h \cdot x := Ad^*_{A(h)^{-1}} x \ . \tag{6.38}$$

Proof: We first recall the adjoint and co-adjoint actions $Ad_g : \mathfrak{g} \to \mathfrak{g}$ and $Ad^*_g : \mathfrak{g}^* \to \mathfrak{g}^*$; in particular, eq. 4.76. So the action in eq. 6.38 is:

$$\forall x \in \mathfrak{g}^*, \forall \xi \in \mathfrak{g} \ : \ < h \cdot x; \xi > = < x; Ad_{A(h)^{-1}} \xi > \ . \tag{6.39}$$

As usual, we compute for $\eta \in \mathfrak{h}$, the infinitesimal generator $\eta_{\mathfrak{g}^*}$ at $x \in \mathfrak{g}^*$ by differentiating eq. 6.39 with respect to h at e in the direction $\eta \in \mathfrak{h}$. We get (cf. eq. 4.83):

$$< \eta_{\mathfrak{g}^*}(x); \xi > = \; - < x; ad_{\alpha(\eta)}\xi > = < ad^*_{\alpha(\eta)}(x); \xi > \; . \tag{6.40}$$

We define $\mathbf{J}(x) := \alpha^*(x)$: that is,

$$J(\eta)(x) \; \equiv < \mathbf{J}(x); \eta > := < \alpha^*(x); \eta > \equiv < x; \alpha(\eta) > ; \tag{6.41}$$

which implies

$$\nabla_x J(\eta) \; = \; \alpha(\eta) \; . \tag{6.42}$$

Now we recall that Hamilton's equations for $J(\eta)$ as the Hamiltonian are (cf. eq. 5.30)

$$\dot{x} \equiv X_{J(\eta)}(x) = ad^*_{\nabla_x J(\eta)}(x) \; . \tag{6.43}$$

Combining eq. 6.40 to eq. 6.43, we get:

$$X_{J(\eta)}(x) = ad^*_{\alpha(\eta)}(x) = \eta_{\mathfrak{g}^*}(x) \; ; \tag{6.44}$$

proving that $\mathbf{J}(x) := \alpha^*(x)$ is a momentum map. QED.

(4): *Momentum maps for subgroups* :—
Assume that $\mathbf{J} : M \rightarrow \mathfrak{g}^*$ is a momentum map for a canonical left action of G on M; and let $H < G$ be a subgroup of G. Then H also acts canonically on M, and this action has as a momentum map the restriction of $\mathbf{J}$'s values to $\mathfrak{h} \subset \mathfrak{g}$. That is: the map

$$\mathbf{J}_H : M \rightarrow \mathfrak{h}^* \quad \text{given by} \quad \mathbf{J}_H(x) := \mathbf{J}(x) \mid_{\mathfrak{h}} \; . \tag{6.45}$$

For the canonical action of G ensures that if $\eta \in \mathfrak{h} \subset \mathfrak{g}$, then $\eta_M = X_{J(\eta)}$. Then $J_H(\eta) := J(\eta) \forall \eta \in \mathfrak{h}$ defines a momentum map for H's action. That is

$$\forall x \in M, \; \forall \eta \in \mathfrak{h} : \;\; < \mathbf{J}_H(x); \eta > = < \mathbf{J}(x); \eta > \; . \tag{6.46}$$

6.4 Equivariance of momentum maps

In (1) of Section 4.2, we defined the general notion of an equivariant map $f : M \rightarrow N$ between manifolds as one that respects the actions of a group G on M and on N: eq. 4.29. We now develop an especially important case of this notion: the equivariance of momentum maps $\mathbf{J} : M \rightarrow \mathfrak{g}^*$, where the action on $\mathfrak{g}^*$ is the co-adjoint action, eq. 4.77.

For us, this notion will have two main significances:—
(i): many momentum maps that occur in examples are equivariant in this sense;
(ii): equivariance has various theoretical consequences: in particular, momentum maps for cotangent lifted actions are always equivariant (Section 6.5), and equivariance is crucial in theorems about reduction (Section 7).

In this Section, we will glimpse these points by:
(i): defining the notion, and remarking on a weakened differential version of the notion (Section 6.4.1);
(ii): proving that equivariant momentum maps are Poisson (Section 6.4.2).

6.4.1 Equivariance and infinitesimal equivariance

Let Φ be a canonical left action of G on M, and let $\mathbf{J} : M \rightarrow \mathfrak{g}^*$ be a momentum map for it. We say $\mathbf{J}$ is *equivariant* if for all $g \in G$

$$\mathbf{J} \circ \Phi_g \ = \ Ad^*_{g^{-1}} \circ \mathbf{J} \ ; \tag{6.47}$$

cf. eq. 4.29 and the definition of co-adjoint action, eq. 4.78:

$$\begin{array}{ccc} M & \stackrel{\mathbf{J}}{\longrightarrow} & \mathfrak{g}^* \\ \uparrow \Phi_g & & \uparrow Ad^*_{g^{-1}} \\ M & \stackrel{\mathbf{J}}{\longrightarrow} & \mathfrak{g}^* \end{array} \tag{6.48}$$

An equivalent formulation arises by considering that we can add to the commutative square in eq. 6.48 the two commutative triangles:

$$M \stackrel{J(\xi)}{\longrightarrow} \mathrm{I\!R} \quad \text{is} \quad M \stackrel{\mathbf{J}}{\longrightarrow} \mathfrak{g}^* \stackrel{|_\xi}{\longrightarrow} \mathrm{I\!R} \ ; \tag{6.49}$$

representing the fact that $J(\xi)(x) = \mathbf{J}(x)(\xi)$; and

$$\mathfrak{g}^* \stackrel{|_\xi}{\longrightarrow} \mathrm{I\!R} \quad \text{is} \quad \mathfrak{g}^* \stackrel{Ad^*_{g^{-1}}}{\longrightarrow} \mathfrak{g}^* \stackrel{|_{Ad_g \xi}}{\longrightarrow} \mathrm{I\!R} \ ; \tag{6.50}$$

representing the fact that for all $\eta \in g^*$

$$< Ad^*_{g^{-1}}(\eta); Ad_g(\xi) > = < \eta; Ad_{g^{-1}} Ad_g(\xi) > \equiv < \eta; \xi > \ . \tag{6.51}$$

Eq.s 6.49 and 6.50 imply that an equivalent formulation of equivariance is that for all $x \in M, g \in G$ and $\xi \in \mathfrak{g}$ (and with $g \cdot x \equiv \Phi_g(x)$)

$$\mathbf{J}(g \cdot x)(Ad_g \xi) \equiv J(Ad_g \xi)(g \cdot x) = J(\xi)(x) \equiv \mathbf{J}(x)(\xi) \ . \tag{6.52}$$

In (2) of Section 4.4, we differentiated the general notion of an equivariant map, and got the weaker differential notion that the infinitesimal generators ξ_M and ξ_N of the actions of G on M and on N are f-related.

Here also we can differentiate equivariance, and get the notion of *infinitesimal equivariance*. But I will not go into details since:

(i): we will not need the notion, not least because (as mentioned above), many momentum maps are equivariant;

(i): under certain common conditions (e.g. the group G is compact, or is connected) an infinitesimally equivariant momentum map can always be replaced by an equivariant one.

So let it suffice to say that infinitesimal equivariance *is* theoretically important. In particular, the result eq. 6.7, viz.

$$X_{J([\xi,\eta])} = X_{\{J(\xi),J(\eta)\}_M} \tag{6.53}$$

implies that

$$\Sigma(\xi, \eta) := J([\xi, \eta]) - \{J(\xi), J(\eta)\}_M \tag{6.54}$$

is a distinguished function on the Poisson manifold M, and so constant on every symplectic leaf.

This makes it natural to ask when $\Sigma \equiv 0$. After all, cf. eq. 6.12. Both $\xi \mapsto \xi_M$ and $F \mapsto X_F$ are Lie algebra anti-homomorphisms. So it is natural to ask whether J is a Lie algebra homomorphism, i.e. whether $\Sigma = 0$. And it turns out that infinitesimal equivariance is equivalent to $\Sigma = 0$.

6.4.2 Equivariant momentum maps are Poisson

The following result is important, both as a general method of finding canonical maps between Poisson manifolds, and for the Lie-Poisson reduction theorem of Section 7.

> **Equivariant momentum maps are Poisson** Let $\mathbf{J} : \mathbf{M} \rightarrow \mathfrak{g}^*$ be an equivariant momentum map for a canonical left action of G on a Poisson manifold M. Then $\mathbf{J}$ is a Poisson map: for all $F_1, F_2 \in \mathcal{F}(\mathfrak{g}^*)$,
>
> $$\mathbf{J}^*\{F_1, F_2\}_{\mathfrak{g}^*} = \{\mathbf{J}^*F_1, \mathbf{J}^*F_2\}_M \ \text{; i.e.} \ \ \{F_1, F_2\}_{\mathfrak{g}^*} \circ \mathbf{J} = \{F_1 \circ \mathbf{J}, F_2 \circ \mathbf{J}\}_M \ . \tag{6.55}$$

Proof:— We will relate (i) the left hand side, then (ii) the right hand side of eq. 6.55 to J; and finally we will use the fact that the Poisson bracket on M depends only on the values of the first derivatives.

(i): Let $x \in M, \alpha = \mathbf{J}(x) \in \mathfrak{g}^*$; and let $\xi = \nabla F_1$ and $\eta = \nabla F_2$ evaluated at α, so that $\xi, \eta \in \mathfrak{g}^{**} = \mathfrak{g}$. Then

$$\{F_1, F_2\}_{\mathfrak{g}^*}(\mathbf{J}(x)) \equiv < \alpha; [\nabla F_1, \nabla F_2] > = < \alpha; [\xi, \eta] > = J([\xi, \eta])(x) = \{J(\xi), J(\eta)\}(x) \ ; \tag{6.56}$$

where the third equation just applies the definition of $\mathbf{J}$, eq. 6.13, and the fourth equation uses (infinitesimal) equivariance.

(ii): We show that $(F_1 \circ \mathbf{J})(x)$ and $J(\xi)(x)$ have equal x-derivatives. For any $x \in M$ and $v_x \in T_xM$

$$\mathbf{d}(F_1 \circ \mathbf{J})(x) \cdot v_x = \mathbf{d}F_1(\alpha) \cdot T_x\mathbf{J}(v_x) = < T_x\mathbf{J}(v_x); \nabla F_1 > = \mathbf{d}J(\xi)(x) \cdot v_x \ ; \tag{6.57}$$

where the first equation uses the chain rule, and the last uses the definition of $\mathbf{J}$, eq. 6.13 and the fact that $\xi = \nabla F_1$.

Finally, since the Poisson bracket on M depends only on the values of the first derivatives, we infer from eq. 6.57 that

$$\{F_1 \circ \mathbf{J}, F_2 \circ \mathbf{J}\}(x) = \{J(\xi), J(\eta)\}(x) \ . \tag{6.58}$$

Combining this with (i), the result follows. QED.

6.5 Momentum maps on cotangent bundles

Let a Lie group G act on a manifold ("configuration space") Q. We saw in Section 4.1.A that this action can be *lifted* to the cotangent bundle T^*Q; cf. eq.s 4.6, 4.9 and 4.11. In this Section, we focus on momentum maps for such *cotangent lift* actions. We shall see that any such action has an equivariant momentum map, for which there is an explicit general formula. The general theory (Sections 6.5.1, 6.5.2) will need just one main new notion, the *momentum function*. We end with some examples (Section 6.5.3).

6.5.1 Momentum functions

Given a manifold Q and its vector fields $\mathcal{X}(Q)$, we define the map

$$\mathcal{P} : \mathcal{X}(Q) \to \mathcal{F}(T^*Q) \quad \text{by :} \quad (\mathcal{P}(X))(\alpha_q) \ := \ < \alpha_q ; X(q) > \tag{6.59}$$

for $q \in Q, X \in \mathcal{X}(Q)$ and $\alpha_q \in T^*_q Q$. Here, α_q is, strictly speaking, a point in the cotangent bundle above the base-point $q \in Q$: so α_q can be written as (q, α) with α a covector at q, i.e. $\alpha \in T^*_q Q$. But as we mentioned just before defining cotangent lifts (eq. 4.6): it is harmless to (follow many presentations and) conflate a point in T^*Q, i.e. a pair $(q, \alpha), q \in Q, \alpha \in T^*_q Q$, with its form α, provided we keep track of the q by writing the form as α_q.

$\mathcal{P}(X)$, as defined by eq. 6.59, is called the *momentum function* of X. In coordinates, $\mathcal{P}(X)$ is given by

$$\mathcal{P}(X)(q^i, p_i) = X^j(q^i) p_j \tag{6.60}$$

where we sum on $j = 1, ..., n := \dim Q$. (So NB: This $\mathcal{P}$ is different from that in $\mathcal{P}(M)$, the infinitesimal Poisson automorphisms of M, discussed in Section 6.1.1.)

We also denote by $\mathcal{L}(T^*Q)$ the space of smooth functions $F : T^*Q \to \mathrm{I\!R}$ that are linear on fibres of T^*Q: i.e. writing the bundle points $\alpha_q, \beta_q \in T^*_q Q$ as (q, α) and (q, β), we have for $\lambda, \mu \in \mathrm{I\!R}$

$$F(q, (\lambda\alpha + \mu\beta)) = \lambda F((q, \alpha)) + \mu F((q, \beta)) \ . \tag{6.61}$$

So functions F, H that are in $\mathcal{L}(T^*Q)$ can be written in coordinates as (summing on $i = 1, ..., n$)

$$F(q,p) = X^i(q) p_i \quad \text{and} \quad H(q,p) = Y^i(q) p_i \tag{6.62}$$

for functions X^i and Y^i; and so any momentum function $\mathcal{P}(X)$ is in $\mathcal{L}(T^*Q)$.

One readily checks that the standard Poisson bracket (from T^*Q's symplectic structure, Section 2.1.1) of such an F and H is also linear on the fibres of T^*Q. In fact, eq. 6.62 implies

$$\{F, H\}(q,p) := \frac{\partial F}{\partial q^j}\frac{\partial H}{\partial p_j} - \frac{\partial H}{\partial q^j}\frac{\partial F}{\partial p_j} = \left(\frac{\partial X^i}{\partial q^j} Y^j - \frac{\partial Y^i}{\partial q^j} X^j \right) \ . \tag{6.63}$$

So $\mathcal{L}(T^*Q)$ is a Lie subalgebra of $\mathcal{F}(T^*Q)$.

The next result summarizes how momentum functions relate $\mathcal{X}(Q)$ and Hamiltonian vector fields on T^*Q to $\mathcal{L}(T^*Q)$.

Three (anti)-isomorphic Lie algebras The two Lie algebras
(i) $(\mathcal{X}(Q), [,])$ of vector fields on Q;
(ii) Hamiltonian vector fields X_F on T^*Q with $F \in \mathcal{L}(T^*Q)$
are isomorphic. And each is anti-isomorphic to
(iii) $(\mathcal{L}(T^*Q), \{,\})$.
In particular, the map $\mathcal{P}$ is an anti-isomorphism from (i) to (iii), so that we have

$$\{\mathcal{P}(X), \mathcal{P}(Y)\}_{T^*Q} = -\mathcal{P}([X,Y]) \ . \tag{6.64}$$

Proof: Since $\mathcal{P}(X) : T^*Q \to \mathbb{R}$ is linear on fibres, $\mathcal{P}$ maps $\mathcal{X}(Q)$ into $\mathcal{L}(T^*Q)$. $\mathcal{P}$ is also onto $\mathcal{L}(T^*Q)$: given $F \in \mathcal{L}(T^*Q)$, we can define $X(F) \in \mathcal{X}(Q)$ by

$$< \alpha_q; X(F)(q) > := \ F(\alpha_q) \quad \forall \alpha_q \in T^*_q Q \tag{6.65}$$

so that $\mathcal{P}(X(F)) = F$. $\mathcal{P}$ is linear and $\mathcal{P}(X) = 0$ implies that $X = 0$. Also, eq. 6.64 follows immediately by comparing eq. 6.63 with the Lie bracket of $X, Y \in \mathcal{X}(Q)$; cf. eq. 3.27. So $\mathcal{P}$ is an anti-isomorphism from $(\mathcal{X}Q, [,])$ to $(\mathcal{L}(T^*Q), \{,\})$.

The map

$$F \in (\mathcal{L}(T^*Q), \{,\}) \mapsto X_F \in (\{X_F \mid F \in \mathcal{L}(T^*Q)\}, [,]) \tag{6.66}$$

is surjective by definition. It is a Lie algebra anti-homomorphism, by eq. 3.32 (i.e. result (2) in Section 3.2.2). And if $X_F = 0$, then F is constant on T^*Q; and hence $F \equiv 0$ since F is linear on the fibres (cf. eq. 6.61). QED.

6.5.2 Momentum maps for cotangent lifted actions

We begin this Subsection with a result relating the Hamiltonian flow on T^*Q induced by the momentum function $\mathcal{P}(X)$ to the Hamiltonian flow on X induced by X. From this result, our main result—the guarantee of an equivariant momentum map for a cotangent lifted action, and an explicit formula for it—will follow directly.

The Hamiltonian flow of a momentum function Let $X \in \mathcal{X}(Q)$ have flow ϕ_τ on Q; cf. Section 3.1.2.B. Then the flow of $X_{\mathcal{P}(X)}$ on T^*Q is $T^*\phi_{-\tau}$. That is: the flow of $X_{\mathcal{P}(X)}$ is the cotangent lift (Section 4.1.A) of $\phi_{-\tau}$, as given by the diagram, with π_Q the canonical projection:

$$\begin{tikzcd} Q \arrow[r, "\phi_\tau"] & Q \\ T^*Q \arrow[u, "\pi_Q"] \arrow[r, "T^*\phi_{-\tau}"] & T^*Q \arrow[u, "\pi_Q"] \end{tikzcd} \tag{6.67}$$

Proof: We differentiate the relation in eq. 6.67, i.e.

$$\pi_Q \circ T^*\phi_{-\tau} = \phi_\tau \circ \pi_Q \tag{6.68}$$

at $\tau = 0$ to get

$$T\pi_Q \circ Y = X \circ \pi_Q \quad \text{with} \quad \forall \alpha_q \in T_q^*Q, \; Y(\alpha_q) = \frac{d}{d\tau}\mid_{\tau=0} T^*\phi_{-\tau}(\alpha_q) \; ; \tag{6.69}$$

i.e. $T^*\phi_{-\tau}$ is the flow of Y.

Now we will show that $Y = X_{\mathcal{P}(X)}$, using eq. 6.69 and the geometrical formulation of Hamiltonian mechanics of Section 2.1, especially Cartan's magic formula, eq. 2.20, applied to the canonical one-form $\theta \equiv \theta_H$ (defined by eq. 2.8 and 2.9).

We reported (at the start of (2) of Section 4.1.A) that the cotangent lift $T^*\phi_{-\tau}$ preserves $\theta \equiv \theta_H$ on T^*Q. So $\mathcal{L}_Y\theta = 0$. Then the definition of ω as the negative exterior derivative of θ, and Cartan's magic formula, eq. 2.20, yields

$$\mathbf{i}_Y\omega = -\mathbf{i}_Y\mathbf{d}\theta = \mathbf{d}\mathbf{i}_Y\theta \; . \tag{6.70}$$

On the other hand, we also have

$$\mathbf{i}_Y\theta(\alpha_q) \equiv < \theta(\alpha_q); Y(\alpha_q) > = < \alpha_q; T\pi_Q(Y(\alpha_q)) > = < \alpha_q; X(q) > = \mathcal{P}(X)(\alpha_q) \tag{6.71}$$

where the second equation applies the definition of the canonical one-form (eq. 2.8), the third applies eq. 6.69, and the fourth applies the definition eq. 6.59 of momentum functions.

Combining eq. 6.70 and 6.71, we have:

$$\mathbf{i}_Y\omega = \mathbf{d}\mathcal{P}(X) \tag{6.72}$$

which is Hamilton's equations (eq. 2.15) telling us that $Y = X_{\mathcal{P}(X)}$. QED.

Accordingly the Hamiltonian vector field $X_{\mathcal{P}(X)}$ on T^*Q is called the *cotangent lift* of $X \in \mathcal{Q}$ to T^*Q. In local coordinates, we can write, by combining eq. **??** and 6.60

$$X_{\mathcal{P}(X)} = \frac{\partial \mathcal{P}(X)}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial \mathcal{P}(X)}{\partial q^i}\frac{\partial}{\partial p_i} = X^i\frac{\partial}{\partial q^i} - \frac{\partial X^i}{\partial q^j}p_i\frac{\partial}{\partial p_j} \; . \tag{6.73}$$

Note in particular that, combining the usual sign-change between Lie algebras and Poisson brackets (eq. 3.32) with the sign-change for momentum functions (eq. 6.64), we have

$$[X_{\mathcal{P}(X)}, X_{\mathcal{P}(Y)}] = -X_{\{\mathcal{P}(X),\mathcal{P}(Y)\}} = -X_{-\mathcal{P}([X,Y])} = X_{\mathcal{P}([X,Y])} \; . \tag{6.74}$$

We can now readily prove our main result guaranteeing, and giving a formula for, equivariant momentum maps.

Equivariant momentum maps Let G act on the left on Q and so by cotangent lift on T^*Q. The cotangent lifted action has an equivariant momentum map $\mathbf{J} : T^*Q \rightarrow \mathfrak{g}^*$ given by

$$< \mathbf{J}(\alpha_q); \xi > = < \alpha_q; \xi_Q(q) > \equiv \ \mathcal{P}(\xi_Q)(\alpha_q) \ . \tag{6.75}$$

In coordinates q^i, p_i on T^*Q and ξ^a on $\mathfrak{g}$, and with $\xi^i_Q = \xi^a A^i_a$ the components of ξ_Q, this reads

$$J_a \xi^a = p_i \xi^i_Q = p_i A^i_a \xi^a \tag{6.76}$$

so that $J_a(q,p) = p_i A^i_a(q)$.

Proof: The preceding result tells us that for any $\xi \in \mathfrak{g}$, the infinitesimal generator of the cotangent lifted action on T^*Q is $\xi_{T^*Q} \equiv X_{\mathcal{P}(\xi_Q)}$. So a momentum map for this action is given by

$$J(\xi) = \mathcal{P}(\xi_Q) \ . \tag{6.77}$$

This gives eq. 6.75, just by applying the definitions of the momentum map $\mathbf{J}$ (eq. 6.13) and of momentum function (eq. 6.59).

To prove equivariance, we argue as follows:

$$< \mathbf{J}(g \cdot \alpha_q); \xi > = < (g \cdot \alpha_q); \xi_Q(g \cdot q) > \tag{6.78}$$

$$= < \alpha_q; (T\Phi_{g^{-1}})\xi_Q(g \cdot q) > \equiv < \alpha_q; (T_{g \cdot q}\Phi_{g^{-1}} \circ \xi_Q \circ \Phi_g)(q). \tag{6.79}$$

$$= < \alpha_q; (\Phi^*_g \xi_Q)(q) > \tag{6.80}$$

$$= < \alpha_q; (Ad_{g^{-1}}\xi)_Q(q) > \tag{6.81}$$

$$= < \mathbf{J}(\alpha_q); Ad_{g^{-1}}\xi > = < Ad^*_{g^{-1}}(\mathbf{J}(\alpha_q)); \xi > \ . \tag{6.82}$$

Here we have applied in succession: (i) eq. 6.75; (ii) the fact that $g \cdot \alpha_q$ is short for $T^*(\Phi_{g^{-1}})(\alpha_q)$, cf. eq. 4.11 and 4.6; (iii) the definition of pullback, cf. eq. 4.57; (iv) result [2], eq. 4.52, of Section 4.5.1; (v) eq. 6.75 again; and finally, (vi) the fact that Ad^* is the adjoint of Ad, cf. eq. 4.76. QED.

6.5.3 Examples

We discuss first our familiar examples, linear and angular momentum i.e. (1) and (2) from Section 6.3; and then the cotangent lift of left and right translations on G—an example motivated by Section 4.6's description of kinematics on a Lie group G.

(1): *Total linear momentum of N particles*:—
Since the translation group $\mathbb{R}^3$ acts on $Q := \mathbb{R}^{3N}$ by $\Phi(\mathbf{x}, (\mathbf{q}_i)) = (\mathbf{q}_i + \mathbf{x})$, the infinitesimal generator on Q is

$$\xi_{\mathbb{R}^{3N}}(\mathbf{q}_i) = (\xi, \ldots, \xi)(\xi \ N \ \text{ times}) \tag{6.83}$$

Applying eq. 6.75, the equivariant momentum map is given by

$$J(\xi)(\mathbf{q}_i, \mathbf{p}^i) = \left(\Sigma_{i=1}^N \mathbf{p}^i\right) \cdot \xi \ , \quad \text{i.e.} \quad \mathbf{J}(\mathbf{q}_i, \mathbf{p}^i) = \Sigma_{i=1}^N \mathbf{p}^i \ ; \tag{6.84}$$

agreeing with our previous solution, eq. 6.31, based on the differential equation eq. 6.24.

(2): *Angular momentum of a single particle*:—
$SO(3)$ acts on $\mathbb{R}^3$ by $\Phi(A, \mathbf{q}) = A\mathbf{q}$. Writing $\xi \in \mathfrak{so}(3)$ as $\xi = \Theta\omega$ (cf. eq. 3.19, 3.23 and 3.77), the infinitesimal generator is

$$\xi_{\mathbb{R}^3}(\mathbf{q}) = \xi\mathbf{q} = \omega \wedge \mathbf{q} \ . \tag{6.85}$$

So applying eq. 6.75, the equivariant momentum map $\mathbf{J} : T^*\mathbb{R}^3 \to \mathfrak{so}(3) \cong \mathbb{R}^3$ is given by

$$< \mathbf{J}(\mathbf{q}, \mathbf{p}); \omega > = < \mathbf{p}; \omega \wedge \mathbf{q} > = \mathbf{p} \cdot (\omega \wedge \mathbf{q}) = \omega \cdot (\mathbf{q} \wedge \mathbf{p}) \ , \quad \text{i.e.} \quad \mathbf{J}(\mathbf{q}, \mathbf{p}) = \mathbf{q} \wedge \mathbf{p} \ ; \tag{6.86}$$

agreeing with our previous solution, eq. 6.37, based on the differential equation eq. 6.24.

(3): *The cotangent lift of left and right translations on G*:—
Recalling eq. 4.42, viz. that the infinitesimal generator of left translation is

$$\xi_G(g) = (T_e R_g)\xi \ , \tag{6.87}$$

a *right*-invariant vector field, and applying eq. 6.75, we see that the momentum map $\mathbf{J}_L : T^*G \to \mathfrak{g}^*$ for the cotangent lift of left translation is given by

$$< \mathbf{J}_L(\alpha_g); \xi > \; = \; < \alpha_g; \xi_G(g) > \; = \; < \alpha_g; (T_e R_g)\xi > \; = \; < (T_e^* R_g)(\alpha_g); \xi > \tag{6.88}$$

where the last equation applies the definition of the cotangent lift eq. 4.6. That is: the equivariant momentum map is

$$\mathbf{J}_L(\alpha_g) = T_e^* R_g(\alpha_g) \ . \tag{6.89}$$

In words: the momentum map $\mathbf{J}_L$ of the cotangent lift of left translation is the cotangent lift of right translation.

In a similar way, we could consider right translation: $R_g : h \mapsto hg$. Right translation defines a right action on G, has $\xi_G(g) = (T_e L_g)\xi$ as its infinitesimal generator, and so has

$$\mathbf{J}_R : T^*G \to \mathfrak{g}^* \ \ ; \ \ \mathbf{J}_R(\alpha_g) := T_e^* L_g(\alpha_g) \tag{6.90}$$

as the momentum map of its cotangent lift. Note that this momentum map is equivariant with respect to Ad_g^*: which, as discussed after eq. 4.76, is a *right* action.

7 Reduction

7.1 Preamble

In this final Section, the themes of Section 2 onwards come together—at last! As announced in Section 5.1, we will concentrate on proving what is nowadays called the *Lie-Poisson reduction theorem*: that is, the isomorphism of Poisson manifolds

$$T^*G/G \cong \mathfrak{g}^* \ . \tag{7.1}$$

Here the quotient of T^*G is by the cotangent lift of G's action on itself by left translation.

As it happens, this Chapter's main sources (i.e. Abraham and Marsden (1978), Arnold (1989), Olver (2000) and Marsden and Ratiu (1999)) do not contain what is surely the most direct proof of this result. So we give it in Section 7.2. The result will follow directly from four previous main results, one from Section 5 and three from Section 6.

'Directly', but for one wrinkle! This relates to "flipping" between left and right translation, and their various lifts. In short: the four previous results show that T^*G/G is isomorphic as a Poisson manifold, *not* to $\mathfrak{g}^*$ with the Lie-Poisson bracket familiar since eq. 5.24 and 5.27, but instead to $\mathfrak{g}^*$ equipped with this bracket's *negative*, i.e. equipped with

$$\{F, H\}_-(x) \ := \ - < x; [\nabla F(x), \nabla H(x)] > \ , \ \ x \in \mathfrak{g}^* \ . \tag{7.2}$$

But we shall (mercifully!) not reproduce, with minus signs appropriately added, our entire discussion of the Lie-Poisson bracket that ensued after eq. 5.24; (exercise for the reader!).

To avoid ambiguity, we shall sometimes write $\mathfrak{g}^*_+$ for $\mathfrak{g}^*$ equipped with the positive Lie-Poisson bracket of eq. 5.27; and $\mathfrak{g}^*_-$ for $\mathfrak{g}^*$ equipped with the negative Lie-Poisson bracket of eq. 7.2.

In fact, it will be clearest from now on, to treat right actions on a par with left actions; despite our previous emphasis on the latter. This will mean that we will also treat right-invariant vector fields (and another notion of right-invariance defined in Section 7.3.1) on a par with left-invariant vector fields (and Section 7.3.1's corresponding new notion of left-invariance). Indeed, we have already glimpsed this would be necessary in:

(i): Section 4.4's result that the infinitesimal generator of left translation is a right-invariant vector field, and vice versa (eq. 4.42, 4.43); and its corollaries in Example (3) of Section 6.5.3, that

(ii): the momentum map $\mathbf{J}_L$ of the cotangent lift of left translation is the cotangent lift of right translation; (eq. 6.89); and

(iii): the momentum map $\mathbf{J}_R$ of the cotangent lift of right translation is the cotangent lift of left translation; (eq. 6.90).

So by the end of Section 7.2, we will have a short proof of the Lie-Poisson reduction theorem. But (as often happens), the most direct proof does not give very much information about the situation. So in Section 7.3 we give more information (following Marsden and Ratiu (1999)). Then in Section 7.4, we discuss the reduction of dynamics (as against Poisson structure) from T^*G to $\mathfrak{g}^*$.

Finally, in Section 7.5 we state another reduction theorem, which is cast in terms of symplectic, not Poisson, manifolds—but which uses several notions from Section 3, such as free and proper actions, and isotropy groups. But we do not prove this theorem: we include it mostly in order to emphasize our previous remark, that (despite its length!) this Chapter just scratches the surface of the subject. We also discuss the relation between it and the Lie-Poisson reduction theorem.

7.2 The Lie-Poisson Reduction Theorem

First we recall from the end of Section 4.6.2 (eq. 4.112) that $\bar{\lambda} : T^*G \to G \times \mathfrak{g}^*$ is an equivariant map relating the cotangent lifted left action of left translation on T^*G to the G-action on $G \times \mathfrak{g}^*$ given just by left translation on the first component. So we passed to the quotients, and defined $\hat{\bar{\lambda}} : T^*G/G \to (G \times \mathfrak{g}^*)/G$ by eq. 4.116, viz.

$$\hat{\bar{\lambda}} : \mathrm{Orb}(\alpha) \equiv \{\beta \in T^*G \mid \beta = T^*L_{h^{-1}}(\alpha), \text{ some } h \in G\} \mapsto \tag{7.3}$$

$$\mathrm{Orb}(\bar{\lambda}(\alpha)) \equiv \{(hg, (T_e^*L_g)(\alpha)) \mid \text{some } h \in G\} \equiv \{(h, (T_e^*L_g)\alpha) \mid \text{some } h \in G\}\ . \tag{7.4}$$

where $\alpha \in T_g^*G$, so that $T^*L_{h^{-1}}\alpha \in T_{hg}^*G$. Finally, we identified $(G \times \mathfrak{g}^*)/G$ with $\mathfrak{g}^*$, so that the diffeomorphism $\hat{\bar{\lambda}}$ maps T^*G/G to $\mathfrak{g}^*$, as in eq. 4.117:

$$\hat{\bar{\lambda}} : \mathrm{Orb}(\alpha) \equiv \{\beta \in T^*G \mid \beta = T^*L_{h^{-1}}(\alpha), \text{ some } h \in G\} \in T^*G/G \mapsto (T_e^*L_g)(\alpha) \in \mathfrak{g}^*. \tag{7.5}$$

So now, we are to show that the diffeomorphism $\hat{\bar{\lambda}} : T^*G/G \to \mathfrak{g}^*$ is a Poisson map, in the sense of eq. 5.52 (Section 5.3.2). So we need to show:

(i): T^*G/G is a Poisson manifold;

(ii): $\hat{\bar{\lambda}}$ maps (i)'s Poisson structure on T^*G/G to that of $\mathfrak{g}^*$. In fact, as announced in Section 7.1, $\hat{\bar{\lambda}}$ maps on to the Poisson structure of $\mathfrak{g}^*_-$, i.e. as given by eq. 7.2.

Prima facie, there could be a judicious choice to be made about (i), i.e. about how to define the Poisson structure on T^*G/G, so as to secure (ii), i.e. so that $\hat{\bar{\lambda}}$ respects the Poisson structure. But in fact our previous work gives a pre-eminently obvious choice—which works. Namely: we use the Poisson structure induced on T^*G/G by the Poisson reduction theorem of Section 5.5. The result follows directly by combining with this theorem, three results from Section 6:

(i): that equivariant momentum maps are Poisson; eq. 6.55 in Section 6.4.2;

(ii): that a cotangent lifted left action has an equivariant momentum map; eq. 6.75 in Section 6.5.2;

(iii): that the momentum maps of the cotangent lifts of left and right translation on G are $\mathbf{J}_L = T_e^* R_g$ and $\mathbf{J}_R = T_e^* L_g$; eq. 6.89 and 6.90 in Section 6.5.3.

In particular, combining (i)-(iii): one deduces (exercise!) that $\mathbf{J}_R = T_e^* L_g$ is equivariant with respect to Ad_g^*, and so Poisson with respect to the negative Lie-Poisson bracket (eq. 7.2's bracket) on $\mathfrak{g}^*$. That is: it is Poisson with the codomain $\mathfrak{g}^*_-$.

Thus we have the

Lie-Poisson reduction theorem The diffeomorphism $\hat{\hat{\lambda}} : T^*G/G \to \mathfrak{g}^*$:

$$\hat{\hat{\lambda}} : \mathrm{Orb}(\alpha) \equiv \{\beta \in T^*G \mid \beta = T^*L_{h^{-1}}(\alpha), \text{ some } h \in G\} \in T^*G/G \mapsto (T_e^* L_g)(\alpha) \in \mathfrak{g}^* \tag{7.6}$$

is Poisson.

Proof: First, eq. 7.6 means we have a commutative triangle. For with $\pi : T^*G \to T^*G/G$ the canonical projection, the momentum map $\mathbf{J}_R : T^*G \to \mathfrak{g}^*, \alpha_g \mapsto (T_e^* L_g)\alpha_g$ is equal to $\hat{\hat{\lambda}} \circ \pi$:

$$T^*G \xrightarrow{\pi} T^*G/G \xrightarrow{\hat{\hat{\lambda}}} \mathfrak{g}^* \ . \tag{7.7}$$

Since left translation is a diffeomorphism of G, and the cotangent lift of any diffeomorphism of a manifold to its cotangent bundle is symplectic (cf. after eq. 4.5 in Section 4.1.A), the Poisson reduction theorem of Section 5.5 applies. That is, there is a unique Poisson structure on T^*G/G such that π is Poisson. We also know from eq. 6.75, 6.55 and 6.90 that $\mathbf{J}_R = T_e^* L_g$ is Poisson with respect to eq. 7.2's bracket on $\mathfrak{g}^*$.

We can now deduce that $\hat{\hat{\lambda}}$ is Poisson, i.e. that for all $x \in T^*G/G$ and all $F, H \in \mathcal{F}(\mathfrak{g}^*_-)$

$$(\{F, H\}_{\mathfrak{g}^*_-} \circ \hat{\hat{\lambda}})\,(x) \;=\; \{F \circ \hat{\hat{\lambda}}, H \circ \hat{\hat{\lambda}}\}_{T^*G/G}\,(x) \ . \tag{7.8}$$

We just use (in order) the facts that:

(i): π is surjective, so that for all $x \in T^*G/G$ there is an $\alpha_g \in T^*G$ with $x = \pi(\alpha_g) \equiv \mathrm{Orb}(\alpha_g)$;

(ii): $\mathbf{J}_R = \hat{\hat{\lambda}} \circ \pi$;

(iii): $\mathbf{J}_R$ is Poisson; and

(iv): π is Poisson:

$$(\{F, H\}_{\mathfrak{g}^*_-} \circ \hat{\hat{\lambda}})\,(x) \;=\; \{F, H\}_{\mathfrak{g}^*_-} \circ (\hat{\hat{\lambda}} \circ \pi)\,(\alpha_g) \tag{7.9}$$

$$=\; \{F, H\}_{\mathfrak{g}^*_-} \circ \mathbf{J}_R\,(\alpha_g) \;=\; \{F \circ \mathbf{J}_R, H \circ \mathbf{J}_R\}_{T^*G}\,(\alpha_g) \tag{7.10}$$

$$=\; \{F \circ \hat{\hat{\lambda}}, H \circ \hat{\hat{\lambda}}\}_{T^*G/G}\,(\pi(\alpha_g)) \;\equiv\; \{F \circ \hat{\hat{\lambda}}, H \circ \hat{\hat{\lambda}}\}_{T^*G/G}\,(x) \ . \ \text{QED.} \tag{7.11}$$

7.3 Meshing with the symplectic structure on T^*G: invariant functions

We turn to giving more information about the situation described by the Lie-Poisson reduction theorem. The general idea will be that the Lie-Poisson bracket on $\mathfrak{g}^*$ meshes with the canonical symplectic structure on T^*G. This will be made precise in two ways: the first is discussed in the first two Subsections, the second is discussed in the third Subsection.

The first discussion will have three stages:

(i): we show that scalars on $\mathfrak{g}^*$, $F \in \mathcal{F}(\mathfrak{g}^*)$, are in one-one correspondence with scalars on T^*G that are constant on the orbits of the cotangent lift of left translation, which will be called *left-invariant* functions; and similarly, for the cotangent lift of right translation (a correspondence with *right-invariant* functions);

(ii): we take the usual canonical Poisson bracket in T^*G of these left-invariant or right-invariant scalars; and restrict this bracket to $\mathfrak{g}^*$ regarded as the cotangent space T^*_eG at the identity $e \in G$; and then

(iii): we show that this restriction is the Lie-Poisson bracket on $\mathfrak{g}^*$: the familiar positive one for *right*-invariant functions, and the new negative one of eq. 7.2 for the *left*-invariant functions.

We do stages (i) and (ii) in Section 7.3.1. These stages will not involve the choice between the positive and negative Lie-Poisson brackets. But stage (iii), in Section 7.3.2, will involve this choice. It will be a one-liner corollary of Section 6.4.2's result that equivariant momentum maps are Poisson maps, eq. 6.55; (unsurprisingly, in that we also used this result in Section 7.2's proof of the reduction theorem).

In the third Subsection, we use invariant functions to show a different sense in which the Lie-Poisson bracket on $\mathfrak{g}^*$ meshes with the symplectic structure on T^*G. Namely, we derive the Lie-Poisson bracket on $\mathfrak{g}^*$ from the Poisson reduction theorem of Section 5.5, by using the ideas of invariant functions and momentum functions.

7.3.1 Left-invariant and right-invariant functions on T^*G

We say that a function $F : T^*G \to \mathbb{R}$ is *left-invariant* if for all $g \in G$, and all $\alpha_g \in T^*_gG$

$$(F \circ T^*L_g)(\alpha_g) = F(\alpha_g) \tag{7.12}$$

where T^*L_g is the cotangent lift of $L_g : G \to G$. Similarly, $F : T^*G \to \mathbb{R}$ is called *right-invariant* if for all $g \in G$

$$(F \circ T^*R_g) = F \; . \tag{7.13}$$

So if $F : T^*G \to \mathbb{R}$ is left-invariant or right-invariant, it is determined by its values for arguments in $T^*_eG = \mathfrak{g}^*$.

Since any $\alpha \in \mathfrak{g}^*$ is mapped by $T^*L_{g^{-1}} \equiv (T^*L_g)^{-1}$ to an element of T^*_gG, a function is left-invariant iff it is constant on the orbits of the various T^*L_g for $g \in G$, i.e. constant on the orbits of the cotangent lift of left translation. Similarly, a function is right-invariant iff it is constant on the orbits of the cotangent lift of right translation.

So left-invariant functions induce well-defined functions on the quotient space T^*G/G; and so, by Section 7.2, on its diffeomorphic (indeed Poisson manifold) copy $\mathfrak{g}^*$. Similarly for right-invariant functions.

But let us for the moment consider the smooth left-invariant (or right-invariant) functions on T^*G, rather than the induced maps on the quotient space. We will denote the space of all smooth left-invariant functions on T^*G by $\mathcal{F}_L(T^*G)$, and similarly the space of smooth right-invariant functions by $\mathcal{F}_R(T^*G)$.

Recalling (from the discussion after eq. 4.5) that cotangent lifts are symplectic maps, i.e. T^*L_g and T^*R_g are symplectic maps on T^*G, it follows immmediately that $\mathcal{F}_L(T^*G)$ and $\mathcal{F}_R(T^*G)$ are each closed under the canonical Poisson bracket on T^*G. So they are each a Lie algebra with this bracket.

Now we can use the momentum maps $\mathbf{J}_L$ and $\mathbf{J}_R$ of Example (3) of Section 6.5.3 to extend any scalar $F : \mathfrak{g}^* \to \mathbb{R}$, i.e. $F \in \mathcal{F}(\mathfrak{g}^*)$, to a left-invariant, or right-invariant, scalar on T^*G.

Thus, given $F : \mathfrak{g}^* \to \mathbb{R}$ and $\alpha_g \in T^*_g G$, we define $F_L \in \mathcal{F}_L(T^*G)$ by

$$F_L(\alpha_g) := (F \circ \mathbf{J}_R)(\alpha_g) \equiv (F \circ T^*_e L_g)(\alpha_g) \ . \tag{7.14}$$

So F_L is by construction left-invariant, and is called the *left-invariant extension* of F from $\mathfrak{g}^*$ to T^*G.

One similarly defines the *right-invariant extension* $F_R \in \mathcal{F}_R(T^*G)$ of any $F \in \mathcal{F}(\mathfrak{g}^*)$ by

$$F_R(\alpha_g) := (F \circ \mathbf{J}_L)(\alpha_g) \equiv (F \circ T^*_e R_g)(\alpha_g) \ . \tag{7.15}$$

Then the maps

$$F \in \mathcal{F}(\mathfrak{g}^*) \mapsto F_L \in \mathcal{F}_L(T^*G) \ \text{ and } \ F \in \mathcal{F}(\mathfrak{g}^*) \mapsto F_R \in \mathcal{F}_R(T^*G) \tag{7.16}$$

are vector space isomorphisms (exercise for the reader!) whose inverse is just restriction to the fiber $T^*_e G = \mathfrak{g}^*$.

This completes what we called 'stages (i) and (ii)': describing a correspondence between scalars on $\mathfrak{g}^*$ and scalars on T^*G that are constant on the orbits of the cotangent lifts of left and right translation; and considering the canonical Poisson bracket (on T^*G) of these scalars, i.e. the Lie algebras $\mathcal{F}_L(T^*G)$ and $\mathcal{F}_R(T^*G)$.

7.3.2 Recovering the Lie-Poisson bracket

We now do stage (iii): we show that the restriction of the canonical Poisson bracket on T^*G of the right/left invariant functions, to $\mathfrak{g}^*$ regarded as the cotangent space $T^*_e G$ at the identity $e \in G$, is the positive/negative Lie-Poisson bracket.

Since the inverses of the maps eq. 7.16 are just restriction to the fiber $T^*_e G = \mathfrak{g}^*$, it suffices to show that the maps eq. 7.16 are Lie algebra isomorphisms. More precisely:

> **Recovery of the Lie-Poisson bracket** Using the positive Lie-Poisson bracket on $\mathfrak{g}^*$ (we write $\mathfrak{g}^*_+$): $F \mapsto F_R$ is a Lie algebra isomorphism.
> Similarly: using the negative Lie-Poisson bracket on $\mathfrak{g}^*$ (we write $\mathfrak{g}^*_-$): $F \mapsto F_L$ is a Lie algebra isomorphism.

That is: for all $F, H \in \mathcal{F}(\mathfrak{g}^*)$

$$\{F, H\}_+ = \{F_R, H_R\}_{T^*G} \mid_{\mathfrak{g}^*} \;\; ; \;\; \{F, H\}_- = \{F_L, H_L\}_{T^*G} \mid_{\mathfrak{g}^*} \qquad (7.17)$$

Proof: Consider $\mathbf{J}_L : T^*G \to \mathfrak{g}^* \equiv \mathfrak{g}^*_+$, $\mathbf{J}_L = T^*_e R_g$. $\mathbf{J}_L$ is an equivariant momentum map. So, by the result eq. 6.55 of Section 6.4.2, it is Poisson. That is:

$$\{F, H\}_+ \circ \mathbf{J}_L = \{F \circ \mathbf{J}_L, H \circ \mathbf{J}_L\}_{T^*G} = \{F_R, H_R\}_{T^*G} \,. \qquad (7.18)$$

Restricting eq. 7.18 to $\mathfrak{g}^*$ gives the first equation of eq. 7.17.

Similarly, one proves the second equation by using the fact that $\mathbf{J}_R : T^*G \to \mathfrak{g}^* \equiv \mathfrak{g}^*_-$, $\mathbf{J}_R = T^*_e L_g$ is an equivariant momentum map and so is Poisson. That is:

$$\{F, H\}_- \circ \mathbf{J}_R = \{F \circ \mathbf{J}_R, H \circ \mathbf{J}_R\}_{T^*G} = \{F_L, H_L\}_{T^*G} \,. \qquad (7.19)$$

We then restrict eq. 7.19 to $\mathfrak{g}^*$. QED.

7.3.3 Deriving the Lie-Poisson bracket

Our discussion so far, in both Section 7.2 and the two previous Subsections, has taken the Lie-Poisson bracket (whether positive or negative) as given. We now show, using invariant functions and Section 6.5.1's idea of momentum functions, how to *derive* the Lie-Poisson bracket on $\mathfrak{g}^*$.

So this derivation will amount to another, more "constructive", proof of the Lie-Poisson reduction theorem. As in Section 7.2's proof, two main ingredients will be:

(a): the diffeomorphism $\hat{\bar{\lambda}}$ between T^*G/G and $\mathfrak{g}^*$ (eq. 4.117 or 7.5 or 7.6), and

(b): the Poisson reduction theorem of Section 5.5, applied to G's action on T^*G.

But instead of Section 7.2's proof's using the facts that (i) the momentum maps $\mathbf{J}_R \equiv T^*_e L_g$ and $\mathbf{J}_L \equiv T^*_e R_g$ are equivariant and (ii) equivariant momentum maps are Poisson, we will now use the ideas of invariant functions and momentum functions.

We begin by recalling that (since left translation is a diffeomorphism of G, and the cotangent lift of any diffeomorphism of a manifold to its cotangent bundle is symplectic), the Poisson reduction theorem implies that there is a unique Poisson structure on T^*G/G such that $\pi : T^*G \to T^*G/G$ is Poisson. We now use the diffeomorphism $\hat{\bar{\lambda}} : T^*G/G \to \mathfrak{g}^*$ to transfer this Poisson structure to $\mathfrak{g}^*$. Let us call the result $\{,\}_-$. Though this is not to be read (yet!) as the negative Lie-Poisson bracket, our aim now is to calculate that it is in fact this bracket.

Notice first that since the momentum map $\mathbf{J}_R : T^*G \to \mathfrak{g}^*, \alpha_g \mapsto (T^*_e L_g)\alpha_g$ is equal to $\hat{\bar{\lambda}} \circ \pi$ (eq. 7.7), we know that $\mathbf{J}_R$ is Poisson with respect to this induced bracket on $\mathfrak{g}^*$. That is

$$\{F, H\}_- \circ \mathbf{J}_R\,(\alpha_g) \;=\; \{F \circ \mathbf{J}_R, H \circ \mathbf{J}_R\}_{T^*G}\,(\alpha_g) \;=\; \{F_L, H_L\}_{T^*G}\,(\alpha_g) \,. \qquad (7.20)$$

To calculate the right hand side, we will apply the ideas of invariant functions and momentum functions to each argument of the bracket; in particular to the first:

$$F_L(\alpha_g) = F(T_e^* L_g \cdot \alpha_g) \ . \tag{7.21}$$

We observe that since a Poisson bracket depends only on the values of first derivatives, we can replace $F \in \mathcal{F}(\mathfrak{g}^*)$ by its linearization. That is, we can assume F is linear, so that at any point $\alpha \in \mathfrak{g}^*$, $F(\alpha) =< \alpha; \nabla F >$, where ∇F is a constant in $\mathfrak{g} \equiv \mathfrak{g}^{**}$. Applying this, and the definition of a momentum function eq. 6.59, to eq. 7.21, we get:

$$F(T_e^* L_g \cdot \alpha_g) \ =< T_e^* L_g \cdot \alpha_g; \nabla F > = < \alpha_g; T_e L_g \cdot \nabla F > = \ \mathcal{P}(X_{\nabla F})(\alpha_g) \ , \tag{7.22}$$

where the last equation applies the definition of a momentum function to the left-invariant vector field on G, $X_\xi(g) \equiv T_e L_g(\xi)$, for the case $\xi = \nabla F$.

Now we apply to eq. 7.22, in order: eq. 6.64, the definition of the Lie algebra bracket (cf. eq. 3.46), eq. 6.59 again, and the definition of left-invariant vector fields. We get:

$$\{F_L, H_L\}_{T^*G} \, (\alpha_g) \ = \ \{\mathcal{P}(X_{\nabla F}), \mathcal{P}(X_{\nabla H})\}_{T^*G} \, (\alpha_g) \ = \ -\mathcal{P}([X_{\nabla F}, X_{\nabla H}])(\alpha_g) \tag{7.23}$$

$$= \ -\mathcal{P}(X_{[\nabla F, \nabla H]})(\alpha_g) \ = \ - < \alpha_g; X_{[\nabla F, \nabla H]} > \tag{7.24}$$

$$= \ - < \alpha_g; T_e L_g([\nabla F, \nabla H]) > = \ - < T_e^* L_g(\alpha_g); [\nabla F, \nabla H] > \ . \tag{7.25}$$

Combining eq. 7.20 and eq. 7.25, and writing $\alpha \in \mathfrak{g}^*$ for $(T_e^* L_g)\alpha_g \equiv \mathbf{J}_R(\alpha_g)$, we have our result:

$$\{F, H\}_-(\alpha) = - < \alpha; [\nabla F, \nabla H] > \ . \tag{7.26}$$

One similarly derives the positive Lie-Poisson bracket by considering right-invariant extensions of linear functions. The minus sign coming from eq. 6.64 is cancelled by the sign reversal in the Lie bracket of right-invariant vector fields. That is, it is cancelled by a minus sign coming from eq. 3.58.

7.4 Reduction of dynamics

We end our account of the Lie-Poisson reduction theorem by discussing the reduction of dynamics from T^*G to $\mathfrak{g}^*$.

We can be brief since we have already stated the main idea, when discussing the Poisson reduction theorem; cf. (2)(A) in Section 5.5. Thus recall that (under the conditions of the theorem) a G-invariant Hamiltonian function on a Poisson manifold M, $H : M \to \mathbb{R}$, defines a corresponding function h on M/G by $H = h \circ \pi$, where π is the projection $\pi : M \to M/G$; and since π is Poisson, and so pushes Hamiltonian flows forward to Hamiltonian flows, π pushes X_H on M to X_h on M/G:

$$T\pi \circ X_H = X_h \circ \pi \ . \tag{7.27}$$

Applying this, in particular eq. 7.27, to the Lie-Poisson reduction theorem, we get

Reduction of dynamics Let $H : T^*G \to \mathbb{R}$ be left-invariant. That is: the function $H^- := H\mid_{\mathfrak{g}^*}$ on $\mathfrak{g}^*$ satisfies

$$H(\alpha_g) = H^-(\mathbf{J}_R(\alpha_g)) \equiv H^-(T_e^* L_g \cdot \alpha_g) \ , \quad \alpha_g \in T_g^* G \ . \tag{7.28}$$

Then $\mathbf{J}_R$ pushes X_H forward to X_{H^-}. Or in terms of the flows $\phi(t)$ and $\phi^-(t)$ of X_H and X_{H^-} respectively:

$$\mathbf{J}_R(\phi(t)(\alpha_g)) \ = \ \phi^-(t)(\mathbf{J}_R(\alpha_g)) \ . \tag{7.29}$$

Similar statements hold for a right-invariant function $H : T^*G \to \mathbb{R}$, its restriction $H^+ := H\mid_{\mathfrak{g}^*}$ and $\mathbf{J}_L \equiv T_e^* R_g$.

Besides, we already know the vector field of H^- on $\mathfrak{g}^*$. For eq. 5.30 in (3) of Section 5.2.4 gave a basis-independent expression of Hamilton's equations on $\mathfrak{g}^*$ in terms of ad^*. We just need to note that since we are now using the negative Lie-Poisson bracket on $\mathfrak{g}^*$, all terms in the deduction (eq. 5.29) apart from the left hand side, get a minus sign. So writing $\alpha \in \mathfrak{g}^*$, eq. 5.30 for the vector field X_{H^-} becomes:

$$\frac{d\alpha}{dt} \ = \ - \ ad^*_{\nabla H^-(\alpha)}(\alpha) \ . \tag{7.30}$$

On the other hand, we can go in the other direction, reconstructing the dynamics on T^*G from eq. 7.30 on $\mathfrak{g}^*$. The statement of the main result, below, is intuitive, in that the "reconstruction equation" for $g(t) \in T^*G$ is

$$g^{-1}\dot{g} = \nabla H^- \ . \tag{7.31}$$

This is intuitive since it returns us to the basic idea of mechanics on $\mathfrak{g}$ and $\mathfrak{g}^*$, viz. that the map

$$\lambda_g : \dot{g} \in T_g G \mapsto \lambda_g(\dot{g}) := (T_g L_{g^{-1}})\dot{g} \in \mathfrak{g} \tag{7.32}$$

maps the generalized velocity to its body representation; cf. eq. 4.90. However, the proof of this result is involved (Marsden and Ratiu (1999: theorems 13.4.3, 13.4.4, p. 423-426); so we only state the result. It is:—

Reconstruction of dynamics Suppose given a Lie group G, a left-invariant $H : T^*G \to \mathbb{R}$, its restriction $H^- := H\mid_{\mathfrak{g}^*}$, and an integral curve $\alpha(t)$ of the Lie-Poisson Hamilton's equations eq. 7.30 on $\mathfrak{g}^*$, with the initial condition $\alpha(0) = T_e^* L_{g_0}(\alpha_{g_0})$. Then the integral curve in T^*G of X_H is given by

$$T^*_{g(t)} L_{g(t)^{-1}}(\alpha(t)) \ ; \tag{7.33}$$

where $g(t)$ is the solution of the reconstruction equation

$$g^{-1}\dot{g} = \nabla H^- \tag{7.34}$$

with initial condition $g(0) = g_0$.

7.5 Envoi: the Marsden-Weinstein-Meyer theorem

I emphasize that our discussion of reduction has only scratched the surface: after all this Section has been relatively short! But now that the reader is armed with the long and leisurely exposition from Section 3 onwards, they are well placed to pursue the topic of reduction; e.g. through this Chapter's main sources, Abraham and Marsden (1978), Arnold (1989), Olver (2000) and Marsden and Ratiu (1999).

In particular, the reader can now relate the Lie-Poisson reduction theorem to another main theorem about symplectic reduction, usually called the Marsden-Weinstein-Meyer or Marsden-Weinstein theorem (after these authors' papers in 1973 and 1974).

This theorem concerns a symplectic action of a Lie group G on a symplectic manifold (M, ω). For the sake of completeness, and to orient the reader to Landsman's discussion of this theorem (this vol., ch. 5, especially Section 4.5), it is worth stating it (as usual, for the finite-dimensional case only), together with the lemma used to prove it, and the ensuing reduction of dynamics. These statements will also round off our discussion by illustrating how some notions expounded from Section 3 onwards, but not used in this Section, are nevertheless useful—e.g. in stating the hypotheses of this theorem.

So suppose the Lie group G acts symplectically (eq. 6.1) on the symplectic manifold (M, ω); and that $\mathbf{J} : M \to \mathfrak{g}^*$ is an Ad^*-equivariant momentum map for this action (eq. 6.47 and 6.52). Assume also that $\alpha \in \mathfrak{g}^*$ is a regular value of $\mathbf{J}$, i.e. that at every point $x \in \mathbf{J}^{-1}(\alpha)$, $T_x\mathbf{J}$ is surjective. So the submersion theorem of (1) of Section 3.3.1 applies; in particular, $\mathbf{J}^{-1}(\alpha)$ is a sub-manifold of M with dimension $\dim(M)$ - $\dim(\mathfrak{g}^*)$ $\equiv \dim(M)$ - $\dim(G)$.

Let G_α be the isotropy group (eq. 4.33) of α under the co-adjoint action, i.e.

$$G_\alpha := \{g \in G \mid Ad^*_{g^{-1}}\alpha = \alpha\}. \tag{7.35}$$

So since $\mathbf{J}$ is Ad^*-equivariant under G_α, the quotient space $M_\alpha := \mathbf{J}^{-1}(\alpha)/G_\alpha$ is well-defined.

Now assume that G_α acts freely and properly on $\mathbf{J}^{-1}(\alpha)$, so that (Section 4.3.B) the quotient space $M_\alpha = \mathbf{J}^{-1}(\alpha)/G_\alpha$ is a manifold. M_α is the *reduced phase space* (corresponding to the momentum value α).

Now we assert:

Marsden-Weinstein-Meyer theorem M_α has a natural symplectic form ω_α induced from (M, ω) as follows. Let u, v be two vectors tangent to M_α at some point $p \in M_\alpha$: so p is an orbit of G_α's action on $\mathbf{J}^{-1}(\alpha)$, and $u, v \in T_pM_\alpha$. Then u and v are obtained, respectively, from some vectors u' and v' tangent to $\mathbf{J}^{-1}(\alpha)$ at some point $x \in \mathbf{J}^{-1}(\alpha)$ of the orbit p, by the projection $\pi_\alpha : \mathbf{J}^{-1}(\alpha) \to M_\alpha$. That is:

$$T\pi_\alpha(u') = u \ ; \ T\pi_\alpha(v') = v \ . \tag{7.36}$$

It turns out that the value assigned by M's symplectic form ω is the same whatever choice of x, u', v' is made. So we define the symplectic form ω_α on M_α as assigning this value. In other words: writing π_α for the projection, $i_\alpha : \mathbf{J}^{-1}(\alpha) \to M$ for the inclusion, and * for pullback:

$$\pi_\alpha^* \omega_\alpha \; = \; i_\alpha^* \omega \; . \tag{7.37}$$

The proof of this theorem uses the following Lemma. Let us write $G \cdot x$ for the orbit $\mathrm{Orb}(x)$ of x under the action of all of G, and similarly $G_\alpha \cdot x$ for the orbit under G_α, i.e. $\{\Phi(g, x) \mid g \in G_\alpha\}$. Then the Lemma states:

For any $x \in \mathbf{J}^{-1}(\alpha)$:—
(i): $T_x(G_\alpha \cdot x) = T_x(G \cdot x) \cap T_x(\mathbf{J}^{-1}(\alpha))$; and
(ii): $T_x(G \cdot x)$ and $T_x(\mathbf{J}^{-1}(\alpha))$ are ω-orthogonal complements of one another in TM. That is: for all $u' \in T_xM$:
$u' \in T_x(\mathbf{J}^{-1}(\alpha))$ iff $\omega(u', v') = 0$ for all $v' \in T_x(G \cdot x)$.

Both the Lemma and the theorem are each proven in some dozen lines. For details, cf. Abraham and Marsden (1978: Theorems 4.3.1-2, p. 299-300), or Arnold (1989: Appendix 5.B, p. 374-376).

Two final remarks. (1): The reduction of dynamics secured by the Marsden-Weinstein-Meyer theorem is similar to what we have seen before, for both the Poisson reduction theorem ((2) of Section 5.5), and the Lie-Poisson reduction theorem (Section 7.4). One proves, again in a few lines (Abraham and Marsden (1978: Theorems 4.3.5, p. 304):

Marsden-Weinstein-Meyer reduction of dynamics Let $H : M \to \mathbb{R}$ be invariant under the action of G on M, so that by Noether's theorem for momentum maps (Section 6.2) $\mathbf{J}$ is conserved, i.e. $\mathbf{J}^{-1}(\alpha)$ is invariant under the flow $\phi(t)$ of X_H on M. Then $\phi(t)$ commutes with the action of G_α on $\mathbf{J}^{-1}(\alpha)$ (i.e. $\phi(t) \circ \Phi_g = \Phi_g \circ \phi(t)$ for $g \in G_\alpha$), and so defines a flow $\hat{\phi}(t)$ on M_α such that $\pi_\alpha \circ \phi(t) = \hat{\phi}(t) \circ \pi_\alpha$, i.e.

$$\begin{array}{ccc} \mathbf{J}^{-1}(\alpha) & \xrightarrow{\pi_\alpha} & M_\alpha \\ \uparrow{\scriptstyle \phi(t)} & & \uparrow{\scriptstyle \hat{\phi}(t)} \\ \mathbf{J}^{-1}(\alpha) & \xrightarrow{\pi_\alpha} & M_\alpha \end{array} \tag{7.38}$$

The flow $\hat{\phi}(t)$ is Hamiltonian with the Hamiltonian H_α defined by $H_\alpha \circ \pi_\alpha = H \circ i_\alpha$.

(2): I said at the start of this Subsection that the reader can now relate the Lie-Poisson reduction theorem to the Marsden-Weinstein-Meyer theorem. It is not hard to show that the former is an example of the latter. As the symplectic manifold M

one takes T^*G, acted on symplectically by the cotangent lift of left translation. So we know (from (3) of Section 6.5.3) that $\mathbf{J}_L := T_e^* R_g$ is an Ad^*-equivariant momentum map ... and so on: I leave this as an exercise for the reader! The answer is supplied at Arnold (1989: 377, 321) and Abraham and Marsden (1978: 302). (Abraham and Marsden call it the 'Kirillov-Kostant-Souriau theorem'.)

Suffice it to say here that this exercise gives another illustration of one of our central themes, that $\mathfrak{g}^*$'s symplectic leaves are the orbits of the co-adjoint representation. For the reduced phase space M_α is naturally identifiable with the co-adjoint orbit $\mathrm{Orb}(\alpha)$ of $\alpha \in \mathfrak{g}^*$, with the symplectic forms also naturally identified; (cf. also result (2) at the end of Section 5.4).

Acknowledgements:— I am grateful to audiences in Irvine, Oxford, Princeton and Santa Barbara; to several colleagues for encouragement; and to Gordon Belot, Klaas Landsman, Barrie Tonkinson, David Wallace, and especially Graeme Segal, for helpful, and patient!, conversations, correspondence and corrections.

8 References

R. Abraham and J. Marsden (1978), *Foundations of Mechanics*, second edition: Addison-Wesley.

V. Arnold (1973), *Ordinary Differential Equations*, MIT Press.

V. Arnold (1989), *Mathematical Methods of Classical Mechanics*, Springer, (second edition).

J. Barbour and B. Bertotti (1982), 'Mach's principle and the structure of dynamical theories', *Proceedings of the Royal Society of London* **A 382**, p. 295-306.

G. Belot (1999), 'Relationism rehabilitated', *International Studies in Philosophy of Science* **13**, p. 35-52.

G. Belot (2000), 'Geometry and motion', *British Journal for the Philosophy of Science* **51**, p. 561-596.

G. Belot (2001), 'The principle of sufficient reason', *Journal of Philosophy* **98**, p. 55-74.

G. Belot (2003), 'Notes on symmetries', in Brading and Castellani (ed.s) (2003), pp. 393-412.

G. Belot (2003a), 'Symmetry and gauge freedom', *Studies in the History and Philosophy of Modern Physics* **34**, p. 189-225.

G. Belot (this volume).

G. Belot and J. Earman (2001), 'Pre-Socratic quantum gravity', in C. Callender and N. Huggett (ed.s) *Physics meets Philosophy at the Planck Scale*, Cambridge University Press, pp. 213-255.

R. Bishop and S. Goldberg (1980), *Tensor Analysis on Manifolds*, New York: Dover.

J. Boudri (2002), *What was Mechanical about Mechanics: the Concept of Force between Metaphysics and Mechanics from Newton to Lagrange*, Dordrecht: Kluwer Academic.

K. Brading and E. Castellani (ed.s) (2003), *Symmetry in Physics*, Cambridge University Press.

K. Brading and E. Castellani (this volume).

J. Butterfield (2004), 'Some Aspects of Modality in Analytical mechanics', in *Formal Teleology and Causality*, ed. M. Stöltzner, P. Weingartner, Paderborn: Mentis. Available at Los Alamos arXive: http://arxiv.org/abs/physics/0210081 ; and at Pittsburgh archive: http://philsci-archive.pitt.edu/archive/00001192.

J. Butterfield (2004a), 'Between Laws and Models: Some Philosophical Morals of Lagrangian Mechanics'; available at Los Alamos arXive: http://arxiv.org/abs/physics/0409030 ; and at Pittsburgh archive: http://philsci-archive.pitt.edu/archive/00001937/.

J. Butterfield (2005), 'Between Laws and Models: Some Philosophical Morals of Hamiltonian Mechanics', in preparation.

J. Butterfield (2006), 'On Symmetry and Conserved Quantities in Classical Mechanics', forthcoming in a *Festschrift* for Jeffrey Bub, ed. W. Demopoulos and I. Pitowsky, Kluwer: University of Western Ontario Series in Philosophy of Science. available at Los Alamos arXive: http://arxiv.org/abs/physics/ ; and at Pittsburgh archive: http://philsci-archive.pitt.edu/archive/00002362/

R. Darling (1994), *Differential Forms and Connections*, Cambridge University Press.

M. Dickson (this volume).

E. Desloge (1982), *Classical Mechanics*, John Wiley.

J. Earman (2003), 'Tracking down gauge: an ode to the constrained Hamiltonian formalism', in Brading and Castellani (ed.s) (2003), pp. 140-162.

H. Goldstein (1950), *Classical Mechanics*, Addison-Wesley; (1966 third printing).

I. Grattan-Guinness (2000), 'A sideways look at Hilbert's twenty-three problems of 1900', *Notices of the American Mathematical Society* **47**, p. 752-757.

I. Grattan-Guinness (2006), 'Classical mechanics as a formal(ised) science', in B. Loewe (ed.), *Foundations of the Formal Sciences*, Kluwer, to appear.

T. Hawkins (2000), *Emergence of the Theory of Lie Groups: an essay in the history of mathematics 1869-1926*, New York: Springer.

M. Jammer (1957), *Concepts of Force*, Harvard University Press.

M. Jammer (1961), *Concepts of Mass in Classical and Modern Physics*, Harvard University Press; republished by Dover in 1997.

M. Jammer (2000), *Concepts of Mass in Contemporary Physics and Philosophy*, Princeton University Press.

O. Johns (2005), *Analytical Mechanics for Relativity and Quantum Mechanics*, Oxford University Press, forthcoming.

A. Kirillov (1976), *Elements of the Theory of Representations*, Grunlehren Math. Wiss., Springer-Verlag.

H. Kragh (1999), *Quantum Generations*, Princeton University Press.

N. Landsman (this volume), 'Between Classical and Quantum'.

S. Lie (1890). *Theorie der Transformationsgruppen: zweiter abschnitt*, Leipzig: B.G.Teubner.

J. Lutzen (1995), *Denouncing Forces; Geometrizing Mechanics: Hertz's Principles of Mechanics*, Copenhagen University Mathematical Institute Preprint Series No 22.

J. Lutzen (2003), 'Between rigor and applications: developments in the concept of function in mathematical analysis', in *Cambridge History of Science, vol. 5: The modern physical and mathematical sciences*, ed. M.J. Nye, p. 468-487.

J. Lutzen (2005) *Mechanistic Images in Geometric Form: Heinrich Hertz's 'Principles of Mechanics'*, Oxford University Press.

J. Marsden and T. Hughes (1982), *Mathematical Foundations of Elasticity*, Prentice-Hall; Dover 1994.

J. Marsden and T. Ratiu (1999), *Introduction to Mechanics and Symmetry*, second edition: Springer-Verlag.

E. McMullin (1978), *Newton on Matter and Activity*, University of Notre Dame Press.

I. Newton (1687), *Principia: Mathematical Principles of Natural Philosophy*, trans. I.B. Cohen and A. Whitman, Cambridge University Press (1999).

P. Olver (2000), *Applications of Lie Groups to Differential Equations*, second edition: Springer-Verlag.

C. Rovelli (this volume), 'Quantum Gravity'.

T. Ryckman (2005), *The Reign of Relativity: Philosophy in Physics 1915-1925* Oxford University Press.

S. Singer (2001), *Symmetry in Mechanics: a Gentle Modern Introduction*, Boston: Birkhauser.

E. Slovik (2002), *Cartesian Spacetime: Descartes' Physics and the Relational Theory of Space and Motion*, Dordrecht: Kluwer Academic.

M. Stöltzner (2003), 'The Principle of Least Action as the Logical Empiricist's Shibboleth', *Studies in History and Philosophy of Modern Physics* **34B**, p. 285-318.

D. Wallace (2003), 'Time-dependent Symmetries: the link between gauge symmetries and indeterminism', in Brading and Castellani (ed.s) (2003), pp. 163-173.