Diffeomorphism groups, hydrodynamics and relativity

Jerrold E. Marsden,[1] David G. Ebin,[2] and Arthur E. Fischer[3]

Notes by Murray Cantor (through Part II) and David Lerner (Part III)

CONTENTS

[1] Part of this author's work was done while at the University of Toronto and The Institute for Advanced Study. Present address: Department of Mathematics, University of California, Berkeley, California 94720. Partially supported by the National Research Council of Canada and NSF grant GP-15735.
[2] Part of this author's work was done at the Institute for Advanced Study. Present address: Department of Mathematics, SUNY at Stony Brook, Stony Brook, N.Y. 11790.
[3] University of California, Berkeley, California, 94720. Partially supported by AEC contract AT(04-3)-34).

1. Introduction

The goal of these lectures[4] was to present some applications of global analysis to physical problems, specifically to hydrodynamics and general relativity.

Parts I and II form a unit. Only a small amount of material from Part I is needed in Part III—an acquaintance with the rudiments of the diffeomorphism groups. The sort of global analysis used in hydrodynamics is developed in Part I. The machinery needed in relativity—infinite dimensional Hamiltonian systems—is developed as it is used. This organization should make it possible for one to read Part III separately if desired.

Because of inevitable time restrictions, it was necessary to make some selection with regard to the topics and their depth of discussion. There was an effort made to include topics and points of view that would be of interest to geometers and global analysts. Nevertheless, we hope the overall perspective presented gives a reasonably correct picture of some questions of interest to specialists in theoretical hydrodynamics and general relativity.

As far as prerequisites go, we assume the following: A knowledge of the general facts about manifolds, differential forms and Riemannian geometry. For example Lang [1] and Bishop–Crittenden [1] contain more than enough background. No special knowledge of fluid mechanics is needed. In this regard we have included a certain amount of introductory material in § 2 below. One can supplement this by consulting standard texts; see for instance Feynman [1] and Landau–Lifschitz [1]. For general relativity we assume, in addition to the geometry above, some familiarity with the basic ideas of relativity; see for instance Taylor–Wheeler [1]. (This is mainly to motivate several points in the discussion.)

Much of the material on hydrodynamics is taken from Ebin–Marsden [1]. However, our exposition here is more informal and gets at several points from a different direction. The exposition regarding turbulence is largely influenced by Ruelle–Takens [1]. The material on relativity is mostly taken from Fischer-Marsden [1].

We wish to thank the many people at the congress who offered encouragement and suggestions, some of which are incorporated into these notes. In this regard, special thanks are due to E. Calabi, H. Cohen, J. Eells, D. Graff, F. Quinn, D. Norman, and R. Vanstone. We especially thank Murray Cantor and Dave Lerner for their enthusiasm and excellent notes of the lectures. The notes proved to be most useful during the preparation of the final draft. We should especially like to thank David Ruelle for many useful remarks and references on turbulence and stability, T. Lam for several related comments and S. T. Chin for assistance in proof reading. Some helpful points have also been made by A. Chorin, G. Duff, H. Hartman, T. Kato, R. Sachs, C. Simon, S. Smale, and A. Taub. Finally the organizers of the congress, especially Ray Vanstone, deserve much of the credit for the success of the seminar.

[4] The lectures were delivered by J. Marsden.

2. Basic ideas in hydrodynamics

Throughout, let M be a fixed compact, oriented, Riemannian, n-manifold, possibly with a C^∞ boundary. Intuitively, M is the space in which the fluid moves. For example, M might be the unit ball in $\mathbb{R}^3$. As an aside, for the general theory there seems to be no particular advantage to assuming M is open in $\mathbb{R}^n$. This is because the spaces of mappings of M to M that we will shortly discuss are still very nonlinear.

2.1 Some notation

A *diffeomorphism* on M is a C^∞ bijective map $\eta: M \to M$ such that η^{-1} is also C^∞.

We let $\mathscr{D} = \{$orientation preserving diffeomorphisms on $M\}$.

If the Riemannian structure is given locally by $g_{ij}: M \to \mathbb{R}$, then the *volume element* μ on M is the n-form which, in a (positively oriented) coordinate chart, is given by

$$\mu = \sqrt{\det(g_{ij})}\, dx^1 \wedge \ldots \wedge dx^n$$

or, intrinsically, $\mu(v_1, \ldots, v_n) = \sqrt{\det\langle v_i, v_j\rangle}$ for $v_1, \ldots, v_n$ oriented tangent vectors. We say a diffeomorphism η is *volume preserving* if $\eta^*\mu = \mu$. Here $\eta^*\mu$ is the pull back of μ under η or, if you prefer, it is μ transformed by η. (See Lang [1].) The condition $\eta^*\mu = \mu$ means that the Jacobian of η is one.

By the change of variables formula, it follows that a diffeomorphism η is volume preserving if and only if for every measurable set $A \subset M$, $\mu(A) = \mu(\eta(A))$. Here we also use μ to stand for the measure defined by μ (cf. Abraham [2], § 12).

Set $\mathscr{D}_\mu = \{\eta \in \mathscr{D} \mid \eta$ is volume preserving$\}$.

For technical reasons it will be convenient to enlarge $\mathscr{D}$, $\mathscr{D}_\mu$ to slightly larger spaces. Namely let $\mathscr{D}^s$ (resp. $\mathscr{D}^s_\mu$) be the completion of $\mathscr{D}$, (resp. $\mathscr{D}_\mu$) under the Sobolev H^s topology; this will be discussed in detail later. The point is that $\mathscr{D}^s_\mu$ is modeled on a Hilbert space while $\mathscr{D}_\mu$ is merely locally Fréchet. We also remark that for technical questions involving partial differential equations, one must use Sobolev spaces rather than C^k spaces in order for the analysis to work.

2.2 Perfect fluids and Geodesics on $\mathscr{D}_\mu$

At least in the beginning, we will be discussing perfect fluids; i.e., nonviscous, homogeneous and incompressible. We also ignore external forces for simplicity.

Consider, then, our manifold M whose points are supposed to represent the fluid particles at $t = 0$. Let us look at the fluid moving in M (Figure 2.1). As t increases, call $\eta_t(m)$ the curve followed by the fluid particle

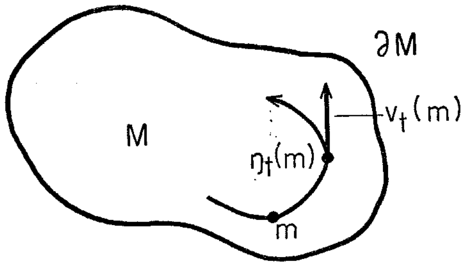


FIGURE 2.1

which is initially at $m \in M$. For fixed t, each η_t will be a diffeomorphism of M. In fact, since the fluid is incompressible, we have $\eta_t \in \mathscr{D}_\mu$. The function $t \mapsto \eta_t$ is thus a curve in $\mathscr{D}_\mu$ (they are easily seen to be orientation preserving since they are connected to η_0, the identity function on M). Note that if M has a fixed boundary the flow will be parallel to ∂M.

We will make $\mathscr{D}^s_\mu$ into an infinite dimensional (Hilbert) "Riemannian" manifold. The metric on $\mathscr{D}^s_\mu$ will correspond to the total kinetic energy of the fluid

$$\text{Energy} = \frac{1}{2}\int_M \|v\|^2 \, d\mu$$

where v is the velocity field of the fluid:

$$v_t(\eta_t(m)) = \frac{d}{ds}\eta_s(m)\bigg|_{s=t}$$

(see Figure 2.1). Note that v_t is a time dependent vector field on M.

With this metric, the curve $t \mapsto \eta_t$ will turn out to be a geodesic on $\mathscr{D}_\mu$.

Also this metric is *right invariant;* $\mathscr{D}_\mu$ being a group under composition.

In general the equations of motion for a geodesic are given by

$$\frac{d^2x^i}{dt^2} + \Gamma^i_{jk}\frac{dx^j}{dt}\frac{dx^k}{dt} = 0$$

where the Γ^i_{jk} are the Christoffel symbols. However, below we will use the equivalent notion of a spray (see Lang [1] for a review of this concept). This is the more natural way of dealing with geodesics in the infinite dimensional case. The spray of a metric is just a coordinate independent way of handling the Christoffel symbols.

2.3 The Euler equations

The motions of a perfect fluid are given by the Euler equations which are as follows

$$\text{(Euler equations)}\begin{cases} \dfrac{\partial v_t}{\partial t} + \nabla_{v_t}v_t = -\text{grad}\, p_t \\ \text{div}\, v_t = 0 \\ v_t \text{ is tangent to } \partial M. \end{cases}$$

In this, $\nabla_{v_t}v_t$ is the covariant derivation and its ith component is given in a coordinate chart by

$$(\nabla_{v_t}v_t)^i = \sum_j v^j_t \frac{\partial v^i_t}{\partial x^j} + \sum_{j,k}\Gamma^i_{jk}v^j_t v^k_t$$

and $p_t = p(t)$ is some (unknown) real valued function on M called the *pressure.*

In the case of Euclidean space, each $\Gamma^i_{jk} = 0$ and then we get, using vector analysis notation

$$\nabla_v v = (v \cdot \vec{\nabla})v$$

Note. We shall always use a subscripted variable to denote that the variable is held fixed, as in v_t. It will never denote differentiation.

The physical derivation of these equations is quite simple in $\mathbb{R}^n$. We use Newton's Law $F = ma$. We can ignore the mass because of homogeneity (i.e., constant mass density) and we are assuming there are no external forces, so the only forces result from the internal pressure. We wish to deal with conservative force fields and therefore one assumes these internal forces arise as the gradient of a real valued function, the pressure. So we have

$$\text{acceleration} = -\text{grad}\, p_t.$$

(The negative sign is a convention of Physics.)

To compute the acceleration, consider Figure 2.2.

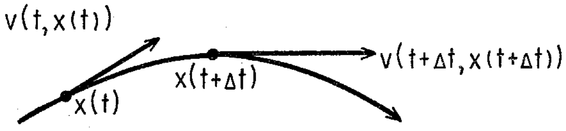


FIGURE 2.2

Clearly the acceleration is given by

$$a = \lim_{\Delta t \to 0} \frac{v(t + \Delta t, x(t + \Delta t)) - v(t, x(t))}{\Delta t}$$

$$= \frac{\partial v}{\partial t} + \sum_i \frac{\partial v}{\partial x^i}\frac{\partial x^i}{\partial t} = \frac{\partial v}{\partial t} + \sum_i \frac{\partial v}{\partial x^i} v^i.$$

Here we have just used the chain rule. This gives us the correct equation for $\partial v/\partial t$. Now div $v = 0$ is the same as assuming η_t is volume preserving, and v parallel to ∂M just corresponds to particles not moving across ∂M.

In subsequent lectures we will prove that *η_t is a geodesic on $\mathscr{D}_\mu$ iff v_t satisfies the Euler equations.* Thus the two problems (finding geodesics on $\mathscr{D}_\mu$ and solving the Euler equations) are equivalent.

There are two standard coordinate systems used in classical fluid mechanics. In the first, one describes the fluid as "seen" from one of the particles of the fluid. The observer follows the fluid. This is Lagrangian coordinates. The other system describes the fluid from the viewpoint of a fixed observer. This is termed Eulerian coordinates. Working on $\mathscr{D}_\mu^s$ corresponds to using Lagrangian coordinates. When one is just working with the equations in v_t, then one is working in Eulerian coordinates. Note the Euler equations are written out in Eulerian coordinates. For further information on this, and the derivation of the equations, see Serrin [1].

2.4 Stability and geodesics

The idea of studying geodesics in $\mathscr{D}_\mu$ in order to do hydrodynamics is due to V. Arnold (see Arnold [1]), although the basic idea seems to go back to P. Ehrenfest's thesis in 1904. (This interesting historical fact was pointed out by D. Ruelle.)

One of the problems Arnold studied was the stability of solutions under small changes of initial data (this perhaps has applications to weather prediction). See Arnold [1] or Chernoff–Marsden [1]. The basic idea is

group theoretical. This can alternatively be studied by using the equation of geodesic deviation, which involves the sectional curvature. (Actually this has certain other technical advantages over the Lie group approach.) If geodesics are diverging, (sectional curvature <0) small changes in initial data will evolve to large changes after a certain time (Figure 2.3). If the sectional curvature is >0, nearby geodesics will remain nearby and we have stability. V. Arnold [1] computes the curvature of $\mathscr{D}_\mu$ in case M is a flat 2-torus. The general case seems harder to get hold of explicitly. Later on in Part II we shall study some questions of stability in greater depth.

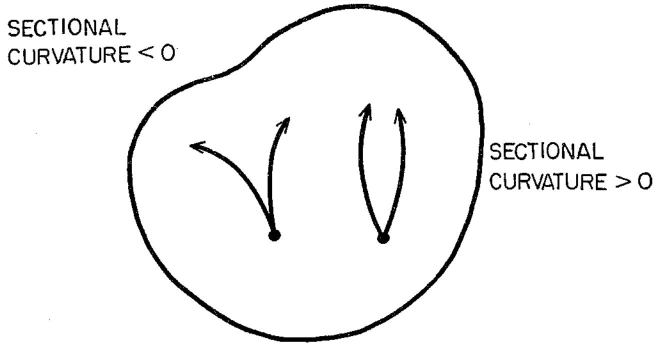


FIGURE 2.3

2.5 Evolution equations in hydrodynamics

On our manifold M as above, let $\mathbb{E}$ be the linear space of C^∞ divergence free vector fields on M and which are parallel to ∂M. If X is any C^∞ vector field on M (not necessarily parallel to ∂M) then as we shall explain later, X has a unique decomposition $X = Y + \operatorname{grad} p$ where Y is divergence free and parallel to M. Define a linear map P from $\mathscr{X}(M)$ = vector fields on M to $\mathbb{E}$ by setting $P(X) = Y$. P is called the *projection onto the divergence free part*. Define $T: \mathbb{E} \to \mathbb{E}$ by $T(v) = -P(\nabla_v v)$. Note that

$$-P(\nabla_v v) = -(\nabla_v v - \operatorname{grad} p) = -\nabla_v v + \operatorname{grad} p$$

and therefore we can rewrite the Euler equation (modulo a trivial sign convention on p) as a differential equation on the linear space $\mathbb{E}$:

$$\begin{cases} \dfrac{\partial v_t}{\partial t} = T(v_t) \\ v_0 \text{ is given.} \end{cases}$$

This looks like an initial value problem from the theory of ordinary differential equations. There are various approaches that might work to prove existence of solutions and to study their properties:

1. If T is locally Lipschitz and $\mathbb{E}$ is Banach then one can apply the Picard method (see Lang [1]). This generally fails for partial differential equations. If $\mathbb{E}$ consists of C^∞ vector fields then $\mathbb{E}$ is not Banach and it is well known that this method fails for such Frechet spaces (for instance let $\mathbb{E} = C^\infty$ maps of $[0, 1]$ to $\mathbb{R}$ which vanish to all orders at 0, 1 and let $T = \partial/\partial x$; the solution of $\partial f/\partial t = Tf$ is $f(t, x) = f_0(x + t)$ which generally leaves $\mathbb{E}$ for $t \neq 0$). On the other hand if $\mathbb{E}$ consists of C^k (or H^s) vector fields, then T does not map $\mathbb{E}$ into $\mathbb{E}$ since it involves derivations. In other words, *one cannot use the techniques of ordinary differential equations for partial differential equations like the Euler equations.*

2. If T were linear and densely defined, we could apply classical results of Hille and Yosida (see Yosida [1] ch. IX for instance). Of course in our case T is *not* linear.

3. If T were Quasi-Linear (i.e., linear in the top derivatives), there would be applicable theorems (see for instance Courant-Hilbert [1], Fischer-Marsden [2] and Chernoff-Marsden [1], chapter VII). If it were not for the projection P, we could use this theory. Hence, this is appropriate for the compressible case. The difficulty with P is that it is a *nonlocal* operator. That is, knowing X in a neighborhood of a point does not tell us $P(X)$ near that point.

4. A generalization of 2 to certain types of nonlinear situations has been developed by, amongst many others, Browder, Komura, and Kato (cf. Browder [1]). Unfortunately this does not apply to the Euler equations (for those who know the terms, the semigroup defined by the Euler equations is not globally quasi-contractive).

5. The oldest nonlinear existence theory, the Cauchy–Kowalewski theorem (cf. Courant-Hilbert [1]) is not appropriate; not merely because of the nonlocal nature of P, but because it fails to deal with Hadamard's basic criterion of being well posed—the solutions should vary continuously with the initial data.

Because of a surprising fact, we *can* use the Picard method on $\mathscr{D}^s_\mu$. The fact is that the spray on $\mathscr{D}^s_\mu$ is C^∞ and everywhere defined. Hence one can use the Picard method to get existence and uniqueness of geodesics on $\mathscr{D}^s_\mu$ for a short time, and as mentioned above this is equivalent to finding solutions to the Euler equations.

This should be surprising in view of our remarks (1)–(5) concerning the Euler equations. Indeed, the Picard method generally will *fail* when applied to the operator T. Thus there is a *real* technical difference between working on the linear space $\mathbb{E}$ and on the nonlinear space $\mathscr{D}^s_\mu$. This seemingly strange phenonenon was discovered by D. Ebin and J. Marsden [1]. However, it seems to have been known by some previous authors for special cases (although this is, to a large extent, buried in their proofs). In the case of the Euler equations with no pressure term, this phenomenon appears to have been common knowledge; it was first pointed out to us by T. Kato.

We would now like to try to give the essence of this idea. The key thing is that in Lagrangian coordinates, the equations change their character completely. Let us ignore the pressure term and consider

$$\frac{\partial v}{\partial t} + (v \cdot \vec{\nabla})v = 0$$

on $\mathbb{R}^3$. We let η_t be the flow of v and look at the new variables η_t, $X = v_t \circ \eta_t$ instead of v itself. Now

$$\begin{aligned}\frac{\partial X}{\partial t} &= \frac{\partial v}{\partial t} \circ \eta_t + \sum \frac{\partial v_t}{\partial x^i} \frac{\partial \eta_t^i}{\partial t} \\ &= \frac{\partial v}{\partial t} \circ \eta_t + \sum \left(\frac{\partial v_t}{\partial x^i} \cdot v_t^i\right) \circ \eta_t = 0\end{aligned}$$

since v satisfies $(\partial v/\partial t) + (v \cdot \vec{\nabla})v = 0$. Thus the velocity is *constant* in Lagrangian coordinates. Hence $\eta_t(x) = x + tv_0(x)$ (since $v_0 = X$) and so $v_t = X \circ \eta_t^{-1}$, an explicit solution (assuming η_t^{-1} exists).

The point of this is that in Lagrangian coordinates the derivative part $(v \cdot \vec{\nabla})v$ cancelled out. A similar thing happens on manifolds. Now a crucial point is that this property is not destroyed when a pressure term is added. This aspect is explained on p. 190. This cancellation of the term $(v \cdot \vec{\nabla})v$ is the basic fact which accounts for the smoothness of the spray in Lagrangian coordinates and consequently why the Picard method works on $\mathscr{D}^s_\mu$ and not in $\mathbb{E}$.

It still seems necessary to introduce $\mathscr{D}^s_\mu$ because η_t will satisfy $(\partial \eta_t/\partial t) = X_t$ and the appropriate function space for η is $\mathscr{D}^s_\mu$. That is, it does not seem possible to completely carry out the above basic program without introducing nonlinear manifolds of mappings. The fact that these nonlinear function spaces had not been developed until the 1960's perhaps explains why this program has not been carried out earlier.

2.6 The global existence problem

This is an outstanding problem in global analysis and fluid mechanics. It has been open since the early 1930's when J. Leray considered such questions seriously.

PROBLEM. *Let M be a (compact) 3-manifold. Then is $\mathscr{D}^s_\mu$ geodesically complete? (That is, do geodesics exist for all time?)*

The following simple lemma bears on the problem (the lemma is standard and is proven later).

LEMMA. *Let G be a finite dimensional Lie group with a right invariant metric. Then G is geodesically complete.*

The lemma also holds if G is a "Hilbert group-manifold," but unfortunately, it does not apply to our problem because the topology of our metric (recall it gives the L^2 norm) does not coincide with the topology on $\mathscr{D}^s_\mu$ (as we shall see $s = 0$, which corresponds to L^2, is not allowed).

The problem is thus equivalent to showing that the solutions to the Euler equations exist, and remain smooth, for all time. As we shall see later, this has bearing on rival theories of turbulence. We shall discuss this problem in detail later, including when viscosity is added. For now we just mention that in case $\dim M = 2$ the problem was answered affirmatively by Leray [3] (in case some viscosity is present) and by Wolibner [1] in 1933 for the Euler equations; slicker proofs have been given by Judovich [2] and Kato [1].

If the P were not there, and correspondingly the requirement $\operatorname{div} v = 0$ dropped, the result is definitely false—this is the phenomenon of shock waves in compressible flow. (For example the solution of $(\partial u/\partial t) + u(\partial u/\partial x) = 0$ in one dimension is $u(t, x) = u_0(y)$ where $x = y + tu_0(y)$. One can see as soon as $x \mapsto y$ becomes non-invertible, that derivatives of u blow up; we shall discuss this compressible case in more detail later).

Many people have tried to settle the problem for $\dim M = 3$, but with little success. We feel that it is one of the most important open problems in hydrodynamics and global analysis. It is probably one of the most difficult as well.

3. Motivation for general relativity

We now shall briefly indicate the sort of thing we will be doing in relativity. A more detailed outline is given in Part III. In relativity, one is interested in Lorentz manifolds (4-manifolds with a symmetric two tensor or "metric" of signature $(+, +, +, -)$), whose Ricci curvature Ric is zero. The equations $\text{Ric} = 0$ are the *Einstein field equations* for an empty space.

In actually solving for the metric g in question, proving existence theorems, and studying properties of solutions, one generally converts the problem to an initial value problem with g prescribed at $t = 0$ as a Riemannian metric on a given fixed three manifold M. One regards the 3 metric as then evolving in the space $\mathscr{M}$ of all metrics on M; $\mathscr{M}$ is an open cone in the linear space of all symmetric two tensors. The motion turns out to be Hamiltonian, specifically geodesic motion in the presence of a potential. These ideas go back to Arnowitt, Deser and Misner [1] and to De Witt [1]

To reconstruct the full coordinate transformations of spacetime, one introduces the "lapse" and "shift" functions of Wheeler [1]. One of our aims will be to describe these objects and to their effect on the above dynamical picture.

It turns out that the diffeomorphism group $\mathscr{D}$ of our 3 manifold M plays an important role. One can think of $\mathscr{D}$ as the totality of ("active") coordinate transformations of space. Furthermore, we know from the general theory of Hamiltonian systems, that associated with any invariance group, there are conserved quantities. It turns out that $\mathscr{D}$ is an invariance group, and we shall work out these conservation laws; for those who know the jargon, it is the conservation of $\delta\pi = 0$.

The existence theory in general relativity is of quite a different nature than that in hydrodynamics. In this case the quasi-linear theory applies. We shall briefly sketch this out later. Moreover, the constraints are of a rather different type, although they are somewhat analogous to the pressure in fluid mechanics.

General relativity has a very particular feature not shared by other classical field theories like electromagnetism (Maxwells equations). This feature is that the energy is *pointwise* constant in time—as opposed to the conservation of the total, or integrated, energy which one normally expects. Misner [1] has argued this on physical grounds. We shall present a general theorem which establishes the necessity of this condition in any "relativistic" theory ("external" fields are not allowed). This pointwise conservation of energy $\mathscr{H}$, expressed in terms of the reconstructed four geometry follows directly from the equations $\text{Ric} = 0$. Similarly for $\delta\pi = 0$. The necessity of these extra equations comes about because $\text{Ric} = 0$ represents 10 equations; there are 6 evolution equations and 4 constraint equations: $\mathscr{H} = 0$, $\delta\pi = 0$.

The approach here follows that of Fischer–Marsden [1].

I. SOME TOOLS FROM GLOBAL ANALYSIS

In this part we shall develop appropriate tools that will be needed later. For the most part we shall just sketch out the ideas of proofs, although complete proofs are given in some cases.

In section one we study Hodge theory. This is motivated by the need for the projection operator P introduced above. In fact the Hodge theory can be regarded as a generalization of the classical result that every vector field Y on $\mathbb{R}^3$ can be uniquely decomposed as follows:

$$Y = \nabla \times W + \nabla p$$

into a divergence free and gradient part.

In section two we study manifolds of maps. This generalizes the classical functions spaces, such as $L^2(\mathbb{R}^n)$ or $C^k(\mathbb{R}^n)$, to maps between manifolds. In general these spaces of maps are not linear spaces, but if the domain is a compact manifold, the sets of maps may be made into infinite dimension manifolds modeled on Hilbert or Banach spaces.

As we pointed out in § 2 above, for hydrodynamics, using these non-linear spaces seems to have a real technical advantage, as well as an aesthetic one. This is true even when the underlying manifold is flat, say an open region in $\mathbb{R}^3$.

Other applications of manifold of maps include:

(a) The Smale-Palais treatment of Calculus of Variations and infinite dimensional Morse theory as explained in Professor Klingenberg's lectures;

(b) Moser's proof of structural stability of Anosov diffeomorphisms (see Mather **[1]** and Robbin **[3]**);

(c) New proofs that a C^k vector field has a C^k flow, Robbin **[1]**; and proofs of the Frobenius theorem, Penot **[4]**.

In section three we shall introduce the Hopf bifurcation and centre manifold theory. Basically a bifurcation indicates that some sudden qualitative change takes place, usually at some critical value of a parameter. Ruelle and Takens **[1]** have shown how the center manifold theory can be used in a very elegant manner to study bifurcation problems. These ideas bear on the ideas of stability and turbulence in hydrodynamics and we shall be going into that in some detail in § 3 of Part II.

1. Hodge theory

1.1 Sobolev spaces

Let $\Omega \subset \mathbb{R}^n$ be an open bounded set with C^∞ boundary. Let $\bar{\Omega}$ be the closure of Ω. Define $C^\infty(\Omega, \mathbb{R}^n)$ to be the set of functions from Ω into $\mathbb{R}^n$ that can be extended to a C^∞ function on some open set in $\mathbb{R}^n$ containing $\bar{\Omega}$. Let $C_0^\infty(\Omega, \mathbb{R}^m) = \{f \in C^\infty(\Omega, \mathbb{R}^m) \mid$ the support of f is contained in a compact subset of $\Omega\}$.

To describe the Sobolev spaces in an elementary fashion, we temporarily

introduce some more notation. An n multi-index is an ordered set of n non-negative integers. If $k = (k_1, \ldots, k_n)$ is an n multi-index, then put $|k| = k_1 + k_2 + \ldots + k_n$. If $u \in C^\infty(\Omega, \mathbb{R}^m)$, define $D^k u$ by the formula

$$D^k u = (\partial^{|k|} u / \partial x_1^{k_1} \ldots \partial x_n^{k_n})$$

and $D^0(u) = u$. For $u \in C^\infty(\Omega, \mathbb{R}^m)$ (or $C_0^\infty(\Omega, \mathbb{R}^n)$), define

$$\|u\|_s^2 = \int_\Omega \sum_{0 \leq |k| \leq s} |D^k u(x)|^2 \, dx.$$

Now $H^s(\Omega, \mathbb{R}^m)$ (resp. $H_0^s(\Omega, \mathbb{R}^m)$) is defined to be the completion of $C^\infty(\Omega, \mathbb{R}^m)$ (resp. $C_0^\infty(\Omega, \mathbb{R}^m)$) under the $\| \ \|_s$ norm. These H^s spaces are called the *Sobolev spaces*. Note that $H_0^0(\Omega, \mathbb{R}^m) = H^0(\Omega, \mathbb{R}^m) = L_2(\Omega, \mathbb{R}^m) \supset H^s(\Omega, \mathbb{R}^m)$; but for $s \geqq 1$, $H_0^s(\Omega, \mathbb{R}^m) \neq H^s(\Omega, \mathbb{R}^m)$ as we shall see below.

There is another equivalent, and perhaps better, definition of the Sobolev norm. Let $d^k u$ be the kth total derivative of u so that $d^k u : \Omega \to L^k(\mathbb{R}^n, \mathbb{R}^m)$ where $L^k(\mathbb{R}^n, \mathbb{R}^m)$ denotes the k-linear maps on $\underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \ldots \times \mathbb{R}^n}_{k \text{ times}} \to \mathbb{R}^m$ with the standard norm. Then if we set

$$|u|_s^2 = \int_\Omega \sum_{0 \leq k \leq s} \|d^k u(x)\|^2 \, dx$$

the $| \ |_s$ and $\| \ \|_s$ norms are equivalent. This is a simple exercise.

Also note that $H^s(\Omega, \mathbb{R}^n)$ and $H_0^s(\Omega, \mathbb{R}^n)$ are Hilbert spaces with the inner product

$$\langle u, v \rangle = \int_\Omega \sum_{0 \leq |k| \leq s} D^k u(x) \cdot D^k v(x) \, dx.$$

1.1.1 Sobolev Theorem

(a) *Let $s > (n/2) + k$. Then $H^s(\Omega, \mathbb{R}^m) \subset C^k(\Omega, \mathbb{R}^m)$ and the inclusion map is continuous (in fact is compact) when $C^k(\Omega, \mathbb{R}^m)$ has the standard C^k topology, (the sup of the derivatives of order $\leqq k$).*

(b) *If $s > (n/2)$ then $H^s(\Omega, \mathbb{R}^m)$ is a ring under pointwise multiplication of components. (This is often called the Schauder ring.)*

(c) *If $s > \frac{1}{2}$ and $f \in H^s(\Omega, \mathbb{R}^m)$ then $f \mid \partial\Omega \in H^{s-1/2}$.*

(d) *(Calderon Extension Theorem) If $f \in H^s(\Omega, \mathbb{R}^m)$ then f has an extension $\tilde{f} \in H^s(\mathbb{R}^n, \mathbb{R}^m)$.*

Regarding (c), see Palais **[1]** for a discussion of continuous Sobolev chains; i.e., the definition of H^s for s not an integer; basically one can use the Fourier transform. (d) means that f can be extended across $\partial\Omega$ in an H^s way.

Differentiability properties at the boundary presents some technical problems but are very important in hydrodynamics. Thus it is important to distinguish H_0^s from H^s.

The proof of Theorem 1.1.1 can be found in Nirenberg **[1]** and Palais **[1]**; see also Sobolev **[1]**.

For most of hydrodynamics we will need $s > (n/2) + 1$. One of the outstanding problems in the field is determining to what extend we can relax this condition on s. For many problems, one would like to allow corners and discontinuities in such things as the density of the fluid or the velocity field.

1.2 H^s Spaces of sections

Let M be a compact manifold, possibly with boundary. Also, let E be a finite dimensional vector bundle over M. For example E may be the tangent bundle, or a tensor bundle over M. Let $\pi: E \to M$ be the canonical projection. The following fact is useful.

1.2.1 PROPOSITION. *Suppose for each $x \in M$, we have $\pi^{-1}(x) \cong \mathbb{R}^m$. Then there is a finite open cover $\{U_i\}$ of M such that each U_i is a chart of M and $\pi^{-1}(U_i) \cong U_i \times \mathbb{R}^m$ for each i.*

Such a cover is called *trivializing*. (See Lang [1]). Recall that a *section* of E is a map $h: M \to E$ such that $\pi \circ h = id_M$. For example, a vector field is a section of the tangent bundle and a differential one-form is a section of the cotangent bundle. Informally, we define, for $s \geqq 0$, $H^s(E)$ to be the set of sections of E whose derivatives up to order s are in L_2.

This makes sense since in view of the proposition, a section of E can locally be thought of as a map from $\mathbb{R}^n$ to $\mathbb{R}^m$ where n is the dimension of M. Similarly, we can put a Hilbert structure on $H^s(E)$ by using a trivializing cover. However, since this Hilbert space structure depends on the choice of charts, the norm on $H^s(E)$ is not canonical, so we call $H^s(E)$ a *Hilbertible Space* (i.e., it is a space on which some complete inner product exists).

One has to check that the definition of $H^s(E)$ is independent of the trivialization and this can be done by virtue of compactness of M.

Of course the Sobolev theorems have analogues for $H^s(E)$. In particular if $s \geqq 1$ it makes sense to restrict a section $h \in H^s(E)$ to ∂M. This is by part (c) of Theorem 1.1.1. Of course if $s > (n/2)$, h will be continuous and so

this will be clear. For $s = 0$, we have $L_2(E)$ and restriction to ∂M does not make sense.

One defines $H_0^s(E)$ in a similar way. For $s > \frac{1}{2}$, when we restrict $h \in H_0^s(E)$ to ∂M, h will vanish, as will its derivatives to order $s - \frac{1}{2}$.

Much of the theory goes over for *M noncompact, but we must specify* a metric on M and a connection on E; further M must be complete and obey some curvature restriction such as sectional curvature bounded above; cf. Cantor **[2]** and Fischer-Marsden **[2, II]**.

1.3 Some operations on differential forms

Now, let M be a compact oriented Riemannian n-manifold without boundary.

Let Λ^k be the vector bundle over M whose fiber at $x \in M$ consists of k-linear skew-symmetric maps from T_xM, the tangent space to M at $x \in M$, to $\mathbb{R}$. For each x, $\bigoplus_{k=0}^n \Lambda_x^k$ forms a graded algebra with the wedge product. Then $H^s(\Lambda^k)$ is a space of H^s differential k-forms. The exterior derivative d then is an operator:

$$d: H^{s+1}(\Lambda^k) \to H^s(\Lambda^{k+1}).$$

It drops one degree of differentiability because d differentiates once; i.e., is a first order operator. Recall that if α, β are k and l forms respectively, then $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k\alpha \wedge d\beta$; cf. Abraham **[2]**.

The star operator $*: H^s(\Lambda^k) \to H^s(\Lambda^{n-k})$ is given on Λ^k at $x \in M$ by

$$*(1) = \pm dx_1 \wedge \ldots \wedge dx_n, \qquad *(dx_1 \wedge \ldots \wedge dx_n) = \pm 1$$

and

$$*(dx_1 \wedge \ldots \wedge dx_p) = \pm dx_{p+1} \wedge \ldots \wedge dx_n$$

where the "+" is taken if the $dx_1 \wedge \ldots \wedge dx_n$ is positively oriented and "−" otherwise, $x_1, \ldots, x_n$ form a coordinate system orthogonal at x, and $*$ is extended linearly as an operator $\Lambda^k \to \Lambda^{n-k}$. Now if $\alpha \in H^s(\Lambda^k)$ then clearly $*\alpha \in H^s(\Lambda^{n-k})$, so $*$ can be taken as an operator from $H^s(\Lambda^k)$ to $H^s(\Lambda^{n-k})$.

The space Λ^k carries, at each point $x \in M$, an inner product. It is the usual business: the metric converts covariant tensors to contravariant ones (i.e., it raises or lowers indices) and then one contracts. If α_i, β_j are one forms, we have $\langle \alpha_1 \wedge \ldots \wedge \alpha_k, \beta_1 \wedge \ldots \wedge \beta_k \rangle = \det[\langle \alpha_i, \beta_j \rangle]$. It is not hard to check that if μ is the volume form on M then

$$\langle \alpha, \beta \rangle \mu = \alpha \wedge *\beta = \beta \wedge *\alpha.$$

Note that the inner product may be defined by the above formula. See Flanders **[1]** for more details on these matters.

Define the operator $\delta: H^{s+1}(\Lambda^k) \to H^s(\Lambda^{k-1})$ by $\delta = (-1)^{n(k+1)+1} *d*$. There is an inner product on $H^0(\Lambda^k)$ (and hence on $H^s(\Lambda^k)$) given by

$$(\alpha, \beta) = \int_M \langle \alpha, \beta \rangle \, d\mu.$$

1.3.1 PROPOSITION. *For* $\alpha \in H^s(\Lambda^k)$ *and* $\beta \in H^s(\Lambda^{k+1})$

$$(d\alpha, \beta) = (\alpha, \delta\beta).$$

PROOF. Note that $d(\alpha \wedge *\beta) = d\alpha \wedge *\beta + (-1)^k \alpha \wedge d*\beta$

$$= d\alpha \wedge *\beta - \alpha \wedge *\delta\beta,$$

since $** = (-1)^{k(n-k)}$.

Since $\partial M = \varnothing$, by Stokes Theorem, we get

$$\begin{aligned} 0 &= \int_M d(\alpha \wedge *\beta) \\ &= \int_M d\alpha \wedge *\beta - \int_M \alpha \wedge *\delta\beta \\ &= (d\alpha, \beta) - (\alpha, \delta\beta). \quad \blacksquare \end{aligned}$$

Rephrasing 1.3.1, one says that d and δ are *adjoints* in the (,) inner product.

The δ operator corresponds to the classical divergence operator. This is easily seen. Let X be a vector field on M. Then because of the Riemannian structure X corresponds to a 1-form $\tilde{X}$, where $\tilde{X}(v) = \langle X, v \rangle$.

1.3.2 PROPOSITION. $\operatorname{div}(X) = -\delta(\tilde{X})$.

PROOF AND DISCUSSION. Let $L_X\mu$ be the Lie derivative of μ with respect to X. Then by definition, $\operatorname{div}(X)\mu = L_X\mu$ (see Abraham [**2**]). We have the general formula

$$L_X\mu = di_X(\mu) + i_X d(\mu)$$

where $i_X(\mu) = X \lrcorner \mu \in H^s(\Lambda^{n-1})$, the interior product is defined in general as follows: If $\beta \in H^s(\Lambda^k)$ and $v_1, \ldots, v_{k-1} \in TM$, then $i_X(\beta)(v_1, \ldots, v_{k-1}) = k\beta(X, v_1, \ldots, v_{k-1})$. Now $d(\mu) = 0$ since μ is an n-form, so $L_X\mu = d(i_X\mu) = d(*\tilde{X})$ (one easily checks that $i_X\mu = *\tilde{X}$). Hence

$$\operatorname{div}(X) = \operatorname{div}(X)*\mu = *(\operatorname{div}(X)\mu) = *d*\tilde{X} = -\delta\tilde{X},$$

since for $k = 1$, $(-1)^{n(k+1)+1} = -1$. $\blacksquare$

The *Laplace de Rham* operator is defined by $\Delta = \delta d + d\delta$. Note that $\Delta: H^s(\Lambda^k) \to H^{s-2}(\Lambda^k)$. If f is a real valued function on $\mathbb{R}^n$, it is easy to check, using the above expressions for d, δ, that $\Delta(f) = -\nabla^2(f)$ where $\nabla^2 f = \operatorname{div}(\operatorname{grad} f)$ is the usual Laplacian. Note $\delta f = 0$ on functions.

1.3.3 PROPOSITION. *Let* $\alpha \in H^s(\Lambda^k)$, *then* $\Delta\alpha = 0$ *iff*

$$d\alpha = 0 \quad \text{and} \quad \delta\alpha = 0.$$

PROOF. It is obvious that if $d\alpha = 0$ and $\delta\alpha = 0$ then $\Delta\alpha = 0$. To show the converse, assume $\Delta\alpha = 0$. Then $0 = (\Delta\alpha, \alpha) = ((d\delta + \delta d)\alpha, \alpha) = (\delta\alpha, \delta\alpha) + (d\alpha, d\alpha)$, so the result follows. $\blacksquare$

A form α for which $\Delta\alpha = 0$ is called *harmonic*.

1.4 The Hodge theorem (for $\partial M = \varnothing$)

1.4.1 *Let* $\omega \in H^s(\Lambda^k)$. *Then there is* $\alpha \in H^{s+1}(\Lambda^{k-1})$, $\beta \in H^{s+1}(\Lambda^{k+1})$ *and* $\gamma \in C^\infty(\Lambda^k)$ *such that* $\omega = d\alpha + \delta\beta + \gamma$ *and* $\Delta(\gamma) = 0$. *Here* $C^\infty(\Lambda^k)$ *denotes the* C^∞ *sections of* Λ^k. *Furthermore* $d\alpha$, $\delta\beta$, *and* γ *are mutually* L_2 *orthogonal and so are uniquely determined.*

1.4.2 *If* $\mathscr{H}^k = \{\gamma \in C^\infty(\Lambda^k) \mid \Delta\gamma = 0\}$, *then*

$$H^s(\Lambda^k) = d(H^{s+1}(\Lambda^{k-1})) \oplus \delta H^{s+1}(\Lambda^{k+1}) \oplus \mathscr{H}^k.$$

Note that 1.4.2 is just a rewriting of 1.4.1. The fact that the Harmonic forms $\mathscr{H}^k$ are all C^∞, follows from regularity theorems on the Laplacian. This fact is also called Weyl's lemma or, its generalization, Friedrich's theorem. (See Yosida [**1**].)

The Hodge theorem goes back to V. W. D. Hodge [**1**], in the 1930's. Substantial contributions have been made by many authors, leading up to the present theorem. See for example Weyl [**1**], and Morrey-Eells [**1**].

We can easily check that the spaces in 1.4.2 are orthogonal. For example

$$(d\alpha, \delta\beta) = (dd\alpha, \beta) = 0$$

since δ is the adjoint of d and $d^2 = 0$.

The basic idea in 1.4.1 can be abstracted as follows. We consider a linear operator T on a Hilbert space $\mathbb{E}$ with $T^2 = 0$. In our case $T = d$ and $\mathbb{E}$ is the L_2 forms. (We ignore the fact that T is only densely defined, etc.) Let T^* be the adjoint of T. Let $\mathscr{H} = \{x \in \mathbb{E} \mid Tx = 0 \text{ and } T^*x = 0\}$. We assert

$$\mathbb{E} = \overline{\text{Range } T} \oplus \overline{\text{Range } T^*} \oplus \mathscr{H}$$

which, apart from technical points on differentiability and so on is the essential content of 1.4.1, 1.4.2.

To see this, note, as before that the ranges of T and T^* are orthogonal because

$$\langle Tx, T^*y\rangle = \langle T^2x, y\rangle = 0.$$

Let $\mathscr{C}$ be the orthogonal complement of $\overline{\text{Range } T} \oplus \overline{\text{Range } T^*}$. Certainly $\mathscr{H} \subset \mathscr{C}$. But if $x \in \mathscr{C}$,

$$\langle Ty, x\rangle = 0 \quad \text{for all} \quad y \Rightarrow T^*x = 0.$$

Similarly $Tx = 0$, so $\mathscr{C} \subset \mathscr{H}$ and hence $\mathscr{C} = \mathscr{H}$.

The complete proof of the theorem may be found in Morrey [1]. For more elementary expositions, also consult Flanders [1] and Warner [1].

An interesting consequence of this theorem is that $\mathscr{H}^k$ is isomorphic to the kth de Rham cohomology class (the closed k-forms mod the exact ones). This is clear since over M, each closed form ω may be written $\omega = d\alpha + \gamma$. (One can check that the $\delta\beta$ term drops out when ω is closed; indeed we get $0 = d\,\delta\beta$ so $(d\,\delta\beta, \beta) = 0$ or $(\delta\beta, \delta\beta) = 0$ or $\delta\beta = 0$.)

Now let $(\mathscr{H}^k)^\perp$ be the L_2 orthogonal complement to $\mathscr{H}^k$ in $H^s(\Lambda^k)$. Define the *Green's Operator* $G: H^s(\Lambda^k) \to (\mathscr{H}^k)^\perp$ by letting $G(\alpha)$ equal the unique solution ω of $\Delta\omega = \alpha - H(\alpha)$ in $(\mathscr{H}^k)^\perp$, where $H(\alpha)$ is the harmonic part of α. The following is easy to check.

1.4.3 PROPOSITION. *G commutes with d, δ, and Δ.*

Note that if $\alpha \in H^s(\Lambda^k)$ then $\alpha = d\,\delta G(\alpha) + \delta\, dG(\alpha) + H(\alpha)$.

REMARKS. 1. The Hodge theorem fails for the C^k topologies. This is due to the fact that $\Delta f \in C^k$ does not imply that $f \in C^{k+2}$.

2. The Hodge theorem is very awkward and sometimes false for non-positive definite metrics. See Avez [1].

3. One can obtain a weighted Hodge decomposition, using a density function together with the measure. This is important for inhomogeneous fluid flow (see Ebin-Marsden [2]).

For us, one of the important consequences of the theorem is the following.

1.4.4 COROLLARY. *Let X be a H^s vector field on M, $s \geqq 0$. Then there is a unique divergence free H^s vector field Y and a gradient field* grad p *such that*

$$X = Y + \operatorname{grad} p.$$

Setting $P(X) = Y$, P is a bounded linear operator in L_2 and in H^s.

PROOF. In terms of the corresponding one form $\tilde{X}$, we write $\tilde{X} = (d\alpha + \delta\beta + \gamma)$ and set $d\alpha = dp$, $\tilde{Y} = \delta\beta + \gamma$. Since $\delta\delta = 0$, $\delta\tilde{Y} = 0$. ∎

The above theorem is a generalization of a classical theorem of Helmholtz, for which one can give a direct argument.

1.4.5 THEOREM. *Let V be a vector field on $\mathbb{R}^3$. Then $V = -\nabla\varphi + \nabla \times \omega$ where φ is a scalar function and ω is a vector field.*

SKETCH OF THE PROOF. It is well known that in the sense of distributions, $\nabla^2(1/|x - x'|) = -4\pi\,\delta(x - x')$ from which we get

$$V(x) = \frac{\Delta}{4\pi}\int_{\mathbb{R}^3} \frac{V(x')}{|x - x'|}\,dx'.$$

What we have done here is found the Green's operator in $\mathbb{R}^3$ (that this is so can be seen easily and is found in any elementary book on differential equations; see for instance Duff and Naylor [1]). Using this and the identity corresponding to $\Delta = \delta d + d\delta$:

$$-\nabla^2 = \nabla(\nabla \cdot\ ) - \nabla \times (\nabla \times\ )$$

we get

$$\begin{aligned} V(x) &= \frac{\Delta}{4\pi}\int_{\mathbb{R}^3} \frac{V(x')}{|x - x'|}\,dx \\ &= -\nabla\left(\int_{\mathbb{R}^3} \nabla \cdot \frac{V(x')}{4\pi\,|x - x'|}\,dx\right) \\ &\quad + \nabla \times \left(\nabla \times \int_{\mathbb{R}^3} \frac{V(x')}{4\pi\,|x - x'|}\,dx'\right). \quad ∎ \end{aligned}$$

Since $\mathbb{R}^n$ is topologically trivial, there are no harmonic forms in L_2.

In this theorem, the integrals are singular and so the smoothness of the summands is not obvious.

1.5 Hodge theory for manifolds with boundary

This theory was worked out by Kodaira [1], Duff-Spencer [1], and Morrey [1]. (See Morrey [2], Chapter 7.) Differentiability across the boundary is very delicate, but important. The best possible results in this regard were worked out by Morrey.

Also note that d and δ may not be adjoints in this case, because boundary terms arise when we integrate by parts. Hence we must impose certain boundary conditions.

Let $\alpha \in H^s(\Lambda^k)$. Then α is *parallel* or *tangent* to ∂M if the normal part, $n\alpha = i^*(*\alpha) = 0$ where $i: \partial M \to M$ is the inclusion map. Analogously α is *perpendicular* to ∂M if $t\alpha = i^*(\alpha) = 0$.

Let X be a vector field on M. Using the metric, we know when X is tangent or perpendicular to ∂M. X corresponds to the one-form $\tilde{X}$ and also to the $n - 1$ form $i_X\mu$ (μ is, as usual, the volume form). Then X is tangent to ∂M if and only if $\tilde{X}$ is tangent to ∂M iff $i_X\mu$ is normal to ∂M. Similarly X is normal to ∂M iff $\tilde{X}$ is normal to ∂M iff $i_X\mu$ is tangent to ∂M. Set

$$H^s_t(\Lambda^k) = \{\alpha \in H^s(\Lambda^k) \mid \alpha \text{ is tangent to } \partial M\}$$

$$H^s_n(\Lambda^k) = \{\alpha \in H^s(\Lambda^k) \mid \alpha \text{ is perpendicular to } \partial M\}$$

and

$$\mathscr{H}^s(\Lambda^k) = \{\alpha \in H^s(\Lambda^k) \mid d\alpha = 0,\ \delta\alpha = 0\}.$$

The condition that $d\alpha = 0$ and $\delta\alpha = 0$ is now stronger than $\Delta\alpha = 0$. Following Kodaira [1], one calls elements of $\mathscr{H}^s$, *harmonic fields.*

1.5.1 The Hodge Theorem

$$H^s(\Lambda^k) = d(H^{s+1}_t(\Lambda^{k-1})) \oplus \delta(H^{s+1}_n(\Lambda^{k+1})) + \mathscr{H}^s(\Lambda^k)$$

One can easily check from the formula

$$(d\alpha, \beta) = (\alpha, \delta\beta) + \int_{\partial M} \alpha \wedge *\beta$$

that the summands in this decomposition are orthogonal.

There are two other decompositions that are not derivable directly from 1.5.1.

1.5.2 Theorem

(a) $H^s(\Lambda^k) = d(H^{s+1}(\Lambda^{k-1})) \oplus D^s_t$

where

$$D^s_t = \{\alpha \in H^s_t(\Lambda^k) \mid \delta\alpha = 0\}$$

and dually

(b) $H^s(\Lambda^k) = \delta(H^{s+1}(\Lambda^{k+1})) + C^s_n$

where C^s_n are the closed forms normal to ∂M. Again one can easily check orthogonality.

We can apply this to get a general Helmholtz theorem.

1.5.3 Corollary. *Let X be an H^s vector field on M. Then $X = Y +$ grad p where* $\operatorname{div}(Y) = 0$ *and Y is H^s and tangent to ∂M.*

Proof. Identify X with $\tilde{X}$ and take its hodge decomposition: 1.5.2(a).

$$\tilde{X} = dp + \alpha.$$

Again using the metric we identify dp with grad p and α with $\tilde{Y}$. By remarks above and 1.3.2, Y is tangent to ∂M and $\operatorname{div}(Y) = 0$. ∎

Again we let P denote the projection onto the divergence free part. In this corollary, X does not have to be parallel to ∂M. The decomposition will automatically "straighten it out" to be parallel to the boundary.

2. Manifolds of maps and diffeomorphism groups

2.1 History

The basic idea was first laid down by Eells [1] in 1958. He constructed a smooth manifold out of the continuous maps between two manifolds. In 1961, Smale and Abraham worked out the more general case of C^k mappings. Their notes are pretty much unavailable, but the 1966 survey article by Eells [2] is a good reference. The H^s case is found in a 1967 article by Elliasson [1]. This is also found in Palais [4] where it is done in the more general context of fiber bundles.

Making the manifold out of the C^k diffeomorphism group on a compact manifold without boundary was done independently by Abraham (see Eells [2]) and Leslie [1] around 1966. The H^s case is found in a paper by Ebin [1] and one by Omori [1] around 1968. Ebin also showed that the volume preserving diffeomorphisms form a manifold. Finally Ebin-Marsden [1] worked out the manifold structure for the H^s diffeomorphisms, the symplectic and volume preserving diffeomorphisms for a compact manifold with smooth boundary.

Other papers on manifolds of maps include those of Saber [1], Leslie [2, 3], Omori [2], Gordon [1], Penot [2, 3], and Graff [1]. Some further references are given below.

2.2 Local structure

Let M and N be compact manifolds and assume N is without boundary. Let n be the dimension of M, and l the dimension of N. Say $f \in H^s(M, N)$ if for any $m \in M$ and any chart (U, φ) containing m and any chart (V, ψ)

at $f(m)$ in N, the map $\psi \circ f \circ \varphi^{-1}: \varphi(U) \to \mathbb{R}^l$ is in $H^s(\varphi(U), \mathbb{R}^l)$. This can be shown to be a well defined notion, independent of charts for $s > (n/2)$. The basic fact one needs is that by the Sobolev Theorem we have $H^s(M, N) \subset C^0(M, N)$. Things are not as nice, however, for $s < (n/2)$. It is possible for a map to have a (derivative) singularity which is L_2 integrable in one coordinate system on N and not be integrable in another. So for $s < (n/2)$, $H^s(M, N)$ cannot be defined invariantly. Hence, from now on we assume $s > (n/2)$.

In order to find charts in $H^s(M, N)$ we first need to determine the appropriate modeling space. Let $f \in H^s(M, N)$. The modeling space, should it exist, must be isomorphic to $T_f H^s(M, N)$, whatever that is. So a way to begin is to find a plausible candidate for $T_f H^s(M, N)$. If P is any manifold and $p \in P$ then $T_p P$ can be constructed by considering any smooth curve c in P such that $c(0) = p$; then $c'(0) \in T_p P$ (Figure 2.1).

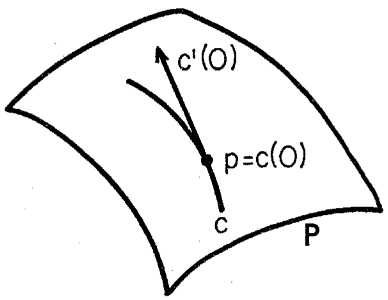


FIGURE 2.1

With this in mind, let us consider a curve $c_f: (-1, 1) \to H^s(M, N)$ such that $c_f(0) = f$. Now if $m \in M$, then the function $t \mapsto c_f(t)(m)$ is a curve in N (i.e., for each $t \in (-1, 1)$, $c_f(t) \in H^s(M, N)$ and therefore $c_f(t): M \to N$.) Now $c_f(0)(m) = f(m)$, so the derivative of this curve at 0, $(d/dt)c_f(t)(m)\,|_{t=0}$ is an element of $T_{f(m)}N$. So the map $m \mapsto (d/dt)c_f(t)(m)\,|_{t=0}$ maps M to TN and covers f, i.e., if $\pi_N: TN \to N$ is the canonical projection, this diagram commutes:

$$\begin{tikzcd} & TN \arrow[d, "\pi_N"] \\ M \arrow[ur, "c_f'(0)"] \arrow[r, "f"'] & N \end{tikzcd} \qquad c_f'(0) = \frac{d}{dt} c_f(t)\bigg|_{t=0}$$

where

$$c_f'(0)(m) = \frac{d}{dt} c_f(t)(m)\bigg|_{t=0}.$$

Making the identification

$$\left(\frac{d}{dt} c_f(t)\bigg|_{t=0}\right)(m) = \frac{d}{dt} c_f(t)(m)\bigg|_{t=0},$$

$c_f'(0)$ is a good candidate as the tangent to c_f at f.

With the above motivation, let us define

$$T_f H^s(M, N) = \{X \in H^s(M, TN) \mid \pi_N \circ X = f\}.$$

Note this is a linear space, for if V_f and X_f are in $T_f H^s(M, N)$, we can define $aV_f + X_f$ ($a \in \mathbb{R}$) as the map $m \mapsto aV_f(m) + X_f(m)$ where $V_f(m)$ and $X_f(m)$ are in $T_{f(m)}N$. It is this space which we use as a model for $H^s(M, N)$ near f.

To show this we need the map $\exp_N: T_p N \to N$ for $p \in N$. Recall that if $v_p \in T_p N$ there is a unique geodesic σ_{v_p} through p whose tangent vector at p is v_p. Then $\exp_p(v_p) = \sigma_{v_p}(1)$. In general $\exp_p$ is a diffeomorphism from some neighborhood of 0 in $T_p N$ onto a neighborhood p in N. However, since N is compact and without boundary, it is geodesicly complete and hence $\exp_p$ is defined on all of $T_p N$. This map can be extended to a map $\exp: TN \to N$ such that if $v_q \in TN$ then $\exp(v_q) = \exp_q(v_q)$. With this map we define the map $\overline{\exp}_f: T_f H^s(M, N) \to H^s(M, N)$

$$X \mapsto \exp \circ X.$$

We assert that $\overline{\exp}_f$ maps the linear space $T_f H^s(M, N)$ *onto* a neighborhood of f in $H^s(M, N)$ taking 0 to f and hence is a candidate for a chart in $H^s(M, N)$. It should be remarked that in spite of the use of the map exp, the structure is independent of the metric on N. The assertion is easy to check in case things are C^∞ or C^s, by using standard properties of exp; Milnor [1].

For the H^s case and to show that the change of charts is well defined (i.e., maps into the right spaces) and is smooth, one needs the following lemma.

2.2.1 (LOCAL) ω-LEMMA (LEFT COMPOSITION OF MAPS). *Let U be a bounded open set in $\mathbb{R}^p$, and $h: \mathbb{R}^n \to \mathbb{R}^m$ be C^∞. Then $\omega_h: H^s(U, \mathbb{R}^n) \to H^s(U, \mathbb{R}^m)$ defined by $\omega_h(f) = h \circ f$ is a C^∞ map.*

This conclusion is not true if h is merely an H^s or C^s map. The problem can be seen in this way. If M and N are manifolds and $g: M \to N$ is C^1 then for $p \in M$ and $v_p \in T_p M$, we have $T_p g: T_p M \to T_{g(p)}N$, which is determined in this manner: Let $c: (-1, 1) \in M$ be a curve such that $c(0) = p$ and $c'(0) = v_p$. Then $T_p g(v_p) = (d/dt)g(c(t))\,|_{t=0}$. Applying this procedure to ω_h, and using the chain rule, we find for $X \in T_f H^s(U, \mathbb{R}^n)$

that the tangent of ω_h is the map $T_f\omega_h: X \mapsto Th \circ X$. But since Th is only H^{s-1}, $Th \circ X$ is, at best, in $H^{s-1}(U, \mathbb{R}^m)$ and $T\omega_h$ does *not* map into the tangent space of $H^s(U, \mathbb{R}^m)$ at $\omega_h(f)$.

This necessity of differentiating h is a crucial difference between composition on the left and composition on the right.

The exact proof of 2.2.1 may be found in Ebin **[1]** and the other references above. In fact, the result essentially goes back to Sobolev **[1]** p. 223. See also Marcus-Mizel **[1]**, and Brezis **[1]**.

Using 2.2.1, it is now routine to check that $\overline{\exp}_f$ yields smooth charts on $H^s(M, N)$. For other methods of obtaining charts, see Palais **[4]**, Penot **[3]** and Krikorian **[1]**.

2.3 Groups of diffeomorphisms

These objects have a very interesting yet complicated structure. For this section we let M be a compact manifold without boundary. Let $\mathscr{D}^s(M) = \{f \in H^s(M, M) \mid f$ is one-one, orientation preserving and $f^{-1} \in H^s(M, M)\}$. The fact that $\mathscr{D}^s(M)$ is a manifold is a trivial consequence of the fact that $H^s(M, M)$ is a manifold and the following proposition;

2.3.1 Proposition. *If $s > (n/2) + 1$, then $\mathscr{D}^s(M)$ is open in $H^s(M, M)$.*

Proof. Since $s > (n/2) + 1$, we have a continuous inclusion $H^s(M, M) \subset C^1(M, M)$ (by the Sobolev Theorem 1.1.1(a)). So it is sufficient to show that if a map g on M is C^1 close to a diffeomorphism, then g is a diffeomorphism. To show this, note that $G: f \mapsto \inf_{x \in M} Jf(x)$ is a continuous real valued map on $C^1(M, M)$, where $Jf(x)$ is the Jacobian of $f: f^*\mu = (Jf)\mu$. Also, since M is compact, if $f \in \mathscr{D}^s(M)$, then $G(f) \neq 0$. By continuity of G, there is a neighborhood U of f in $C^1(M, M)$ such that if $g \in U$ then $G(g) \neq 0$. By the inverse function theorem U consists of local diffeomorphisms. It is easy to show that if $g \in U$ then g is an onto map. This is because $g(M)$ is open in M, as g is a local diffeomorphism and since g is continuous and M is compact, then $g(M)$ is closed. Hence if M is connected $g(M) = M$. (If M is not connected, one need just remark that g maps into each component of M since f does and g is uniformly close to f.) It remains to show there is a neighborhood of f containing only 1-1 functions. (It is not true that a local diffeomorphism on a compact set is a diffeomorphism. Consider the map which wraps S^1 around itself twice.) It is an easy exercise in point set topology to show that if M is connected then any local diffeomorphism on M is a covering map; that is, is globally k to 1 for some integer k. Also, the function that assigns to a local diffeomorphism f the number of elements in $f^{-1}(x)$ for any $x \in M$ is continuous

in the C^1 topology onto the integers. In particular there is a neighborhood of a diffeomorphism containing only diffeomorphisms. ■

Because of the differentiability condition in the above proposition, we will henceforth assume $s > (n/2) + 1$.

It is unknown whether, in general, the composition of two H^s maps is again H^s. In all known proofs one needs that one of the maps is a diffeomorphism or is C^∞. Hence composition in $\mathscr{D}^s$ presents no problem. The main composition properties are stated in the following.

2.3.2 Theorem

(a) *$\mathscr{D}^s$ is a group under composition.*

(b) *(α-Lemma) If $\eta \in \mathscr{D}^s$ the map $R_\eta: \mathscr{D}^s \to \mathscr{D}^s$, $\zeta \mapsto \zeta \circ \eta$ is a C^∞ map (in fact R_η is clearly "formally linear" and continuous).*

(c) *(ω-Lemma-Global) If $\eta \in \mathscr{D}^s$. Then $L_\eta: \zeta \mapsto \eta \circ \zeta$ is C^0. (This map is definitely not smooth, in fact it is not even a locally Lipschitz map.)*

(c)′ *More generally, the map*

$$\mathscr{D}^{s+l} \times \mathscr{D}^s \to \mathscr{D}^s$$
$$(\eta, \zeta) \mapsto \eta \circ \zeta$$

is C^l.

(d) *$\mathscr{D}^s$ is a topological group.*

Remark. (d) follows from the other parts of the theorem because of the following lemma of Montgomery **[1]**:

2.3.3 Lemma. *Let G be a group that is also a topological space. Assume further that G is a separable, metrizable, Baire space and multiplication in G is separately continuous. Then G is a topological group.*

We shall not prove (a), (b), (c)′, here since we have already given the basic ideas involved. The proof may be found in Ebin **[1]**. Another useful fact proved by Ebin is that if η is an H^s map with a C^1 inverse, then the inverse is H^s. This is analogous to what one has in the C^k inverse function theorem (Lang **[2]**).

2.4 $\mathscr{D}^s$ as a "Lie group"

$\mathscr{D}^s$ is not precisely a Lie group, (since a left multiplication is continuous, but not smooth) but it shares some important Lie group properties. If we were to work with $\mathscr{D} = \mathscr{D}^\infty$, we would have Lie group, but not a Banach manifold.

In general if G is a Lie group and $e \in G$ is the unit element, then the Lie Algebra $\mathscr{J}$ of G may be identified with T_eG. Hence, $T_e\mathscr{D}^s(M) = \mathscr{X}^s(M) = H^s(TM) = H^s$ vector fields on M (recall members of $T_e\mathscr{D}^s(M)$ cover the identity map on M) serve as the Lie algebra for $\mathscr{D}^s$. Since right multiplication is smooth, we can talk about right invariant vector fields on $\mathscr{D}^s$. By the ω-lemma, if $X \in \mathscr{X}^{s+l}$, the map $\tilde{X}: \eta \mapsto X \circ \eta$ is a C^l map from $\mathscr{D}^s$ to $T\mathscr{D}^s$ $(l \geqq 0)$; in particular $X \circ \eta \in T_\eta\mathscr{D}^s$ and so it is a vector field on $\mathscr{D}^s$. In fact $\tilde{X}$ is a right invariant C^l vector field on $\mathscr{D}^s$ (i.e., $(R_\eta) * (\tilde{X}_\zeta) = \tilde{X}_\zeta \circ \eta$ for $\eta \in \mathscr{D}^s$ and $\tilde{X}_\zeta = X(\zeta) \in T_\zeta\mathscr{D}^s$). Conversely if $\tilde{X}$ is a right invariant C^l vector field, then $\tilde{X}(e) \in \mathscr{X}^{s+l}$. In fact the right invariant C^l vector fields are isomorphic to $\mathscr{X}^{s+l}$ by evaluation at e, and in particular $T_e\mathscr{D}^s$ is isomorphic to the C^0 right invariant vector fields.

For $l \geqq 1$, there is a natural Lie bracket operation on the C^l right invariant vector fields on $\mathscr{D}^s$. This defines the bracket operation on the corresponding members of $T_e\mathscr{D}^s(M)$. We now establish that the Lie algebra structure of $\mathscr{D}^s$ is the usual Lie algebra structure on the vector fields.

2.4.1 THEOREM. *Let $l \geqq 1$ and for X, $Y \in H^{s+l}(TM)$, let $\tilde{X}$ and $\tilde{Y}$ be the corresponding right invariant vector fields on $\mathscr{D}^s$. Then $[\tilde{X}, \tilde{Y}]_e = [X, Y]$, the usual Lie bracket of vector fields on M.*

PROOF. Recall that locally $[X, Y] = DX \cdot Y - DY \cdot X$ (where DX is the derivative of X; cf. Lang [1]). However, as shown above, for $\eta \in \mathscr{D}^s$, $\tilde{X}(\eta) = X \circ \eta$ and $\tilde{Y}(\eta) = Y \circ \eta$, so in particular since $T\tilde{Y} \cdot X = TY \cdot X$ we get $[\tilde{X}, \tilde{Y}]_e = (D\tilde{X} \cdot \tilde{Y} - D\tilde{Y} \cdot \tilde{X})_e = D\tilde{X}(e) \cdot \tilde{Y}(e) - D\tilde{Y}(e)\tilde{X}(e) = DX \cdot Y - DY \cdot X$. ■

Note since $DX \cdot Y \in H^{s+l-1}(TM)$, we really cannot put this bracket on $T_e\mathscr{D}^s = \mathscr{X}^s$ and none of the $\mathscr{D}^{s+l}(TM)$ are Lie algebras since they are not closed under the bracket operation; one would have to pass to $\mathscr{D} = \mathscr{D}^\infty$.

For any Lie group G, there is a standard exp map from $\mathscr{J}$ onto a neighborhood of the identity e in G. If $X \in \mathscr{J}$, there is a unique one parameter smooth subgroup c in G (i.e., $c(t + s) = c(t) \cdot c(s)$ and $c(0) = e$) such that $c'(0) = X$. In this case X is the infinitesimal generator of c; c is the solution of $c'(t) = \tilde{X}(c(t))$ where $\tilde{X}$ is the right invariant vector field equaling X at e. Define $\exp(X) = c(1)$.

If G has a Riemannian structure, then there is another $\overline{\exp}: \mathscr{J} \to G$ defined (as above) by following geodesics instead of subgroups. If the metric is bi-invariant (i.e., if $g = (g_{ij})$ is the Riemannian metric, then for $a \in G$, $(R_a)^*(g) = (L_a)^*(g) = (g)$) then it is easy to show the two exp maps coincide. A compact group always has a bi-invariant metric. (cf. Milnor [1] for further information.)

In the case of $\mathscr{D}^s$, we will construct a metric that is right invariant, but not left invariant, and so the two exp maps will in general be different. It seems unlikely that $\mathscr{D}^s(M)$ has any bi-invariant metrics. (Any bi-invariant $B: \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ would have to satisfy $B([X, Y], X) = 0$ for all X, $Y \in \mathscr{X}(M)$ as is easy to see.)

Let $X \in T_e\mathscr{D}^s$. Then $X \in H^s(TM) = \mathscr{X}^s$ and therefore has a flow F_t ($F_t(m)$ is the integral curve of X starting at m). This is a one parameter group since $F_{s+t} = F_s \circ F_t$. Since M is compact, F_t is defined on all of M for all $t \in \mathbb{R}$.

Let us argue that we should have

$$\exp X = F_1$$

where exp is the (right) exponential map on $\mathscr{D}$. Indeed we need to show F_t is an integral curve of $\tilde{X}$ defined above. But

$$\frac{d}{dt} F_t(m) = X(F_t(m))$$

or

$$\frac{d}{dt} F_t = X \circ F_t = \tilde{X}(F_t).$$

Hence F_t is an integral curve in $\mathscr{D}^s$ of $\tilde{X}$. This justifies us in saying $\exp X = F_1$.

Actually it is not obvious that $F_t \in \mathscr{D}^s$; i.e., the flow of an H^s vector field is H^s. (This, of course, is well-known in the C^k case Lang [1].) However we do have this theorem:

2.4.2 THEOREM. *If X is in $\mathscr{X}^s(M) = T_e(\mathscr{D}^s)$ and its flow is F_t, and if $s > (n/2) + 2$, then F_t is a C^1 curve in $\mathscr{D}^s(M)$.*

Interestingly, the classical proof for the C^k case from ordinary differential equations does not seem to work. The following proof, taken from Ebin-Marsden [1] was inspired by Robbin [1]. This theorem is true for $s > (n/2) + 1$ as has been shown by H. Brezis in an unpublished communication, and by Fischer and Marsden [2].

2.4.3 LEMMA. *Let X be an H_s vector field, $s > (n/2) + 2$ and, F_t the flow of X. Then $F_t \in \mathscr{D}^{s-1}$ for short $|t| < \epsilon$ and $t \mapsto F_t$ is a C^1 curve in $\mathscr{D}^{s-1}$.*

PROOF. Since X is C^1, it has a C^1 flow F_t defined for all $t \in \mathbb{R}$. Consider the Hilbert manifold $\mathscr{D}^{s-1}(M)$ and its tangent bundle, elements of

$H^{s-1}(M, TM)$ which cover some $\eta \in \mathscr{D}^{s-1}$. Thus as X is H^s,

$$\omega_X : \mathscr{D}^{s-1} \to T\mathscr{D}^{s-1}$$
$$\eta \mapsto X \circ \eta$$

is a C^1 map by the omega lemma. Thus ω_X has a unique integral curve $F_t \in \mathscr{D}^{s-1}$ with $F_0 =$ identity, and $t \mapsto F_t \in H^{s-1}(M, M)$ is C^1, $|t| < \varepsilon$. Thus as $s > (n/2) + 2$, F_t is a C^1 diffeomorphism, and $dF_t/dt = X \circ F_t$, so F_t is the flow of X. ■

We shall also need the fact that $F_t \in \mathscr{D}^{s-1}$ depends continuously on $X \in H^s(TM)$. Indeed, the vector field ω_X depends continuously on X, so integral curves do as well, for $|t| < \varepsilon$.

Note. If one knew that C^0 vector fields on $\mathscr{D}^s$ had integral curves, we would have $F_t \in \mathscr{D}^s$ in the above. However this is not the case; but we still get our result by the following procedure.

2.4.4 LEMMA. *Under the hypotheses of* 2.4.3, $F_t : M \to M$ *is an* H^s *map for small* t *(so* $F_t \in \mathscr{D}^s$ *for* $|t| < \varepsilon$*), and* $t \mapsto F_t$ *is* C^1 *as a map from* $\mathbb{R}$ *to* $\mathscr{D}^s$.

PROOF. Since X and F_t are of class C^1, we have $(d/dt)DF_t = (DX \cdot F_t) \cdot DF_t$. Consider this as an equation

$$\frac{d}{dt} u(t) = B_t \cdot u(t)$$

in a space of H^{s-1} maps with $u(0) =$ identity. Here $B_t = DX \circ F_t$ is a continuous linear operator on H^{s-1} and H^{s-2}. Thus there is a unique solution in H^{s-1}. But DF_t is a solution in H^{s-2}. Hence DF_t is in fact in H^{s-1} so F_t is H^s and the lemma follows. ■

PROOF OF 2.4.2 Let F_t be the flow of X, so F_t is a C^1 map defined for all $t \in \mathbb{R}$. We want to show $F_t \in \mathscr{D}^s$ for all $t \in \mathbb{R}$. By the argument in 2.4.4, $F_t \in \mathscr{D}^s$ if t is small. For general t write, for each integer $n > 0$,

$$F_t = (F_{t/n})^n = F_{t/n} \circ F_{t/n} \circ \ldots \circ F_{t/n}.$$

For n large, $F_{t/n} \in \mathscr{D}^s$, and since $\mathscr{D}^s$ is a group, $F_t \in \mathscr{D}^s$. ■

For an alternative proof, see Fischer-Marsden **[2]**. It has been remarked by Brezis **[1]** that these arguments also apply to the spaces $W^{p,k}$. They also hold for the Holder spaces $C^{k+\alpha}$.

So via this theorem and the remark that $F_0 = id$, we have a sort of Lie group exponential map from $T_e\mathscr{D}^s(M)$ *into* a neighborhood of identity, $X \mapsto F_1$. It is natural to ask why not use this exp map to directly define charts on $\mathscr{D}^s(M)$. We cannot do this because it is a fact that exp does *not* map onto any neighborhood of the identity in $\mathscr{D}^s(M)$. This is equivalent to saying that there are diffeomorphisms near e not embeddable in a flow. In other words for any neighborhood U of e in $\mathscr{D}^s$, there is $\eta \in U$ such that there is no flow F_t with $F_1 = \eta$. In fact η will not, in general, have a square root. Explicit examples have been given by several people such as Eells and Smale. One is written down in Omori **[1]** and in Friefeld **[1]**.

A consequence of this is that the exp map on $T_e\mathscr{D}^s$ is not C^1, for if it were, it would be locally onto by the inverse function theorem.

2.5 Volume preserving diffeomorphisms

For now let M be a compact Riemannian manifold without boundary. (The boundary case is done below.) Let μ be the volume form given by the metric on M. Recall from the introduction that $\mathscr{D}^s_\mu = \{f \in \mathscr{D}^s \mid f^*(\mu) = \mu\}$. We shall show that $\mathscr{D}^s_\mu$ is a smooth submanifold of $\mathscr{D}^s$.

Recall that if $f: P \to Q$ is a smooth map between manifolds, f is a *submersion* on a set $A \subset P$ if $T_x f: T_x P \to T_{f(x)} Q$ is a surjection, for each $x \in A$. We shall need these lemmas:

2.5.1 LEMMA. *Let* P, Q *be Hilbert manifolds and* $f: P \to Q$ *a* C^∞ *map, then for* $g \in Q$, $f^{-1}(g)$ *is a* C^∞ *submanifold of* P, *if* f *is a submersion on* $f^{-1}(g)$.

This is simply a corollary of the Implicit function theorem (see Lang **[1]**). In the Banach case, one also assumes ker $T_x f$ has a closed complement.

2.5.2 LEMMA. *Let* λ *be an n-form on* M *such that* $\int_M \lambda = 0$. *Then* λ *is exact;* $\lambda = d\alpha$ *for an* $n - 1$ *form* α.

This is a special case of de Rham's theorem, stating that a closed form is exact if all its periods vanish. For the proof, see for example Warner **[1]**. A discussion is also found in Flanders **[1]**.

For $\partial M = \varnothing$, the following theorem is due to Ebin **[1]**.

2.5.3 THEOREM. Let $s > (n/2) + 1$. *Then* $\mathscr{D}^s_\mu$ *is a closed* C^∞ *submanifold of* $\mathscr{D}^s$.

REMARK. This seems to be false for C^k diffeomorphisms since it depends on the Hodge theorem.

PROOF. Let μ be the volume form on M. Now by the Hodge theorem, $[\mu]_s = \mu + d(H^{s+1}(\Lambda^{n-1}))$ is a closed affine subspace of $H^s(\Lambda^n)$, being the translate of the closed subspace $d(H^{s+1}(\Lambda^{n-1}))$ by μ. Define the map

$$\psi:\mathscr{D}^{s+1}(M) \to [\mu]_s$$
$$\eta \mapsto \eta^*\mu.$$

Now $\eta^*\mu \in [\mu]_s$ since

$$\int_M (\mu - \eta^*\mu) = \int_M \mu - \int_M \eta^*\mu = 0.$$

Hence $\mu - \eta^*\mu = d\alpha$ by Lemma 2. By the ω-Lemma, one can easily see that ψ is a C^∞ map. Now $\mathscr{D}_\mu^{s+1}(M) = \psi^{-1}(\mu)$, so by Lemma 1, if ψ is a submersion then $\mathscr{D}_\mu^{s+1}(M)$ is a C^∞ submanifold of $\mathscr{D}^{s+1}(M)$.

We shall show this at $e \in \mathscr{D}^s(M)$ (e is the identity map). It turns out that $T_e\psi(X) = L_X\mu$ where $X \in T_e\mathscr{D}^{s+1}(M)$. Indeed let η_t be a curve tangent to X, such as its flow. Then $T_e\psi(X) = (d/dt)\eta_t^*\mu\,|_{t=0}$ which is well known to be the Lie derivative (Abraham [2].) Using the standard formula $L_X\mu = di_X\mu + i_Xd\mu$ for the Lie derivative and the fact that $d\mu = 0$, we get

$$T_e\psi(X) = L_X\mu = di_X\mu.$$

Hence to show $T_e\psi$ is a surjection, we only need show that

$$\{i_X\mu : X \in T_e\mathscr{D}^{s+1}\} = H^{s+1}(\Lambda^{n-1}).$$

But $i_X\mu = *\tilde{X}$ and $*$ is a bijection between $n-1$ forms and 1-forms. Hence $T_e\psi$ is onto. Similarly $T_\eta\psi$ is onto. ■

However this last step in the proof only holds if μ is a (nowhere zero) n-form or a *closed* nondegenerate 2-form. This remark allows us to show that the diffeomorphisms that preserve a symplectic form form a submanifold of the diffeomorphism group using the same sort of argument.

It is well known that a vector field generates volume preserving diffeomorphisms if and only if it is divergent free (see Abraham [2]). In our context this means $T_e\mathscr{D}_\mu^s(M) = \{X \in \mathscr{D}^s(M) \mid \delta\tilde{X} = 0\}$. This is clearly a subspace of $T_e\mathscr{D}^s(M)$ and in fact $T_e\mathscr{D}_\mu^s(M)$ is closed under the bracket operation, in the same sense as $T_e\mathscr{D}^s(M)$ (see above).

2.6 Manifolds with boundary

Suppose M is a compact, oriented, Riemannian Manifold with smooth boundary. Let $\tilde{M}$ be the double of M, i.e., $\tilde{M}$ is two copies of M with the boundaries identified, with the obvious differential structure. Now $\tilde{M}$ is a compact, oriented, Riemannian manifold without boundary and M has a natural imbedding in $\tilde{M}$. We have the manifold structure of $H^s(M, \tilde{M})$ by our above work. Clearly $\mathscr{D}^s(M) \subset H^s(M, \tilde{M})$ and in fact:

2.6.1 THEOREM. *$\mathscr{D}^s(M)$ is a C^∞ submanifold of $H^s(M, \tilde{M})$.*

SKETCH OF PROOF. Briefly, we put a metric on $\tilde{M}$ such that $\partial M \subset \tilde{M}$ is totally geodesic. Then let $E: TH^s(M, \tilde{M}) \to H^s(M, \tilde{M})$ be the exponential map associated with this metric.

Let $\eta \in \mathscr{D}^s(M) \subset H^s(M, \tilde{M})$ and choose an exponential chart $E: U \subset T_\eta H^s(M, \tilde{M}) \to H^s(M, \tilde{M})$ about η. Also we should have

$$T_\eta\mathscr{D}^s(M) = \{X \in H^s(M, T\tilde{M}) \mid X \text{ covers } \eta \text{ and}$$
$$X(x) \in T_{\eta(x)}\,\partial M \text{ for all } x \in \partial M\}$$

which is a closed subspace of

$$H^s(M, T\tilde{M}) = TH^s(M, \tilde{M}).$$

Since ∂M is totally geodesic, E takes $U \cap T_\eta\mathscr{D}^s(M)$ onto a neighborhood of η in $\mathscr{D}^s(M)$. See Ebin-Marsden [1] for details. ■

By inspecting the above argument we see $T_e\mathscr{D}^s(M) = \{H^s$ vector fields on M that are tangent to $\partial M\}$. Formally, this is a Lie algebra in the same sense as we had when M had no boundary.

2.6.2 THEOREM. *If μ is the volume on M and $\mathscr{D}_\mu^s(M)$ is the set of volume preserving diffeomorphisms, then $\mathscr{D}_\mu^s(M) \subset \mathscr{D}^s(M)$ is a smooth submanifold.*

This is proven as in the case that M has no boundary. This proof works here because we have the Hodge theorems for manifolds with boundary.

The rest of the material from the no boundary case (such as the α and ω-lemmas) carries over to the case when M has a boundary. For the noncompact case, see Cantor [1, 2].

2.7 Topology of the diffeomorphism group

For topological theorems we can work in $\mathscr{D}(M) = \mathscr{D}^\infty(M)$. Indeed it follows from very general results of Cerf [1] and Palais [3] that the topology of $\mathscr{D}^s$ and $\mathscr{D}$ are the same; one uses the fact that the injection of $\mathscr{D}$ into $\mathscr{D}^s$ is dense. The first theorem in this field was proven by Smale [1] in 1959. He showed that $\mathscr{D}(S^2)$ is contractable to $SO(3)$; here S^2 is the 2-sphere, and $SO(3)$ is the special orthogonal group on $\mathbb{R}^3$, which we can regard as the (identity component of the) isometry group of S^2. This theorem was extended to all compact 2 manifolds by Earle and Eells **[1]** and to the boundary case by Earle and Schatz **[1]**.

It is fairly simple to show that *$\mathscr{D}(S^1)$ is contractable to $SO(2)$*. The following argument is based on a suggestion of J. Eells.

First fix $s \in S^1$. Let $\theta: [0, 1] \to S^1$ be a parameterization of S^1 such that $\theta(0) = \theta(1) = s$. Now let f be a diffeomorphism that leaves s fixed. Then the map

$$h_f(t, x): [0, 1] \times S^1 \to S^1$$

$$(t, x) \to \begin{cases} \theta(t\theta^{-1}(x) + (1 - t)\theta^{-1}(f(x))) & x \neq s \\ s & x = s \end{cases}$$

is an homotopy from f to id_{S^1}.

Suppose $g: S^1 \to S^1$ maps s to $g(s) \neq s$; then there is a rotation $r: S^1 \to S^1$ that carries $g(s)$ to s and therefore $r \circ g(s) = s$. Hence, by the above argument $r \circ g$ is homotopic to the identity. Therefore g is homotopic to r^{-1}, which is, naturally, also a rotation. ■

For dimension 3 the situation is much more complicated and little is known. The work of Cerf **[2]** seems indicative of the complexity. Antonelli et al. **[1]** have shown that if M has high dimension $\mathscr{D}(M)$ will not have the homotopy type of a finite cell complex. Various people have also been working towards showing $\mathscr{D}(M)$ is a simple group. cf. Herman **[1]**, Epstein **[1]** and Herman-Sergeraert **[1]**.

Another important result in this field is that of Omori **[1]**. He proved that for any compact Riemannian manifold without boundary $\mathscr{D}(M)$ is contractable to $\mathscr{D}_\mu(M)$, the set of volume preserving diffeomorphisms. In fact if $\mathscr{V} = \{\nu \in C^\infty(\Lambda^n) \mid \nu$ is nondegenerate, positively oriented and $\int_M \nu = \int_M \mu\}$ ($C^\infty(\Lambda^n)$ are the C^∞ n-forms) then $\mathscr{D}(M)$ is diffeomorphic to $\mathscr{D}_\mu \times \mathscr{V}$. This implies $\mathscr{D}(M)$ is contractable to $\mathscr{D}_\mu(M)$ since $\mathscr{V}$ is contractable to μ. (In fact $\mathscr{V}$ is convex.) The proof that $\mathscr{D}(M) \approx \mathscr{D}_\mu(M) \times \mathscr{V}$ uses an important result of Moser **[1]**.

2.7.1 Theorem [Moser]. *If on a compact manifold M, there are 2 volume elements μ and ν such that $\int_M \nu = \int_M \mu$, then there is map $f \in \mathscr{D}(M)$ such that $f^*(\nu) = \mu$.*

We formulate the results more precisely following the proof of Ebin-Marsden **[1]**.

2.7.2 Theorem. *Let M be compact without boundary with a smooth volume element μ. Let*

$$\mathscr{V} = \left\{\nu \in C^\infty(\Lambda^n) \mid \nu > 0, \int_M \nu = \int_M \mu\right\}.$$

Then $\mathscr{D}$ is diffeomorphic to $\mathscr{D}_\mu \times \mathscr{V}$. In particular (since $\mathscr{V}$ is convex), $\mathscr{D}_\mu$ is a deformation retract of $\mathscr{D}$.

For the proof, we begin by proving Moser's result.

2.7.3 Lemma. *There is a map $\chi: \mathscr{V}^s \to \mathscr{D}^s$, $s > (n/2) + 1$ such that $\psi_\mu: \mathscr{D}^s \to \mathscr{V}^{s-1}$, $\psi_\mu(\eta) = \eta^*(\mu)$ satisfies $\psi_\mu \circ \chi =$ identity. Further, $\chi: \mathscr{V} \to \mathscr{D}$ is a C^∞ map.*

Proof. For $\nu \in \mathscr{V}_s$, let $\nu_t = t\nu + (1 - t)\mu$, so that $\nu_t \in \mathscr{V}^s$. Since $\int \mu = \int \nu$, we can write, as before, $\mu - \nu = d\alpha$. Define X_t by $i_{X_t}\nu_t = \alpha$ so that $X_t \in H^s(TM)$. Let η_t be the flow of X_t, so $\eta_t \in \mathscr{D}^s$. Define $\chi(\nu) = \eta_1^{-1}$ We want to show that $\eta_t^*(\nu_t) = \mu$ by showing $d/dt(\eta_t^*(\nu_t)) = 0$. Indeed, we have, from the basic fact about Lie derivatives (Abraham **[2]**, p. 52):

$$\begin{aligned} \frac{d}{dt}(\eta_t^*(\nu_t)) &= \eta_t^*\left(L_{X_t}\nu_t + \frac{d}{dt}\nu_t\right) \\ &= \eta_t^*(di_{X_t}\nu_t - (\mu - \nu)) \\ &= \eta_t^*(d\alpha - \mu + \nu) = 0. \quad ■ \end{aligned}$$

Note that χ is canonically defined, given the Riemannian metric on M.

Proof of 2.7.2. Define $\Phi: \mathscr{D}_\mu \times \mathscr{V} \to \mathscr{D}$ by $\Phi(\zeta, \nu) = \zeta \circ \chi(\nu)$. Then $\Phi^{-1}(\eta) = (\eta \circ (\chi\eta^*)(\mu))^{-1}, \eta^*(\mu))$ as is easily checked. ■

This can be generalized to the boundary case as well.

The basic technique in 2.7.3 has found considerable applications recently. See for example Weinstein **[1]** and Palais **[5]**.

It is also possible to study other groups of diffeomorphisms. (See also Leslie **[3]**.) For example, let M be a compact manifold and let G be a compact group. Let $\Phi: G \times M \to M$ be a group action, and let $\Phi_g(m) = \Phi(g, m)$. Set

$$\mathscr{D}^s_\Phi(M) = \{\eta \in \mathscr{D}^s(M) \mid \eta \circ \Phi_g = \Phi_g \circ \eta\}.$$

This is a subgroup of $\mathscr{D}^s(M)$, a C^∞ submanifold and has "Lie algebra"

$$T_e\mathscr{D}^s_\Phi(M) = \{V \in T_e\mathscr{D}^s(M) \mid V \text{ commutes with all infinitesimal generators of } \Phi\}.$$

Of course, we can also take $\mathscr{D}^s_{\mu,\Phi}(M) = \mathscr{D}^s_\mu(M) \cap \mathscr{D}^s_\Phi(M)$. Since this intersection is not in general transversal, it is not obvious that $\mathscr{D}_{\mu,\Phi}(M)$ is a submanifold. It is true, but requires some argument (see Ebin-Marsden **[2]**). The group $\mathscr{D}^s_{\mu,\Phi}(M)$ is important in the study of flows with various symmetries (e.g., a flow in $\mathbb{R}^3$ that is symmetric with respect to a given axis). Also, in general we find that $\dim(\mathscr{D}_\Phi(M))$ and $\operatorname{codim}(\mathscr{D}_\Phi(M))$ are both infinite so Frobenius methods do not work (Leslie [2] and Omori [1, 3] have shown that if $\mathscr{J}$ is a Lie subalgebra of $T_e\mathscr{D}$ with finite dimension or codimension, then $\mathscr{J}$ comes from a smooth subgroup of $\mathscr{D}$).

3. Bifurcation and the centre manifold theorem

3.1 Hopf bifurcation

This theorem was invented by Hopf [1] in 1942 apparently with fluid mechanics in mind. See also Hopf [2]. Just recently, David Ruelle and Floris Takens [1] published a paper in which this theorem was used in a discussion of turbulence and stability. Later in Part II, § 3 we shall consider such questions in detail.

In order to understand Hopf's theorem, let us review some standard material in ordinary differential equations. For a complete discussion of this material, see Coddington-Levinson [1] and Abraham-Robbin [1]. Let $X: \mathbb{R}^n \to \mathbb{R}^n$ be a linear map. Then regarding X as a vector field on $\mathbb{R}^n$, its flow is given by $F_t(a) = e^{tX}(a)$, where $a \in \mathbb{R}^n$ and $e^{tX} = \sum_{n=0}^{\infty} (t^n X^n/n!)$; in this expression $X^0 = I$ and multiplication is as matrices. Let $\lambda_1, \ldots, \lambda_k$ be the (possibly complex) eigenvalues of X. Since X has only real entries when considered as a matrix, the λ_i appear in conjugate pairs. Clearly $e^{t\lambda_1}, \ldots, e^{t\lambda_k}$ are the eigenvalues of F_t.

Now suppose that for all i, we have $\operatorname{Re}(\lambda_i) < 0$. Then as t increases $|e^{\lambda_i t}|$ is decreasing and hence the orbit of a point $a \in \mathbb{R}^n$ i.e., the curve $t \mapsto F_t(a)$, is approaching zero. (This is clear if X is diagonalizable; for the general case one uses the Jordan canonical form.) Since F_t is linear, for each t we have $F_t(0) = 0$. In this situation, we say 0 is an *attracting* or *stable fixed point*.

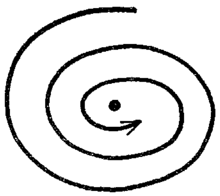


FIGURE 3.1

Now if all $\operatorname{Re}(\lambda_i) > 0$, it is clear that each $|e^{t\lambda_i}|$ is increasing with t, and so the orbit of a point under the flow is away from 0. Here, we say 0 is a *repelling* or *unstable fixed point*.

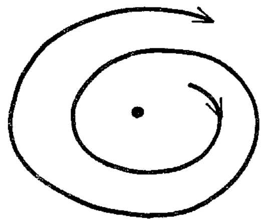


FIGURE 3.2

For the nonlinear case, we linearize and apply the above results as follows. Let X be a vector field on some manifold M. Suppose there is a point $m_0 \in M$ such that $X(m_0) = 0$. Then F_t, the flow of X leaves m_0 fixed; $F_t(m_0) = m_0$. It makes sense to consider $DX(m_0): T_{m_0}M \to T_{m_0}M$. If $y_1, \ldots, y_n$ is a coordinate system for M at m_0, the coordinate matrix expression for $DX(m_0)$ is just $DX(m_0) = (\partial X^i/\partial y^j)(m_0)$. Now, $DX(m_0)$ can be treated as a linear map on $\mathbb{R}^n$ and the same analysis as above applies. Hence m_0 is an attracting or repelling fixed point (or neither) for the flow of X depending on the sign of the real part of the eigenvalues of $(\partial X^i/\partial y^j)(m_0)$. However if m_0 is attracting (when the real parts of the eigenvalues are <0), it is only *nearby* points which $\to m_0$ as $t \to \infty$.

To begin our study of the Hopf theorem, let us consider some physical examples of the general phenomenon of *bifurcation*. The idea in each case is that the system depends on some real parameter, and the system undergoes a sudden qualitative change as the parameter crosses some critical point. (For research in a slightly different direction and for more examples, consult the papers in Antman-Keller [1] and Zarantonello [1].)

EXAMPLE 1. Consider a rigid rod with force F applied equally at both ends (Figure 3.3). If the force F is small, we observe no change in the rod

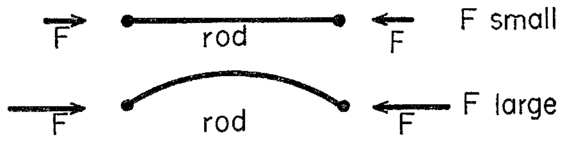


FIGURE 3.3

(except a small compression perhaps). However as the force increases, eventually it will reach a point where the rod suddenly bends.

EXAMPLE 2. (Couette Flow). Suppose we have a viscous fluid between two concentric cylinders (Figure 3.4). Suppose further we forcibly rotate

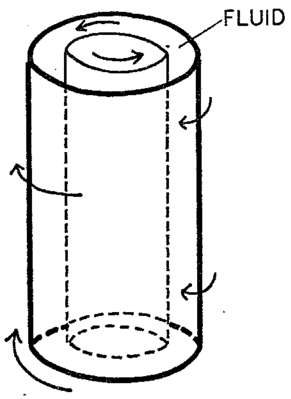


FIGURE 3.4

the cylinders in opposite directions at some constant angular velocity ρ (which is our parameter). For ρ near 0, we get a steady horizontal laminar flow in the fluid. However as ρ reaches some critical point, the fluid breaks up into what are called *Taylor cells* and the fluid moves radially in cells from the inner cylinder to the outer one and vice versa; Figure 3.5. (Note, that the directions of flow are such that flow is continuous.) We shall discuss this flow at greater length later.

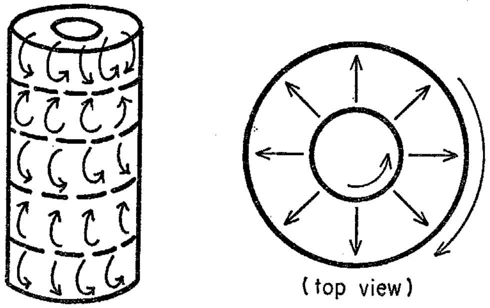


FIGURE 3.5

In both of the above examples, we have a situation described by differential equations and at some critical point of the parameter, the given solution becomes unstable and the system shifts to a "stable" solution. This sharp division of solutions is the sort of bifurcation we shall encounter in Hopf's theorem.

For simplicity, let us consider the case where the underlying space is simply $\mathbb{R}^2$. Let X_μ be a vector field on $\mathbb{R}^2$ depending smoothly on some real parameter μ. Actually it is convenient to put X_μ in $\mathbb{R}^3$ by considering

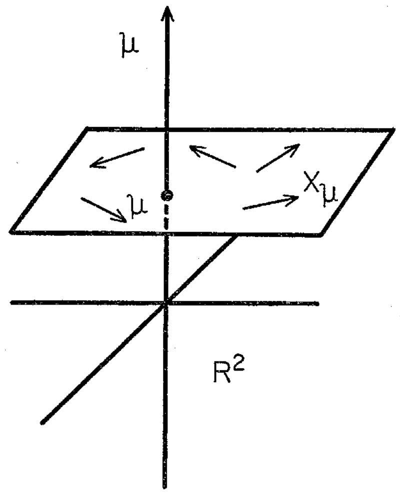


FIGURE 3.6

the map $\tilde{X}:(x, y, \mu) \to (X_\mu(x, y), 0)$. This way we can graph the flow F_t^μ of X_μ and keep track of the parameter μ. The flow G_t of $\tilde{X}$ is $G_t(x, y, \mu) = (F_t^\mu(x, y), \mu)$. Similarly, we consider X_μ acting on the plane $\mu =$ const as in Figure 3.6.

Now suppose $X_\mu(0, 0) = (0, 0)$ for each μ; more generally one could consider a curve (x_μ, y_μ) of critical points of X_μ. We can apply the analysis we developed for vector fields, i.e., for each μ, we look at the eigenvalues of $DX_\mu(0, 0)$ say $\lambda(\mu)$ and $\overline{\lambda(\mu)}$. (They are complex conjugate.) Note that the eigenvalues depend on μ and by our earlier analysis of flows, we know the qualitative behaviour of the flow depends on the sign of $\mathrm{Re}(\lambda(\mu))$ and $\mathrm{Re}(\overline{\lambda(\mu)})$ (which are equal in case $\lambda(\mu)$ itself is not real). So if we know how $\lambda(\mu)$ depends on μ then we can hope to extract some information about the flow near (0, 0) as μ increases. We make these hypotheses:

Suppose $\mathrm{Re}(\lambda(\mu)) < 0$ *for* $\mu < 0$ *and* $\mathrm{Re}(\lambda(0)) = 0$ *and* $\mathrm{Re}(\lambda(\mu))$ *is increasing as* μ *increases across* 0 (Figure 3.7). *Also assume that*

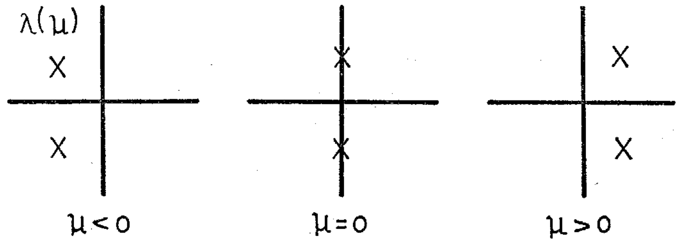


FIGURE 3.7

$\lambda(\mu)$ is not real and that for $\mu = 0$, $(0, 0)$ is an attracting fixed point for X
(perhaps with a weaker or slower attraction than when $\mathrm{Re}(\lambda(\mu)) < 0$).

Now for $\mu < 0$, we know from the above that the flow is "stable," i.e., points near $(0, 0)$ are carried towards $(0, 0)$ by the flow, as is the case for

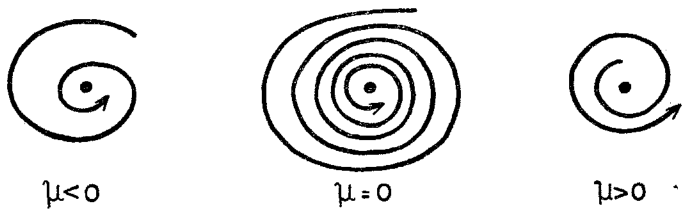


FIGURE 3.8

$\mu = 0$ (only slower) by assumption. The surprising case is the behavior for $\mu > 0$.

3.1.1 THEOREM (E. HOPF). *In the situation described above, there is a stable periodic orbit for X_μ when $0 < \mu < \varepsilon$ for some $\varepsilon > 0$. (Stable here means points near the periodic orbit will remain near and eventually be carried closer to the orbit by the flow.)*

So as in the examples we get a qualitative change in the stable solutions as μ crosses 0, from an attracting fixed point at $(0, 0)$ to a periodic solution away from $(0, 0)$. (Figures 3.9, 3.10)

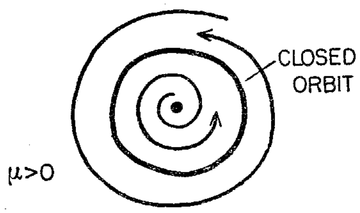


FIGURE 3.9

This theorem does generalize to $\mathbb{R}^n$ where we can get tori forming as the stable solutions (instead of closed orbits) as further bifurcations take place; see Ruelle-Takens [1] for details.

The proof of the theorem occurs in many places besides Hopf [1]. See, for instance Andronov and Chaikin [1], or Bruslinskaya [3]. We have followed Ruelle-Takens [1], correcting a minor error in their 6.1.

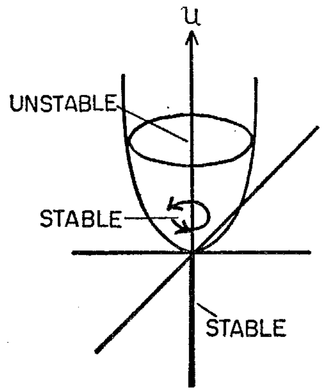


FIGURE 3.10

Hopf's theorem is closely related to a linear model used in physics known as the "Turing model." As D. Ruelle, S. Smale, and P. Hartman have remarked, these sort of phenomena may be basic for understanding a large variety of qualitative changes which occur in nature, including biological and chemical systems. See for instance Turing [1], Selkov [1]. We have examined here only one of many types of possible bifurcations. There are many others which occur in Thom's theory of morphogenesis (see articles in Chillingworth [1] and Abraham [4] for more details and bibliography). Meyer [1] is representative of the Hamiltonian case.

3.2 The centre manifold theorem

This is an important existence theorem which perhaps has been underrated as to its importance. We will first consider the case of flows. Let $\mathbb{E}_1$, $\mathbb{E}_2$, and $\mathbb{E}_3$ be Banach spaces and $X_i: \mathbb{E}_i \to \mathbb{E}_i$ be vector fields such that for each i, $X_i(0) = 0$. Suppose:

$$\text{if } \lambda \text{ is in the spectrum of } DX_1(0),\ \mathrm{Re}(\lambda) < -\delta < 0$$

$$\text{if } \lambda \text{ is in the spectrum of } DX_2(0),\ \mathrm{Re}(\lambda) = 0$$

$$\text{if } \lambda \text{ is in the spectrum of } DX_3(0),\ \mathrm{Re}(\lambda) > \delta > 0$$

where $\delta > 0$. For example, let $\mathbb{E}_1 = \mathbb{E}_2 = \mathbb{E}_3 = \mathbb{R}$ with $\mathbb{E}_1$ the x-axis, $\mathbb{E}_2$ the y-axis and $\mathbb{E}_3$ the z-axis. We get the picture as shown in Figure 3.11 for the linearized fields (i.e., the fields $DX_i(0)$).

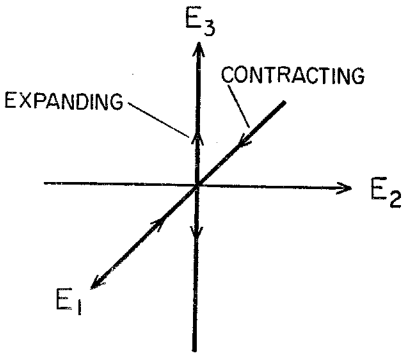


FIGURE 3.11

3.2.1 THEOREM. *In the above situation there is an invariant submanifold C of (X_1, X_2, X_3) in $\mathbb{E}_1 \times \mathbb{E}_2 \times \mathbb{E}_3$ as shown in figure* 3.12. (*Here invariant means that if the flow of (X_1, X_2, X_3) has initial conditions in C, it remains in C for all time.*)

There similarly are *stable* and *unstable manifolds* for the vector field, the points of which converge to 0 as $t \to \infty$ and $t \to -\infty$ respectively.

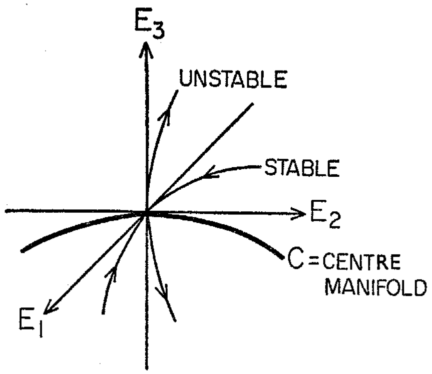


FIGURE 3.12

The proof of the theorem for the case of Euclidean space may be found in Appendix C by A. Kelley of Abraham-Robbin [1]. Some discussion is also found in Abraham [2], and Hartman [1]. As far as we know, the general proof has not appeared, but is forthcoming in Hirsch-Pugh-Shub [2].

There is a similar theorem for a map V (not vector fields) on $\mathbb{E}_1 \times \mathbb{E}_2 \times \mathbb{E}_3$. Here, as above, we assume $V_i(0) = 0$ or 0 is a fixed point for V and

assume that

if e^λ is in the spectrum of $DV_1(0): \mathbb{E}_1 \to \mathbb{E}_1$ then $\mathrm{Re}(\lambda) < -\delta < 0$

if e^λ is in the spectrum of $DV_2(0): \mathbb{E}_2 \to \mathbb{E}_2$ then $\mathrm{Re}(\lambda) = 0$

if e^λ is in the spectrum of $DV_3(0): \mathbb{E}_3 \to \mathbb{E}_3$ then $\mathrm{Re}(\lambda) > \delta > 0$,

then there are, as above, appropriate invariant submanifolds for the map V. Again the reference is Hirsch-Pugh-Shub [2]. Note that the map V need not be a diffeomorphism (that the theorem is valid for general maps was pointed out by C. Pugh, who we thank for discussions on this point).

The difference between the map and the vector field theorems is made clearer if one remembers that if $V = F_t$ where F_t is the flow of X, then the aforementioned conditions on V and X are equivalent to each other, since the spectra are related by exponentiating, as was explained previously.

3.3 Bifurcation theorem for maps

An application of the centre manifold theory is the following basic bifurcation theorem. We follow a formulation in Ruelle-Takens [1]. Theorems of this type have been discovered by many different authors in various contexts.

3.3.1 THEOREM. *Let $\mathbb{H}$ be a Hilbert space (or manifold) and $\Phi_\mu: \mathbb{H} \to \mathbb{H}$ a map defined for each $\mu \in \mathbb{R}$ such that the map $(\mu, x) \mapsto \Phi_\mu(x)$ is a C^k map, $k \geqq 1$, from $\mathbb{R} \times \mathbb{H}$ to $\mathbb{H}$, and for all $\mu \in R$, $\Phi_\mu(0) = 0$. Define $L_\mu = D\Phi_\mu(0)$ and suppose the spectrum of L_μ lies inside the unit circle for $\mu < 0$. Assume further there is a real, simple, isolated eigenvalue $\lambda(\mu)$ of L_μ such that $\lambda(0) = 1$, $(d\lambda/d\mu)(0) > 0$, and L_0^* has the eigenvalue 1; then there is a C^{k-1} curve l of fixed points of $\Phi: (x, \mu) \mapsto (\Phi_\mu(x), \mu)$ near $(0, 0) \in \mathbb{H} \times \mathbb{R}$. The curve is tangent to $\mathbb{H}$ at $(0, 0)$ in $\mathbb{H} \times \mathbb{R}$* (Figure 3.14). *These points and the points $(0, \mu)$ are the only fixed points of Φ in a neighborhood of $(0, 0)$. Moreover these new fixed points are stable (figure* 3.15).

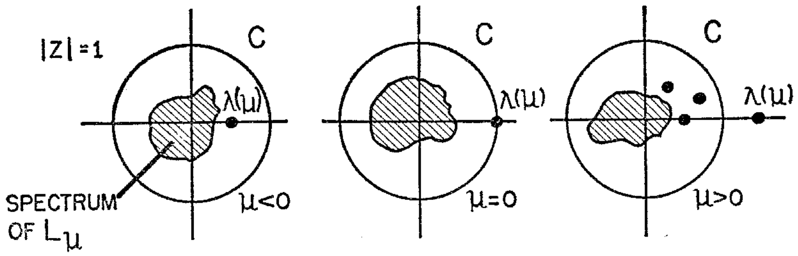


FIGURE 3.13

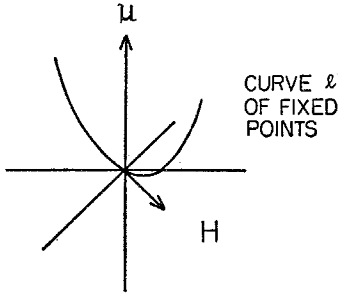


FIGURE 3.14

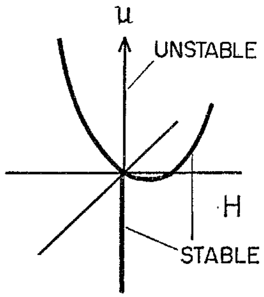


FIGURE 3.15

Before sketching the proof of 3.3.1, we give an example which is instructive here (kindly pointed out by Professor Calabi). Consider a circular track free to rotate about a vertical axis and a ball in the track. Suppose we rotate the ring around this vertical axis with angular velocity μ. Then if μ is small there is just one stable rest solution, namely at the bottom. As μ increases past a critical value, there appear two additional stable rest solutions on either side of the solution at the bottom which has now become unstable (Figure 3.16). If the stable solution is a height h_μ from the bottom, observe that h_μ increases with μ.

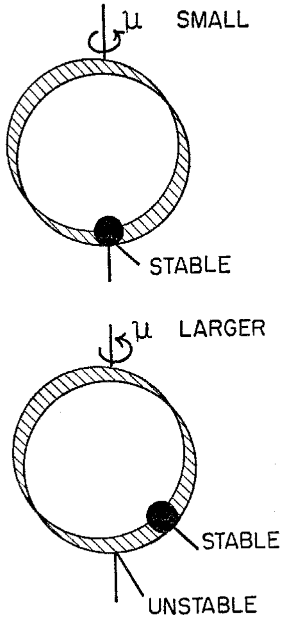


FIGURE 3.16

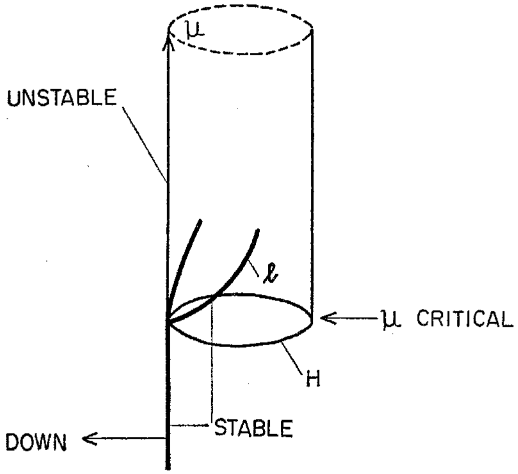


FIGURE 3.17

In this example we can let $\mathbb{H}$ = a portion of the loop parametrized by the distance s from the bottom, and $\Phi_\mu(s)$ = the position of the ball after 1 second if starting at s. Let 0 = "bottom of the circle," so $\Phi_\mu(0) = 0$. We then have the picture shown in Figure 3.17. Here the points on l correspond to the fixed point of Φ_μ when the ball is at height h_μ.

The idea involved in the proof of the result 3.3.1 is the following. We pick an eigenvector z for $(L_0, 0)$ in $\mathbb{H} \times \{0\}$ with eigenvalue one. Thus we get a two dimensional centre manifold as shown in Figure 3.18. Choose coordinates α, μ on C determined by: μ is the coordinate as above, and α is obtained by projection on the eigenvector $z(\mu)$ for L_μ. Write $\Phi(\alpha, \mu) = (f(\alpha, \mu), \mu)$, α, $\mu \in \mathbb{R}$ so that $f(0, \mu) = 0$, $(\partial f/\partial\alpha)(0, \mu) = \lambda(\mu)$. Set

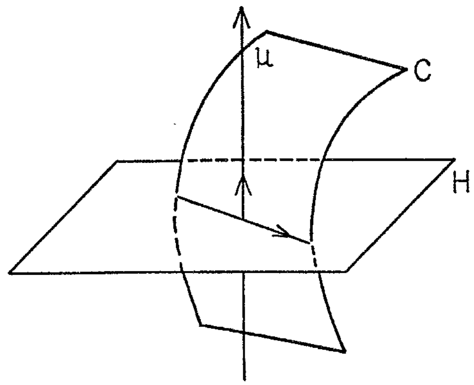


FIGURE 3.18

$g(\alpha, \mu) = [f(\alpha, \mu)/\alpha] - 1$ which is C^{k-1}. The fixed points are zeros of g. But $(\partial g/\partial \mu)(0, 0) = (d\lambda/d\mu)(0) > 0$ so we get a curve of zeros of g in C by the implicit function theorem. ■

Let us see briefly how one can recover the Hopf theorem from this result. Indeed, let $\mathscr{P}_\mu : \mathbb{R} \to \mathbb{R}$ be the Poincaré map determined by the flow as shown in Figure 3.19.

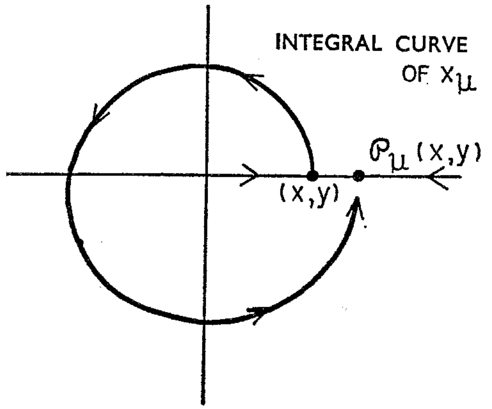


FIGURE 3.19

Then the appearance of two stable fixed points for $\mathscr{P}_\mu$ as μ increases (as the theorem gives us) corresponds exactly to an appearance for X_μ of a stable closed orbit. This is also geometrically clear from Figure 3.19.

For systems with symmetries, such as Couette flow, it is important to take these symmetries into account when dealing with these bifurcation theorems. See Ruelle [**4**].

II. APPLICATION TO HYDRODYNAMICS

We now shall apply the machinery of part I to some problems in fluid flow. The first section is devoted to showing that a curve $\eta_t \in \mathscr{D}^s_\mu$ is a geodesic if and only if the corresponding velocity field satisfies the Euler equation (see the introductory § 2). The method we use is differential geometric as opposed to Arnold's original method which was group theoretical (cf. Arnold [**1**], and Chernoff-Marsden [**1**]). We shall then explain the existence theorem and why we gain a technical advantage by working on $\mathscr{D}^s_\mu$; i.e., in Lagrangian coordinates (see remarks in the introductory § 2). We also shall prove a number of miscellaneous properties of the Euler flow.

In the second section we outline what is known concerning the existence of solutions for all time. We give a sufficient condition for the Euler flow and a weaker condition for the Navier–Stokes flow which essentially goes back to Leray [**3**]. A fairly complete proof of Leray's theorem is included.

Finally in the last section we discuss the interrelationship between the concepts of turbulence, bifurcation and stability. Here one studies changes which occur as the viscosity $\nu \to 0$ or equivalently the Reynolds number $\mathscr{R} \to \infty$. We study both the analytical aspects of the limit $\nu \to 0$, which in the boundary case is related to boundary layer theory, and the qualitative aspects using the bifurcation theory developed in Part II, § 3.

1. Geodesics on $\mathscr{D}^s_\mu$ and the Euler equations

1.1 The metric on $\mathscr{D}^s_\mu$

It follows from the section on manifolds of maps that the tangent space to $\mathscr{D}^s_\mu(M)$ at a point $\eta \in \mathscr{D}^s_\mu$ is given by $T_\eta \mathscr{D}^s_\mu(M) = \{X \in H^s(M, TM) \mid X$ covers η, $\delta(X \circ \eta^{-1}) = 0$, and X is parallel to $\partial M\}$. Note that if $X \in T_\eta \mathscr{D}^s_\mu(M)$ then $X \circ \eta^{-1}$ is a vector field on M. If we are working on $\mathscr{D}^s$ then the divergence condition $\delta(X \circ \eta^{-1}) = 0$ is dropped, so $T_\eta \mathscr{D}^s$ consists of H^s sections parallel to ∂M which cover η.

Let M be a compact Riemannian manifold $m \in M$ and let $\langle\ ,\ \rangle_m$ be the inner product on $T_m M$. Now we put a metric on $\mathscr{D}^s(M)$ as follows: Let $\eta \in \mathscr{D}^s(M)$ and $X, Y \in T_\eta \mathscr{D}^s(M)$. Then $X(m)$ and $Y(m)$ are in $T_{\eta(m)}M$. Now define:

$$(X, Y)_\eta = \int_M \langle X(m), Y(m)\rangle_{\eta(m)}\, d\mu(m).$$

This is a symmetric bilinear form on each tangent space $T_\eta \mathscr{D}^s$ of $\mathscr{D}^s(M)$. By restriction it also defines a symmetric bilinear form on each tangent space of $\mathscr{D}^s_\mu$.

The norm induced by this inner product is clearly an L^2 norm and hence the topology it induces is weaker than the H^s topology on each $T_\eta \mathscr{D}^s(M)$. Thus, we call $(\ ,\ )$ a *weak metric*. For manifolds of maps it is important to allow weak metrics although most definitions of Riemannian manifolds exclude this (as in Lang [**1**]). Also, it should be mentioned that for physical reasons this is the appropriate metric for hydrodynamics, since for $X \in T_\eta \mathscr{D}^s(M)$, $\frac{1}{2}(X, X)$ represents the total kinetic energy of a fluid in state η and velocity field $v = X \circ \eta^{-1}$ (see below). So finding geodesics is formally the same as finding a flow satisfying a least energy condition.

This metric $(\,,\,)_\eta$ just constructed is smooth in this sense: If $B(T_\eta\mathscr{D}^s_\mu, T_\eta\mathscr{D}^s_\mu)$ is the vector bundle of bilinear maps over the tangent spaces of $\mathscr{D}^s(M)$ (i.e., if $g_\eta \in B(T_\eta\mathscr{D}^s_\mu, T_\eta\mathscr{D}^s_\mu)$ then $g_\eta: T_\eta\mathscr{D}^s_\mu(M) \times T_\eta\mathscr{D}^s_\mu(M) \to R$ is bilinear), then the map $\eta \to (\,,\,)_\eta$ is a section of this bundle, and to say the metric is smooth is to say this section is smooth. (Here each fiber of $B(T_\eta\mathscr{D}^s_\mu, T_\eta\mathscr{D}^s_\mu)$ has the standard topology put on bilinear maps on banach spaces, and one constructs the bundle as in Lang **[1]**, Ch. III, § 4.)

Note. It is not always true that a weak metric yields goedesics. For example, suppose $\partial M \neq \varnothing$. Then on $\mathscr{D}^s(M)$, this weak metric would yield geodesics which would try to cross the boundary of M. We shall see this in more detail below.

1.2 The spray on $\mathscr{D}^s_\mu(M)$

In coordinates, one finds geodesics by solving the equation

$$\ddot{x}^i = -\sum_{j,k} \dot{x}^j\dot{x}^k\Gamma^i_{jk}(x)$$

where the $\Gamma^i_{jk}(x)$ are the Christoffel symbols. This is awkward in infinite dimensions as the geodesic equation would involve infinite sums. So, instead, we use the equivalent but more easily handled notion of a spray. To do this, recall there is a canonical manifold structure on the tangent bundle of a manifold. Specifically let Q be a C^∞ n-manifold, let (U, φ) be a chart at m, $V = \varphi^{-1}(U)$, $V \subset \mathbb{R}^n$, and $x_0 = \varphi^{-1}(m)$. Now $\Phi: U \times \mathbb{R}^n \to TQ$ defined by the expression $\Phi = (\varphi, D\varphi(x))$ is a chart on TQ since $D\varphi(x): \mathbb{R}^n \to T_mQ \approx \mathbb{R}^n$ is an isomorphism. It is easy to check that the collection of such maps form an atlas, and this determines a C^∞ manifold structure on TQ.

Define $T^2Q = T(TQ)$ to be the tangent bundle of the manifold TQ. Informally the *spray* associated to a metric on Q is a vector field $Z: TQ \to T^2Q$ such that the integral curves of Z are the curves in TQ that are the tangent curves to geodesics. (For a precise definition see Lang **[1]**.) This makes sense since integral curves of Z lie in TQ. The spray can also be defined in terms of Hamiltonian systems, as in Prof. Klingengburg's lectures; see also Abraham **[2]**, and Part III, § 1 below.

Now we will define a spray for the metric on $\mathscr{D}^s(M)$, and give plausible arguments that it is the correct choice. For a more rigorous treatment, see Ebin-Marsden **[1]**. For the rest of this paragraph suppose that $\partial M = \varnothing$. The following result is essentially contained in Eliasson **[1]**.

1.2.1 THEOREM. *Let Z be the spray of the metric on M. Then the spray of $(\,,\,)$ on $\mathscr{D}^s(M)$ is given by*

$$\bar{Z}: T\mathscr{D}^s \to T^2(\mathscr{D}^s);\; X \mapsto Z \circ X.$$

Note. As with $T\mathscr{D}^s$, it is not hard to see that $T_X(T\mathscr{D}^s)$ consists of H^s maps $Y: M \to T^2M$ which cover X; i.e., $\pi_1 \circ Y = X$, where $\pi_1: T^2M \to TM$ is the projection. The spray Z satisfies $\pi_1 \circ Z =$ identity, since Z is a vector field. Thus $\bar{Z}(X) \in T_XT\mathscr{D}^s$ so $\bar{Z}$ is indeed a vector field on $T\mathscr{D}^s$.

The idea behind the proof of 1.2.1 is to realize that we can explicitly write down what should be the geodesics on $\mathscr{D}^s(M)$. From the construction of charts on $\mathscr{D}^s(M)$, there is the map $\overline{\exp}: T\mathscr{D}^s(M) \to \mathscr{D}^s(M)$ where $\overline{\exp}(X) = \exp \circ X$ and $\exp: TM \to M$ is the Riemannian exponential map on M. First we assert that for $X \in T_e\mathscr{D}^s(M)$, the geodesic on $\mathscr{D}^s(M)$ through e in the direction X is given by $t \mapsto \overline{\exp}(tX)$. What this geodesic looks like is seen by considering any $m \in M$. Then $t \mapsto \overline{\exp}(tX)(m) = \exp(tX_m)$ is the geodesic starting at m in the direction X_m. So $\overline{\exp}(tX)$ represents all of the geodesics on M in the direction of the vector field X evaluated at $m \in M$. As in the picture, as t increases it is likely that some pair of geodesics will intersect. Say this happens at $t = t_0$. Then $\exp(t_0X)$ is not a diffeomorphism. Hence even if M is a simple manifold (like the flat 2-torus), $\mathscr{D}^s(M)$ is *not* geodesicly complete. We shall come back to this point later.

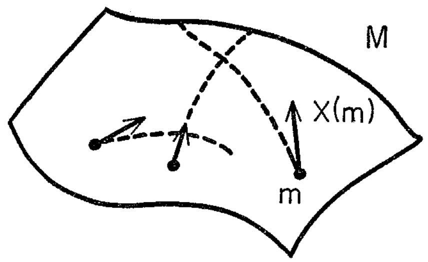


FIGURE 1.1

If we can show $t \mapsto \exp tX$ is a geodesic on $\mathscr{D}^s$, then the formula for $\bar{Z}$ in 1.2.1 follows at once, since for each $m \in M$, $v(t) = (d/dt)\exp(tX(m))$ satisfies $(d/dt)v(t) = Z(v(t))$, and $v(0) = X(m)$. Hence to establish 1.2.1 it suffices to establish our assertion concerning the geodesics on $\mathscr{D}^s$.

Of course a fundamental property of geodesics is that they locally minimize length. Suppose we have a family of geodesic curves $t \mapsto \eta(t)(m)$, starting at $m \in M$, where for $t_0 \in \mathbb{R}$, t_0 near 0, the map $m \mapsto \eta(t_0)(m)$ is a diffeomorphism so that $t \mapsto \eta_t$ is a curve in $\mathscr{D}^s$. Then since the length of a curve in $\mathscr{D}^s(M)$ given by our weak metric is the integral over M of the lengths of each curve, $t \mapsto \eta_t(m)$, this integrated length is also minimized. Hence it is reasonable that $t \mapsto \eta(t)$ should be a geodesic on $\mathscr{D}^s(M)$. The curves $t \mapsto \overline{\exp}(tX)(m)$ have all the above properties so $t \mapsto \overline{\exp}(tX)$ should be a geodesic on $\mathscr{D}^s(M)$. This concludes our justification of 1.2.1.

1.2.2 Corollary. *Z is a C^∞ vector field on $T\mathscr{D}^s$.*

This is a consequence of the omega lemma since Z is a C^∞ map.

Let us consider a simple example. Let $\mathbf{T}^2$ be the flat 2-torus. Then $T(\mathbf{T}^2) \simeq \mathbf{T}^2 \times \mathbb{R}^2$ is also a flat 4-manifold and $T(T\mathbf{T}^2) \simeq (\mathbf{T}^2 \times \mathbb{R}^2) \times (\mathbb{R}^2 \times \mathbb{R}^2)$. In this case the spray for the flat metric is given by

$$Z: T(\mathbf{T}^2) \to T(T(\mathbf{T}^2)): (x, v) \mapsto ((x, v), (v, 0)).$$

The x in the first coordinate is just the base point of the tangent vector in $T\mathbf{T}^2$. The v in the third coordinate is an important formal property

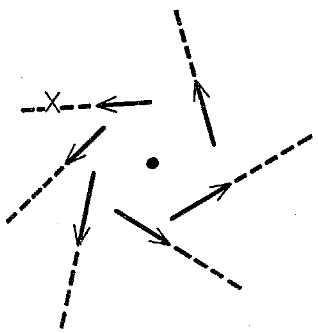


Figure 1.2

of sprays reflecting the fact that the geodesic equations are "second order" (see Lang [1]) and the 0 in the last coordinate reflects the fact that the metric is flat, hence each $\Gamma^i_{jk} = 0$. In this case the geodesics are of the form $\eta(t)(m) = m + tX(m)$ (where $X \in T_e\mathscr{D}^s(\mathbf{T}^2)$ and using the obvious identification). These are straight lines and hence $\mathscr{D}^s(\mathbf{T}^2)$ is essentially flat. In general, in coordinates $x = (x^1, \ldots, x^n)$ on a manifold M, we have $Z(x, v) = ((x, v), (v, -\Gamma^i_{jk}v^j v^k))$.

We now consider the metric for $\mathscr{D}^s_\mu(M) \subset \mathscr{D}^s(M)$. Even if $\mathscr{D}^s(M)$ is geometrically relatively simple, as above for $\mathbf{T}^2$, $\mathscr{D}^s_\mu(M)$ may be geometrically very complicated. Consider the above example. It should be clear that the diffeomorphism specified by having each point moving along straight lines is generally not volume preserving. See Figure 1.2. So requiring each point on a geodesic in $\mathscr{D}^s$ to be volume preserving must introduce some curvature.

Suppose S is a submanifold of a Riemannian manifold Q such that we

have an orthogonal projection of T_pQ onto T_pS for each $p \in S$. This gives us a bundle map $P: TQ \restriction S \to TS$ (where $TQ \restriction S = \{v_p \in T_pQ \mid p \in S\}$). This is of course the situation we have for $\mathscr{D}^s_\mu(M)$ as a submanifold of $\mathscr{D}^s(M)$ where the projection is given by the Hodge theorem (i.e., we project onto the divergent free part of X for $X \in T_e\mathscr{D}^s(M)$). In this situation, the following tells us how to put the spray on the submanifold.

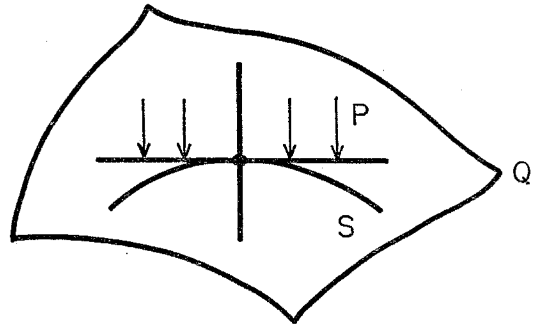


Figure 1.3

1.2.3 Lemma. *If Z is the spray on Q then $TP \circ Z$ is the spray on S.*

This is a standard result in Riemannian geometry. A proof using Hamiltonian theory may be found in Chernoff-Marsden [1].

Now Z is a vector field on TQ as is $TP \circ Z$ on TS. However their difference, say h, can be identified (technically by means of the vertical lift—see below) with a map of TS into $TQ \restriction S$, which turns out to be (the quadratic part of) the second fundamental form of S as a submanifold. Specifically for $v \in TS$, $h(v)$ is the normal component of $\nabla_v v$; cf. Chernoff-Marsden [1] for details. Thus this difference h in the sprays tells us, how curved S is in Q. (More exactly the curvatures on Q and on S are related through this second fundamental form by the Gauss-Codazzi equations; cf. Yano [1], p. 94.)

Define

$$P_e: T_e\mathscr{D}^s(M) \to T_e\mathscr{D}^s_\mu(M)$$

by carrying a vector field to its divergent free part. As we mentioned above, this is an L^2 orthogonal projection so it is orthogonal for the weak inner product on $T_e\mathscr{D}^s(M)$. We define for $X \in T_\eta\mathscr{D}^s(M)$; $\eta \in \mathscr{D}^s_\mu(M)$

$$P_\eta(X) = (P_e(X \circ \eta^{-1})) \circ \eta.$$

This makes P right invariant and is correct since the metric on $\mathscr{D}^s_\mu(M)$ is *right invariant* as we now show

1.2.4 PROPOSITION

(i) *Let* $\eta \in \mathscr{D}^s(M)$; *then* $(R_\eta)^* X_\zeta = X_\zeta \circ \eta$ (*where* $\zeta \in \mathscr{D}^s(M)$, $(R_\eta)^*: T_\zeta \mathscr{D}^s(M) \to T_{\zeta\circ\eta}\mathscr{D}^s(M)$).

(ii) *If* $\eta \in \mathscr{D}^s_\mu(M)$ *then* $((R_\eta)^* X, (R_\eta)^* Y)_{\zeta\circ\eta} = (X, Y)_\zeta$, *where* $X, Y \in T_\zeta \mathscr{D}^s(M)$.

PROOF. Part (i) has been used before and is easily seen. We will show the second part. Let $\eta \in \mathscr{D}^s_\mu(M)$; then:

$$((R_\eta)^* X, (R_\eta)^* Y)_{\zeta\circ\eta} = (X \circ \eta, Y \circ \eta)_{\zeta\circ\eta}$$

$$= \int_M \langle X \circ \eta(m), Y \circ \eta(m)\rangle_{\zeta\circ\eta(m)}\, d\mu$$

$$= \int_{\eta^{-1}(M)} \langle X(m), Y(m)\rangle_{\zeta(m)} (\eta^{-1})^*(d\mu).$$

But, since η^{-1} is volume preserving, $(\eta^{-1})^*(d\mu) = d\mu$ and $\eta^{-1}(M) = M$. Hence

$$((R_\eta)^* X, (R_\eta)^* Y) = \int_M \langle X(m), Y(m)\rangle_{\zeta(m)}\, d\mu$$

$$= (X, Y)_\zeta. \quad \blacksquare$$

Note that the metric on $\mathscr{D}^s$ is *not* right invariant.

Putting all this together we can write down the spray S on $\mathscr{D}^s_\mu(M)$ using 1.2.3. Namely, for $X \in T\mathscr{D}^s_\mu(M)$ we have $S(X) = TP(\bar{Z}(X)) = TP(Z \circ X)$. There is a major assumption in writing down this formula. When we write TP, we assume P is a C^∞ map. This is not at all obvious since if $X \in T_\eta \mathscr{D}^s_\mu(M)$, we compose X with η^{-1}, project, and then compose with η. As we have seen, composition of H^s maps is not generally smooth (see Part II, § 2). However, we have this surprising fact.

1.2.5 THEOREM. *P is a C^∞ bundle map. That is $P: T\mathscr{D}^s(M) \restriction \mathscr{D}^s_\mu(M) \to T\mathscr{D}^s_\mu(M)$ is C^∞. Hence the spray S on $\mathscr{D}^s_\mu$, $S(X) = TP(Z \circ X)$, is also a C^∞ vector field on $T\mathscr{D}^s_\mu$.*

For a proof see Ebin-Marsden [1]. There is an alternative and perhaps simpler proof to the one in the aforementioned paper. In this proof one defines another metric on $T\mathscr{D}^s(M)$; namely for $X, Y \in T_e\mathscr{D}^s(M)$, set

$$(X, Y)_s = (X, Y) + (\Delta^{s/2} X, \Delta^{s/2} Y)$$

where (,) is the L^2 metric on $T_e\mathscr{D}^s(M)$, and Δ is the Laplacian. Then extend $(,)_s$ to make it right invariant.

It turns out that this metric is smooth and by regularity properties of Δ is equivalent to the H^s metric. Smoothness facts like this again are not obvious but are proven in Ebin [1]. These facts are also useful for other purposes. The Hodge decomposition is then easily seen to be orthogonal in this *strong* metric $(,)_s$ and hence it follows automatically that the projection P is smooth. Below in § 1.4 we shall give further justification for why the map P is smooth.

This result 1.2.5 is important for we are going to apply the Picard theorem from ordinary differential equations to the equation:

$$\frac{dX_t}{dt} = S(X_t) = TP(Z \circ X_t)$$

and this requires that S is at least a Lipshitz map.

In case M has boundary, we do not get a spray on $\mathscr{D}^s$, but we do get one on $\mathscr{D}^s_\mu$. This is basically because P projects from vector fields sticking out of M, onto vector fields parallel to ∂M. We shall just accept as plausible that this extension can be made.

As mentioned earlier, it is unknown whether $\mathscr{D}^s_\mu(M)$ is geodesically complete. (By results below, this is the same thing as saying solutions to the Euler equations go for all time, and remain in H^s). Note that this is not equivalent to saying the induced distance metric is complete since the metric is only weak. In fact $\mathscr{D}^s_\mu(M)$ is not complete in this distance sense since the completion of $\mathscr{D}^s_\mu(M)$ under an L^2 topology is much larger than $\mathscr{D}^s_\mu(M)$. (Presumably it consists of a class of measure preserving maps from M to M.)

1.3 Derivation of the Euler equations

To show geodesics in $\mathscr{D}^s_\mu(M)$ satisfy the Euler equations, we need to know a bit more about T^2M. Let $\pi: TM \to M$ be the projection so that $T\pi: T^2M \to TM$. An element $w \in T^2M$ is called *vertical* if $T\pi(w) = 0$ (in coordinates this means the third component is 0). Now let $v, w \in T_mM$; define the *vertical left* of w with respect to v to be

$$(w)^l_v = \frac{d}{dt}(v + tw)\Big|_{t=0} \in T^2_vM = T_v(TM).$$

In coordinates this is simply

$$(w)^l_v = (m, v, 0, w).$$

The proof that geodesics in $\mathscr{D}^s_\mu$ yield solutions to the Euler equations essentially is calculations. The idea is to show that if a curve $X_t \in T\mathscr{D}^s_\mu$ satisfies the spray equation

$$\frac{dX_t}{dt} = S(X_t), \qquad X_t \in T\mathscr{D}^s_\mu(M).$$

Then X_t gives rise to a solution to the Euler equations in a sense explained below. For alternative proofs, see Arnold **[1]**, Marsden-Abraham **[1]**, or Chernoff-Marsden **[1]**; see also Hermann **[1]**.

1.3.1 LEMMA. *$\bar{Z}(X) = Z \circ X = TX \circ X - (\nabla_X X)^l_X$ for $X \in T_e\mathscr{D}^s$.*

PROOF. In coordinates

$$(\nabla_X X)^i = \sum_j X^j \frac{\partial X^i}{\partial x^j} + \sum_{j,k} \Gamma^i_{jk} X^j X^k.$$

Now

$$(TX \circ X)^i = \sum_j X^j \frac{\partial X^i}{\partial x^j} \quad \text{so} \quad (TX \circ X - (\nabla_X X))^i = -\sum_{jk} \Gamma^i_{jk} X^j X^k.$$

This then puts the right expressions in the fourth component. ■

Note that both $TX \circ X$ and $(\nabla_X X)^l_X$ are elements of $T_X\mathscr{D}^s_\mu$. The latter is by construction of the vertical lift. To see this for $TX \circ X$, let $\pi_1 : T^2M \to TM$ be the projection; then since $\pi_1 \circ TX = X \circ \pi$ we have

$$\pi_1 \circ TX \circ X = X \circ \pi \circ X = X$$

since $\pi \circ X$ is the identity.

As we observed in 1.2.2, the map $X \mapsto Z \circ X$ (for $X \in T\mathscr{D}^s(M)$) is C^∞. Hence even though $TX \circ X$ and $\nabla_X X$ are only H^{s-1}, their difference must be H^s.

1.3.2 LEMMA. *Let σ and X be in $T_\eta\mathscr{D}^s_\mu(M)$ then $TP[(\sigma)^l_X] = (P(\sigma))^l_X$.*

PROOF. Since P is linear on each fiber and $P(X) = X$, we get

$$\begin{aligned}(P(\sigma))^l_{P(X)} &= \frac{d}{dt}(P(X) + tP(\sigma))\Big|_{t=0}\\ &= \frac{d}{dt}P(X + t\sigma)\Big|_{t=0}\\ &= TP\left(\frac{d}{dt}(X + t\sigma)\right)\Big|_{t=0} \quad \text{(chain rule)}\\ &= TP[(\sigma^l_X]. \quad ■\end{aligned}$$

1.3.3 LEMMA. *Let $\eta \in \mathscr{D}^s_\mu$ and $X \in T_\eta\mathscr{D}^s(M)$; then $TP(T(X \circ \eta^{-1}) \circ X) = \{T(P_e[X \circ \eta^{-1}])\} \circ X$.*

PROOF. $X \circ \eta^{-1}$ is an H^s vector field on M. Let F_t be its flow (or any curve tangent to it). Let $G_t = (X \circ \eta^{-1}) \circ F_t$. Then $G_0 = X \circ \eta^{-1}$ and $(dG_t/dt) = T(X \circ \eta^{-1}) \circ (X \circ \eta^{-1})$. Thus we get

$$\begin{aligned}TP(T(X \circ \eta^{-1}) \circ (X \circ \eta^{-1})) &= \frac{d}{dt}P(G_t)\Big|_{t=0} \quad \text{(chain rule)}\\ &= \frac{d}{dt}(P_e(X \circ \eta^{-1}) \circ F_t)\Big|_{t=0}\\ &= T(P_e(X \circ \eta^{-1})) \circ (X \circ \eta^{-1}).\end{aligned}$$

But by right invariance $TP(T(X \circ \eta^{-1}) \circ (X \circ \eta^{-1})) = TP(T(X \circ \eta^{-1}) \circ X) \circ \eta^{-1}$. ■

1.3.4 PROPOSITION. *The spray on $T\mathscr{D}^s_\mu$ is given by*

$$S(X) = T(X \circ \eta^{-1}) \circ X - (P_e\nabla_{X\circ\eta^{-1}}X \circ \eta^{-1})^l_X \circ \eta \quad \textit{where} \quad X \in T_\eta\mathscr{D}^s_\mu(M).$$

PROOF. This follows directly from the above lemmas. ■

So now that we have an explicit formula for the spray, let us inspect the Euler equations. Recall from the introductory § 2 that these describe the time evolution of the velocity vector field on M. The equations are written out in Eulerian coordinates and are equations involving elements of $T_e\mathscr{D}^s(M)$. The spray on the other hand is a map on all of $T\mathscr{D}^s(M)$. The integral curves of the spray are the velocities written in Lagrangian coordinates. So if $X_t \in T_{\eta(t)}\mathscr{D}^s(M)$ is an integral curve of the spray, one would hope the pullback of X_t, i.e., $X_t \circ \eta_t^{-1} \in T_e\mathscr{D}^s(M)$, would be a solution of the Euler equations.

The vector field $v(t) = X_t \circ \eta_t^{-1}$ is more carefully justified as follows. We want η_t to be the flow of v, so this means that

$$\frac{d}{dt}\eta_t(m) = v_t(\eta_t(m)).$$

Since we are dealing with geodesics and hence $(d\eta/dt) = X$, we get the desired relation $v_t = X_t \circ \eta_t^{-1}$.

It turns out, as we shall see momentarily, that the derivative loss of the Euler equations occurs in this pullback operation (or "coordinate change").

We are interested in computing (dv/dt), and so we need this lemma.

1.3.5 LEMMA. *We have:*

$$\frac{dv(t)}{dt} = \frac{d}{dt}(X_t \circ \eta_t^{-1}) = \frac{dX_t}{dt} \circ \eta_t^{-1} - TX_t \circ T\eta_t^{-1} \circ X_t \circ \eta_t^{-1}.$$

PROOF. This follows by differentiating both places t occurs, using the chain rule and the formula

$$\frac{d}{dt}(\eta_t^{-1}) = -T\eta_t^{-1} \circ \frac{d\eta}{dt} \circ \eta_t^{-1}.$$

The last formula follows from the chain rule applied to $\eta_t \circ \eta_t^{-1} = id.$ ■

So, putting this together, we get:

$$\frac{dv_t}{dt} = S(X_t) \circ \eta_t^{-1} - T(X_t \circ \eta_t^{-1}) \circ X_t \circ \eta_t^{-1}.$$

Now using 1.3.4, this becomes $= -(P_e \nabla_v v_t)_v^l$. Note especially the cancellation of the $Tv \circ v$ terms which has occurred. But as we recall $P_e(\nabla_{v_t} v_t) = \nabla_{v_t} v_t - \nabla p_t$ where p is a smooth function. We can identify $P_e(\nabla_{v_t} v_t)_v^l$ with $P_e(\nabla_{v_t} v_t)$ (since dv/dt really stands for its vertical lift) and hence get the Euler equations

$$\frac{dv_t}{dt} = -\nabla_{v_t} v_t + \nabla p_t$$

or

$$\frac{dv_t}{dt} + \nabla_{v_t} v_t = \nabla p_t.$$

(The minus sign on the pressure can be recovered by using $-p_t$.) Thus we have proved:

1.3.6 THEOREM. *If X_t is an integral curve of the spray on $\mathscr{D}_\mu^s$, its pullback $v_t = X_t \circ \eta_t^{-1}$ does satisfy the Euler equations. In other words, η_t is a geodesic on $\mathscr{D}_\mu^s$ iff its velocity field satisfies the Euler equations.*

By inspecting the above calculation it becomes clear where the derivative loss occurs. If X is an H^s vector field on M, we know $S(X)$ is an H^s vector field on TM. However it is the sum of two H^{s-1} vector fields on TM. The top derivatives cancel, but when this is pulled back to Eulerian coordinates one of these terms disappears, namely $TX_t \circ X$ and so what we are left with is one of the H^{s-1} summands.

All of the above goes through for manifolds with boundary since the

Hodge theorem projects vector fields at the boundary onto those which are tangent to the boundary as mentioned before.

As a consequence of these calculations we have this theorem.

1.3.7 THEOREM. *Given $v_0 \in T_e\mathscr{D}_\mu^s$ there is an $\varepsilon > 0$ and a unique vector field $v(t) \in T_e\mathscr{D}^s$ for $-\varepsilon < t < \varepsilon$ which satisfies the Euler equation. Moreover, these solutions v_t depend continuously on the initial data v_0.*

PROOF. From 1.3.6 we see that the theorem is equivalent to finding short-time solutions to the geodesic spray on $\mathscr{D}_\mu^s(M)$. But since $\mathscr{D}_\mu^s(M)$ is a Hilbert manifold and the spray is smooth (1.2.4) the existence follows immediately from standard results on ordinary differential equations (see Lang [1]).

The continuous dependence on initial conditions follows from the fact that the pullback $v_t = X_t \circ \eta_t^{-1}$ involves left composition so it is continuous but not smooth. The initial condition for the spray on $\mathscr{D}_\mu^s(M)$ is an element of $T_e\mathscr{D}_\mu^s(M)$ since we are interested in flows in $\mathscr{D}_\mu^s(M)$ starting at the identity. ■

This existence theorem has been proved in weaker forms by Lichtenstein **[1]** and Guynter **[1]**. The general case of manifolds with boundary is due to Ebin-Marsden **[1]**.

The flow in Lagrangian coordinates is C^∞. In Euler coordinates, let $E_t(v_0) = v_t$ be the solution flow. Then for *fixed t*, E_t is a smooth map, although this is not obvious from what we have proven so far. However it is not difficult to prove: TE_t is just the pull back of the tangent to the flow of S, which we know is smooth. Such smoothness properties are sometimes important in applications; cf. Bardos-Tartar **[1]** and Foias-Prodi **[1]**. For general properties of flows smooth for fixed t, see Chernoff-Marsden **[2]**. See also Marsden **[3, 4]**.

1.4 Discussion of the existence theorem

The proof of Theorem 1.3.7 is based on the existence of integral curves for the spray S. This in turn follows from the fundamental existence theorem for ordinary differential equations. Recall that this theorem is proven by showing an iteration (called Picard iteration) always yields solutions. So, by inspecting the above proof it should be possible to find an approximation procedure which converges to solutions.

This in fact, points out an essential difference between working with the whole spray and working with its pullback $P_e(\nabla_v v)$. The Picard method will not in general converge for the pullback even though it works for the spray.

As an aside, this particular trick of using a spray on $\mathscr{D}^s_\mu$ finding solutions and then pulling back will only work when the nonlinear term is of the form $\nabla_v v$. For other systems one would have to modify the procedure.

One of the crucial steps in the proof was the smoothness of the projection $P: T\mathscr{D}^s(M) \restriction \mathscr{D}^s_\mu \to T\mathscr{D}^s_\mu(M)$, which entailed smoothness of the spray S on $\mathscr{D}^s_\mu$. Let us try to see why this is. Recall that $S(X) = TX \circ X - P_e(\nabla_X X)$ is the spray on $\mathscr{D}^s_\mu$ while $Z(X) = TX \circ X - \nabla_X X$ is the spray on $\mathscr{D}^s$.

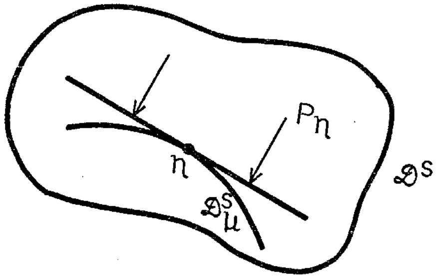


FIGURE 1.4

Since Z is smooth, then (smoothness of S implies that) $S - Z$ is smooth. Now this difference is

$$\nabla_X X - P_e(\nabla_X X) = \operatorname{grad} p.$$

So if X is an H^s vector field, this implies in particular that grad $p =$ gradient part of $\nabla_X X$ is H^s; this of course is not obvious since $\nabla_X X$ is only H^{s-1}.

1.4.1 PROPOSITION. *Let $X \in T_e\mathscr{D}^s_\mu$, so that X is H^s, $\delta X = 0$. Now let $\nabla_X X = Y +$ grad p be the Hodge decomposition of $\nabla_X X$. Then* grad p *is H^s and not merely H^{s-1}.*

We will show this over the flat n torus $\mathbf{T}^n$ (the general case is similar). Note $T(\mathbf{T}^n) = \mathbf{T}^n \times \mathbb{R}^n$ so we have a global orthonormal frame field on $\mathbf{T}^n$. Then we have the coordinate expression

$$(\nabla_X X)^i = (X \cdot \nabla)X^i = \sum_{j=1}^n X^j \frac{\partial X^i}{\partial x^j}$$

and so

$$\begin{aligned}\Delta p = \delta\, dp &= \delta \nabla_X X \\ &= \sum_{i,j} \frac{\partial}{\partial x^i}\left(X^j \frac{\partial X^i}{\partial x^j}\right) \\ &= \sum_{i,j} \frac{\partial X^j}{\partial x^i}\frac{\partial X^i}{\partial x^j} + X^j \frac{\partial^2 X^i}{\partial x^i \partial x^j}.\end{aligned}$$

But $\sum_i \partial^2 X^i/\partial x^i\, \partial x^j = 0$ since $\nabla \cdot X = 0$ by assumption. Thus $\Delta p = \sum (\partial X^i/\partial x^j) \cdot (\partial X^j/\partial x^i)$ which is an H^{s-1} function on $\mathbf{T}^n$ since it only involves only first derivatives of X. This implies, by the regularity of the Laplacian, that p is an H^{s+1} function on $\mathbf{T}^n$. Hence grad p is an H^s vector field. The crucial point, of course, was that the top order term in Δp dropped out. ■

This shows the map $X \mapsto$ grad p maps $T_e\mathscr{D}^s_\mu(M)$ to itself. This does not prove the map is smooth, but it does make it reasonable since the range space does not have a weaker topology than the domain space. This proof can also be made to work for manifolds with boundary.

The grad p_t term has geometric significance. It is, being the difference of the sprays, essentially the second fundamental form of $\mathscr{D}^s_\mu$ as a submanifold of $\mathscr{D}^s$. It therefore tells us (via the Codazzi equation) how curved a submanifold $\mathscr{D}^s_\mu$ is.

Suppose $Q \subset R$ is a submanifold. A geodesic in Q may not be one in R since the sprays are different. Intuitively it takes extra force to constrain a geodesic to Q. This force is clearly given by the difference of the sprays (since force is acceleration). So the grad p term measures the "force" the fluid experiences by being constrained to $\mathscr{D}^s_\mu$. If there were no pressure term, the particles of the fluid would follow geodesics on M and the resulting motion would not necessarily be volume preserving. Thus grad p may be viewed as a force of constraint necessary to keep things volume preserving in the same way that the centrifugal force is the constraint force ($\simeq$ second fundamental form) for circular motion in the plane.

1.5 Kelvin Circulation Theorem

This is a standard classical theorem of hydrodynamics that is very easy to prove in our context. It says the amount of circulation about any closed loop is constant in time.

1.5.1 KELVIN CIRCULATION THEOREM. *Let M be a manifold and $l \subset M$ a smooth closed loop i.e., a compact one manifold. Let u_t be a solution to the Euler Equations on M and $l(t)$ be the image of l at time t when each* *particle moves under the flow η_t of u_t i.e., $l(t) = \eta_t(l)$. Then*

$$\frac{d}{dt}\int_{l(t)} \tilde{u}_t = 0 \quad (\tilde{u}_t \text{ is the one form dual to } u_t).$$

PROOF. We have the identity $L_u\tilde{u} = \nabla_u u + \frac{1}{2}\, d\langle u, u\rangle$ valid for any vector field u on the manifold M. We leave the verification as an exercise.

Then, identifying the differential forms with their dual vector fields, we find $P_e(L_u\tilde{u}) = P_e(\nabla_u u)$ since P_e annihilates exact forms.

So, substituting into the Euler equations, we get the following alternative form:

$$\frac{\partial\tilde{u}}{\partial t} + P_e(L_u\tilde{u}) = 0.$$

Let η_t be the flow of u_t. Then $l(t) = \eta_t(l)$, and so changing variables,

$$\int_{\eta_t(l)} \tilde{u}_t = \int_l \eta_t^*(\tilde{u}_t)$$

which becomes, on carrying out the differentiation,

$$\frac{d}{dt}\int_{\eta_t(l)} \tilde{u}_t = \int_l \eta_t^*(L_u\tilde{u}) + \eta_t^*\frac{\partial\tilde{u}}{\partial t}.$$

Let $P_e(L_u\tilde{u}) = L_u\tilde{u} - \operatorname{grad} q$. By Stokes theorem $\int_l \operatorname{grad} q = 0$, so

$$\frac{d}{dt}\int_{l(t)} \tilde{u}_t = \int_l \eta_t^*\left(L_u\tilde{u} + \frac{\partial\tilde{u}}{\partial t} - \operatorname{grad} q\right) = 0. \quad ■$$

In practical fluid mechanics, this is an important theorem. One can obtain a lot of qualitative information about specific flows by following a closed loop throughout time and using the fact the circulation is constant. In a similar way, one can show *directly* that the energy $\int\langle u, u\rangle\, d\mu$ is constant in time.

1.6 Steady Flows

A flow is steady if its vector field satisfies $(\partial u/\partial t) = 0$; i.e., u is constant in time. This condition means that the "shape" of the fluid flow is not changing. Even if each particle is moving under the flow, the global configuration of the fluid does not change.

Not much is really known about steady flows, their stability, or what initial conditions result in steady flows. We should mention, however, that for viscous flow quite a bit more is known. See for example Ladyzhenskaya [2] and Finn [1]. There are some elementary equivalent formulations of the Euler problem.

1.6.1 PROPOSITION. *Let u_t be a solution to the Euler equations on a manifold M and η_t its flow. Then the following are equivalent:*

(1) $u_0 \in T_e\mathscr{D}^s_\mu$ *yields a steady flow* (i.e., $(\partial u/\partial t) = 0$)
(2) η_t *is a one parameter subgroup of* $\mathscr{D}^s_\mu(M)$
(3) $L_{u_0}\tilde{u}_0$ *is an exact form*
(4) $i_{u_0}\, du_0$ *is an exact form.*

The proof should be clear; cf. the proof of 1.5.1.

It follows at once from (4) that if $u_0 \in T_e\mathscr{D}^s_\mu(M)$ is a *harmonic* vector field; i.e., u_0 satisfies $\delta u_0 = 0$ and $d\tilde{u} = 0$ then it yields a stationary flow. Also it is known there are other steady flows for manifolds with boundary. For example, on a closed 2-disc, with polar coordinates (r, θ), $v = f(r)(\partial/\partial\theta)$ is the velocity field of a steady flow because

$$\nabla_v v = -\nabla p, \qquad \text{where } p(r, \theta) = \int_0^r f^2(s)s\, ds.$$

Clearly such a v need not be harmonic.

It would also be interesting to see if $\mathscr{D}^s_\mu$ has any closed geodesics. Perhaps the methods in Klingenberg's lectures could be modified. Actually of more physical interest would be periodic solutions to the Euler equations: $u_{t+\tau} = u_t$. In terms of the geodesic η_t this means not that η_t is periodic, but only $\eta_{t+\tau} = \eta_t \circ \eta_\tau$ for all $t \in \mathbb{R}$.

1.7 Compressible flow

From the previous discussions of geodesics on $\mathscr{D}^s(M)$ and $\mathscr{D}^s_\mu(M)$ it follows that the pull back of the geodesics on $\mathscr{D}^s(M)$ yields the equation

$$\frac{\partial u}{\partial t} + \nabla_u u = 0.$$

Here we assume $\partial M = \varnothing$. Also we have from our explicit formula for geodesics on $\mathscr{D}^s$, an explicit solution to this equation:

$$u(t, x) = \left(\frac{d}{dt}\exp_x tu_0(x)\right) \circ [\exp_x tu_0(x)]^{-1}$$

where the $[\exp_x(tu_0(x))]^{-1}$ term is the pullback to $T_e\mathscr{D}^s(M)$ and the inverse is with respect to the map in x. This of course is undefined where the map

$F_t: x \mapsto \exp_x(tu_0(x))$ fails to be a diffeomorphism. In general there will be a t where this will happen (as we discussed earlier) and so compressible flow will not in general be a solution to the geodesic pullback equation for all time.

Where F_t fails to be a diffeomorphism we have a "shock or rarefaction wave." Classically a shock wave is described by a large pressure gradient (where ∇p is the irrotational part of $\nabla_u u$). F_t can fail to be a diffeomorphism by mapping several points to the same point (a shock) or by failing to be onto (a rarefaction). In either case a large pressure gradient would develop. When a shock or rarefaction develops one has to prescribe other rules for its subsequent development, such as entropy conditions.

Shock waves probably do not occur in $\mathscr{D}^s_\mu(M)$. This is because the volume preserving condition tends to prevent the above phenomena which make the diffeomorphisms go bad. Also note that in $H^s(M, M)$, we will not notice any singularities because even when the F_t fail to be diffeomorphisms, it still is an H^s map (physically it is unreasonable though, because other effects come in when a shock wave develops). For more information on compressible flow, see Friedrichs [2]. For recent research and further references, consult for instance *B.* Quinn [1].

1.8 Other theorems using manifolds of maps

Let M be a manifold and let X and Y be time independent vector fields. Now let H_t be the flow of $X + Y$, F_t the flow of X and G_t the flow of Y. Then we have the formula

$$H_t = \lim_{n\to\infty}(F_{t/n} \circ G_{t/n})^n$$

where the power n means we iterate the map n times. Note this implies that if $F_t \circ G_t = G_t \circ F_t$, i.e., $[X, Y] = 0$, then $H_t = F_t \circ G_t$. See for instance Nelson [1] or Chernoff-Marsden [1] for a proof of this result. The result is usually referred to as the Lie-Trotter product formula. More generally we have the formula of Chernoff which states that if $K(t): M \to M$ is a family of maps, and $X = K'(0)$, the flow of X is

$$H_t = \lim_{n\to\infty} K(t/n)^n$$

One recovers the above case by taking $K(t) = F_t \circ G_t$. See Chernoff [1], Ebin-Marsden [1, 2], Marsden [2] and Brezis-Pazy [1] for more details.

We will use this type of theorem to get an expression for the Euler flow. This makes sense since the Euler equations are derived from a vector field on $T\mathscr{D}^s_\mu(M)$, which is a modification the geodesic spray on $\mathscr{D}^s$. In particular, let F_t be the flow of the geodesic spray on $T_e\mathscr{D}^s(M)$, i.e., the solution to the equation $(\partial u_t/\partial t) + \nabla_{u_t} u_t = 0$, given *explicitly above*, and let E_t be the Euler flow on $T_e\mathscr{D}^s_\mu(M)$, P_e is as usual the Hodge projection on $T_e\mathscr{D}^s(M)$ to $T_e\mathscr{D}^s_\mu(M)$.

1.8.1 Theorem. *We have the formula*

$$E_t(u_0) = \lim_{n\to\infty}(P_e \circ F_{t/n})^n(u_0)$$

as long as $E_t(u_0)$ exists.

As we noted above, we have an explicit formula for F_t, and P_e is a bounded linear operator which can be written down using Green's operators, as in Part I, § 1, so this formula gives us real information about the Euler flow.

This theorem is hard to prove directly but is fairly easy if we transfer to $\mathscr{D}^s_\mu(M)$, since the sprays are C^∞ there. One need first consider the spray $TP \circ Z$ on $T\mathscr{D}^s_\mu(M)$ and pullback to $T_e\mathscr{D}^s_\mu(M)$. Here $K(t) = P \circ F_t$.

Product formulas like the above have several technical applications. For example, they can be used to study viscous flow and the limit of zero viscosity. See Ebin-Marsden [1], [2] and § 3 below. These formulas have been found to be efficient in actual numerical computations; see Chorin [2]. See also Temam [1] for some related ideas.

Smoothness of the spray on $\mathscr{D}^s_\mu$ implies that the geodesic exponential map on $\mathscr{D}^s_\mu$ is C^∞; $\exp_e: T_e\mathscr{D}^s_\mu(M) \to \mathscr{D}^s_\mu(M)$. Now $D\exp_e(0)$ maps $T_e\mathscr{D}^s_\mu(M)$ to $T_e\mathscr{D}^s_\mu(M)$ (this is identifying $T_0T_e\mathscr{D}^s_\mu(M) \approx T_e\mathscr{D}^s_\mu(M)$ which is correct since $T_e\mathscr{D}^s_\mu(M)$ is a linear space). In fact $D\exp_e(0)$ is the identity and so is an isomorphism (compare Lang [1]). Hence we can apply the inverse function theorem to $\exp_e$ and conclude that it is a diffeomorphism from a neighborhood of 0 onto a neighborhood of e in $\mathscr{D}^s_\mu$. In other words, *there is a neighborhood U of e in $\mathscr{D}^s_\mu(M)$ and $\varepsilon > 0$ such that for $\eta \in U$, there is a unique vector field u_0 on M with $\delta\tilde{u}_0 = 0$, u_0 tangent to ∂M and $\|u_0\|_{H^s} < \varepsilon$ such that if η_t is the Euler flow of u_0 then $\eta = \eta_1$.*

Intuitively, this says one can go from one state of a fluid to another nearby state by going through a uniquely determined Euler flow. Some ideas related to this and regularity discussed in § 2 below were abstracted by Omori [2].

2. The global (all time) existence problem and Leray's criterion

As discussed above, the all time existence problem for the Euler equations is equivalent to the problem of whether or not $\mathscr{D}^s_\mu(M)$ is geodesically complete. This problem is trivial if the dimension of M equals one since divergence free vector fields are then constant. If $\dim(M) = 2$, the problem has been solved and is nontrivial. In 1933, Wolibner [**1**] proved that the Euler equations have global solutions and hence that $\mathscr{D}^s_\mu(M)$ is geodesically complete when $\dim(M) = 2$. Judovich [**2**] made this earlier result more precise. In 1967, Kato [**1**] found a shorter, more elegant proof of this fact using the Schauder fixed point theorem. However, all of the later proofs are based on a key estimate already found in Wolibner's 1933 paper. The viscous case was treated by Leray in the early 1930's and will be discussed below. As stated above, it is unknown whether $\mathscr{D}^s_\mu(M)$ is geodescially complete when M is a 3-manifold. The basic difference between the 2 and 3 dimensional cases will be discussed below. We shall now consider all these problems in a bit more detail.

2.1 Sufficient conditions for extendability of Euler solutions

Recall from the standard extendability of solutions theorem in ordinary differential equations that the length of time that solutions can be extended varies inversely with the norm of the derivative of the vector field and if the derivative is bounded globally then the solutions will go for all time; see Lang [**1**]. This is the same sort of argument used for extending solutions of the Euler equations for all time. This is not surprising since the geodesic equations on $\mathscr{D}^s_\mu(M)$ are ordinary differential equations.

We first want to recall the following.

2.1.1 LEMMA. *Let G be a (finite dimensional) Lie group with $\langle\ ,\ \rangle$ a right invariant Riemannian metric. Then G is geodesically complete.*

PROOF. It is enough to consider geodesics starting at the identity e, by right invariance. Now from the local existence theorem given $\delta > 0$, there is $\epsilon > 0$ such that if $v_0 \in T_eG$ and $\|v_0\| < \delta$, then the geodesic starting in direction v_0 exists for a time interval $[-\varepsilon, \varepsilon]$. (cf. Lang [**1**], Ch. 4, § 3). Now let $v(t)$ be the tangent to this geodesic. By conservation of energy, $\|v(t)\| = \|v_0\|$. Let $v(t) \in T_{g(t)}G$. By right invariance, if we look at the geodesic starting in direction $TR_{g(\varepsilon)^{-1}}v(\varepsilon) \in T_eG$ where $R_{g^{-1}}$ is right translation by g^{-1}, and translate it back out to $g(\varepsilon)$, we get the geodesic starting in direction $v(\varepsilon)$. But by our above remarks, this geodesic exists for a time ε as well. Continuing the process, we see that the geodesic can be infinitely extended. ■

This does not work for the Euler equations because the Euler equations have an existence time dependent on the H^s structure, whereas the metric is an L_2 metric. Nevertheless, the argument can be used as the basis for the following.

2.1.2 THEOREM. *Let v_t be an H^s Euler vector field on a manifold M defined for $t \in [0, T)$ (since the flow is reversible we could use $(-T, T)$ here). Then if there is a $K \in \mathbb{R}$ such that the C^1 norm of v_t is bounded by K: $\|v_t\|_{C^1} \leqq K$ for all $t \in [0, T)$ then v_t can be smoothly extended past T in H^s.*

PROOF. For simplicity we will use $\|v_t\|_{H^s} \leqq K$ $(s > n/2 + 1)$. This is a slightly weaker result. (The general case will be discussed later.) Let $v_0 \in T_e\mathscr{D}^s_\mu(M)$, and let X_t be the geodesic velocity on $T\mathscr{D}^s_\mu(M)$, and η_t be its base point. As in 2.1.1, given K there is an $\epsilon > 0$ such that η_t exists for time at least ϵ provided $\|v_0\|_{H^s} \leqq K$. Now use the same right translation argument as above to conclude that $v(t)$ can be extended beyond time T. ■

It should be noted that if we had the same theorem using the L^2 norm, we could immediately conclude that $\mathscr{D}^s_\mu(M)$ is geodesically complete. This is because conservation of energy tells us the L^2 norm is constant and hence uniformly bounded for all time. This fails because $\mathscr{D}^s_\mu$ is known to be a nice manifold only if s is large.

As an obvious corollary to the above theorem we know that if, for an Euler vector field v_t, $\|v_t\|_{C^1}$ is uniformly bounded for all time then it gives an Euler flow for all time (or equivalently the geodesic flow η_t, based on v_t goes for all time). Another corollary of the C^1 theorem is the following *regularity theorem.*

2.1.2 COROLLARY. *If $v \in T_e\mathscr{D}^s_\mu(M)$ is actually a C^∞ vector field and v_t is an H^s Euler vector field, $s > (n/2) + 1$, such that $v_0 = v$, then v_t is a C^∞ vector field.*

PROOF. One simply applies the theorem to any $s' > (n/2) + 1$ to show v_t is an $H^{s'}$ vector field for each s', since the C^1 norm is bounded on any finite t-interval of existence. Also note that the time of existence of v_t is independent of s' and in particular does not approach 0 as s' approaches infinity. ■

Note. A different proof of this appears in Ebin-Marsden [**1**], § 12.

If M is a manifold with boundary, one must be careful about what happens at the boundary, but the result remains valid in that case (again see below).

Because of special properties of 2-manifolds, we can essentially find C^1 bounds on the vector field and thereby prove that solutions do go for all time in this case. The basic reasoning behind this is as follows. Let v_t be an Euler vector field and η_t its flow. Now $d\tilde{v}_t = \omega(t)$ corresponds to the curl of v_t (which is the *vorticity* of the flow). From $(\partial\tilde{v}/\partial t) + L_v\tilde{v} = -dp$, we get $(\partial\omega/\partial t) + L_v\omega = 0$ by applying the exterior derivative d. It follows from this that $\omega(t) = (\eta_t^{-1})^*\omega(0)$. Now since η_t is volume preserving and $\omega(t)$ is a (time dependent) 2 form identifiable with a function, we get $\omega(t) = (\eta_t^{-1})^*(\omega(0)) = \omega(0) \circ \eta_t^{-1}$. This last calculation shows that the C^0 norm and all L^p norms of $\omega(t)$ are constant in time. In particular, since $\omega = d\tilde{v}$, this allows us to make global estimates on the *derivatives* of v_t which is what is needed for the all-time problem. This is just the basic idea, although the details are more involved. Indeed one can really only directly estimate the C^α norm for $\alpha < 1$, the case $\alpha = 1$ being delicate. The conservation of vorticity is simply the two dimensional case of the Kelvin Circulation Theorem, since vorticity at a point p in a two manifold can be thought of as the mean circulation about a circular loop centered at p.

It seems unlikely that this sort of approach, (i.e., finding bounds on the derivatives) will work for dimension $\geqq 3$. In fact there are examples of 3-dimensional flow for which the vorticity grows without bound. A good deal of effort has gone into this without significant results.

2.2 Viscous flow

Viscosity is internal friction of the fluid and friction between the fluid at the boundary. When one stirs up some fluid, it is its viscosity which makes it eventually slow down. Since friction absorbs energy, it is often easier to prove existence of viscous flows, since this allows us to more easily find bounds on the H^s norms of the vector fields. Intuitively, since viscosity makes fluid effects dissipate, viscous flows are, in a sense, bounded by Euler flows. Viscous flow is often described by the Navier–Stokes equations:

$$\text{N.S.}\begin{cases} \dfrac{\partial v_t}{\partial t} - \nu\nabla^2 v_t + (v_t \cdot \nabla)v_t = -\text{grad } p_t \\ \text{div } v_t = 0 \\ v_t = 0 \text{ on boundary of } M \end{cases}$$

Here $\nu > 0$ and is called the viscosity of the fluid and is usually assumed

to be constant. Because of the Laplacian term, one only expects solutions for $t \geqq 0$. Also since the Euler part conserves the L_2 norm, it is easy to check that for a solution of the Navier-Stokes equations the L_2 norm is decreasing.

These equations are usually defined for a region $\Omega \subset \mathbb{R}^n$. If the metric is not flat, we need a modified Laplacian which involves the curvature of the metric. Note also that the boundary conditions are different than those of Euler flow. Here we specify that the fluid does not move at the boundary (this is the "no-slip" condition). This is what actually happens in practice. Consider a situation where a viscous fluid (in fact, water will work) is flowing down a pipe and suppose when the water is not moving we introduce a dye orthogonal to the flow as shown in Figure 2.1

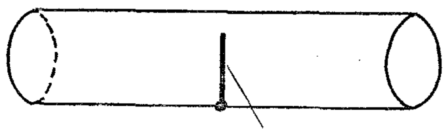


FIGURE 2.1

When the fluid is moving, one observes the situation in Figure 2.2.

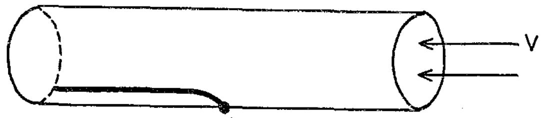


FIGURE 2.2

The Navier-Stokes equation $[(\partial v_t/\partial t) - \nu\nabla^2 v_t + (v_t \cdot \nabla)v_t = -\nabla p_t)]$ is derived from the assumption that the internal stress exerted on the fluid depends linearly on the deformation and is independent of position and direction (see Serrin [**1**]). If one only assumes that the stress depends continuously on the deformation (rather than linearly) then one can show this dependence is at worst quadratic. Engineers, when working with the Navier-Stokes equations on a specific problem pick a distance scale d (which may be the diameter of a pipe, say) a velocity scale V like the max of the initial velocity field and define $\mathscr{R} = (Vd/\nu)$; this is the *Reynolds number* and is dimensionless. Then they use $(\partial v_t/\partial t) - (1/\mathscr{R})\nabla^2 v_t + (v_t \cdot \nabla)v_t = -\text{grad } p_t$. Note that if we rescale v_t by V, the coordinates by

d and the time by d/V we get another solution with the new value of $\mathscr{R}$. All of this is important in order to understand much of the following.

As stated earlier, the Navier-Stokes equations are sometimes easier to work with than the Euler equations. This is because the $-\nabla^2 v_t$ term tends to slow the fluid down and thereby makes it less likely it will run into a singularity. In fact in order to keep a viscous fluid moving, one has to keep pumping in energy (in Couette flow, for example, we move the boundaries). It is surprising, but true, that steady pumping of energy does not really affect the difficulty in dealing with the possible development of singularities of the flow in the mathematical theory.

It is known that the Navier-Stokes equations have short-time solutions, but it is unknown whether these solutions extend for all time over 3-manifolds for all choices of initial data. There are three basic results on the Navier-Stokes equations in three dimensions.

(a) *There are weak (as in the sense of distributions) solutions which go for all time* $t \geqq 0$. These solutions are only L^2 flows and they are known not to be unique! (Ladyzhenskaya [3]) So this theorem fails to specify the flow of the fluid. This theorem is due to E. Hopf [3] in 1950. A proof may be found in Ladyzhenskaya [2].

(b) *If*

$$\|v_0\|_{L^2} \left\| \frac{\partial v_t}{\partial t}\bigg|_{t=0} \right\|_{L_2} \leqq C\nu^3$$

(*where C is a constant depending on the region* Ω) *then these initial conditions yield smooth solutions for all* $t \geqq 0$. This theorem is due to Kiselev-Ladyzhenshaya [1].

Intuitively, it says if the fluid is not flowing very fast initially then the flow will not have a singularity. (It is probably too slow for the interesting case of turbulence.) This is not surprising considering our earlier observations about the Navier-Stokes equations. It might be mentioned that this theorem only works for dimension $\leqq 4$; cf. Serrin [3].

The above two results may lead one to speculate that if $\|v_0\|_{L^2}$ is too large, the Navier-Stokes equations quickly break down and this is the source of the nonuniqueness in the weak solutions, as well as turbulence. Actually this seems unlikely, but is not fully settled yet. This is discussed more fully under the section on turbulence.

(c) *For fixed initial data, the time of existence is* $\geqq C\nu$ *where C is a constant.* (Ladyzhenskaya [2].)

Now this seems to imply that as $\nu \to 0$, the equations do not have solutions. But, at least for manifolds without boundary, when $\nu = 0$, the equations reduce to the Euler equations, which we know have solutions.

Later however, we shall show that if $\partial M = \varnothing$, strong, i.e., sufficiently

smooth, solutions of the Navier-Stokes equations have an existence time independent of ν and do converge to Euler flows as ν approaches 0.

The two dimensional situation is more satisfactory. There one does have global existence of smooth solutions. This is due to Leray [2] and Ladyzhenskaya [1].

In case of no boundaries, one can use the Trotter product formula to study solutions of the Navier-Stokes equations (see p. 194 above). One considers the equations as the sum of the Euler equations and the heat equation. For example, this can be used to show convergence of the solutions as $\nu \to 0$ (see Ebin-Marsden [1, 2] and below), and might be useful in trying to show that the set of time T solutions is dense (see Marsden [2] and Bardos-Tartar [1]).

2.3 Some examples

Some three dimensional examples may help one understand some of the difficulties and why the three dimensional case is essentially different than the two dimensional case.

1. *The tornado*. Here we can assume $\nu = 0$. A "tornado" is a rotational (spiral) flow whose radius is decreasing and whose height is increasing. By conservation of angular momentum (or Kelvin's Circulation Theorem) the decrease in radius leads to an increase in angular velocity (this is what seems to cause the high wind speeds in these storms). The height increasing causing a diameter decrease occurs because of incompressibility.

A possible singularity in this flow would arise if the radius went to zero

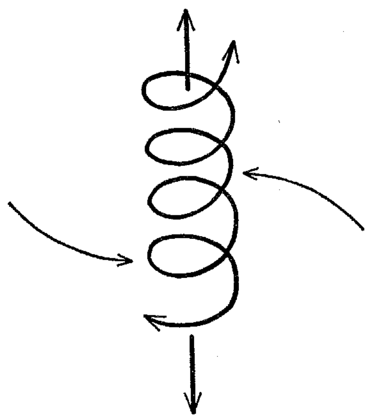


FIGURE 2.3

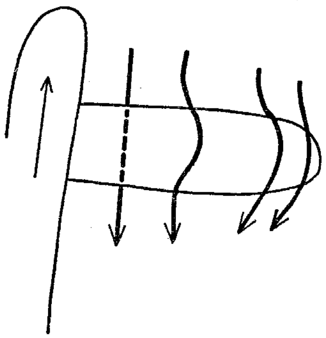


FIGURE 2.4

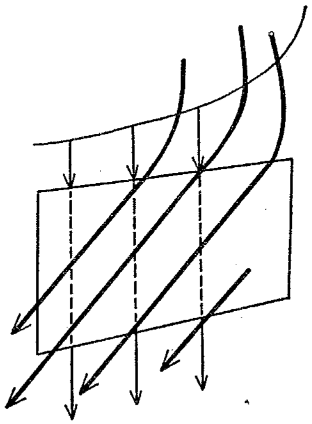


FIGURE 2.5

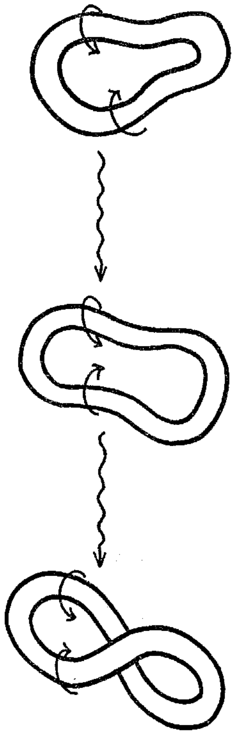


FIGURE 2.6

(and the velocity to infinity) in a finite time. However, one can prove the radius cannot reach zero in a finite time.

2. *The airplane wing*. If an airplane is flying at a constant velocity, assume the flow across the wing is time independent (i.e., it is a steady flow). One finds experimentally that that flow goes over the middle of the wing, but around the edge. This is called the "edge effect." See Figure 2.4. If we look on a sheet behind the airplane wing (possibly with smooth, rounded edges) one finds an apparent discontinuity in the flow between that below and above the sheet. It is unknown whether this is a real discontinuity in the solution to the Navier-Stokes equations. This may lead to a counter-example to the all-time existence problem. More precisely, if one sets up the initial conditions corresponding to the wind blowing towards the wing such a discontinuity surface might develop even if everything is initially C^∞. This is a genuinely three dimensional effect.

3. *Smoke rings*. As an example of the possible complexity of three dimensional flows, consider the configurations of smoke rings as they move in time. They tend to twist up and become more complicated as time progresses, as shown in Figure 2.6; cf. von Neuman **[1]**, where it is argued that this may be responsible for "turbulent like" flows.

2.4 Leray's theorem

Leray **[3]** found that for viscous three dimensional flows, having an $L^{3+\epsilon}$ bound on the velocity is sufficient to guarantee all-time smooth solutions. Here ϵ is any positive number. In the proof of this theorem, we use the following inequality of Sobolev-Nirenberg-Gagliardo; cf. Nirenberg **[1]**.

2.4.1 LEMMA. *Let $\Omega \subset \mathbb{R}^n$ be a bounded open set with smooth boundary and $f:\Omega \to \mathbb{R}^k$. Then if*

$$\frac{1}{p} = \frac{j}{n} + a\left(\frac{1}{r} - \frac{m}{n}\right) + (1-a)\frac{1}{q}$$

where $(j/m) \leqq a \leqq 1$ (if $m - j - n/r$ is an integer $\geqq 1$, only $a < 1$ is allowed) then

$$\| D^j(f)\|_{L^p} \leqq C_0 \| D^m f\|^a_{L^r} \| f\|^{1-a}_{L^q}$$

where C_0 depends on Ω and n, as well as on j, m, r, q, a, and $D^j f$ denotes a derivative of order j.

2.4.2 THEOREM (LERAY). *For three dimensions and $\nu > 0$, if v_t is a smooth solution of the Navier-Stokes equations on $[0, T)$ and if $\|v_t\|_{L^{3+\epsilon}} \leqq$*

$K < \infty$ for all $t \in [0, T)$, then v_t can be extended beyond T to a smooth solution of the Navier-Stokes equations.

(For a more complete analysis of theorems of this type, see Serrin [3] and Duff [2].)

SKETCH OF PROOF. We shall show here that $\|v_t\|_{L^{3+\epsilon}} \leqq K$ implies the H^1 norm of v_t is uniformly bounded. One can then proceed in a similar way to show that all H^s norms are bounded on $[0, T)$. This implies the result by general theory as in the Euler case.

Let $\Delta = \nabla^2$ be the usual Laplacian operator and let

$$(1) \qquad (\|v_t\|_1)^2 = -(\Delta v_t, v_t)_{L^2} = -\int_\Omega \langle \Delta v_t(x), v_t(x)\rangle \, dx$$

where $\Omega \subset R^3$ is our domain. By ellipticity of Δ and our boundary condition $v_t = 0$ on Q, one can show that this norm is equivalent to the usual H^1 norm, modulo the L_2 part (which we already have control over).

Now:

$$(2) \qquad \frac{1}{2}\frac{d}{dt}\|v_t\|_1^2 = \frac{1}{2}\frac{d}{dt}\int_\Omega \langle \Delta v_t(x), v_t(x)\rangle \, dx$$

$$= -\int_\Omega \left\langle \Delta v_t(x), \frac{\partial v_t}{\partial t}(x)\right\rangle dx$$

this calculation uses the bilinearity of $\langle\ ,\ \rangle$ and the fact that partials commute. But

$$\frac{\partial v_t}{\partial t} = -(v_t \cdot \nabla)v_t + \nu\,\Delta v_t - \operatorname{grad} p_t$$

so

$$(3) \quad \frac{1}{2}\frac{d}{dt}\|v_t\|_1^2$$

$$= -\int_\Omega \langle \Delta v_t(x), \nu\,\Delta v_t(x)\rangle \, dv(x) - \int_\Omega \langle \Delta v_t(x), -v_t(x)\cdot \Delta v_t(x) - \operatorname{grad} p_t\rangle \, dx$$

Note that

$$(4) \qquad -\int_\Omega \langle \Delta v_t(x), \nu\,\Delta v_t(x)\rangle \, dv(x) \leqq -\nu \cdot (\text{constant}) \cdot \|D^2 v_t\|_{L^2}^2$$

again using ellipticity of Δ, and ignoring L_2 parts. Clearly, it is important

to estimate the second integral in (3). Using integration by parts, we get

(5)

$$\int_\Omega \langle \nabla^2 v_t, -(v_t\cdot\nabla)v_t - \operatorname{grad} p_t\rangle = -\int_\Omega \langle \nabla v_t, \nabla(v_t\cdot\nabla)v_t\rangle - \int_\Omega \langle \nabla^2 v_t, \operatorname{grad} p_t\rangle$$

If there were no boundary we would have $\int_\Omega \langle \nabla^2 v_t, \operatorname{grad} p_t\rangle = 0$. The boundary term arises because $\nabla^2 v_t$ need not be parallel to $\partial\Omega$. This point is somewhat delicate and we shall ignore it, by pretending there is no boundary.[1] Writing out (5) in coordinates, we have

$$\int_\Omega \langle \nabla v_t, \nabla(v_t\cdot\nabla v_t)\rangle \, dx = \int_\Omega \sum_{i,j,k} \frac{\partial v_t^i}{\partial x^j}\frac{\partial}{\partial x^i}\left(v_t^k \frac{\partial v_t^j}{\partial x^k}\right) dx$$

$$(6) \qquad = \int_\Omega \sum_{i,j,k} \frac{\partial v_t^i}{\partial x^j}\left(\frac{\partial v_t^k}{\partial x^i}\frac{\partial v_t^j}{\partial x^k} + v_t^k \frac{\partial^2 v_t^j}{\partial x^k \partial x^i}\right) dx$$

$$= \int_\Omega \sum_{i,j,k} \frac{\partial v_t^i}{\partial x^j}\frac{\partial v_t^k}{\partial x^i}\frac{\partial v_t^j}{\partial x^k}\, dx + \int_\Omega \sum_{i,j,k}\left(\frac{\partial v_t^i}{\partial x^j}\right) v_t^k \frac{\partial^2 v_t^j}{\partial x^k \partial x^i}\, dx.$$

The second term is zero as one sees by integration by parts, using $\operatorname{div} v_t = 0$. Hence we find from (6) that

$$(7) \qquad \int_\Omega \langle \nabla v_t, \nabla(v_t\cdot\nabla v_t)\rangle = \|Dv_t\|_{L^3}^3$$

Putting things together we have from (4) and (7) the estimate

$$(8) \qquad \frac{1}{2}\frac{d}{dt}\|v_t\|_1^2 \leqq -\nu\cdot\text{constant}\cdot\|D^2 v_t\|_{L^2}^2 + \|Dv_t\|_{L^3}^3$$

Now we apply the lemma 2.4.1, letting $j = 1$, $p = 3$, and $r = m = 2$. We get

$$\|Dv_t\|_{L^3} \leqq C_0 \|D^2 v_t\|_{L^2}^a \|v_t\|_{L^q}^{1-a}$$

We have to find q and a. From the lemma, we know they satisfy this relationship;

$$\tfrac{1}{3} = \tfrac{1}{3} + a(\tfrac{1}{2} - \tfrac{2}{3}) + (1-a)\frac{1}{q}$$

[1] This point can be dealt with by rewriting the boundary term (using the equations and boundary conditions) and using the sort of argument in Proposition 1.4.1 to estimate this boundary part.

which implies that $a = [6/(q+6)]$. Hence

$$\|Dv_t\|_{L^3} \leqq C_0 \|v_t\|_{H^2}^{6/(q+6)} \|v_t\|_{L^q}^{q/(q+6)} \qquad (\text{since } \|D^2 v_t\|_{L^2} \leqq \|v_t\|_{H^2})$$

and so from (8),

$$\frac{1}{2}\frac{d}{dt}\|v_t\|_1^2 \leqq -\nu C_1 \|v_t\|_{H^2}^2 + C_0(\|v_t\|_{H^2}^2)^{9/(q+6)} \|v_t\|_{L^q}^{3q/(q+6)}$$

for any q we choose. If we let $q = 3 + \varepsilon$ ($\varepsilon > 0$) so that $[9/(q+6)] < 1$, we find:

$$\frac{1}{2}\frac{d}{dt}\|v_t\|_1^2 \leqq -\nu C_1 \|v_t\|_{H^2}^2 + C_0(\|v_t\|_{H^2}^2)^{9/(9+\varepsilon)} \|v_t\|_{L^{3+\varepsilon}}^{(9+3\varepsilon)/(9+\varepsilon)}$$

by assumption, the norm $\|v_t\|_{L^{3+\varepsilon}}$ is bounded, so if the H^2 norm of v_t is unbounded, the first term eventually dominates the inequality and the right hand side becomes $\leqq 0$, so for the H^2 norm large, the H^1 norm is decreasing, so bounded. If the H^2 norm of v_t is bounded, the H^1 norm is certainly also bounded. So in either case we get the H^1 norm of v_t is bounded. ∎

If we could lower $3 + \varepsilon$ to 2, we would have solved the problem. This is because the L^2 norm is globally bounded since the energy is decreasing in time. However in the L^p spaces there is an awfully large gap between $p = 2$ and $p = 3 + \varepsilon$! But nevertheless the assumption that the $L^{3+\varepsilon}$ norm is bounded seems to be quite weak, since it does not involve any derivatives. Compare with what we could get for the Euler equations in Theorem 2.1.2 above.

One can put the Navier-Stokes equations on $\mathscr{D}_\mu^s$ as we did the Euler flow to get a dissipative system. Although this converts the difficult term $\nabla_v v$ to an algebraic one, it doesn't seem to help in lowering Leray's $L^{3+\varepsilon}$ to L_2, because of other difficulties which arise.

It should be stressed that many aspects of this global existence question remain glaringly open; for example even if we accept a priori that a solution exists with no external forces and $v = 0$ on $\partial\Omega$, it is known that $v \to 0$ in L^2 as $t \to \infty$ but not in L^p or higher Sobolev norms (see 3.4.1). Intuitively, of course, we expect that v must go to zero quite strongly. In cases where global existence can be proved, this condition can be established; of course the question is closely related to the global existence of v.

We also mention that when the more complete viscosity term is used, including nonlinear terms, then the situation is improved but not decisively (for example in Ladyzhenskaya [2], appendix, the coefficient of the quadratic part cannot be too small).

2.5 Discussion of the Kolmogorov's law

We would like to briefly discuss a "physical law" which bears on the all time existence problem discussed above. The law is, we believe, some evidence for thinking that the equations do *not* break down if turbulence sets in, at least turbulence of a certain type commonly observed. (It does not cover the type of example indicated above by the airplane wing.) We shall discuss the "nature" of turbulence in more detail below.

For now let us think of turbulence intuitively as chaotic motion with collections of vortices of varying sizes. We can imagine again Couette flow *with the cylinders rotating rapidly in opposite directions. The motion* becomes chaotic, or fully turbulent when the cylinder speeds are large enough. In such a situation, we are continually pumping in energy at a constant rate, say ε. This energy is believed to pass from the larger to the smaller eddies and finally to dissipate due to viscosity. (In the energy spectrum the energy flows out to the tail of the spectrum, i.e., to large frequencies.)

The Kolmogorov law states that when the turbulence is locally isotropic, as seems to be usually the case, the velocity v_λ of eddies of size $\sim\lambda$ is proportional to $(\varepsilon\lambda)^{1/3}$. One also thinks of v_λ as the variation of v over distances of size λ. All this is in regions of $\mathbb{R}^3$.

The "derivation" of this law is done purely by dimensional analysis and elementary physical reasoning (see Landau-Lifschitz [1], p. 120–121). In other words it is a heuristic principle which does seem to be valid in a large number of situations. The law is more extensively discussed in Batchelor [1] where further references are given. (We remark that the law is the same as the "$k^{-5/3}$ law" for the energy spectrum.)

Now the Kolmogorov law appears to give a bound on the velocity field. In fact viewed directly the law *indicates* that for $0 < \alpha < \frac{3}{2}$, the αth power of Dv should be integrable. In other words,

$$v \in H_1^\alpha \quad \text{for} \quad 0 < \alpha < \tfrac{3}{2}.$$

Now if we use the Sobolev Lemma 2.4.1 above we see this implies that

$$\|v\|_{L^p} \quad \text{is bounded for} \quad p < 3$$

(here we take $j = 0$, $m = 1$, $n = 3$, $r = \frac{3}{2}$, $q = 2$, $a = 1$, $p = 3$).

Note that the critical value of p which occurs here is exactly the same as that which occurs in Leray's theorem. In other words, in this form the Kolmogorov law apparently is not quite decisive in deciding whether or not global solutions exist.

However, it has been pointed out to us by A. Chorin that when we translate the $k^{-5/3}$ law on the spectrum into a law on v, with some additional physical assumptions (basically that the turbulence consists of vortices, and so on, which are time repetitive) then one obtains a bound on the $H^{2/3}$ norm of v. Now by 2.4.1 ($j = 0$, $m = \frac{2}{3}$, $n = 3$, $q = 2$, $a = 1$, $p = 3\frac{3}{5}$, $r = 2$), we would then find that

$$\text{the } L_{3\frac{3}{5}} \text{ norm is bounded}$$

which *is* decisive for Leray's theorem. This analysis is, however, still the subject of research and it is not yet complete; cf. Chorin [**1**].

We note not all flows need have the $H^{2/3}$ norm bounded; for example on the three torus with coordinates (x, y, z),

$$v(x, y, z) = (\sin y, 0, \sin(x - ty))$$

solves the Euler equations, but its H^α norms for $\alpha > 0$ are unbounded as $t \to \infty$; a similar phenomena is undoubtedly true for the Navier-Stokes equations. But this example does not conform to the sort of hypotheses in Kolmogorov's law.

The moral of all this is that if the physical assumptions concerning turbulence are correct then Kolmogorov's law *may* indeed show, via Leray's theorem, that global smooth solutions do indeed exist for turbulent flow. What would be interesting would be a proof of Kolmogorov's law (in this $H^{2/3}$ form) under the assumption of "turbulence," say as defined in § 3 below, plus some isotropy assumption. This would provide a very important step towards the complete understanding of Leray's problem.

3. Turbulence, stability, and the limit of zero viscosity

3.1 Nature of turbulence

Intuitively, turbulence is very complicated fluid motion. There are two rival theories as to the mathematical structure of turbulence

(a) Leray believed, perhaps because he could not prove that smooth solutions to the Navier-Stokes equations exist for all time, that solutions can become singular in a finite time. *Turbulence therefore is supposed to occur when Navier-Stokes equations break down and fail to describe the actual flow of the fluid.* Although this point of view is not widely accepted, there is some evidence for it, which was developed by Leray. There is a *formal* counter-example, i.e., an example of a Navier-Stokes flow where the equations eventually do not hold (Leray [3], p. 224). Namely, it can be shown that if $u(x)$ is a solution of an equation $Au = 0$ where A is a specific nonlinear elliptic operator then

$$u(t, x) = \frac{1}{\sqrt{T - t}}\, u(x/\sqrt{T - t})$$

is a solution to the Navier-Stokes equations with a singularity at $t = T$. However, so far no one has found a solution to $Au = 0$ which is also square integrable (i.e., has finite energy) and $\operatorname{div} u = 0$. So this is not a "real" counter-example. As Professor G. Duff points out, in case one is willing to accept infinite energies, one can easily do this just using $u(x, y, z) = (x, y, -2z)$ with $u(t, x, y, z)$ as above and $p(t, x, y, z) = 3z^2/(T - t)$.

(b) *The equations do hold for all time, but in case of turbulence the solutions necessarily get complicated.*

This seems to be the current belief among most theoretical fluid dynamicists, and the one we will favor below.

There is, in conjunction with (b), a well studied statistical theory of turbulence. This involves computing correlation functions, spectra and so on. One fundamental "law" in this field is the Kolmogorov relation discussed in § 2.5 above. This theory, while it is useful, does not explain what turbulence "really" is or how it arises. It would be very nice if the statistical theorems could be made a consequence of the Navier-Stokes equations or perhaps equivalent to them (as Brownian motion is related to the Laplace equation). For this statistical aspect, see for example, Landau-Lifschitz [**1**] Batchellor [**1**] and the recent survey article S. Orszag [**1**]. Much important work has been done recently by Kraichnan [**1, 3**]. As was argued in § 2.5, knowledge from statistical theory does indicate that it is alternative (b) which is correct.

There is some reason to believe that mathematically turbulence is related to the bifurcation theorem. The parameter is now the Reynolds number ($\mathscr{R} = Vd/\nu$) defined earlier. To vary $\mathscr{R}$, one can either vary ν or V. As mentioned above, we may write the main Navier-Stokes equation as

$$\frac{\partial v_t}{\partial t} + (v_t \cdot \nabla)v_t - \frac{1}{\mathscr{R}} \nabla^2 v_t = -\operatorname{grad} p_t.$$

Physically, one would expect complexity as one decreased ν or increased V (which is the same as increasing $\mathscr{R}$). In fact as $\mathscr{R}$ is increased one often observes a sudden qualitative change in the solutions. This means the flow represents a "stable" solution for small $\mathscr{R}$, which becomes unstable as $\mathscr{R}$ increases and consequently the flow moves to another stable solution.

This behavior explains a standard hydrodynamics "paradox." (cf. Ladyzhenskaya [**2**], Birkhoff [**1**].) One would expect that symmetric initial data would yield symmetric flows ("symmetric causes lead to symmetric effects"). And, indeed, this is observed for small $\mathscr{R}$. However, as $\mathscr{R}$ increases asymmetric flows develop. This can be explained by noting that the initial data cannot be perfectly symmetric, and at the critical point these symmetric solutions become unstable and the solution "falls" to an asymmetric stable one. In fact sometimes the more careful one is to have the initial data symmetric, the longer the symmetries are preserved in the flow.

For fixed Reynolds number, if we have turbulence, one generally observes the turbulence to occur in well defined regions, such as the wake behind an airplane. The flow in the turbulent region is time dependent and highly rotational. Viscosity effects occur only in the smallest eddies, so in some respects the Euler equations are valid in the turbulent regions! See Landau-Lifshitz [1], § 34.

3.2 The limit of zero viscosity

Because turbulence and bifurcation phenomena take place as $\mathscr{R}$ increases, the detailed study of the Navier-Stokes equations as $\mathscr{R} \to \infty$ is of considerable importance.

We can study the nature of Navier-Stokes flows as $\mathscr{R}$ increases by considering $\nu \to 0$ (rather than $V \to \infty$). So we shall consider V and d fixed and write the Navier-Stokes equation like this:

$$\text{N.S.}\begin{cases} \dfrac{\partial v_t}{\partial t} - \nu\,\Delta v_t + \Delta_{v_t} v_t = -\operatorname{grad} p_t + f \\ \operatorname{div} v_t = 0 \\ v_t = 0 \quad \text{on } \partial M \text{ (or } v \text{ specified on } \partial M) \end{cases}$$

where f is an external force. Investigating $\nu \to 0$ is very delicate since the equations are non-linear, and the Δv_t term is of highest order. When $\nu = 0$ the equation is in a sense "hyperbolic," and when $\nu > 0$ the equation is "parabolic" or dissipative. When $\partial M \neq \varnothing$, one cannot even expect C^0 convergence of solutions as $\nu \to 0$ since the boundary conditions of the Euler equations are different than those of the Navier-Stokes equations. In Euler flow, the fluid slips by the boundary (indeed one can prove that in general even if $v_0 = 0$ on ∂M in Euler flow, v_t will not vanish on ∂M but merely be parallel to ∂M), in the Navier-Stokes equations the fluid is always motionless at the boundary.

The first question then is whether the boundary is the only source of trouble. As the next theorem shows, the answer is, in a sense, yes.

First some notation. For $\nu > 0$, call v_t^ν the solution of the Navier-Stokes equations with viscosity ν. Denote by $v_t^0 = v_t$, an Euler flow. Assume $v_0^\nu = v_0$.

3.2.1 THEOREM. *If $\partial M = \varnothing$ and the initial condition v_0 is a C^∞ vector field (or H^{s+2} will do) then v_t^ν converges to v_t in the H^s topology as $\nu \to 0$ on a t-interval* $[0, T]$, $T > 0$.

This was proved by Ebin-Marsden [1]. We shall give below the main steps of a somewhat different proof. One can, it seems, relax the compactness of M. One can show, also, that the convergence takes place on the largest t-interval for which v_t exists in H^s.

Swann [1] and Kato [2] also proved this theorem for $\mathbb{R}^3$ using more general arguments. The simplicity of the proof given below is due to the cancellation of the highest order terms trick used earlier. This proof requires the following lemma. (This lemma, with essentially the same proof gives the C^1 regularity theorem 2.1.1 stated earlier.) In this lemma we can assume $\partial M \neq \varnothing$.

3.2.2 LEMMA. *Let v_t satisfy the Euler equations, $v_t \in T_e\mathscr{D}^{s-1}(M)$ on $[0, T)$, and $\|v_t\|_{H^{s-1}} \leqq K$ on $[0, T)$. Then if $v_0 \in T_e\mathscr{D}^s(M)$, $\|v_t\|_{H^s}$ is also bounded on $[0, T)$. (The bound depends on K and v_0.)*

PROOF. Define $\nabla^s v$ to be the iterated sth covariant derivative. Then if we let $\|v\|_s^2 = (\nabla^s v, \nabla^s v)$ this norm is equivalent to the H^s norm modulo the L_2 part, which we can safely ignore. In particular, consider

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}(\nabla^s v_t, \nabla^s v_t) &= \left(\nabla^s v_t, \nabla^s \frac{\partial v_t}{\partial t}\right) \\ &= -(\nabla^s v_t, \nabla^s P_e(\nabla_{v_t} v_t)) \end{aligned}$$

where P_e is the Hodge projection. Recall that from the discussion of Green's operators that $P_e(\nabla_{v_t} v_t) = \nabla_{v_t} v_t - d\Delta^{-1}\,\delta\nabla_{v_t} v_t$. Hence

$$\frac{1}{2}\frac{d}{dt}(\nabla^s v_t, \nabla^s v_t) = -(\nabla^s v_t, \nabla^s \nabla_{v_t} v_t) + (\nabla^s v_t, d\Delta^{-1}\,\delta\nabla_{v_t} v_t).$$

Now

$$\begin{aligned} &(\nabla^s v_t, \nabla^s_{v_t} v_t) \\ &\quad = \frac{1}{2}\int_M v_t\langle \nabla^s v_t, \nabla^s v_t\rangle + \int_M \langle \nabla^s v_t, \nabla^{s-1}_{v_t} v_t\rangle + \text{lower order terms} \end{aligned}$$

This computation uses the formula $v\langle w, y\rangle = \langle \nabla_v w, y\rangle + \langle w, \nabla_v y\rangle$ and the fact that ∇_v and ∇ commute up to lower order terms involving the curvature. (In fact, $\nabla_X\nabla_Y Z - \nabla_Y\nabla_X Z = \nabla_{[X,Y]}Z + R(X, Y)Z$.)

Now letting $f = \langle \nabla^s v_t, \nabla^s v_t\rangle$ we have

$$\begin{aligned} \int_M v_t\langle \nabla^s v_t, \nabla^s v_t\rangle &= \int_M v_t(f) = \int_M \langle \nabla f, v_t\rangle \\ &= \int_M f\cdot \delta v_t + \int_M f\,{*v_t} = 0 + 0 = 0 \end{aligned}$$

since $\delta v_t = 0$ and saying v_t is tangent to ∂M means $*v_t = 0$ on ∂M.

It is evident that $\int_M \langle \nabla^s v_t,\ \nabla_v^{s-1} v_t\rangle$ is dominated by $\|v_t\|_{H^s}^2\,\|v_t\|_{H^{s-1}}$, since the H^s functions form a ring.

Now, using the fact $\delta v_t = 0$, one can show that if v_t is H^{k+1} then $\delta\nabla_{v_t} v_t$ is H^k. This calculation has been done before. Hence $(\nabla^s v_t,\ d\Delta^{-1}\,\delta\nabla_{v_t} v_t)$ will be dominated by $\|v_t\|_{H^s}^2\,\|v_t\|_{H^{s-1}}$ as well. Putting this together we get

$$\frac{1}{2}\frac{d}{dt}(\|v_t\|_{H^s}^2) \leqq L\|v_t\|_{H^s}^2\,\|v_t\|_{H^{s-1}}$$

$$\leqq LK\,\|v_t\|_{H^s}^2$$

Hence

$$\|v_t\|_{H^s}\frac{d}{dt}\|v_t\|_{H^s} \leqq LK\,\|v_t\|_{H^s}^2$$

and so

$$\frac{d}{dt}\|v_t\|_{H^s} \leqq LK\,\|v_t\|_{H^s}$$

which implies (by Gronwall's inequality) that $\|v\|_{H^s} \leqq \|v_0\|_{H^s}\,e^{LKt}$, which proves the lemma. ∎

Note that this regularity lemma holds for manifolds with boundary. We now use this argument in one more lemma.

3.2.3 Lemma. *Let $\partial M = \varnothing$ and v_t^ν be a solution of the Navier-Stokes equation with $v_0^\nu = v_0$ a H^s vector field. Then v_t^ν exists on an interval* $[0, T]$ *where T is independent of ν, and there is a constant K and* $\Lambda > 0$ *such that:*

$$\|v_t^\nu\|_{H^s} \leqq K \quad \textit{for} \quad t \in [0, T], \qquad 0 \leqq \nu \leqq \Lambda.$$

Proof. We argue as above by computing the rate of change of the H^s norm of v_t^ν. Since $\partial M = \varnothing$ we can integrate the Laplacian term by parts to get something negative which can then be discarded without destroying the inequality. Keeping track of the constants in the above proof more carefully, we get

$$\frac{1}{2}\frac{d}{dt}(\|v_t^\nu\|_{H^s}^2) \leqq L\|v_t^\nu\|_{H^s}^3$$

or

$$\frac{d}{dt}\|v_t^\nu\|_{H^s} \leqq 2L\|v_t^\nu\|_{H^s}^2.$$

Thus from the dominating ordinary differential inequality $(d/dt)x(t) \leqq 2Lx(t)^2$ we see that on a sufficiently small t-interval $\|v_t^\nu\|_{H^s}$ is a priori bounded independent of ν. The lemma then follows (see the existence statements in § 2). ∎

Now we are ready to prove the theorem.

Proof of Theorem 3.2.1. Suppose v_t, the Euler vector field with initial conditions $v_0 \in T_e\mathscr{D}_\mu^\infty(M)$ exists on $[0, T]$. Since $[0, T]$ is closed, we see that $\|v_t\|_{H^s}$ is bounded on $[0, T]$ for any s. For any $\nu > 0$ we have by Lemma 3.2.3 that $\|v_t^\nu\|_{H^s}$ is bounded, the bound being *independent* of ν. Using the sort of calculations as in the above lemma we find

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|v_t - v_t^\nu\|_{H^s}^2 &= (\nabla^s(v_t - v_t^\nu),\ \nabla^s(P(\nabla_{v_t}v_t) - P(\nabla_{v_t^\nu}v_t^\nu) - \nu\,\Delta v_t^\nu))\\ &\leqq \nu\,\|v_t - v_t^\nu\|_s\,\|v_t^\nu\|_{s+2} + (v_t - v_t^\nu,\ \nabla_{v_t - v_t^\nu}v_t + \nabla_{v_t^\nu}(v_t - v_t^\nu))\\ &\quad + \text{pressure terms of lower order.}\end{aligned}$$

Note that as in 3.2.3 this situation is a bit different than the lemma. If M had boundary, the Δv_t^ν term need not be tangent to ∂M, so we need to suppose $\partial M = \varnothing$.

So we want to estimate the terms

$$(\nabla^2(v_t - v_t^\nu),\ \nabla^2\nabla_{v_t - v_t^\nu}v_t) \text{ and } (\nabla^2(v_t - v_t^\nu),\ \nabla^s\nabla_{v_t^\nu}(v_t - v_t^\nu))$$

The second term's highest order part is zero, as in the lemma:

$$\int_M v_t^\nu\langle\nabla^2(v_t - v_t),\ \nabla^2(v_t - v_t^\nu)\rangle = 0$$

and the first term is dominated by $\|v_t - v_t^\nu\|_{H^s}^2\,\|v_t\|_{H^{s+1}}$.

Now by lemma 3.2.2 $\|v_t\|_{H^{s+1}}$ and $\|v_t\|_{H^{s+2}}$ are bounded on $[0, T]$. Also, by 3.2.3 $\|v_t^\nu\|_{H^{s+2}}$ is bounded on $[0, T]$, if we choose T sufficiently small (which depends on v_0). Thus we have

$$\frac{1}{2}\frac{d}{dt}\|v_t - v_t^\nu\|_{H^s}^2 \leqq \nu K_1\,\|v_t - v_t^\nu\|_{H^s} + K_2\,\|v_t - v_t^\nu\|_{H^s}^2.$$

So, by Gronwall's inequality (after differentiating as in the lemma)

$$\|v_t - v_t^\nu\|_{H^s} \leqq \nu K_1 e^{K_2 t}$$

which approaches 0 as ν approaches 0. ∎

As a corollary we observe that the convergence is linear in ν. That is, for fixed $t \in [0, T]$,

$$\|v_t - v_t^\nu\|_{H^s} = 0(\nu).$$

3.3 Boundary layer theory

Boundary layer theory relates to the study of viscous flows on manifolds with boundary. As mentioned above, v_t^ν cannot converge to v_t at the boundary, because of the no-slip condition which is present when $\nu > 0$. Hence one might expect there would be layers near the boundary where the derivatives of v_t^ν get large as ν approaches 0 (Figure 3.1).

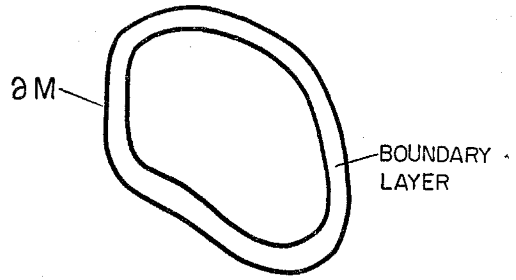


FIGURE 3.1

It seems, at first, a reasonable conjecture that v_t^ν converges to v_t on compact subsets contained in the interior of the manifold. However, on physical grounds we can argue that this conjecture is false. Consider the situation shown in Figure 3.2.

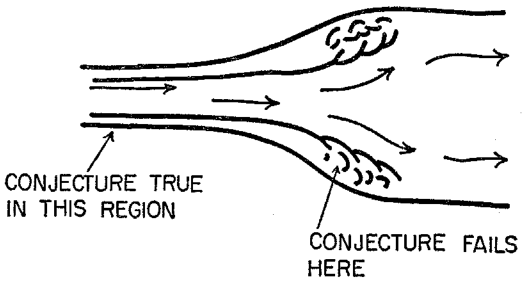


FIGURE 3.2

In the figure one sees fluid flowing from a narrow pipe into a wider one. The boundary layer is carried by the stream into the *interior* of the larger portion of the pipe. So because of the downstream wake of the boundary layer, v_t^ν does not converge to v_t in the interior of the wider pipe; v_t does not have any downstream wake (and is a stationary potential flow). This effect will be seen for any $\nu > 0$ and will *not* die out as ν approaches 0. So it seems the convergence of v_t^ν depends on the geometry of the boundary.

There is an extensive literature on boundary layers (see Serrin [4] and the bibliography), but the theory remains somewhat unsatisfactory. In particular we would like to mention the work of Lions [2] who has several related theorems, (but as yet his hypotheses seem to break down for the Navier-Stokes equations). In Ebin-Marsden [2] one can find results which perhaps isolate the source of the difficulties. For example, one relevant fact already noted above is that Δv_t^ν need not be parallel to the boundary, even if v_t^ν is. In the aforementioned paper it is shown, under a certain condition that, as in the no boundary case, the time of existence of the equations is independent of the viscosity, and we have convergence in the C^1 norms. This is discussed in the appendix to this section.

3.4 Stability and turbulence

Shortly we shall explain more fully the Ruelle-Takens [1] theory of turbulence. For now we just wish to stress the point that turbulence appears to be some complicated flow which sets in after sucessive bifurcations have occurred (see Part I, § 3). In this process, stable solutions become unstable, as the Reynolds number $\mathscr{R}$ is increased. Hence turbulence is supposed to be a necessary consequence of the equations and in fact of the "generic case" and just represents a complicated solution. For example in Couette flow as one increases the angular velocity Ω_1 of the inner cylinder one finds a shift from laminar flow to Taylor cells or related patterns (discussed later) at some bifurcation value of Ω_1. Eventually turbulence sets in. In this scheme, as has been realized for a long time, one first looks for a stability theorem and for when stability fails (Chandresekar [1], Lin [1] etc.). For example, if one stayed close enough to laminar flow, one would expect the flow to remain approximately laminar. Serrin [2] has a theorem of this sort which we present as an illustration:

3.4.1 STABILITY THEOREM. *Let $D \subset \mathbb{R}^3$ be a bounded domain and suppose the flow v_t^ν is prescribed on ∂D (this corresponds to having a moving boundary, as in Couette flow). Let $V = \max_{\substack{x \in D \\ t \geq 0}} \|v_t^\nu(x)\|$, d = diameter of D and ν equals the viscosity. Then if the Reynolds number $\mathscr{R} = (Vd/\nu) \leqq 5.71$, v_t^ν is universally L^2 stable.*

Universally L^2 stable means that if $\bar{v}_t^\nu$ is *any* other solution to the equations and with the same boundary conditions, then the L^2 norm (or energy) of $\bar{v}_t^\nu - v_t^\nu$ goes to zero as $t \to 0$.

The proof of 3.4.1 is really very simple and we recommend reading Serrin [2] for the argument.

Chandresekar [1], Serrin [2], and Velte [1] have analyzed criteria of this sort in some detail for Couette flow, to which we shall return shortly.

As a special case, we recover something that we expect. Namely if $v_t^\nu = 0$ on ∂M is any solution for $\nu > 0$ then $v_t^\nu \to 0$ as $t \to \infty$ in L^2 norm, since by 3.4.1, the zero solution is universally stable.

3.5 Discussion of Couette flow

Let us consider Couette flow again but now in a bit more detail. Let Ω_1 be the angular velocity of the inner cylinder, and Ω_2 the angular velocity of the outer cylinder. Say that $\Omega_1 > 0$ if the cylinder is going in a counterclockwise direction (this just keeps the signs right). Let R_1 be the radius of the inner cylinder and R_2 the radius of the outer one.

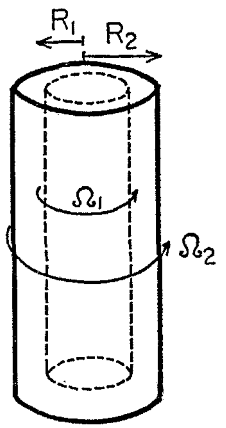


FIGURE 3.3

If one assumes that Ω_1 and Ω_2 are constants, then Couette flow is a particular stationary, or steady-state, solution of the Navier–Stokes equations; that is, it satisfies

$$\begin{cases} 0 = \dfrac{\partial v}{\partial t} = -(v \cdot \nabla)v + \nu \nabla_2 v - \operatorname{grad} p \\ \operatorname{div} v_t = 0 \\ v_t \text{ does not slip at the boundary.} \end{cases}$$

In practice, there are special phenomena at the ends of the cylinders. We will ignore these end effects. This can be done mathematically either by identifying the ends (and thereby work on the annulus crossed with the circle) or by considering the cylinders to be very long with respect to R_2.

For arbitrary Ω_1, Ω_2, Couette flow is an explicit solution, which, in cylindrical coordinates (r, φ, z) is given as follows

$$\begin{cases} v_\varphi = \left(\dfrac{\Omega_2 R_2^2 - \Omega_1 R_1^2}{R_1^2 - R_2^2}\right) r + \left(\dfrac{(\Omega_1 - \Omega_2) R_1^2 R_2^2}{R_1^2 - R_2^2}\right)\dfrac{1}{r} \\ \quad = Ar + \dfrac{B}{r} \\ v_r \equiv 0 \\ v_z \equiv 0. \end{cases}$$

This solution is stable for small velocities and hence is physically observed. In what follows, we are interested in qualitative change of the solutions as Ω_1 and Ω_2 vary.

Fix $\Omega_2 > 0$ and suppose $\Omega_1 > 0$ and imagine Ω_1 increasing. Thus Ω_1 is our bifurcation parameter. When Ω_1 reaches a critical point, Taylor

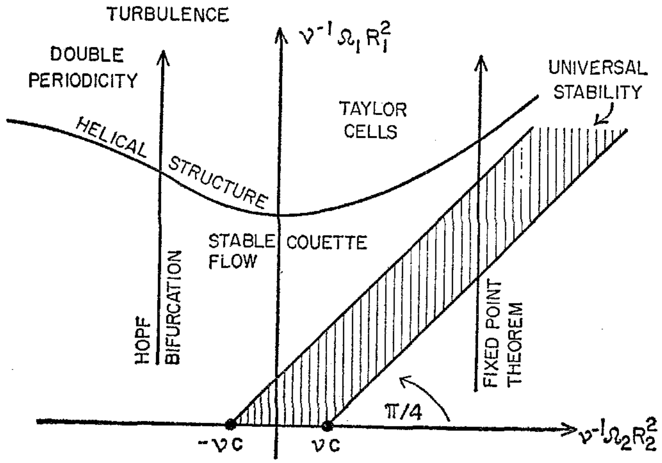


FIGURE 3.4

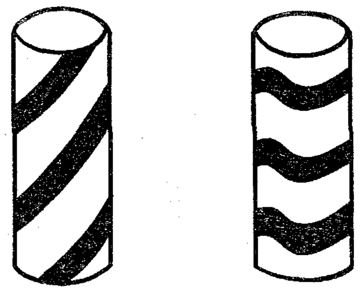
$c = \pi^2(R_2^2 - R_1^2)/[R_1 R_2 \log(R_2/R_1)]^2$

cells develop (see page 170). Intuitively, what is happening is that the particles near the inner cylinder have so great an angular velocity that centrifugal force throws them towards the outer edge. Since the flow is incompressible particles near the outer edge must be pushed back towards the inner cylinder. Hence each particle experiences a circular motion in the $r - z$ plane as well as its motion due to the Couette flow (this was described earlier).

The various qualitative "stable" flows occurring for various values of Ω_1, Ω_2 are shown in the simplified graph in Figure 3.4.

In Figure 3.4, the zone of universal stability has the meaning explained in 3.4.1 and is taken from Serrin [2]. In reality Couette flow is more complicated; cf. Coles [1], and Snyder [1].

For $\Omega_2 > 0$ and Ω_1 sufficiently large, we get a helical (barber pole) pattern of Taylor cells. In this region solutions are periodic in time and presumably correspond to the periodic solutions predicted by the Hopf bifurcation theorem. We also can develop "double periodicity" corresponding to undulations or waves in the Taylor cells which progress around the cylinders. As Ω_1 increases further this pattern breaks down and we find "fully developed" turbulence. This turbulent motion appears as just chaotic, disorganized motion between the cylinders. Now for



(a) HELICAL STRUCTURE (b) DOUBLE PERIODICITY

FIGURE 3.5

$\Omega_2 > 0$ and Ω_1 sufficiently large, we find steady-state Taylor cells described earlier. This bifurcation corresponds to the fixed point theorem discussed earlier (Theorem 3.1). Actually the situation is complicated by the presence of symmetry groups and these groups can affect the generic possibilities for bifurcations; namely whether or not we are in a *real* irreducible representation space for the group will determine if a bifurcation takes place when the eigenvalues generically cross the unit circle on or off the real axis. This seems to distinguish the two cases in Couette flow. See Ruelle [4] for details.

Let us investigate this bifurcation idea further. Fix $\Omega_2 < 0$. Recall that the Navier-Stokes equations can be thought of as a vector field on $T_e\mathscr{D}^s_\mu(M)$ (the space of v_t). So to apply the Hopf Bifurcation theorem in this situation, one has to check that it holds in infinite dimensional spaces. If this could be shown, the Hopf theorem would imply that there would be stable periodic solutions as the eigenvalues of the linearized Navier-Stokes equation crossed the imaginary axis (this involves many nontrivial technicalities due to the presence of not everywhere defined vector fields).

Note. Linearization is done as explained in Part I, § 3 and not by just throwing out the nonlinear terms—the Stokes linearization (cf. Ladyzhenskaya [2]). As more eigenvalues cross the axis the stable periodic solutions lie in higher dimensional tori and thus the periodicity would become more complicated.

In fact, recently two papers have appeared justifying the use of the Hopf theorem in this context: Bruslinskaya [2] in 1965 and Sattinger [1] in 1971, established that the Hopf theorem does indeed hold for the Navier-Stokes equations. Moreover, an important feature is that they can show that when a bifurcation does occur *one retains global existence of smooth solutions near the closed orbit.* This is in fact good evidence in the direction of verifying that the Navier-Stokes equations do *not* break down when turbulence develops. Other important papers in this regard are Joseph and Sattinger [1] and Judovich [3, 4]. These theorems may be more easily proved if we pass to $\mathscr{D}^s_\mu$ and use the characterization of periodic solutions given on p. 193 above. See also Marsden [3, 4].

For $\Omega_2 > 0$ and Ω_1 sufficiently large, the flow develops Taylor cells that are steady-state. Taylor conjectured that Taylor cells are an exact solution to the full (not linearized with respect to v_t) Navier-Stokes equations. But, up to recently no one could even prove these have such a solution. Velte [3] in 1966 showed that Taylor cells do correspond to a solution of the Navier-Stokes equations which bifurcate off from Couette flow. His proof consisted of two main parts:

(a) He checked when the eigenvalues of the linearized equations crossed the imaginary axis.

(b) He used this information together with the Leray-Schauder degree to show a certain operator has a fixed point in a neighborhood of the solution of the linearized equations, and this fixed point corresponds to the desired stationary solution of the full Navier-Stokes equations.

The second part of the proof of this theorem can also be proved using Theorem 3.3.1 in Part I. (In fact this method gives sharper smoothness of the curve of fixed points.) Recall that this theorem involves a Hilbert space H and a smoothly parameterized set of operators $\{\Phi_\mu\}$ on H. Then certain conditions on the eigenvalues of $D\Phi_\mu(0)$ imply that as μ crosses some critical value, the Φ_μ develops another curve of fixed points. In this application we use the map $\Phi_{\Omega_1}: \mathscr{X}^s_\mu(H) \to \mathscr{X}^s_\mu(M)$ where M is the region between the cylinders, and $\mathscr{X}^s_\mu(M)$ is the H^s divergence free vector fields, defined by

$$\Phi_{\Omega_1}(v_t) = \nu^{-1}\Delta^{-1}_{\Omega_1}((v_t \cdot \nabla)v + \operatorname{grad} p)$$

where the $\Delta^{-1}_{\Omega_1}$ means the solution of the Dirichlet problem with the appropriate boundary data determined by Ω_1, and our fixed value of Ω_2. It can be shown that this operator satisfies the hypotheses of the theorem and clearly the fixed points of this operator are steady solutions to the Navier–Stokes equations (Φ maps H^s to H^s because of the smoothing effect of Δ^{-1}). As Ω_1 increases one might expect a bifurcation in the solution of the equations since Theorem 3.1 states there is a bifurcation in the fixed points of Φ_{Ω_1}. Taylor cells occur when the new line of fixed points of Φ_{Ω_1} appear. Strictly speaking one must use a version of the bifurcation theorem in which a symmetry is present, since both Couette flow and Taylor cells are symmetric about the z-axis. As D. Ruelle has pointed out, the theorem must be modified to cover the case when a symmetry group is present.

There is an interesting phenomenon of indeterminacy connected with the doubly periodic solutions (Coles [**1**]). Namely, that the geometry of the waves which develop (their number and location) depends on the past history of the experiment. This is, however, easily understood if we think of a bifurcating vector field on the space $\mathscr{X}_\mu$. Indeed, the vector field can develop many different new attracting sets as $\mathscr{R}$ increases and we can be attracted to any one of these. We thank D. Ruelle for pointing this out, as well as the above proof of Vette's theorem.

Couette flow is not the only situation where this Taylor cell type of phenomenon occurs and where the above analysis is possible. For example, in the Bénard Problem one has a vessel of water heated from below. At a critical value of the temperature gradient, one observes convection currents, which behave like Taylor cells; cf. Rabinowitz [**1**].

This transition from laminar to periodic motion (the Hopf bifurcation) occurs in many other physical situations such as flow behind an obstacle (cf. Scientific American, January, 1970, p. 40).

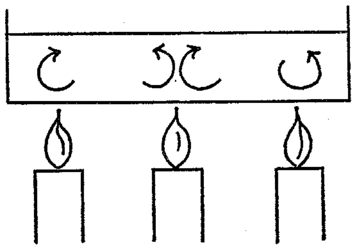


FIGURE 3.6

3.6 A definition of turbulence

A traditional definition (as in Hopf [**2**], Landau–Lifschitz [**1**]) says that turbulence develops when the vector field v_t can be described as $v_t(w_1, \ldots, w_n) = f(tw_1, \ldots, tw_n)$ where f is a quasi-periodic function, i.e., f is periodic in each coordinate, but the periods are not rationally

related. For example, if the orbits of the v_t on the tori given by the Hopf theorem can be described by spirals with irrationally related angles, then v_t would such a flow.

Considering the above example a bit further, it should be clear there are many orbits that the v_t could follow which are qualitatively like the quasi-periodic ones but which fail themselves to be quasi-periodic. In fact a small neighborhood of a quasi-periodic function may fail to contain many other such functions. One might desire the functions describing turbulence to contain most functions and not only a sparse subset. More precisely, say a subset U of a topological space S is *generic* if it is a Baire set (i.e., the countable intersection of open dense subsets). It seems reasonable to expect that the functions describing turbulence should be generic, since turbulence is a common phenomena and the equations of flow are never exact. Thus we would want a theory of turbulence that would not be destroyed by adding on small perturbations to the equations of motion.

The above sort of reasoning lead Ruelle–Takens [**1**] to point out that since quasi-periodic functions are not generic, it is unlikely they "really" describe turbulence. In its place, they propose the use of "strange attractors." (See Smale [**2**] and Williams [**1**].) These exhibit much of the qualitative behavior one would expect from "turbulent" solutions to the Navier–Stokes equations and they are stable under perturbations.

Here is an example of a strange attractor. Let $U \subset \mathbb{R}^n$ be open and $\sigma_t: U \to U$ some flow; suppose further for $x \in U$, there is an $s \in \mathbb{R}$ such that $\sigma_{s+t}(x) = \sigma_t(x)$, i.e., x belongs to a periodic orbit of the flow. Let $(d/dt)\sigma_t(x)\big|_{t=0} = Y_x$ and let V be the affine hypersurface in U orthogonal to Y_x. For a small neighborhood S of x in V, there is a map $P: S \to V$ called the Poincaré map, defined as follows: For $w \in S$, it is easy to show there is a smallest $P_w \in \mathbb{R}$ such that $\sigma_{P_w}(w) \in V$. Call $P(w) = \sigma_{P_w}(w)$. Now of course one can do this for each point of the periodic orbit. By doing this one gets a map on a small "tubular" neighborhood of the periodic orbit in U. (Here one must check that there is a neighborhood N of the orbit such that if $x \in N$ then x belongs to a unique hypersurface orthogonal to the orbit.) Also one can drop the condition that P be defined about a closed orbit by requiring that the vector field be almost parallel and everywhere transversal to a hypersurface V. In this case one can define a Poincaré map P over the entire space U by letting $P(x)$ be the first intersection of the integral curve through x with V.

In particular consider V to be a solid torus in three space and suppose we have a flow σ_t on U such that its Poincaré map wraps the torus around twice (Figure 3.7). Then the attracting set of the flow (i.e., $\{x \in U \mid x = \lim_{t\to\infty} \sigma_t(y)$ for some $y \in U\}$ is locally a Cantor set cross a 2-manifold

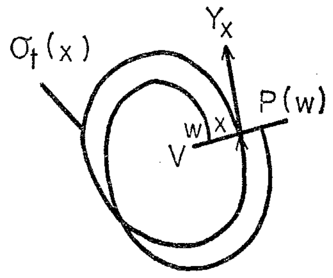


FIGURE 3.7

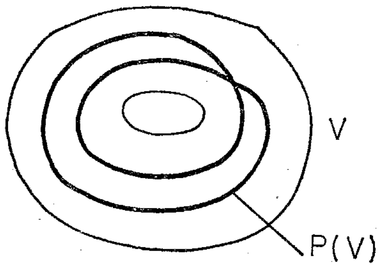


FIGURE 3.8

(see Smale [**2**]). This is certainly a strange attractor! Ruelle–Takens [**1**] have shown if we define a strange attractor to be one which is neither a closed orbit or a point, then there are stable strange attractors on T^4 in the sense that a whole neighborhood of vector fields has a strange attractor as well.

If the attracting set of the flow, in the space of vector fields, which is generated by Navier–Stokes equations is strange, then a solution attracted to this set will clearly behave in a complicated, turbulent manner and since strange attractors are "generic", this sort of behavior should not be uncommon. Thus we have the following reasonable definition of turbulence as proposed by Ruelle–Takens:

"... the motion of a fluid system is turbulent when this motion is described by an integral curve of a vector field X_μ which tends to a set A, and A is neither empty nor a fixed point nor a closed orbit."

This turbulent motion is supposed to occur on one of the tori T^k that occurs in the Hopf bifurcation. This takes place after a *finite* number of successive bifurcations have occurred. However as S. Smale and C. Simon pointed out to us, there may be an infinite number of other qualitative changes which occur during this *onset of turbulence* (such as stable and unstable manifolds intersecting in various ways etc).

Since this sort of phenomena is supposed to be "generic," one would expect it to occur in other similar phenomena such as the Benard problem. (As the temperature gradient becomes very large, the flow becomes "turbulent.")

Recently Ruelle [**1**] has shown how the usual statistical mechanics of ergodic systems can be used to study the case of strange attractors, following work of Bowen [**1**] and Sinai [**1**]. It remains to connect this up with observed statistical properties of fluids, like the time average of the pressure in turbulent flow.

For the analytical nature of turbulent solutions, the work of Bass [**1, 2**] seems to be important.

In summary then, this view of turbulence may be phrased as follows. Our solutions for small μ (= Reynolds number in many fluid problems) are stable and as μ increases, these solutions become *unstable* at certain critical values of μ and the solution falls to a more complicated *stable* solution; eventually, after a certain finite number of such bifurcations, the solution falls to a strange attractor (in the space of all time dependent solutions to the problem). Such a solution, which is wandering close to a strange attractor, is called *turbulent.*

APPENDIX

The limit of zero viscosity in the presence of boundaries

In our work above we saw that we could, by a rescaling process, simulate large Reynolds numbers by taking ν small. Specifically, if $v(x, t)$ is a solution of the Navier–Stokes equations with viscosity ν, then $\bar{v}(x, t) = Vv(x, tV)$ is a solution with $\bar{\nu} = V\nu$. This rescaling argument also makes plausible the fact that for fixed initial data, the time of existence is $\geqq C\nu$; see p. 200.

In the ν-t plane then, our region of existence looks like a wedge:

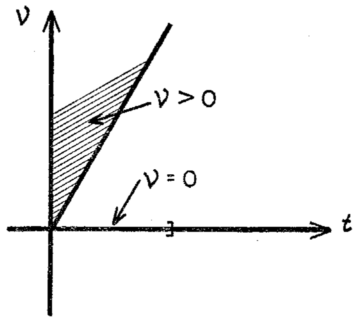


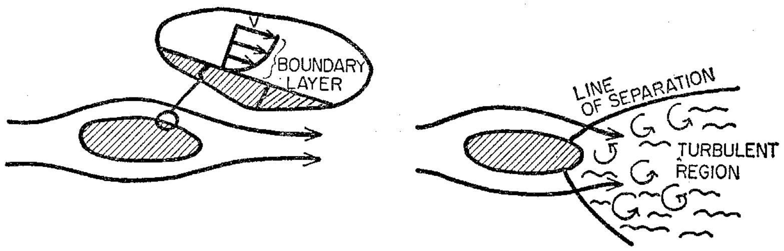
FIGURE 3.9

We saw above that if no boundaries are around, this region could be fattened out to a square and that we get nice convergence to Euler flow as $\nu \to 0$.

As we know (see p. 214), the situation is much different in case boundaries are present. Let us examine another situation. Consider flow around an obstacle as in Figures 3.10(a) and (b).

As $\mathscr{R}$ increases a "line of separation" develops and vorticity gets ripped off the boundary layer and moves downstream to form the turbulent wake. As before, this phenomenon is persistent as $\mathscr{R} \to \infty$, so we do not get convergence to the Euler flow. Not only that, but we can not get convergence to the Euler flow even on the boundary near the leading edge because of the difference in boundary conditions.

This line of separation is discussed in Landau–Lifschitz [1]. There they assume the line is sharp and that various discontinuities of the flow occur.



(a) Low Reynolds Number (b) High Reynolds Number
FIGURE 3.10

This is probably just in the limit $\mathscr{R} \to \infty$, or otherwise we would have a counter example to the global existence questions. This point seems to warrant further investigation. They also suggest that a line of separation must always form (Landau–Lifschitz [1], p. 155). We give a result below which seems to be support for this claim.

It is natural to inquire if we can obtain a time of existence independent of ν even though the flows might not converge. This seems to be a very difficult problem.

A word of caution: for low Reynolds numbers one observes steady flows with the velocities increasing as $\mathscr{R} \to \infty$. When simulated by $\nu \to 0$ and a fixed initial condition note that a time rescaling is also going on, so one must be careful when interpreting the results. Also note that for high Reynolds numbers, turbulent wakes are definitely time dependent flows, so the context of steady flows is certainly not appropriate.

THEOREM. *Let M be a bounded open set in $\mathbb{R}^3$ and let there be no external forces. For fixed initial data, let G denote the region of existence in the ν-t plane. Let* $-\operatorname{grad} p^\nu(x, t)$ *denote the pressure forces due to the inertial terms $\Delta_{v^\nu} v^\nu$, so as $v^\nu = 0$ on ∂M,* $\operatorname{grad} p$ *is tangent to ∂M, $\nu > 0$.*

Suppose that for each t_0, $\operatorname{grad} p^\nu(x, t_0)$ *is bounded in the $L^4_{1\frac{3}{4}}$ norm on ∂M as $\nu \to 0$, $\nu \in G$. (The $L^4_{1\frac{3}{4}}$ norm on ∂M corresponds to the L^4_2 norm on M, i.e., the L_4 norm of derivatives of order $\leqq 2$.)*

Then we have existence on a nontrivial t-interval independent of ν. Moreover v^ν converges to the Euler flow in C^1, uniformly for $t \in [0, T]$.

Actually what is interesting is the contrapositive theorem. Namely, we know that the conclusion is, in interesting cases, false. Therefore, the condition on the pressure must also fail. In particular the pressure condition can fail if, for example, the pressure forces develop a discontinuity, since the L^4_2 norm is at least as strong as the C^0 norm.

Thus, if we have any kind of reasonable boundary layer which prevents the flow v^ν from converging C^0, then the theorem states that the stated norm of the pressure forces must blow up as $\nu \to 0$ or else the time of existence $\to 0$ before ν reaches zero. Colloquially speaking, this means that if we have a boundary layer at all, then a line of separation must develop at very high Reynolds numbers.

The proof of the theorem is too involved to go into here. Suffice it to say that it relies on a coupling of the heat equation terms with the Euler terms via a "product type" formula as was discussed in § 1.8 above. To establish the appropriate estimates, one must pass to $\mathscr{D}^s_\mu$; i.e., Lagrangian coordinates.

As was pointed out by T. Kato, the situation in the case (M unbounded and) external forces are present seems to be different. Consider an infinite cylinder of fluid as shown in Figure 3.11.

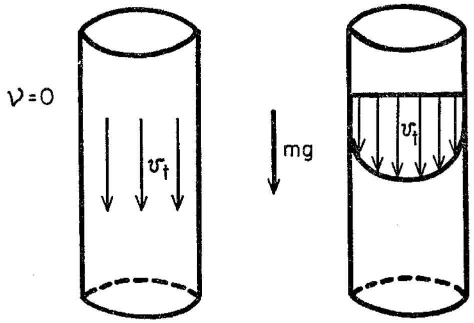


FIGURE 3.11

The fluid starts from rest ($v_0 = 0$) and under gravity, it accelerates down. For $\nu = 0$ it moves uniformly downwards, slipping by the boundary. For $\nu > 0$ a boundary layer develops (one can solve the equations explicitly here) so v^ν will not converge C^0 to v^0. But the pressure forces are zero and no line of separation develops.

III. DYNAMICS OF GENERAL RELATIVITY

In part two above, we saw how the diffeomorphism groups play a fundamental role in hydrodynamics. These groups also enter into the formulation of other theories such as elasticity (this was pointed out to us by H. Cohen; cf. Chernoff–Marsden [1]). In this part we shall investigate general relativity and shall see how the group of diffeomorphisms again enters. In the present context it gives us a clearer understanding of the "coordinate invariance" of the theory.

In hydrodynamics, the idea of a Hamiltonian system in its special case of the geodesic flow also played a key role. For relativity we shall be primarily concerned with geodesics in the presence of a potential term. As before, we are dealing with an infinite dimensional system. We shall begin in § 1 with the basic ideas about Hamiltonian systems. After some motivation for general relativity we shall then apply this Hamiltonian formalism to study the geodesic motion on the space $\mathscr{M}$ of riemannian metrics and its connection with the usual Einstein equations $\mathrm{Ric} = 0$.

After this we shall want to see how the evolution, or Hamiltonian picture changes when we make a space-time coordinate transformation. This is done by introducing what are called the "lapse" and "shift" functions of Wheeler [1]. We shall also investigate the geometry of these quantities. Some additional theorems related to the lapse and shift functions will also be presented.

The basic references for this work are Arnowitt-Deser-Misner [1], Wheeler [1], Dewitt [1], and Fischer–Marsden [1]. The approach in these papers is based on the Lagrangian formalism. However one can also base the development on the Hamilton–Jacobi theory; cf. Gerlach [1].

1. Infinite dimensional Hamiltonian systems

Before beginning relativity, we shall first develop some preliminary machinery. This machinery is a fairly straight forward generalization to infinite dimensions of finite dimensional mechanics; see Abraham [2]. Since we are not presupposing a detailed knowledge of that material, we make the exposition as self contained as possible. The treatment of Lagrangian systems is fairly complete and also allows the degenerate case for later purposes.

We shall begin with the special case of the spray of a metric and motion in a potential.

1.1 The spray of a metric

Let M be an infinite dimensional manifold (typically a manifold of maps) with a symmetric, bilinear form B defined on each tangent space T_xM. Recall that B induces in a natural way, a map $B: T_xM \to T_x^*M$ at each $x \in M$: if $v \in T_xM$, $(B \cdot v)(w)$ is defined to be $B(v, w)$ for any $w \in T_xM$.

The form B is said to be *weakly nondegenerate* if the map $v \mapsto Bv$ is injective at each $x \in M$; this is clearly equivalent to the statement: $B(v, w) = 0$ for all w implies $v = 0$. If the map $v \mapsto Bv$ is an isomorphism, B is said to be *strongly nondegenerate*. The two notions coincide in the finite dimensional case. If B is weakly nondegenerate, (M, B) is called a *weak pseudo-Riemannian manifold.*

It is convenient to denote B by $\langle\ ,\ \rangle_x$ on the fiber over $x \in M$. We now wish to define the *spray* S of the metric $\langle\ ,\ \rangle_x$. This should be a vector field on TM; $S: TM \to T^2M$ whose integral curves project onto geodesics. Locally, if $(x, v) \in T_xM$, write $S(x, v) = ((x, v), (v, \gamma(x, v))$. If M is finite dimensional, the geodesic spray is given by putting $\gamma^i(x, v) = -\Gamma^i_{jk}(x)v^jv^k$. In the general case, the correct definition for γ is

(1) $$\langle\gamma(x, v), w\rangle_x \equiv \tfrac{1}{2}D_x\langle v, v\rangle_x \cdot w - D_x\langle v, w\rangle_x \cdot v$$

where $D_x\langle v, v\rangle_x \cdot w$ means the derivative of $\langle v, v\rangle_x$ with respect to x in the direction of w. In the finite-dimensional case, the right hand side of (1) is given by

$$\frac{1}{2}\frac{\partial g_{ij}}{\partial x^k} v^iv^jw^k - \frac{\partial g_{ij}}{\partial x^k} v^iw^jv^k,$$

which is the same as $-\Gamma^i_{jk}v^jv^kw_i$. So with this definition of γ, S is taken to be the spray. The verification that S is well-defined independent of the charts is not too difficult. Notice that γ is quadratic in v. One can also show that S is just the Hamiltonian vector field on TM associated with the kinetic energy $\frac{1}{2}\langle v, v\rangle$. This will be done in § 1.4 below; cf. Abraham [2] and Chernoff–Marsden [1].

The point is that the definition of γ in (1) makes sense in the infinite as well as the finite dimensional case, whereas the usual definition of Γ^i_{jk} makes sense only in finite dimensions. This then gives us a way to deal with geodesics in infinite dimensional spaces.

1.2 Equations of motion in a potential

Let $t \mapsto (x(t), v(t))$ be an integral curve of S. That is:

(2) $$\dot{x}(t) = v(t); \qquad \dot{v}(t) = \gamma(x(t), v(t)).$$

These are the equations of motion in the absence of a potential. Now let $V: M \to \mathbb{R}$ (the potential energy) be given. At each x, we have the differential of V, $dV(x) \in T_x^*M$, and we define grad $V(x)$ by:

(3) $$\langle \mathrm{grad}\ V(x), w\rangle_x \equiv dV(x) \cdot w.$$

It is a definite assumption that grad V exists, since the map $T_xM \to T_x^*M$ induced by the metric is not necessarily bijective.

The equation of motion in the potential field is given by:

(4) $$\dot{x}(t) = v(t); \qquad \dot{v}(t) = \gamma(x(t), v(t)) - \mathrm{grad}\ V(x(t)).$$

The *total energy*, kinetic plus potential, is given by $H(x, v) = \frac{1}{2}\langle v_x, v_x\rangle + V(x)$. It is easily checked that H is constant along curves $(x(t), v(t))$ satisfying (4), i.e., along integral curves of the vector field Z which, in coordinates, is given by $Z(x, v) = ((x, v), (v, \gamma - \text{grad } V))$.

1.3 Example: The wave equation

A solution to the three-dimensional wave equation is, by definition, a function $f(x, t)$, $x \in \mathbb{R}^3$, $t \in \mathbb{R}$, satisfying $\partial^2 f(x, t)/\partial t^2 = \nabla^2 f$. We can show these to be the equations of motion in a potential field on the appropriate space M. Choose $M = L_2(\mathbb{R}^3, \mathbb{R})$; M is a Hilbert space with a strong metric given by the usual inner product $\langle f, g\rangle = \int fg\, dx$. This makes sense in our context since M is a linear space and so may be identified with its tangent space. Note that $\gamma = 0$, since the right side of (1) vanishes identically. The geodesics are therefore straight lines. Now let the potential function be defined by

$$V(f) = \frac{1}{2}\int \|\nabla f\|^2\, dx$$

We are proceeding formally at the moment, since V is really defined only on $H^1(\mathbb{R}^3, \mathbb{R})$. (See Part I, § 1 for the definition of the Sobolev space H^1.) The equations of motion (4) become $\ddot{f} = -\text{grad } V(f)$, since $\gamma = 0$. We compute grad V as follows: $\langle \text{grad } V(f), h\rangle = dV(f)\cdot h = \int_{\mathbb{R}^3} \nabla f \cdot \nabla h = -\int (\nabla^2 f)(h) = -\langle \nabla^2 f, h\rangle$, where we have integrated by parts. Thus grad $V(f) = -\nabla^2 f$, and we get $\ddot{f} = \nabla^2 f$, which is the wave equation. Since two derivatives have been lost in this process, the gradient will not be everywhere defined. Nevertheless, one can show the existence of a global flow (producing "curves of functions") which is the solution to the wave equation; cf. Yosida **[1]**, and 2.6 below. The total energy is given by:

$$H(f, \dot{f}) = \tfrac{1}{2}\langle \dot{f}, \dot{f}\rangle + \tfrac{1}{2}\langle \nabla f, \nabla f\rangle$$
$$H = \text{K.E.} + \text{P.E.}$$

Note that the vector field associated to the wave equation is only densely defined, so one has to be somewhat careful with derivatives and flows. For our purposes we can ignore these technicalities which, fortunately, do not affect understanding the basic ideas in what follows. (See Chernoff-Marsden **[1]** for a complete discussion of the technical points just alluded to.)

1.4 Lagrangian systems

We now want to generalize the idea of motion in a potential to the idea of a Lagrangian system, which is a special case of a Hamiltonian system.

See Abraham **[2]** for an alternative exposition of the finite dimensional case, and Marsden **[1]**, and Chernoff-Marsden **[1]** for additional results.

We begin with a manifold M and a given function $L: TM \to \mathbb{R}$ called the *Lagrangian*. (Here M is, in general infinite dimensional not to be confused with the notation M which will be used below for a space like hypersurface.) In case of motion in a potential, one takes

$$L(v_x) = \tfrac{1}{2}\langle v_x, v_x\rangle - V(x)$$

which differs from the energy in that we use $-V$ rather than $+V$.

Now L defines a map, called the *fiber derivative*, $FL: TM \to T^*M$ as follows: let $v, w \in T_xM$. Then

$$FL(v)\cdot w \equiv \frac{d}{dt} L(v + tw)|_{t=0}$$

That is, $FL(V)\cdot w$ is the derivative of L along the fiber in direction w.

In case of $L(v_x) = \frac{1}{2}\langle v_x, v_x\rangle_x - V(x)$, we see that $FL(v_x)\cdot w_x = \langle v_x, w_x\rangle_x$ so we recover the usual map of $TM \to T^*M$ associated with the bilinear form $\langle\ ,\ \rangle_x$.

Now T^*M carries a canonical two form ω called the *canonical sympletic form* (see Abraham **[2]**, Lang **[1]**). Using FL one obtains a closed two form ω_L (i.e., $d\omega_L = 0$) on TM by pulling back the canonical symplectic structure on T^*M, $\omega_L = (FL)^*\omega$. All we shall require is the following local formula for ω_L: if M is modeled on a linear space $\mathbb{E}$, so locally TM looks like $U \times \mathbb{E}$ where $U \subset \mathbb{E}$ is open, then $\omega_L(u, e)$ for $(u, e) \in U \times \mathbb{E}$ is the skew symmetric bilinear form on $\mathbb{E} \times \mathbb{E}$ given by

$$2\omega_L(u, e)\cdot((e_1, e_2), (e_3, e_4)) = D_1(D_2L(u, e)\cdot e_3)\cdot e_1$$
$$- D_1(D_2L(u, e)\cdot e_3)\cdot e_1 + D_2D_2L(u, e)\cdot e_4\cdot e_1 - D_2D_2L(u, e)\cdot e_2\cdot e_3$$

where D_1, D_2 denote the indicated partial derivatives of L. Rather than going through the (routine) derivation of this formula from that of ω, let us just adopt this as our basic two form.

It is easy to see that ω_L is (weakly) nondegenerate iff $D_2D_2L(u, e)$ is nondegenerate. But we want to also allow degenerate cases for later purposes. In case of motion in a potential, nondegeneracy of ω_L amounts to nondegeneracy of the metric $\langle\ ,\ \rangle_x$. The *action* of L is defined by $A: TB \to \mathbb{R}$, $A(v) = FL(v)\cdot v$, and the *energy* of L is $E = A - L$. In charts,

$$E(u, e) = D_2L(u, e)\cdot e - L(u, e)$$

and in finite dimensions it is the expression

$$E(q, \dot{q}) = \frac{\partial L}{\partial \dot{q}^i}\dot{q}^i - L(q, \dot{q}),$$

(summation convention!)

Now given L, we say that a vector field Z on TM is a *Lagrangian vector field or a Lagrangian system* for L if the *Lagrangian condition* holds:

$$2\omega_L(v)(Z(v), w) = dE(v)\cdot w$$

for all $v \in T_b M$, and $w \in T_v(TM)$. Here, dE denotes the differential of E. Below we shall see that for motion in a potential, this leads to the same equations of motion which we found above.

If ω_L were a weak symplectic form there would be at most one such Z. The fact that ω_L may be degenerate however means that Z is not uniquely determined by L so that there is some arbitrariness in what we may choose for Z. Also if ω_L is degenerate, Z may not even exist. If it does, we say that we can define *consistent equations of motion*. These ideas have been discussed in the finite dimensional case by Dirac **[1]** and Kunzle **[1]**.

The dynamics is obtained by finding the integral curves of Z; that is the curves $v(t)$ such that $v(t) \in TM$ satisfies $(dv/dt)(t) = Z(v(t))$. From the Lagrangian condition it is trivial to check that energy is conserved even though L may be degenerate:

1.4.1 PROPOSITION. *Let Z be a Lagrangian vector field for L and let $v(t) \in TM$ be an integral curve of Z. Then $E(v(t))$ is constant in t.*

PROOF. By the chain rule,

$$\frac{d}{dt}E(v(t)) = dE(v(t))\cdot v'(t) = dE(v(t))\cdot Z(v(t))$$
$$= 2\omega_L(v(t))(Z(v(t)), Z(v(t)) = 0$$

by the skew symmetry of ω_L. ■

We now want to generalize our previous local expression for the spray of a metric, and the equations of motion in the presence of a potential. In the general case the equations are called "Lagrange's equations".

1.4.2 PROPOSITION. *Let Z be a Lagrangian system for L and suppose Z is a second order equation (that is, in a chart $U \times \mathbb{E}$ for TM, $Z(u, e) = (e, Z_2(u, e))$ for some map $Z_2: U \times \mathbb{E} \to \mathbb{E}$). Then in the chart $U \times \mathbb{E}$, an integral curve $(u(t), v(t)) \in U \times \mathbb{E}$ of Z satisfies Lagrange's equations:*

$$\text{(1)}\qquad \begin{cases} \dfrac{du}{dt}(t) = v(t) \\[2ex] \dfrac{d}{dt}(D_2L(u(t), v(t))\cdot w = D_1L(u(t), v(t))\cdot w \end{cases}$$

for all $w \in \mathbb{E}$. In case L is nondegenerate we have

$$\text{(2)}\qquad \frac{dv}{dt} = \{D_2D_2L(u, v)\}^{-1}\{D_1L(u, v) - D_1D_2L(u, v)\cdot v\}.$$

In case of motion in a potential, (2) reduces readily to the equations we found previously defining the spray and gradient.

PROOF OF 1.4.2. From the definition of the energy E we have $dE(u, e)\cdot(e_1, e_2) = D_1(D_2L(u, e)\cdot e_1 + D_2D_2L(u, e)\cdot e\cdot e_2 - D_1L(u, e)\cdot e_1$. Locally we may write $Z(u, e) = (e, Y(u, e))$ as Z is a second order equation. Using the formula for ω_L, the condition on Y may be written, after a short computation:

$$D_1L(u, e)\cdot e_1 = D_1(D_2L(u, e)\cdot e_1)\cdot e$$
$$+ D_2(D_2L(u, e)\cdot Y(u, e))\cdot e_1 \quad \text{for all} \quad e_1 \in \mathbb{E}.$$

This is the formula (2) above. Then, if $(u(t), v(t))$ is an integral curve of Z we obtain, using dots to denote time differentiation,

$$D_1L(u, e)\cdot e_1 = D_1(D_2L(u, \dot{u})\cdot e_1\cdot \dot{u} + D_2D_2L(u, \dot{u})\cdot \ddot{u}\cdot e_1$$
$$= \frac{d}{dt}D_2L(u, \dot{u})\cdot e_1$$

by the chain rule ■

We remark that if ω_L were nondegenerate Z would automatically be a second order equation (cf. Abraham **[2]**). Also, the condition of being second order is intrinsic; Z is second order if $T\pi \circ Z =$ identity, where $\pi: TM \to M$ is the projection. See Abraham **[2]**, or Lang **[1]**.

Often L is obtained in the form

$$L(u, \dot{u}) = \int \mathscr{L}\left(u, \frac{\partial u}{\partial x^k}, \dot{u}\right) d\mu$$

for a Lagrangian density $\mathscr{L}$ and μ some volume element on some manifold Q. Then M is a space of functions on Q or more generally sections of a vector bundle over Q. In this case, Lagrange's equations may be converted to the usual form of Lagrange's equations for a density $\mathscr{L}$. We shall see how to do this below in a special case.

1.5 Conservation theorems

We now wish to formulate some basic conservation laws which hold in the presence of a group of symmetries. We first state the laws in the case of motion in a potential.

Let M be a pseudo-riemannian manifold acted on by a Lie group G—i.e., for each $a \in G$, there is a diffeomorphism $\Phi_a: M \to M$, such that $\Phi_a \circ \Phi_b = \Phi_{ab}$ and $\Phi_e =$ identity. Let $v \in \mathscr{I}$, the Lie algebra of G. Then v determines the 1-parameter subgroup $\exp(tv)$ and thus a 1-parameter group of diffeomorphisms $\Phi_{\exp(tv)}$ of M. Let $X(m) = (d/dt)\Phi_{\exp(tv)}(m)|_{t=0}$. The vector field X is called an *infinitesimal generator* of the action. For example the infinitesimal generators of $SO(3)$ acting on $\mathbb{R}^3$ are the usual vector fields corresponding to rotations about an axis.

Suppose that a potential function $V: M \to \mathbb{R}$ is given and suppose further that:

(a) Φ_a is an isometry for all $a \in G$
(b) V is invariant under the action of $G: V \circ \Phi_a = V, a \in G$.

Given this, there exist certain functions which remain constant on the integral curves of the equations of motion. Namely for each infinitesimal generator X,

(1) $\quad P(X): TM \to \mathbb{R} \quad$ defined by $\quad P(X)(v_x) = \langle v_x, X(x)\rangle_x$

is such a function. We call $P(X)$ the *momentum function* for X. (P can be regarded as a homomorphism from $\mathscr{I}$ to the Lie algebra of functions on TM under the Poisson bracket.)

We shall now prove a more general theorem in the context of Lagrangian systems. Namely, consider a general Lagrangian $L: TM \to \mathbb{R}$. The assumption in this case is that $T\Phi_a: TM \to TM$ leaves L invariant. The constants of the motion are now defined by

(2) $$P(X): v \mapsto FL(v) \cdot x$$

where $FL(v) \cdot x$ is the fiber derivative, defined above. If $L = K - V$, (2) reduces to (1) immediately.

In this result, observe that we do allow for the possibility that L might be degenerate. The only special assumption needed on Z is that it exist and be second order.

1.5.1 PROPOSITION. *Let Z be a Lagrangian vector field for $L: TM \to \mathbb{R}$ and suppose Z is a second order equation.*

Let Φ_t be a one parameter group of diffeomorphisms of M generated by the vector field $Y: M \to TM$. Suppose that for each real number t, $L \circ T\Phi_t = L$. Then the function $P(Y): TM \to \mathbb{R}$, $P(Y)(v) = FL(v) \cdot Y$ is constant along integral curves of Z.

PROOF. Let $v(t)$ be an integral curve for Z. Then we shall show that $(d/dt)\{P(Y)(v(t))\} = 0$. Indeed, in a coordinate chart, if $(u(t), v(t))$ is the integral curve,

$$\begin{aligned}\frac{d}{dt}\{FL(v(t)) \cdot Y\} &= \frac{d}{dt}\{D_2L(u(t), v(t)) \cdot Y(u(t))\} \\ &= D_1D_2L(u(t), v(t)) \cdot Y(u(t)) \cdot \dot{u}(t) + D_2D_2L(u(t), v(t)) \\ &\quad \cdot Y(u(t)) \cdot \dot{v}(t) + D_2L(u(t), v(t)) \cdot DY(u(t)) \cdot \dot{u}(t).\end{aligned}$$

Now the condition that Z be the Lagrangian vector field of L means exactly that the first two terms equal $D_1L(u(t), v(t)) \cdot Y(u(t))$ (see the proof of 1.4.2 above). However if we differentiate $L \circ T\Phi_t$ with respect to t we obtain for any point (u, v),

$$\begin{aligned}0 &= \frac{d}{dt} L(\Phi_t(u), D\Phi_t(u) \cdot v)|_{t=0} \\ &= D_1L(u, v) \cdot Y(u) + D_2L(u, v) \cdot DY(u) \cdot v.\end{aligned}$$

Comparing this with the above gives $(d/dt)\{FL(v) \cdot Y\} = 0$ and proves the assertion. ■

2. The basic equations of general relativity; dynamics on $\mathscr{M}$

In this section, we develop the dynamics on the space $\mathscr{M}$ of riemannian metrics on a fixed three manifold M and its relationship with the (exterior) Einstein field equations, namely that the resulting four dimensional metric have zero Ricci curvature.

We shall be concentrating in this section on the simplest possible case in which the four metric ${}^4g$ and the three metric ${}^3g$ are related by

$${}^4g_{\alpha\beta}\, dx^\alpha\, dx^\beta = -dt^2 + g_{ij}\, dx^i\, dx^j.$$

Here $x^\alpha = (t, x^i)$ and our conventions are that ${}^4g$ has signature $(-, +, +, +)$. In the next section we shall consider the case of a general ${}^4g$.

We begin with some motivation and explanation of the usual Einstein field equations.

2.1 General relativity; introductory remarks

Einstein was led to the general theory of relativity by his conviction that gravitational forces should be incorporated into the geometry of spacetime in a "natural" way—"natural" meaning natural in a geometrical sense. The idea is that a body moving under the force of gravitation alone (e.g., a satellite circling the earth) should travel along a geodesic in an appropriate differentiable manifold. Such a manifold is certainly not flat 3-space, since the motion of a satellite would not then be geodesic. It is also easy to see that the manifold cannot be a curved three-dimensional Riemannian space: consider the case of two projectiles P_1, P_2 launched at the same time from A with trajectories as indicated in Figure 2.1, both passing through B (this is easily arranged). It is clear that not both P_1 and P_2 can be geodesic with respect to any 3-space metric; since B can be moved arbitrarily close to A, there are no normal neighborhoods of A (in which there are *unique* minimizing geodesics).

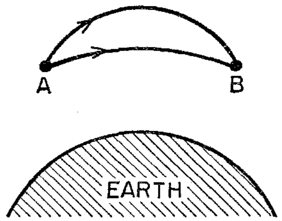


FIGURE 2.1

On the other hand, we *do* get unique trajectories if we require the projectile to pass through B at a given time. So we are compelled to consider a manifold of dimension at least four. Finally, it is almost obvious that this 4-manifold cannot be Riemannian (metric tensor positive definite): Riemannian manifolds are isotropic in the sense that there are no intrinsically defined, distinguished directions. But space time is not isotropic; for example the geodesic connecting (you, now) with (Sirius, 1 second later) could not be traversed by a material particle required to travel at a speed below that of light. One has to distinguish between possible particle trajectories (timelike curves), impossible particle trajectories (spacelike curves), and possible photon trajectories (null curves).

All in all, one is led to consider a four-dimensional pseudo-Riemannian manifold whose metric tensor g has the signature $\{-, +, +, +\}$. This is quite natural since it tells us that locally (in the tangent space, or in a normal neighborhood) the universe looks like Minkowski-space. In this manifold, the "world line" or space time trajectory of a freely falling particle is a geodesic. Furthermore, it is assumed that this geodesic does not depend on the mass of the particle (an orange and a grapefruit behave the same way in the same gravitational field). This is called the *principle of equivalence.*

No matter what the gravitational field, the space-time is required to satisfy the *Einstein field equations* Ric $= 0$ (or $R_{\alpha\beta} = 0$) in the absence of

matter, or in regions exterior to sources. Of course Ric denotes the Ricci tensor of our metric.

We would like to now present the geometrical meaning of the field equations Ric $= 0$. This can be used as motivation for the equations. Compare Pirani [1] and the Rund-Lovelock lectures in this seminar. The following exposition was provided by Dan Norman.

Let u be the tangent to a timelike geodesic $x(t)$ (timelike means $\langle u, u\rangle < 0$), so $\nabla_u u = 0$. Consider the Jacobi field (or deviation vector) η along $x(t)$; it satisfies Jacobi's equation:

$$\nabla_u \nabla_u \eta + R(\eta, u)u = 0$$

where R is the curvature tensor. Regarded as a map $\tilde{R}_u$ in η, Ric(u, u) is its trace.

We are supposing Ric $= 0$. Let e_i, $i = 1, 2, 3$ be vectors orthogonal to u at a point p where $t = 0$. Then extend e_i to be Jacobi fields with initial condition $\nabla_u e_i = 0$ at p. Then $\nabla_u(u \cdot e_i) = \nabla_u u \cdot e_i + u \cdot \nabla_u e_i = u \cdot \nabla_u e_i$ so $\nabla_u \nabla_u (u \cdot e_i) = u \cdot \nabla_u \nabla_u e_i = -u \cdot R(e_i, u)u = 0$ (we have $\langle R(v, u)u, u\rangle = 0$ always by skew symmetry of $\langle R(u, v)w, z\rangle$ in w, z). Hence $u \cdot e_i = 0$ for all time. Choose e_i to be eigenvectors of $\tilde{R}_u$ on the space orthogonal to u. We denote this restriction by $\tilde{R}_u^\perp$. Thus $\tilde{R}_u^\perp e_i = \lambda_i e_i$, and $\lambda_1 + \lambda_2 + \lambda_3 = 0$ because $\tilde{R}_u u = 0$ and the trace of the symmetric operator $\tilde{R}_u$ is zero.

Now these vectors e_i span a three volume. Multiply e_i by ε (so that we can be sure exp maps the field εe_i onto geodesics close to the geodesic through p). A satisfactory approximation of the volume of the cube spanned by these vectors is (vol) $= \varepsilon^3 e_1 \wedge e_2 \wedge e_3$. We compute:

$${}^4 g_{\alpha\beta}\, dx^\alpha\, dx^\beta = -dt^2 + g_{ij}\, dx^i\, dx^j.$$

$$\begin{aligned}\frac{d^2}{ds^2}(\mathrm{vol})\Big|_p &= \nabla_u \nabla_u (\mathrm{vol})\Big|_p \\ &= \varepsilon^3\{(\nabla_u\nabla_u e_1) \wedge e_2 \wedge e_3 + e_1 \wedge (\nabla_u \nabla_u e_2) \wedge e_3 + e_1 \wedge e_2 \wedge \nabla_u \nabla_u e_3 \\ &\quad + \text{First derivative terms}\}\Big|_p.\end{aligned}$$

Since the e_i's are Jacobi fields and eigenvectors (at p) of $\tilde{R}_u^\perp$, and since $\nabla_u e_i|_p = 0$, we have

$$(1) \qquad \frac{d^2}{dt^2}(\mathrm{vol})\Big|_p = -(\lambda_1 + \lambda_2 + \lambda_3)(\mathrm{vol})\Big|_p = 0 \quad \text{if} \quad \frac{d(\mathrm{vol})}{dt}\bigg|_p = 0$$

as the condition equivalent to Ric $= 0$.

We can interpret this more physically as follows. Imagine ourselves in a freely falling elevator and watch a collection of freely falling particles. The particles are initially at rest with respect to each other, but due to motion towards the earth's center, they will pick up a relative motion (see Figure 2.2).

The condition (1) says that the 3-volume (up to second order) is remaining constant during the motion. This geometric property is directly verifiable in the case of the Newtonian gravitational field, so is a reasonable candidate for generalization. Thus we shall adopt Ric $= 0$ as the Einstein field equations in our Lorentz four manifold.

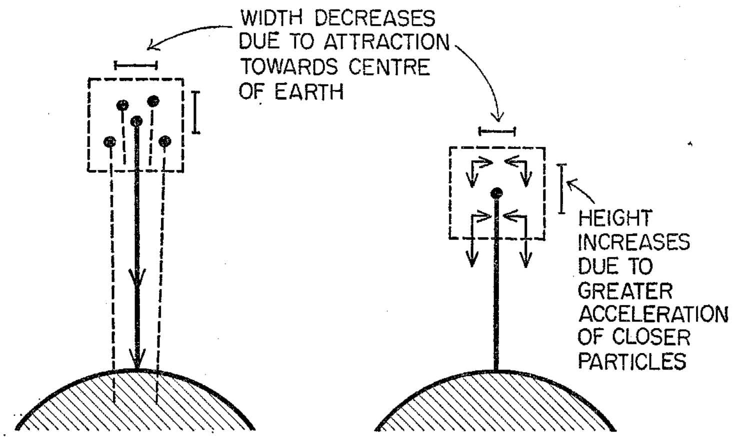


FIGURE 2.2

2.2 The general program

Let V be a spacetime with M a three-dimensional spacelike section without boundary (a spacelike section is a submanifold such that for $0 \neq v \in T_xM$, $\langle v, v\rangle > 0$). Assume for the moment that M is compact, so that there exists a neighborhood U of M in which the timelike geodesics (that is geodesics whose tangent vectors v have $\langle v, v\rangle < 0$) orthogonal to M have no focal points. If we let t measure proper time on these geodesics, with $t = 0$ on $M = M_0$, then the function t is well-defined in U. The surfaces M_t given by $t =$ constant form a one parameter family of space sections, all diffeomorphic to M. Let g_t be the induced Riemannian metric on M_t. Via the aformentioned diffeomorphism, we can regard g_t as a curve in the space of positive definite metrics on M. The fact that V is Ricci flat implies that g_t satisfies certain differential equations. We want to work these out.

We also want to go the other way: given M, a positive definite metric g_0, and a symmetric tensor $k_0 = \dot{g}_0$ (the second fundamental form of M in V) we want to find the curve g_t describing the time evolution of the geometry of M, and then to paste together the resulting 3-manifolds M_t to obtain a piece of spacetime.

2.3 The space of Riemannian metrics and the DeWitt metric

Fix M, a compact, oriented 3-manifold. (It is actually not necessary to assume M compact; the noncompact case is handled by imposing certain asymptotic conditions at infinity.) Let $S_2(M)$ be the linear space of (C^∞) symmetric twice covariant tensor fields on M. The set $\mathcal{M} \subset S_2(M)$ of all positive definite metrics forms an open cone in $S_2(M)$—i.e., $\mathcal{M}$ is invariant under addition and multiplication by positive scalars. The same sorts of technicalities encountered in hydrodynamics arise here, and so one should properly work with the set $\mathcal{M}^s$, the metrics in the Sobolev class H^s; but for simplicity, most of the development will be done directly in $\mathcal{M}$. Since $\mathcal{M}$ is an open cone in $S_2(M)$ (in the C^∞ or H^s topology), we know that $T\mathcal{M} = \mathcal{M} \times S_2$.

We want to construct a Hamiltonian system on $\mathcal{M}$, and the first order of business is to produce a metric—i.e., for each $g \in \mathcal{M}$, we need a symmetric bilinear map $\mathcal{G}_g : S_2 \times S_2 \to \mathbb{R}$. We define $\mathcal{G}$ by:

$$\mathcal{G}_g:(h, k) \mapsto \int_M [h \cdot k - \operatorname{tr}(h)\operatorname{tr}(k)]\, d\mu_g \tag{1}$$

where $h \cdot k = h^{ij}k_{ij}$, $\operatorname{tr}(h) = h^i_i$, all indices being raised with g and μ_g is the volume element associated with g, $\mu_g = \sqrt{\det g_{ij}}\, dx^1 \wedge dx^2 \wedge dx^3$. Observe that $\mathcal{G}$ is symmetric, bilinear, and depends on g; it is called the *DeWitt metric.* $\mathcal{G}$ is not positive definite because of the minus sign, but it *is* nondegenerate:

2.3.1 PROPOSITION. *$\mathcal{G}$ is a weak metric.*

PROOF. Assume that for some $g \in \mathcal{M}$, $\mathcal{G}_g(h, k) = 0$ for all $k \in S_2(M)$. Put $\bar{k} = h - \frac{1}{2}\operatorname{tr}(h)g \in S_2$. Then one sees that $\mathcal{G}_g(h, \bar{k}) = \int_M h \cdot h\, d\mu_g = 0$ so $h = 0$. ∎

Observe that as was the case with hydrodynamics, $\mathcal{G}_g$ is not a strong metric in any of the topologies of interest on $\mathcal{M}$ (such as the C^∞ or H^s topologies).

Next, we wish to compute the spray of $\mathcal{G}$; this is nontrivial because of the dependence on g.

2.3.2 PROPOSITION. *The spray of $\mathscr{G}$ is the map $\tilde{S}: \mathscr{M} \times S_2 \to (\mathscr{M} \times S_2) \times (S_2 \times S_2)$ whose last component is given by*

$$S(g, k) = k \times k - \tfrac{1}{2}\mathrm{tr}(k)k - \tfrac{1}{8}[k \cdot k - [\mathrm{tr}(k)]^2]g \tag{2}$$

where, in coordinates, $k \times k = k_{il}k^l_j$.

To prove this, we must verify that S satisfies the equation (1) of § 1.1 which in this case, reads:

$$\mathscr{G}_g(S(g, k), h) = \tfrac{1}{2}D_g\mathscr{G}_g(k, k) \cdot h - D_g\mathscr{G}_g(k, h) \cdot k. \tag{3}$$

And to verify this, we need to know how to take the appropriate derivatives. This is done in the next lemma.

2.3.3 LEMMA. (a) *The derivative of the map $g \mapsto \mu_g$ in the direction $h \in S_2$ is $\frac{1}{2}\mathrm{tr}(h)\mu_g$.*

PROOF. Let $g(t) = g + th$. We want $(d/dt)\mu_g(t)|_{t=0}$, which we can get once we know $(d/dt)\det(g_{ik} + th_{ik})|_{t=0}$. We have $(g_{ik} + th_{ik}) = (g_{ij}) \times (\delta^j_k + th^j_k)$. So $(d/dt)\det(g_{ik} + th_{ik})|_{t=0} = \det(g_{ij})\{(d/dt)\det(\delta^j_k + th^j_k)|_{t=0}\} = \det(g_{ij})\mathrm{tr}(h_{ij})$. The lemma follows immediately. ■

(b) *For fixed h, k, the derivative of $g \mapsto h \cdot k$ in the direction h_1, is $-2h_1 \cdot (h \times k)$.*

PROOF. Note that we can regard $h \in S_2(M)$ as a map $T_xM \to T^*_xM$ so $g^{-1}h$ is a map of T_xM to T_xM. In coordinates, $g^{-1}h$ is h^i_j, and g^{-1} is g^{ij}. Then $h \cdot k = \mathrm{tr}(g^{-1}hg^{-1}k)$. For $g(t) = g + th_1$, we have

$$(d/dt)g^{-1}(t)|_{t=0} = -g^{-1}h_1g^{-1}.$$

Then $(d/dt)(h \cdot k)|_{t=0} = -\mathrm{tr}(g^{-1}h_1g^{-1}hg^{-1}k) - \mathrm{tr}(g^{-1}hg^{-1}h_1g^{-1}k) = -2h_1 \cdot (h \times k)$. ■

(c) *The derivative of the map $h \mapsto \mathrm{tr}(h)$ in the direction h_1 is $-h_1 \cdot h$.*

PROOF. By definition, $\mathrm{tr}(h)$ is the trace of the linear transformation $g^{-1}h$ of T_xM to itself. If $g(t) = g + th$, then at $t = 0$ the derivative is $-\mathrm{tr}(g^{-1}h_1g^{-1}h) = -h_1 \cdot h$, as in (b). ■

Now to prove the proposition, it is a straightforward job to plug everything in to both sides of (3) and verify the equality. Namely from

(a)–(c), we get

$$D_g\mathscr{G}_g(k, h) \cdot h_1 = \int_M (-2h_1 \cdot (k \times h) + (\mathrm{tr}\, h)h_1 \cdot k + (\mathrm{tr}\, k)h_1 \cdot h)\, d\mu_g$$
$$+ \int_M [h \cdot k - (\mathrm{tr}\, h)(\mathrm{tr}\, k)] \cdot \tfrac{1}{2}\,\mathrm{tr}\, h_1\, d\mu_g.$$

Thus the right hand side of (3) becomes

$$\int_M h \cdot (k \times k) - \tfrac{1}{2}h \cdot k(\mathrm{tr}\, k) - (\mathrm{tr}\, h)k \cdot k$$
$$+ \tfrac{1}{2}(\mathrm{tr}\, k)^2\,\mathrm{tr}\, h + \tfrac{1}{4}[k \cdot k - (\mathrm{tr}\, k)^2]\mathrm{tr}\, h\, d\mu_g$$

while the left side is $\int_M S_g(k) \cdot h - \mathrm{tr}(S_gk)\mathrm{tr}(h)\, d\mu_g$ which becomes, on substituting the stated expression for S_g, using $g \cdot h = \mathrm{tr}\, h$, $\mathrm{tr}\, g = 3$, and $\mathrm{tr}(k \times k) = k \cdot k$,

$$\int_M h \cdot (k \times k) - \tfrac{1}{2}(h \cdot k)\mathrm{tr}\, k - \tfrac{1}{8}\{k \cdot k\,\mathrm{tr}\, h - (\mathrm{tr}\, k)^2\mathrm{tr}\, h\}\, d\mu_g$$
$$- \int_M k \cdot k\,\mathrm{tr}\, h - \tfrac{1}{2}(\mathrm{tr}\, k)^2\,\mathrm{tr}\, h - \tfrac{1}{8}\{k \cdot k \cdot 3\,\mathrm{tr}\, h - 3(\mathrm{tr}\, k)^2\,\mathrm{tr}\, h\}\, d\mu_g$$

which equals the right side of (3) obtained above. ■

2.4 The potential

Now that we have found the spray of the DeWitt metric, the next project is to add on a potential. It turns out that the appropriate object is obtained from the scalar curvature $r(g) \equiv {}^3r(g)$ of the metric g on the 3-manifold M. We define $V: \mathscr{M} \to \mathbb{R}$ by

$$V: g \mapsto -2\int_M r(g)\, d\mu_g. \tag{1}$$

(The factor -2 is conventional here.) To find grad V, we first need to know the derivative of $r(g)$:

2.4.1 LEMMA.

$$D_gr(g) \cdot h = \Delta(\mathrm{tr}\, h) + \delta\delta h - \mathrm{Ric}(g) \cdot h, \tag{2}$$

where $\delta\delta h$ is the double covariant divergence $h^{ij}{}_{|i|j}$ and Δ is the Laplace Beltrami operator on scalars, $\Delta\varphi = -g^{ij}\varphi_{|i|j}$.

This result with further references may be found in Deser [1] and Berger-Ebin [1]. It is straightforward but long, and we shall omit it.

2.4.2 PROPOSITION.

(3) $$\operatorname{grad} V(g) = 2 \operatorname{Ric}(g) - \tfrac{1}{2}r(g)g$$

PROOF. Using 2.3.3 and 2.4.1 and observing that $\int \Delta(\operatorname{tr} h)$ and $\int \delta\delta h$ drop out because M is compact without boundary, we get:

$$dV(g) \cdot h = 2\int_M \operatorname{Ric}(g) \cdot h \, d\mu_g - \int_M r(g)\operatorname{tr}(h) \, d\mu_g$$

It is easy to check that (3) satisfies $\mathscr{G}_g(\operatorname{grad} V(g), h) = dV(g) \cdot h$. Indeed the left hand side becomes

$$\begin{aligned}
&\int(\operatorname{grad} V(g) \cdot h - \operatorname{tr} \operatorname{grad} V(g) \cdot \operatorname{tr} h) \, d\mu_g \\
&\quad = \int(2 \operatorname{Ric}(g) \cdot h - \tfrac{1}{2}r(g)g \cdot h) \, d\mu_g - \int(2r(g)\operatorname{tr}(h) - \tfrac{3}{2}r(g)\operatorname{tr} h) \, d\mu_g \\
&\quad = 2\int \operatorname{Ric}(g) \cdot h \, d\mu_g - \int r(g)\operatorname{tr} h \, d\mu_g
\end{aligned}$$

which equals $dV(g) \cdot h$. Thus (3) holds. ∎

We now summarize what we have done so far:

2.4.3 THEOREM. *The equations for an integral curve $(g(t)k(t))$ for geodesic motion in $\mathscr{M}$ in the presence of the potential $V(g) = -2\int_M r(g)\, d\mu_g$ are:*

$$\begin{cases} \dfrac{d}{dt} g(t) = k(t) \\[2ex] \dfrac{d}{dt} k(t) = k \times k - \tfrac{1}{2}\operatorname{tr}(k)k - 2\operatorname{Ric}(g) - \tfrac{1}{4}\mathscr{H}(g, k)g \end{cases}$$

where

$$\mathscr{H}(g, k) = \tfrac{1}{2}\{k \cdot k - (\operatorname{tr} k)^2\} - 2r(g)$$

PROOF. This follows immediately from equation (4) of § 1.2, 2.3.2, and 2.4.2. ∎

2.5 The constraints

Before we can show that the equations of motion in 2.4.3 lead to a Ricci-flat spacetime, we need to establish certain conservation laws. The first of these will be proved by using the results of § 1.5. The first thing to observe is that there is a natural isometry group for $\mathscr{M}$, the group $\mathscr{D}$ of diffeomorphisms. The action of $\mathscr{D}$ on $\mathscr{M}$ is defined by:

(1) $$\Phi_\eta : g \to (\eta^{-1})^* g.$$

One can easily check that it satisfies the requirements for an action (§ 1.5), consists of isometries of $\mathscr{M}$ and leaves V invariant. This invariance is basically due to the tensorial nature of $\mathscr{G}$ and V. We want to think of $\mathscr{D}$ as the "active" coordinate transformations of M. That the action (1) leaves $\mathscr{G}_g$ and V invariant may therefore be interpreted as the invariance of the theory under (active) coordinate transformations.

To obtain the conservation law, we need to compute the infinitesimal generators of Φ. Let $X \in T_e\mathscr{D}$ and η_t = flow of X, a one parameter group of diffeomorphisms of M. By definition, the infinitesimal generator of the action corresponding to $X \in T_e\mathscr{D}$ is given by $(d/dt)((\eta_t^{-1})^*g)|_{t=0} = -L_X g$, where L_X is the Lie derivative. Thus the infinitesimal generator is $g \mapsto -L_X g$. The constant of the motion is therefore $(g, h) \mapsto \mathscr{G}_g(-L_X g, h)$. To simplify this we need the following version of Stokes's theorem:

2.5.1 LEMMA.

$$\int_M L_X g \cdot k \, d\mu_g = 2\int_M X \cdot \delta k \, d\mu_g.$$

PROOF. One can easily show that $L_X g = X_{i|j} + X_{j|i}$, where $_{|i}$ denotes covariant differentiation. Then

$$\delta(k \cdot X) = (\delta k) \cdot X - k \cdot \nabla X$$

where

$$k \cdot \nabla X = k^{ij} X_{i|j} = \tfrac{1}{2}(k^{ij}X_{i|j} + k^{ji}X_{j|i}) = \tfrac{1}{2}k \cdot L_X g$$

by the symmetry of k. Since, by Stokes's theorem, $\int_M \delta(k \cdot X)\, d\mu_g = 0$, we get the lemma. ∎

2.5.2 THEOREM. *The constant of the motion above is given by*

$$\mathscr{G}_g(k, L_X g) = 2\int_M X \cdot \delta\pi \, d\mu_g$$

where $\pi = k - (\operatorname{tr} k)g$, the "conjugate momentum."

Note. Our notation differs slightly from Fisher-Marsden [1] in the sign of $\mathscr{G}$ and V and the fact that π is not a density.

PROOF OF 2.5.2. Now

$$\begin{aligned}\mathscr{G}_g(k, L_X g) &= \int_M \{k \cdot L_X g - (\operatorname{tr} k)(\operatorname{tr} L_X g)\}\, d\mu_g \\ &= \int_M (L_X g) \cdot (k - (\operatorname{tr} k) g)\, d\mu_g \\ &= \int_M L_X g \cdot \pi \, d\mu_g \\ &= 2\int_M X \cdot \delta\pi \, d\mu_g. \quad \blacksquare\end{aligned}$$

From the arbitrariness of X, it follows that $(\delta\pi)\mu_g$ is conserved and thus *if* $\delta\pi = 0$ *at* $t = 0$, *it remains* 0 on the integral curves of the equations of motion. This is the first basic constraint on the initial data we will wish to impose on the Einstein equations.

Although the result can be proven directly we feel it is important to emphasise that it results just from the invariance of the theory under $\mathscr{D}$. Similarly if one had another theory on a different vector bundle than $S_2(M)$, one would obtain a conservation law merely from the coordinate invariance. Now we come to the second basic constraint.

2.5.3 THEOREM. (*pointwise conservation of energy*) *Let* $(g(t), k(t))$ *be an integral curve of the system in* 2.4.3 *and let* $\delta\pi = 0$ *on this curve. Then the energy density*

$$\mathscr{H}(g, k)\mu_g = \tfrac{1}{2}(k \cdot k - (\operatorname{tr} k)^2)\mu_g - 2r(g)\mu_g$$

is pointwise constant in time.

PROOF. Let $\mathscr{K}_g(k) = \frac{1}{2}(k \cdot k - (\operatorname{tr} k)^2)$, the kinetic energy density. Then

$$\frac{\partial}{\partial t}(\mathscr{K}_g\mu_g) = \hat{\mathscr{G}}_g\left(k, \frac{dk}{dt}\right)\mu_g + D_g\mathscr{K}_g(k) \cdot k\mu_g + \frac{\mathscr{K}_g}{2}\operatorname{tr}(k)\mu_g$$

where $\hat{\mathscr{G}}_g(k_1, k_2) = k_1 \cdot k_2 - (\operatorname{tr} k_1)(\operatorname{tr} k_2)$. Using the formulas in Lemma 2.3.3 we have

$$D_g\mathscr{K}_g(k) \cdot k = -k \cdot (k \times k) + k \cdot k(\operatorname{tr} k)$$

and

$$\begin{aligned}\mathscr{G}_g\left(k, \frac{dk}{dt}\right) &= k \cdot k \times k - \tfrac{1}{2}(\operatorname{tr} k)k \cdot k - \tfrac{1}{4}\mathscr{K}_g(\operatorname{tr} k) \\ &\quad - \{\operatorname{tr} k(k \cdot k) - \tfrac{1}{2}(\operatorname{tr} k)^3 - \tfrac{3}{4}\mathscr{K}_g(\operatorname{tr} k)\} \\ &\quad - 2\operatorname{Ric}(g) \cdot k + \tfrac{1}{2}r(g)\operatorname{tr}(k) \\ &\quad + 2\, \mathrm{r(g)tr(k)} - \tfrac{3}{2}\mathrm{r(g)tr(k)}.\end{aligned}$$

Adding, we get

$$\frac{\partial}{\partial t}(\mathscr{K}_g\mu_g) = (-2\operatorname{Ric}(g) \cdot k + r(g)\operatorname{tr}(k))\mu_g.$$

On the other hand,

$$\begin{aligned}\frac{\partial}{\partial t}(-2r(g)\mu_g) = -r(g)\operatorname{tr}(k)\mu_g - (2\Delta(\operatorname{tr} k) \\ + 2\delta\delta k - 2\operatorname{Ric}(g) \cdot k))\mu_g.\end{aligned}$$

Hence adding,

$$\frac{\partial}{\partial t}(\mathscr{H}\mu_g) = -2(\Delta(\operatorname{tr} k) - \delta\delta k)\mu_g = 2\delta\delta\pi\mu_g = 0. \quad \blacksquare$$

This is a rather surprising result; normally it is the total (integrated) energy which is conserved in dynamical systems. But the pointwise conservation of energy is actually a *necessary* condition for any theory to be relativistically invariant. We shall see this more precisely below in § 3.

To get some idea of what the constraint $\delta\pi = 0$ means, it is instructive to look at the orbits of $\mathscr{D}$ in $\mathscr{M}$ (Figure 2.3). We may think of all metrics in the same orbit as being geometrically the same. The condition $\delta\pi = 0$ means that the curve $g(t)$ must start off perpendicular to the orbit ($\perp$ re. $\mathscr{G}$), while the conservation law states that the motion continues to cut the orbits perpendicularly. This space of orbits $\mathscr{M}/\mathscr{D}$ is known as

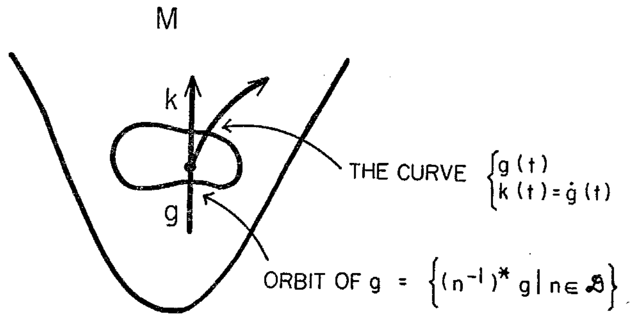


FIGURE 2.3

superspace and has been studied fairly extensively; see Fischer [1]. Superspace is not a manifold, but rather is a stratified set. One of the most important results in this regard is the "slice theorem" proved by Ebin [1].

In order for the Hamiltonian analysis to go through completely, we have to restrict our attention to a subset $\mathscr{C}$ of $T\mathscr{M}$:

$$\mathscr{C} = \{(g, h) \mid \tfrac{1}{2}\{h \cdot h - (\operatorname{tr} h)^2\} - 2r(g) = 0 \quad \text{and} \quad \delta(h - (\operatorname{tr} h)g) = 0\}.$$

Unfortunately, it does not seem that $\mathscr{C}$ is a submanifold of $T\mathscr{M}$; singularities occur at metrics which have nontrivial isometry groups. These singularities are sort of rare since there is a theorem, due to Ebin [1], that almost all metrics (an open dense set) have no nontrivial isometries. However, it is not necessary for our purposes that $\mathscr{C}$ be a submanifold. The important fact is that $\mathscr{C}$ is invariant under the flow—the integral curves starting at points in $\mathscr{C}$, remain in $\mathscr{C}$; this is precisely the meaning of the two constraint theorems just established.

Recall that in hydrodynamics we had the "constraint" $\delta v = 0$, and that the Hodge theorem was basic in this regard. Similarly there are basic decomposition theorems related to the constraint set $\mathscr{C}$ due to Deser [1]; See also Berger-Ebin [1]. These decompositions are basic to analyzing the geometry of $\mathscr{C}$, but we do not go into that here; cf. Fischer-Marsden [4] and Brill-Deser [1].

On $\mathscr{C}$, the equations of motion 2.4.3 simplify to:

$$\text{(2)} \qquad \begin{aligned} \frac{\partial g}{\partial t} &= k \\ \frac{\partial k}{\partial t} &= k \times k - \tfrac{1}{2}(\operatorname{tr} k)k - 2\operatorname{Ric}(g) \end{aligned}$$

since $\mathscr{H} = 0$. Conversely, one can show that a solution of (2) with $\delta\pi$ and $\mathscr{H}$ initially zero maintains these conditions, so 2.4.3 will hold. The difficult term here is the Ricci tensor which is quasilinear (linear in the highest order derivatives); all the other terms are purely algebraic. In some cosmological dust models, it is possible to neglect Ric(g), and in these cases Eardley-Liang-Sachs [1] have obtained explicit solutions.

Warning. One must be careful in interpreting $\mathscr{H}$ as the "physical" energy density. Indeed, in any region where no matter is present we will have $\mathscr{H} = 0$. However, there may in fact be gravitational radiation carrying gravitational energy across such a region. For more on this point, see Arnowitt-Deser-Misner [1, 2], and Brill-Deser [1].

2.6 Remarks on existence of solutions

The original theorem concerning existence of solutions for the Einstein system (which is basically (2) above) is due to Fourès-Bruhat [1]. The result was improved on by Lichnerowicz [1] using Leray systems. See also

Choquet-Bruhat [1] and Dionne [1]. The method involves the theory of *second* order partial differential equations which are quasi-linear and "strictly hyperbolic." Actually, there is a simpler theory of quasi-linear *first* order systems which is applicable here, as was observed by A. Fischer (cf. Fischer-Marsden [2, 3]).

The way this goes is a bit complicated and will not be presented in detail here. However, we will give the intrinsic geometrical definitions, as they should be of some interest to geometers. We will illustrate with the wave equation how one reduces a second order system to a first order one. The method for relativity is more complicated, but the basic idea is the same.

First of all let us consider the linear problem in $\mathbb{R}^n$: Let u be a vector-valued function $u: \mathbb{R}^n \to \mathbb{R}^m$. The system

$$\text{(1)} \qquad \frac{\partial u}{\partial t} = \sum_{i=1}^{n} A^i(x) \frac{\partial u}{\partial x^i} + B(x) \cdot u$$

is said to be *symmetric hyperbolic* if the $m \times m$ matrices A^i are symmetric for all $1 \leqq i \leqq n$. The system is *first* order and *linear* in u. Under fairly mild restrictions, A^i, B should be of class H^s, $s > (n/2) + 1$, there exists a unique solution u_t in H^s (all time) for any initial condition u_0 in H^s. This result is due basically to Petrovsky [1], Friedrichs [1], and others. A proof may be found in Courant-Hilbert [1] Vol. II; see also Kato [3], Dunford-Schwartz [1], and Chernoff-Marsden [1]. Using standard techniques of reducing second order systems to first order, this theorem may be used to solve the wave equation in $\mathbb{R}^n$:

2.6.1 Example. *The wave equation.*
The equation is

$$\frac{\partial^2 f}{\partial t^2} = \nabla^2 f; f = f(x^1, \ldots, x^n, t).$$

Put, formally,

$$\begin{bmatrix} f \\ \dfrac{\partial f}{\partial x^1} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial f}{\partial x^n} \\ \dfrac{\partial f}{\partial t} \end{bmatrix} = \begin{bmatrix} u_0 \\ u_1 \\ \cdot \\ \cdot \\ \cdot \\ u_n \\ u_{n+1} \end{bmatrix}$$

Then the wave equation for f is the same as the following symmetric hyperbolic system for u:

$$\begin{cases} \dfrac{\partial u_0}{\partial t} = u_{n+1} \\ \dfrac{\partial u_1}{\partial t} = \dfrac{\partial u_{n+1}}{\partial x^1} \\ \cdot \\ \cdot \\ \cdot \\ \dfrac{\partial u_n}{\partial t} = \dfrac{\partial u_{n+1}}{\partial x^n} \\ \dfrac{\partial u_{n+1}}{\partial t} = \dfrac{\partial u_1}{\partial x^1} + \dots + \dfrac{\partial u_n}{\partial x^n}. \end{cases}$$

In this case

$$A^1 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ 0 & 1 & \dots & 0 \end{pmatrix} \quad \text{etc.}$$

are symmetric $(n + 2) \times (n + 2)$ matrices.

Thus, using the linear theory for general first order symmetric hyperbolic systems, we get an existence theorem for the wave equation, namely that if $(f_0, (\partial f_0/\partial t)) \in H^{s+1} \times H^s$ there is a unique solution $f_t \in H^{s+1}$, $-\infty < t < \infty$, satisfying the given initial conditions.

The hyperbolicity of $(\partial^2 f/\partial t^2) = \Delta^2 f$ is reflected in the symmetry of the A^i. If we had used $(\partial^2 f/\partial t^2) = -\Delta^2 f$, the A^i would not have come out symmetric—the Cauchy problem in this case is not well posed.

How do we do this invariantly? Well, we consider a vector bundle $\pi: E \to M$ and sections $u: M \to E$. Let $J^1(E)$ be the *first jet bundle*; its fiber over $x \in M$ is

$$J^1(E)_x = L(T_xM, E_x) \oplus E_x.$$

A section u of E gives a section $j(u)$ of $J^1(E)$; the two components of $j(u)$ are Du and u. More precisely, we assume E has a connection and let Du be the horizontal part of Tu; cf. Palais **[1]**.

A first order linear operator may be regarded as a map

$$\tilde{A}: J^1(E) \to E$$

linear on fibers and such that the following diagram commutes

$$\begin{tikzcd} J^1(E) \arrow[rr, "\tilde{A}"] \arrow[dr] & & E \arrow[dl] \\ & M & \end{tikzcd}$$

Then $\tilde{A} \circ j(u)$ reduces, in the case of $\mathbb{R}^n$, to $\sum A_i(\partial u/\partial x^i) + B \cdot u$ as we had before. Let us write, conforming to earlier notation

$$\tilde{A} \circ j(u) = A \cdot Du + B \cdot u$$

(the two components of $\tilde{A}$), so $A: L(TM, E) \to E$.

Suppose now that we have an inner product $\langle\ ,\ \rangle_x$ given on each fiber E_x, as well as the connection.

We call A *symmetric hyperbolic* when the following condition holds:

$$\textit{for any sections } u, v \textit{ of } E, \quad \langle A \cdot j(u), v\rangle = \langle j(u^b), A \cdot v\rangle. \qquad \text{(S.H.)}$$

We now shall explain the symbols appearing in this definition. Here $j(u)$ is the first jet of u (in coordinates, (Du, u)) and u^b is the corresponding section of the dual bundle E^*. Now recalling that A is a map

$$A: L(T_xM, E_x) \to E_x$$

observe that we can also regard A as a map

$$\hat{A}: E_x \to L(E_x^*, T_xM)$$

as follows: let $u_0 \in E_x$, $\gamma \in E_x^*$, $l \in T_x^*M$ and set $l_{u_0} \in L(T_xM, E_x)$, $l_{u_0}(v) = l(v)u_0$. Then let

$$((\hat{A}u_0)\gamma)(l) = \gamma(A(l_{u_0}))$$

Now u^b is a section of E^*, so $j(u^b) \in L(T_xM, E_x^*) \oplus E_x^*$ and $\langle j(u^b), \tilde{A} \cdot V\rangle$ is the natural pairing between

$$L(T_xM, E_x^*) \quad \text{and} \quad L(E_x^*, T_xM)$$

given by

$$\sigma, \tau \mapsto \operatorname{tr}(\tau \circ \sigma).$$

This completes the explanation of condition (S.H.).

One easily sees that in $\mathbb{R}^n$ this condition reduces to the earlier case. The point is that most of the theory works well in this context, provided M is *complete* Riemannian, and the sectional curvature is bounded above.

One requires A to be H^s, a condition which also depends on the metric on M; see Fischer-Marsden [2] and Chernoff-Marsden [1] for details.

Now consider the nonlinear problem in $\mathbb{R}^n$. In this case we have a system of the form

$$\frac{\partial u}{\partial t} = \sum_i A^i(x, t, u)\frac{\partial u}{\partial x^i} + B(x, t, u),$$

where the A^i and B are matrices which are polynomial in u (or more generally, satisfy Sobolev's "condition T"; cf. Sobolev [1]). The system is quasi-linear and the matrices A^i are symmetric. The nonlinear theorem is obtained from the linear theory by adapting the Picard method. In this case also, unique solutions exist in H^s, but only for short time, in contrast to the linear theory. This makes sense on vector bundles in much the same way as the linear case; we just require $\tilde{A}$ to be linear in the first factor Du. The symmetry condition is unchanged.

The Einstein system (2) in § 2.5 above is rather like the wave equation and one can show that on the appropriate bundle, obtained in a way not unlike that for the wave equation, it is symmetric hyperbolic.

The verification that it is symmetric hyperbolic uses "harmonic coordinates"; cf. Lichnerowicz [1].

Thus we get existence and uniqueness of smooth solutions for short time (which can be extended to maximal solutions as well). These solutions depend continuously on the initial data.

In case that M is noncompact, if a section u of the bundle $\pi: E \to M$ is in H^s, then this means, intuitively, that it goes to zero at ∞. For relativity one usually requires an asymptotic condition; for example fix a metric g_0 (such as the spatial part of the Schwartzchild metric on $\mathbb{R}^3$) and require instead that our solution $g(t)$ be *asymptotic* to g_0, that is, $g(t) - g_0$ is in H^s on a neighborhood of ∞. Under certain mild restrictions on g_0, the existence theory goes through in the class of metrics which are asymptotic to g_0.

This existence theory might be useful for finding some non-trival solutions to the Einstein system which exist for all time without singularities (a conjecture of S. Deser). This uses the geometry of $\mathscr{C}$, the slice theorem of Ebin [1], and most importantly the idea of the total mass of a geometry (Brill-Deser [1]); see Fischer-Marsden [4] and Marsden [2].

Such a result is interesting because, in the presence of matter, one expects singularities to develop, as explained in professor Kundt's lectures.

2.7 The Einstein field equations and the dynamical equations

Now we wish to return to the relationship between the four geometry and the evolving three geometry.

Let $g_t = {}^3g_{ij}(x, t)\, dx^i\, dx^j$ be a curve in $\mathscr{M}$. Think of this as sweeping out a 4-manifold as mentioned in the introduction. We want to construct the full spacetime metric $g_{\alpha\beta}(x, t)$ (as usual, Greek indices run from 0 to 3, Latin from 1 to 3). This is straightforward: Put $V = M \times \mathbb{R}$ and on V, define

$$g_{\alpha\beta}(x, t) \equiv g_{ij}\, dx^i\, dx^j - dt^2 \quad (t = x^0).$$

Clearly, this is a special Lorentz metric with $g_{00} = -1$, $g_{0i} = 0$; we shall see how to get the most general ones presently. This construction is hardly remarkable; it can be done for any curve g_t in $\mathscr{M}$. What *is* remarkable is the following:

2.7.1 THEOREM. *$g_{\alpha\beta}$ is Ricci-flat (that is, $g_{\alpha\beta}$ on V satisfies the Einstein vacuum field equations)* $\Leftrightarrow$

(1) *${}^3g$ satisfies the equations* (2) *of* § 5 *and*

(2) $\mathscr{H} = 0$, $\delta\pi = 0$—*the energy and divergence conditions are satisfied: (that is, (g_t, k_t) lies in $\mathscr{C}$)*

REMARK. This theorem (and a more general one to follow), establish an equivalence between a geometrical condition ($R_{\alpha\beta} = 0$) in the 4-geometry and the Cauchy (evolution) problem for ${}^3g$. This is precisely the way in which the field equations are interpreted as equations of motion; i.e., as a dynamical system.

PROOF OF THEOREM 2.7.1 (cf. Wheeler [1]). Given $V = M \times \mathbb{R}$ and the metric $g_{\alpha\beta}(x, t)$ given by (1), $g_{ij}(x, t)$ is the induced hypersurface metric on $M \times \{t\} = M_t$. The necessary additional information comes from the second fundamental form and the Gauss-Codazzi equations, which describe the extrinsic structure (second fundamental form) of M_t as an embedded submanifold. The second fundamental form has a simple expression. K. Yano [1], Ch. 5 § 5 is a good reference for the basic definitions and geometry of embedded hypersurfaces which we shall need. In particular for a hypersurface $M \subset V$ with unit normal Z, the second fundamental form $S(X, Y)$ is defined by $S(X, Y) = \langle H(X), Y\rangle$, $H(X) = \nabla_X Z$, (the Weingarten equation). There is a change in sign over what is in Yano since $\langle Z, Z\rangle = -1$ rather than $+1$. We shall need:

2.7.2 LEMMA. *The second fundamental form of $M_t \subset V$ for M, V, ${}^4g$ as above, is given by*

(1)
$$S_{ij} = \frac{1}{2}\frac{\partial g_{ij}}{\partial t} = \tfrac{1}{2}k_{ij}.$$

PROOF. By definition, $S_{ij} = Z_{i|j}$ where $Z^{\mu} = (0, 1)$ is the unit normal. Now $Z_{\mu} = (0,-1)$ so $S_{ij} = 0 - \Gamma^{\alpha}_{ij}Z_{\alpha} = \Gamma^{0}_{ij}$. But one easily computes from the formula $\Gamma^{\alpha}_{\mu\nu} = \frac{1}{2}g^{\lambda\alpha}(g_{\mu\lambda,\nu} + g_{\nu\lambda,\mu\nu} - g_{\mu\nu,\lambda})$ that $\Gamma^{0}_{ij} = \frac{1}{2}(\partial g_{ij}/\partial t)$. ■

This lemma therefore gives the geometrical significance of the "velocity" k; it is just (twice) the second fundamental form. The tensor S measures the difference between the covariant derivative on V and that on M via the *Gauss equation*:

$$^4\nabla_X Y = {}^3\nabla_X Y + S(X, Y)Z$$

where $^4\nabla$ is the connection of V, $^3\nabla$ that on M (compare the expression for $S(X, X)$ as the difference between the sprays that we used in interpreting the pressure in hydrodynamics earlier).

Now we come to the Gauss-Codazzi equations; they are quite general for any embedded hypersurface. First, the *Gauss equation* states: for X, Y, U, W tangent to M,

$$^{(4)}R(X, Y, U, W) = {}^3R(X, Y, U, W) + \{S(X, W)S(Y, U) - S(Y, W)S(X, U)\}.$$

This follows very easily from the definitions (Yano [1], p. 94). For R we use the standard definition:

$$R(\bar{X}, \bar{Y})\bar{Z} \equiv \nabla_{\bar{X}}\nabla_{\bar{Y}}\bar{Z} - \nabla_{\bar{Y}}\nabla_{\bar{X}}\bar{Z} - \nabla_{[\bar{X},\bar{Y}]}\bar{Z}$$

In coordinates; the Gauss equation reads as follows:

(2) $$^4R_{ijkl} = {}^3R_{ijkl} - \{S_{il}S_{jk} - S_{jl}S_{ik}\}.$$

Similarly the *Codazzi equation* is:

$$^4R(X, Y, U, Z) = ({}^3\nabla S)(X, U, Y) - ({}^3\nabla S)(Y, U, X)$$

where we use $(\nabla S)(Y, U, X) = (\nabla_X S)(Y, U)$ or in coordinates,

(3) $$R_{ijk0} = S_{ik|j} - S_{kj|i}.$$

(Yano [1], p. 95), where the covariant derivative is taken on M, and Z is a unit normal.

This does not yet give a complete decomposition of the curvature tensor because we do not have a formula for $^4R(Z, X, Z, Y)$. This involves more than just M, its metric structure and S. In fact it depends on having a whole family of hypersurfaces M_t and knowing S on each hypersurface. This formula was worked out in general by Schouten and Struik, a convenient reference being Abraham [3]. In our case, the formula may be worked out directly from the definition

$$^4R(Z, X, Z, Y) = \langle {}^4R(Z, X)Z, Y\rangle = \langle \nabla_Z\nabla_X Z, Y\rangle - \langle \nabla_X\nabla_Z Z, Y\rangle + \langle \nabla_{[Z,X]}Z, Y\rangle$$

Taking $X = (\partial/\partial x^i)$, $Y = (\partial/\partial x^j)$, $Z = (\partial/\partial t)$ gives, in our case, using $S^i_j = \nabla_{\partial/\partial x^i}Z^j = Z^j{}_{|i}$, and the formulas

$$\begin{aligned}(\nabla_Z W)^k = W^k{}_{|0} &= \frac{\partial W^k}{\partial t} + \Gamma^k_{0\beta}W^\beta \\ &= \frac{\partial W^k}{\partial t} + \Gamma^k_{00}\ W^0 + \Gamma^k_{0i}W^i \\ &= \frac{\partial W^k}{\partial t} + 0 + \tfrac{1}{2}g^{lk}\frac{\partial g_{il}}{\partial t}W^i\end{aligned}$$

the result:

$$^4R_{0i0j} = \left(\frac{\partial}{\partial t}S^l_i\right)g_{lj} + (S \times S)_{ij}$$

or

$$^4R_{0i0j} = \frac{\partial}{\partial t}(S_{ij}) - S^l_i k_{lj} + (S \times S)_{ij}$$

so by 2.7.2 we finally get:

(4) $$^4R_{0i0j} = \frac{\partial}{\partial t}S_{ij} - (S \times S)_{ij} = \frac{1}{2}\frac{\partial k_{ij}}{\partial t} - \tfrac{1}{4}(k \times k)_{ij}.$$

We now have the information needed to complete the proof of 2.7.1. Assume $^4R_{\alpha\beta} = 0$. Then using (2), (4) and Yano's conventions on Ric,

$$\begin{aligned}0 = {}^4R_{ij} &= {}^4g^{\alpha\beta}\,{}^4R_{\alpha ij\beta} = +{}^4R_{0i0j} - g^{kl}\,{}^4R_{kilj} \\ &= \left\{\frac{1}{2}\frac{\partial k}{\partial t} - \tfrac{1}{4}(k \times k)\right\} + \{\mathrm{Ric}(g) + \tfrac{1}{4}[(\mathrm{tr}\,k)k - (k \times k)]\}\end{aligned}$$

which gives the correct equation for $\partial k/\partial t$.

Similarly using (3),

$$\begin{aligned}0 = {}^4R_{0i} = {}^4g^{\alpha\beta}\,{}^4R_{\alpha 0i\beta} &= {}^3g^{kl}\,{}^4R_{k0il} \\ &= \tfrac{1}{2}(k^j_{i|j} - k^j_{j|i}) \\ &= -\tfrac{1}{2}\,\delta(k - (\mathrm{tr}\,k)g)\end{aligned}$$

which gives $\delta\pi = 0$. Finally

$$\begin{aligned}0 = {}^4R_{00} = {}^4g^{\alpha\beta}\,{}^4R_{\alpha 00\beta} &= -{}^3g^{kl}\,{}^4R_{0k0l} \\ &= -{}^3g^{kl}\left(\frac{1}{2}\frac{\partial k_{kl}}{\partial t} - \tfrac{1}{4}(k \times k)_{kl}\right).\end{aligned}$$

Substituting in the equation for $\partial k/\partial t$ gives $0 = -\frac{1}{2}\mathscr{H}(g, k)$, the energy condition.

The converse is proved by retracing the steps. ■

3. The lapse and shift functions

We now want to generalize the analysis in § 2 to allow more general representations of the four metric than

$$g_{\alpha\beta}\, dx^\alpha\, dx^\beta = -dt^2 + g_{ij}\, dx^i\, dx^j. \tag{1}$$

This may be achieved by subjecting $g_{\alpha\beta}$ to an arbitrary coordinate transformation. Indeed for any Lorentz metric $g_{\alpha\beta}$ and spacelike hypersurface M_0, we can bring $g_{\alpha\beta}$ into the form (1) by choosing Gaussian normal coordinates; i.e., a standard tubular neighborhood defined by the exponential map. This is easy to check and is well known (cf. Synge [**1**]). In other words, *the evolution equations we have derived previously* ((2) in 2.5) *correspond to the use of Gaussian normal coordinates in the space time.* Physically, these coordinates may be described as follows. From M_0 we consider freely falling observers who start out orthogonal to M_0 (i.e., start out "at rest"). The coordinates on M_0 and the proper time as measured by these observers then defines the Gaussian system.

If a set of observers start off in relative motion to the above ones, and if they measure time differently, they will generate an isometric spacetime, but will see the 3-geometry evolving differently. The relative motion is described by the "shift" and the new clock rates by the "lapse." By this procedure we recover the full coordinate invariance of the spacetime.

We can look at the same procedure in a slightly different way. Starting with our 3-manifold $M = M_0$, we can choose a completely *arbitrary* congruence of timelike curves through M_0, subject only to the condition that the new parameter t' be zero on M_0 (this is no restriction). Since M_0 is compact, there will be a finite t'-interval in which the 3-manifolds M_t defined by $t' =$ constant are all spacelike. Thus this new congruence determines a new curve $\tilde{g}_{t'}$ in $\mathscr{M}$, satisfying $\tilde{g}_{0'} = g_0$; and the curve $\tilde{g}_{t'}$ will satisfy a different set of evolution equations than those satisfied by g_t. However, g_t and $\tilde{g}_{t'}$ both represent the same spacetime; thus it is essential to find the relationship between the two curves. To do this, we shall introduce the lapse and shift functions.

Generally, in this dynamical approach we are restricting ourselves to a small interval of $t \in \mathbb{R}$. When a singularity develops one may require other procedures to develop the spacetime further. Such procedures are not well developed; cf. Geroch [**1**]. We refer also to Professor Kundt's lectures for these matters.

We begin by introducing the shift. The lapse will be discussed separately below.

3.1 The shift

A *shift function* is a given time-dependent vector field X_t on M. Let η_t be its flow starting at the identity. The geometric picture for this is that our arbitrary congruence of timelike curves through M_0 is parametrized by requiring the surface $t' =$ constant to be identical to those given by $t =$ the same constant, in Gaussian coordinates. Then the only thing happening is a coordinate transformation within the surface itself via a one-parameter family of diffeomorphisms on M. This is a shift.

Let $g(t)$ be the solution for no shift with initial values g_0, k_0 on M_0. Let $\tilde{g}(t)$ be the solution with the shift.

We have an induced coordinate transformation of $M \times \mathbb{R}$ given by

$$(x, t) \mapsto (\eta_t(x), t) = (\bar{x}, \bar{t}).$$

This transformation (thought of actively) changes the metric

$$g_{\alpha\beta}\, dx^\alpha\, dx^\beta = -dt^2 + g_{ij}\, dx^i\, dx^j$$

into the form

$$\tilde{g}_{\alpha\beta}\, d\bar{x}^\alpha\, d\bar{x}^\beta = -(1 - X \cdot X)\, dt^2 - 2X_i\, d\bar{t}\, d\bar{x}^i + \tilde{g}_{ij}\, d\bar{x}^i\, d\bar{x}^j.$$

Note that we require $X \cdot X < 1$, or the relative velocity of our two observers must be less than that of light $c = 1$. Thus the solution $\tilde{g}_{ij}$ is required to be such that $\tilde{g}_{\alpha\beta}$ will be Ricci-flat. But $\tilde{g} = (\eta_t^{-1})^* g$ so we can easily obtain the evolution equations for $\tilde{g}$ from those of g. Letting $\tilde{k} = (\eta_t^{-1})^* k$, $k = \partial g/\partial t$, we get this system:

$$\begin{aligned} \frac{\partial \tilde{g}}{\partial t} &= \tilde{k} - L_X \tilde{g} \\ \frac{\partial \tilde{k}}{\partial t} &= S_{\tilde{g}}(\tilde{k}) - 2\,\mathrm{Ric}(\tilde{g}) - L_X \tilde{k} \end{aligned} \tag{1}$$

where

$$S_{\tilde{g}}(\tilde{k}) = \tilde{k} \times \tilde{k} - \tfrac{1}{2}(\operatorname{tr} \tilde{k})\tilde{k}.$$

These are the evolution equations for $\tilde{g}$; they differ from those of g by just an additional Lie derivative. In $\mathscr{M}$ we have the picture shown in Figure 3.1. Note that g_t is perpendicular to the orbits, while $\tilde{g}_t$ is not. They both intersect the same orbits at the same t; thus they determine the same curve in superspace.

Notice that (1) is more complicated than the evolution equations with $X = 0$ because of the additional terms $L_X k$, $L_X g$. However, the solution is obtained from the solution for $X = 0$ once we know the flow of X. This is rather like the procedure one uses in the Cauchy method of characteristics (i.e., the solution of $\partial f/\partial t = -\sum X^i(t, x)(\partial f/\partial x^i)$ is $f(t, x) = f_0(\eta_t^{-1}(x))$ where η_t is the flow $X = (X^1, \ldots, X^n)$).

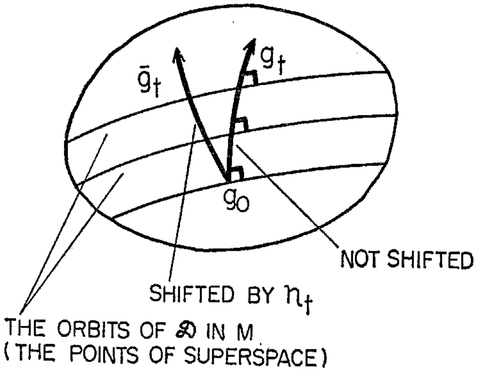


FIGURE 3.1

We wish to stress that the shift X_t is to be prescribed before hand. It is not evolving according to any equations (unless some are imposed for special reasons).

For $X = 0$ we have seen the precise sense in which the evolution equations (and the constraints) arise from a Hamiltonian system (§§ 2.4, 2.5). There remains the same problem for the equations (1) with a shift.

This can be done very simply and naturally as follows. (Compare Arnowitt-Deser-Misner **[1]** for another approach.) Consider the enlarged configuration space $\mathscr{D} \times \mathscr{M}$ and the map

$$\Phi : \mathscr{D} \times \mathscr{M} \to \mathscr{D} \times \mathscr{M}$$

$$(\eta, g) \mapsto (\eta, \eta^* g).$$

Now consider our original system on $\mathscr{D} \times \mathscr{M}$ with the first factor just being ignored (a degenerate variable). We transform this by Φ to obtain another degenerate system on $\mathscr{D} \times \mathscr{M}$; for example the metric is easily worked out to be

(2) $$(\bar{\mathscr{G}}_{\eta,g})(X_1, h_1, X_2, h_2) = \mathscr{G}_g(h_1 + L_{X_1 \circ \eta^{-1}} g,\ h_2 + L_{X_2 \circ \eta^{-1}} g)$$

which is a well defined degenerate metric on $\mathscr{D} \times \mathscr{M}$. Similarly

(3) $$\bar{V}(\eta, g) = V(g)$$

is the potential. Thus we have constructed on $\mathscr{D} \times \mathscr{M}$ a metric and a potential. The following is almost obvious from our discussion:

3.1.1 THEOREM. *The equations of motion* (1) *are consistent equations for the metric $\bar{\mathscr{G}}$ and potential $\bar{V}$ on $\mathscr{D} \times \mathscr{M}$, in the sense defined in* § 1. *The vector field X_t may be arbitrarily specified.*

There is another way of looking at this which is analogous to what happens in hydrodynamics. Namely, let $\mathscr{A}$ denote the set of maps $g : M \to S_2(M)$ which cover some $\eta \in \mathscr{D}$. Then $\mathscr{A}$ becomes a vector bundle over $\mathscr{D}$; $\pi : \mathscr{A} \to \mathscr{D}$ with fiber isomorphic to the sections of $S_2(M)$. In particular we can restrict to $\mathscr{A}_0$ corresponding to the positive definite g's.

Now $\mathscr{A}_0$ is diffeomorphic to $\mathscr{D} \times \mathscr{M}$ in two ways: Let $g \in \mathscr{A}_0$, $\pi(g) = \eta \in \mathscr{D}$.

$$\mathscr{D} \times \mathscr{M} \leftarrow \mathscr{A}_0 \to \mathscr{D} \times \mathscr{M}$$

$$(\eta, (\eta^{-1})^*(g \circ \eta^{-1}) \leftarrow\!\shortmid g \qquad g \mapsto (\eta, g \circ \eta^{-1})).$$

The difference is, of course, just the map Φ above. These ways of realizing $\mathscr{A}_0$ as $\mathscr{D} \times \mathscr{M}$ are analogous to the two ways one realizes TG as $G \times \mathscr{G}$ by left and right translations (G is a Lie group, $\mathscr{G}$ its Lie algebra). The transition is that from "body" to "space" coordinates (cf. Arnold **[1]**), or roughly the difference between a stationary and a moving frame.

Thus $\mathscr{A}_0$ is analogous to $T\mathscr{D}_\mu$ in hydrodynamics; i.e., to Lagrangian coordinates while $\mathscr{D} \times \mathscr{M}$ is analogous to $\mathscr{D}_\mu \times T_e\mathscr{D}_\mu$, or Euler coordinates.

It is a good exercise to write down the metric and spray intrinsically on the manifold $\mathscr{A}_0$, a task which we leave to the reader.

3.2 The lapse

The shift takes care of coordinate changes in M. To rescale the time parameter, one introduces the *lapse function* $N : \mathbb{R} \times M \to \mathbb{R}$. Let $\mathscr{T} = C^\infty(M; \mathbb{R})$, a vector space over $\mathbb{R}$ under pointwise addition. Let ξ_t be a curve in $\mathscr{T}$ with $\xi_0 = 0 =$ identity in $\mathscr{T}$. The idea is that $\xi_t(m)$ is the setting of a clock at $m \in M$. Define the lapse function by $N(t, m) = d\xi_t(m)/dt$, and consider only curves ξ_t such that $N_t(m)$ is positive (time is increasing). Alternatively given the lapse function the clock settings may be obtained by integration.

Define a Lorentzian metric on $M \times \mathbb{R}$ by setting

(1) $$g_{\alpha\beta}\, dx^\alpha\, dx^\beta = -N^2\, dt^2 + g_{ij}\, dx^i\, dx^j$$

(in this section we are ignoring the shift). The evolution equations for this g turn out to be

$$(2)\qquad \begin{cases} \dfrac{\partial g}{\partial t} = Nk = h \\ \dfrac{\partial k}{\partial t} = NS_g(k) - 2N \operatorname{Ric}(g) + 2 \operatorname{Hess} N \\ \text{where} \quad \operatorname{Hess} N = N_{|i|j} \end{cases}$$

and we have the following

3.2.1 THEOREM. *With N as described, the space time (1) is Ricci flat (${}^{(4)}R_{\alpha\beta} = 0$) $\Leftrightarrow$ the evolution equations (2) are satisfied and the constraints $\mathscr{H} = 0$, $\delta\pi = 0$ are satisfied with $k = \dot{g}/N$.*

The proof is a straightforward generalization of the argument given in 2.7.1. We leave it to the reader to make the necessary adjustments. Notice that V is still $M \times \mathbb{R}$, but t is *no* longer the proper time. The relationship between t and the proper time is worked out below.

Two questions immediately arise:

(1) In what sense are the evolution equations (2) Hamiltonian? If we want N to be arbitrarily specifiable, then the appropriate configuration space ought to be $\mathscr{T} \times \mathscr{M}$ rather than $\mathscr{M}$; and the correct Lagrangian L would then live in $T(\mathscr{T} \times \mathscr{M})$ rather than in $T\mathscr{M}$.

(2) Is there a geometrical way to pass from the solutions for one N to those for another N?

3.3 The Lagrangian on $\mathscr{T} \times \mathscr{M}$

Recall that the Lagrangian on $\mathscr{M}$ is given by $L(g, h) = \frac{1}{2}\mathscr{G}_g(h, h) + 2\int_M r(g)\, d\mu_g$. This can be written in the form $L(g, h) = \int_M \mathscr{L}(g, h)\, d\mu_g$; this is obvious from the definition of $\mathscr{G}$. $\mathscr{L}$ is called the *Lagrange density*. Bearing in mind the classical way of obtaining relativistic homogeneous Lagrangians on $T(\mathbb{R} \times M)$ (see Lanczos [1] p. 121, Chernoff-Marsden [1]), we define $\bar{L}: T(\mathscr{T} \times \mathscr{M}) \simeq \mathscr{T} \times \mathscr{T} \times T\mathscr{M} \to \mathbb{R}$ by

$$(3)\qquad \bar{L}(\xi, N, g, h) \equiv \int_M N\mathscr{L}\left(g, \frac{h}{N}\right) d\mu_g$$

(on the subset where $N > 0$). This Lagrangian is necessarily *degenerate*. In the general case of degenerate Lagrangians, there may not be consistent equations of motion. For the classical extension of L to $T(M \times \mathbb{R})$ by $\bar{L}(t, s, v) = sL(v/s)$, there is no problem about the existence of its Lagrangian vector field. One can easily check that the most general such second order vector field is given by

$$\bar{Z}(t, \lambda, v) = \lambda Z(v) \oplus \alpha(t, \lambda)$$

where α is any second order equation on $T\mathbb{R}$ (cf. Abraham [**2**], p. 136). The fact that α can be arbitrarily specified reflects the degeneracy.

We should point out (a remark of A. Taub) that the choice of $\bar{L}$ as defined by (3) is closely related to the 4-dimensional approach embodied in the 4-dimensional variational principle (cf. Adler-Bazin-Schiffer [**1**]).

In case of $\mathscr{T} \times \mathscr{M}$ there will not exist consistent equations of motion because of general considerations (§ 4 below). We must restrict to the subset of constraints.

3.3.1 THEOREM. *Consider the Lagrangian $\bar{L}$ defined by (3) on $T(\mathscr{T} \times \mathscr{M})$. Let $\mathscr{C}$ be the subset of $T(\mathscr{T} \times \mathscr{M})$ on which $\delta\pi = 0$, $\mathscr{H} = 0$. Then at points of $\mathscr{C}$, consistent equations of motion exist and are given by (2) of § 3.2 above. Solution curves map $\mathscr{C}$ to $\mathscr{C}$.*

The fact that the equations leave $\delta\pi = 0$, $\mathscr{H} = 0$ invariant may be proved as before. Surprisingly, the fact that these are consistent equations for $\bar{L}$ is not obvious, but is tied up with the condition $\mathscr{H} = 0$. We explore this aspect next.

3.4 Relativistic Lagrangians

In a "relativistic" theory, one requires that N should be arbitrarily specifiable consistent with the Lagrangian. We shall now prove that this implies that $\mathscr{H} \equiv$ constant in time (where $\mathscr{H}$ is the energy density of the unextended Lagrangian $\mathscr{L}$). That this must be the case has been recognized by Misner [**1**]; see also Gerlach [**1**] and Tomonaga [**1**]. R. Sachs has remarked that the result is analogous to the fact that for a system of relativistic particles, the Hamiltonian which is derived from the Lagrangian of the system is identically zero (cf. Lanczos [**1**], p. 320).

Note that external fields are not allowed—one can "tell time" by looking at the fields. The constant value of $\mathscr{H}$ is often zero, but as Sachs has pointed out, in some dust models, it is nonzero.

3.4.1 THEOREM. (i) *If a Lagrangian vector field $\bar{Z}$ for $\bar{L}$ (defined on $T(\mathscr{T} \times \mathscr{M})$ above) exists at (ξ, N, g, h) then it must be a second order equation provided that it is second order in either ξ or g.*

(ii) *In order that $\bar{Z}$ should exist as a second order equation at (ξ, N, g, h), and that N be arbitrarily specifiable, that is, that the "degenerate direction" is all of $\mathscr{T}$, it is necessary that for any curve $(\xi(t), N(t), g(t), h(t))$ tangent to $\bar{Z}$ we have*

$$\frac{\partial}{\partial t}\left\{\mathscr{H}\left(g, \frac{h}{N}\right)\mu_g\right\} = 0, \quad \text{where} \quad \mathscr{H}\left(g, \frac{h}{N}\right) = \frac{1}{2}\left(\frac{h}{N}\cdot\frac{h}{N} - \left(\operatorname{tr}\frac{h}{N}\right)^2\right) - 2r(g).$$

PROOF. (i) In general the relation between Z and L on a manifold B is the Lagrangian condition

$$2\omega_L(v)(Z(v), w) = dE(v)\cdot w$$

on TB (see § 1.4 above). If we let $Z = (Z_1, Z_2)$ locally on TB, this condition reads as follows: for all e_1, e_2 we have

$$\begin{aligned} D_1D_2L(u, e)\cdot e_1\cdot e - D_1L(u, e)\cdot e_1 + D_2D_2L(u, e)\cdot e_2\cdot e \\ = D_2D_2L(u, e)\cdot e_1\cdot Z_1 - D_1D_2L(u, e)\cdot Z_1\cdot e_1 \\ + D_2D_2L(u, e)\cdot Z_1\cdot e_2 - D_2D_2L(u, e)\cdot Z_2\cdot e_1. \end{aligned}$$

These split up into two conditions:

(1) $$D_2D_2L(u, e)\cdot Z_1\cdot e_2 = D_2D_2L(u, e)\cdot e_2\cdot e$$

and

(2) $$\begin{aligned} D_1D_2L(u, e)\cdot e_1\cdot e - D_1L(u, e)\cdot e_1 = D_1D_2L(u, e)\cdot e_1\cdot Z_1 \\ - D_1D_2L(u, e)\cdot Z_1\cdot e_1 - D_2D_2L(u, e)\cdot Z_2\cdot e_1. \end{aligned}$$

In general, we cannot conclude from (1) that $Z_1(u, e) = e$ because L is degenerate.

Now let us turn to the case at hand. Let us incorporate μ_g into $\mathscr{L}$ so we can briefly just write

$$\bar{L}(\xi, g, N, h) = \int N\mathscr{L}\left(g, \frac{h}{N}\right).$$

We also suppress the fact that $\mathscr{L}$ depends explicitly on Dg, D^2g which is irrelevant for the present discussion.

Using obvious notation, the derivatives of $\bar{L}$ are easily worked out to be the following:

(a) $D_1L(\xi, g, N, h)\cdot(\tilde{\xi}, \tilde{g})$

= derivative of L with respect to (ξ, g) in direction $(\tilde{\xi}, \tilde{g})$

$$= \int N\partial_g\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{g}.$$

(b) $D_2L(\xi, g, N, h)\cdot(\tilde{N}, \tilde{h})$

= derivative with respect to the velocity variables (N, h) in direction $(\tilde{N}, \tilde{h})$

$$= \int \tilde{N}\mathscr{L}\left(g, \frac{h}{N}\right) - \int \tilde{N}\partial_h\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{h}{N} + \int \partial_h\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{h}.$$

(c) $D_1D_2L(\xi, g, N, h)\cdot(\tilde{\xi}, \tilde{g})\cdot(\tilde{N}, \tilde{h})$

$$= \int \tilde{N}\partial_g\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{g} - \int \tilde{N}\partial_g\partial_h\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{g}\cdot\frac{h}{N} + \int \partial_g\partial_h\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{g}\cdot\tilde{h}.$$

(d) $D_2D_2L(\xi, g, N, h)\cdot(\tilde{N}, \tilde{h})\cdot(\tilde{\tilde{N}}, \tilde{\tilde{h}})$

$$\begin{aligned} = \int \tilde{N}\tilde{\tilde{N}}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{h}{N}\cdot\frac{h}{N^2} - \int \tilde{\tilde{N}}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{h}\cdot\frac{h}{N^2} \\ - \int \tilde{N}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{\tilde{\tilde{h}}}{N}\cdot\frac{h}{N} + \int \partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{h}\cdot\frac{\tilde{\tilde{h}}}{N}. \end{aligned}$$

Note that in the computation of the second derivative of L with respect to the velocity variables, two pairs of terms canceled out. Now let us use this expression to write out condition (1). Let us write $Z_1(\xi, g, N, h) = (\dot{\xi}, \dot{g})$ for convenience. Condition (1) splits into two conditions, taking respectively $e_2 = (\tilde{N}, 0)$ and $(0, \tilde{h})$. We get

(1)′ $$0 = \int \tilde{N}\dot{\xi}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{h}{N}\cdot\frac{h}{N^2} - \int \tilde{N}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{\dot{g}}{N}\cdot\frac{h}{N}$$

and

(1)″ $$0 = -\int \dot{\xi}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{h}\cdot\frac{h}{N^2} + \int \partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\tilde{h}\cdot\frac{\dot{g}}{N}.$$

Each of these conditions is equivalent to the single condition $\dot{\xi}h = N\dot{g}$. Thus if $\dot{\xi} = N$ then $h = \dot{g}$ and vice versa. Hence (i) follows.

To establish (ii), we write out condition (2) which now becomes

$$D_1L(u, e)\cdot e_1 = D_1D_2L(u, e)\cdot e\cdot e_1 + D_2D_2L(u, e)\cdot Z_2\cdot e_1.$$

Again we have a split into two separate conditions, taking respectively $e_1 = (\tilde{N}, 0)$ and $e_1 = (0, \tilde{h})$. Letting $Z = (\dot{N}, \dot{h})$ we get:

2)′ $$\begin{aligned} 0 = \int \tilde{N}\,\partial_g\mathscr{L}\left(g, \frac{h}{N}\right)\cdot h - \int \tilde{N}\,\partial_g\,\partial_h\mathscr{L}\left(g, \frac{h}{N}\right)\cdot h\cdot\frac{h}{N} \\ + \int \tilde{N}\dot{N}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{h}{N}\cdot\frac{h}{N^2} - \int \tilde{N}\,\partial_h^2\mathscr{L}\left(g, \frac{h}{N}\right)\cdot\frac{\dot{h}}{N}\cdot\frac{h}{N} \end{aligned}$$

and

$$(2)'' \quad \int N\partial_g\mathscr{L}\left(g,\frac{h}{N}\right)\tilde{h} = \int \partial_g\partial_h\mathscr{L}\left(g,\frac{h}{N}\right)\cdot h\cdot\tilde{h} - \int \dot{N}\partial_h^2\mathscr{L}\left(g,\frac{h}{N}\right)\cdot\tilde{h}\cdot\frac{h}{N^2} + \int \partial_h^2\mathscr{L}\left(g,\frac{h}{N}\right)\cdot\tilde{h}\cdot\frac{\dot{h}}{N}.$$

Condition (2)″ is just the condition for Lagrange's equation for $(h\cdot N) = k$ which we work out in a moment.

For now we want to focus our attention on the nontrivial condition (2)′. Since we are supposed to have complete degeneracy in $\mathscr{T}$, $\tilde{N}$ is arbitrary, so (2)′ is equivalent to

$$0 \equiv \partial_g\mathscr{L}\left(g,\frac{h}{N}\right)\cdot h - \partial_g\partial_h\mathscr{L}\left(g,\frac{h}{N}\right)\cdot h\cdot\frac{h}{N} + \dot{N}\partial_h^2\mathscr{L}\left(g,\frac{h}{N}\right)\cdot\frac{h}{N}\cdot\frac{h}{N^2} - \partial_h^2\mathscr{L}\left(g,\frac{h}{N}\right)\cdot\frac{\dot{h}}{N}\cdot\frac{h}{N}.$$

Setting $k = h/N$, this becomes

$$(2)''' \quad 0 \equiv \partial_g\mathscr{L}\left(g,\frac{h}{N}\right)\cdot h - \partial_g\partial_h\mathscr{L}\left(g,\frac{h}{N}\right)\cdot h\cdot\frac{h}{N} - \partial_h^2\mathscr{L}\left(g,\frac{h}{N}\right)\cdot\dot{k}\cdot\frac{h}{N}.$$

Let us take a curve $(\xi(t), g(t), N(t), h(t))$ tangent to Z which we suppose exists. Then (2)‴ says just that

$$0 \equiv \frac{\partial}{\partial t}\{\partial_h\mathscr{L}(g,k)\cdot k - \mathscr{L}(g,k)\}$$

or

$$\frac{\partial}{\partial t}[\mathscr{H}(g,k)\mu_g] \equiv 0.$$

This proves the theorem. ■

Using condition (2)″ above and the fact that $\mathscr{H}$ is conserved for the equations (2) of § 3.2, one can return to Theorem 3.3.1 and complete its proof. It is straightforward but a bit tedious. We leave the task to the reader. (It boils down to the fact that the gradient of $-2\int_M Nr(g)\,d\mu_g$ in the DeWitt metric is given by $-\text{grad}\,V = -2N\,\text{Ric}(g) + \frac{1}{2}Nr(g)g + 2\,\text{Hess}\,N$; here we pick up the extra term Hess N. This is proved in the same way as 2.4.2.)

3.5 The solution for a general lapse from $N = 1$

Next, we wish to explore the geometry of the Lapse function a little further. Previously we saw a geometrical way to solve the Einstein system

for general shift, if we knew how to solve it for $X = 0$. We want now to do a similar construction for the lapse.

If we use a Gaussian normal coordinate system for $-N^2\,dt^2 + g_{ij}\,dx^i\,dx^j$, then as was pointed out earlier, in the new coordinates the metric takes the form $-d\bar{t}^2 + \bar{g}_{ij}\,d\bar{x}^i\,d\bar{x}^j$. Now finding normal Gaussian coordinates means just that we compute the geodesics normal to $t = 0$ with initial tangent vector of unit length, and use these to define a coordinate system. Thus if we have the solution for a particular N we can construct one for $N = 1$. What about the converse? That is answered by an equally simple construction.

3.5.1 Theorem. *Let g be a solution of the Einstein system for $N = 1$. Construct a four metric on $\mathbb{R}\times M$ by setting*

$$l_{\alpha\beta}\,dx^\alpha\,dx^\beta = \frac{-dt^2}{N^2} + \frac{g_{ij}}{N^2}dx^i\,dx^j$$

and find its Gaussian normal coordinates. In these new coordinates the metric $-dt^2 + g_{ij}\,dx^j\,dx^i$ is transformed to $-N^2\,dt^2 + \bar{g}_{ij}\,dx^i\,dx^j$ which therefore solves the Einstein system for this N.

As one would expect, the proof is very simple. If we examine the condition that new coordinates $\bar{x}^\alpha$ transform N to $N = 1$ (i.e., Gaussian coordinates) we see that it is:

$$-1 = -N^2\left(\frac{\partial x^0}{\partial\bar{x}^0}\right)^2 + g_{kl}\frac{\partial x^k}{\partial\bar{x}^0}\frac{\partial x^l}{\partial\bar{x}^0}$$

$$0 = -N^2\frac{\partial x^0}{\partial\bar{x}^i}\frac{\partial x^0}{\partial\bar{x}^0} + g_{kl}\frac{\partial x^k}{\partial\bar{x}^0}\frac{\partial x^l}{\partial\bar{x}^i} \qquad i = 1, 2, 3.$$

Thus for any g_{kl} we can solve these partial differential equations for $\bar{x}^\alpha(x^\mu)$ with initial conditions $\bar{x}^\alpha(0, x^i) = (0, x^i)$ by using Gaussian coordinates. Similarly the condition that $\bar{x}^\alpha(x^\mu)$ transform $N = 1$ to N is

$$-N^2 = -\left(\frac{\partial x^0}{\partial\bar{x}^0}\right)^2 + g_{kl}\frac{\partial x^k}{\partial\bar{x}^0}\frac{\partial x^l}{\partial\bar{x}^0}$$

$$0 = -\frac{\partial x^0}{\partial\bar{x}^i}\frac{\partial x^0}{\partial\bar{x}^0} + g_{kl}\frac{\partial x^k}{\partial\bar{x}^0}\frac{\partial x^l}{\partial\bar{x}^i} \qquad i = 1, 2, 3.$$

These have the same form as the previous equations with $1/N^2$ for N^2 and g_{ij}/N^2 for g_{ij}. Hence the theorem. ■

3.6 The intrinsic shift

When we discussed the shift above, we saw that prescribing X_t was the same as prescribing arbitrarily a one parameter curve $\eta_t \in \mathscr{D}$. The curve η_t represents the "actual shift" rather than the shift vector field X_t.

Similarly for a lapse N, we can introduce, and also prescribe arbitrarily, the proper time function τ_t, a curve in $\mathscr{T}$. It is just the $\bar{x}^0$ coordinate in Gaussian normal coordinates. The relationship between the curve $\tau_t \in \mathscr{T}$ and the lapse N_t is however more subtle than in the case of the shift X_t. For the case of the shift, $X_t \circ \eta_t$ is just the tangent to the curve $\eta_t \in \mathscr{D}$,

$$\frac{d\eta_t}{dt} = X_t \circ \eta_t.$$

For the case of the lapse, however, we assert that the correct relationship is:

$$N_t = \frac{d\tau_t}{dt} \frac{1}{\sqrt{1 + \|\mathrm{grad}\, \tau_t\|^2}} \tag{1}$$

where $\|\mathrm{grad}\, \tau\|$ is computed with respect to the time-dependent metric. Also τ_0 is taken to be the zero function on M (all clocks start at noon), so that at $t = 0$, $N_0 = (d\tau_t/dt)_{t=0}$. (The meaning of $\xi_t(m)$ introduced earlier is the proper time of the curve $\sigma \mapsto (\sigma, m)$ with m fixed, while $\tau_t(m)$ is the proper time along a geodesic; cf. Wheeler [1].)

To see the relation (1), we write the equations which relate a coordinate system x^α to Gaussian coordinates $\bar{x}^\alpha$ (see § 3.5):

$$\begin{aligned} -1 &= -\frac{1}{N^2}\left(\frac{\partial \bar{x}^0}{\partial x^0}\right)^2 + g^{kl}\frac{\partial \bar{x}^0}{\partial x^k}\frac{\partial \bar{x}^0}{\partial x^l} \\ 0 &= -\frac{1}{N^2}\frac{\partial \bar{x}^0}{\partial x^0}\frac{\partial \bar{x}^i}{\partial x^0} + g^{kl}\frac{\partial \bar{x}^0}{\partial x^k}\frac{\partial \bar{x}^i}{\partial x^l}. \end{aligned} \tag{2}$$

Since $\tau_t = \bar{x}^0$, the first equation yields

$$N_t = \frac{d\tau_t}{dt}\frac{1}{\sqrt{1 + \|\mathrm{grad}\, \tau_t\|^2}}.$$

The presence of the extra factor takes into account the fact that the lapse in general depends on the space coordinates and therefore the proper time function τ_t pushes up the hypersurface M through $\mathbb{R} \times M$ unevenly; see Figure 3.2.

The question naturally arises if we can construct the rest of the Gaussian system from τ_t alone. Let the spatial coordinates be denoted φ_t so that (τ, φ) are the normal Gaussian coordinates. The surprising thing is not that we can find φ (we could just use N for that) but that φ is really a flow of a geometrically interesting vector field on M so that it behaves just like a shift. It describes the "tilting" of the Gaussian normal coordinates, again due to the fact that the lapse N_t depends on space and therefore causes a tilting of the hypersurface. This tilting is described by φ_t; we refer to its generator as the *intrinsic shift of N*.

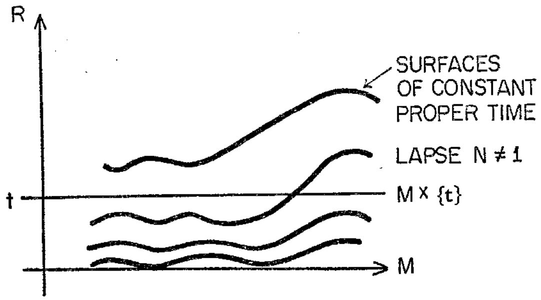


FIGURE 3.2

3.6.1 THEOREM. *Let τ_t be given. Then φ_t described above may be obtained by integrating the time dependent vector field Y_t on M given by*

$$Y_t = -\left(\frac{N_t}{\sqrt{1 + \|\mathrm{grad}\, \tau_t\|^2}}\right)\mathrm{grad}\, \tau_t \quad \textit{as follows:}$$

If ψ_t is the flow of Y_t, then $\varphi_t = \psi_t^{-1}$.

PROOF. The condition on φ_t is, from the second equation of (2)

$$\begin{aligned} \frac{\partial \tau_t}{\partial t}\frac{\partial \varphi_t^i}{\partial t} &= N_t^2 g_t^{kl}\frac{\partial \varphi_t^i}{\partial x^l}\frac{\partial \tau_t}{\partial x^k} \\ &= N_t^2 T\varphi_t \cdot \mathrm{grad}\, \tau_t \\ &= \left(\frac{d\tau}{dt}\right)^2 \frac{1}{1 + \|\mathrm{grad}\, \tau_t\|^2} T\varphi_t \cdot \mathrm{grad}\, \tau_t \end{aligned}$$

so that

$$\begin{aligned} \frac{d\varphi_t}{dt} &= \left(\frac{d\tau_t}{dt}\right)\left(\frac{1}{1 + \|\mathrm{grad}\, \tau_t\|^2}\right) T\varphi_t \cdot \mathrm{grad}\, \tau_t \\ &= -T\varphi_t \cdot Y_t \end{aligned}$$

where

$$-Y_t = \frac{d\tau_t}{dt}\frac{1}{1 + \|\mathrm{grad}\, \tau_t\|^2}\mathrm{grad}\, \tau_t.$$

Thus

$$-Y_t = T\varphi_t^{-1} \circ \frac{d\varphi_t}{dt}$$

$$= T\varphi_t^{-1} \circ \left(\frac{d\varphi_t}{dt} \circ \varphi_t^{-1}\right) \circ \varphi_t$$

$$= (\varphi_t^{-1})_* Z_t \text{ where } Z_t \text{ is the generator of } \varphi_t.$$

It follows that $Y_t = -(\varphi_t^{-1})^* Z_t$ is the generator of the flow φ_t^{-1}. ∎

Note. If Y is time-independent, then φ_t is the flow of $Y = -Z$.

In case the lapse N_t does not depend on space, then $N_t = (d\tau_t/dt)$, and the relation of solutions (g_t, k_t) to Einstein's equations with $N = 1, X = 0$, and $N_t = f(t)$, $X = 0$ is particularly simple. In fact if we define $\tau(t) = \int_0^t N_\lambda \, d\lambda$, then the solutions to the Einstein equations with $N_t = f(t)$, $X = 0$, are just reparameterizations by $\tau(t)$ of the solutions (g_t, k_t) of the Einstein system with $N = 1$, $X = 0$. We check this formally as follows.

3.6.2 PROPOSITION. *Let* $(g, k): \mathbb{R} \to \mathscr{M} \times S_2(M)$, $t \mapsto (g_t, k_t)$ *be a solution of the Einstein with* $N = 1$, $X = 0$. *Then given* $N_t = f(t)$ *we construct* $\tau: \mathbb{R} \to \mathbb{R}$: *by setting* $\tau(t) = \int_0^t N_\lambda \, d\lambda$. *Then the solution of the Einstein system with* $N_t = f(t)$ *and* $X = 0$ *and the same initial conditions* (g_0, k_0) *is given by*

$$\bar{g} = g \circ \tau \quad \text{and} \quad \bar{k} = k \circ \tau.$$

PROOF.

$$\bar{g}'(t) = g'(\tau(t)) \cdot \tau'(t) \quad \left(\tau' = \frac{d\tau}{dt}\right)$$

$$= N_t k(\tau(t))$$

$$= N_t \bar{k}_t$$

and

$$\bar{k}'(t) = N_t k'(\tau(t))$$

$$= N_t(S_{g(\tau(t))}(k(\tau(t))) - 2 \operatorname{Ric}(g(\tau(t)))$$

$$= N_t S_{\bar{g}(t)}(\bar{k}(t)) - 2N_t \operatorname{Ric}(\bar{g}(t)). \quad ∎$$

In this case, where the lapse is independent of the spatial point, we see that $Y_t = 0$ as we expect.

BIBLIOGRAPHY

Abraham, R.
[1] "Lectures of Smale on Differential Topology." (mimeographed, Columbia, 1961).
[2] "Foundations of Mechanics," Benjamin, 1967.
[3] *Piecewise Differentiable Manifolds and the Space-Time of General Relativity*, Journ. Math. and Mech. **11** (1962) 553–592.
[4] *Morphogenesis*, Publications du département de Mathématiques de Lyon (1972).

Abraham, R., and Robbin, J.
[1] "Transversal Mappings and Flows," W. A. Benjamin Inc., N.Y. 1967.

Adler, S., Bazin, R., and Schiffer, M.
[1] "Introduction to General Relativity," McGraw-Hill (1966).

Andrea, S. A.
[1] *The Hodge Existence Theorem* (I, II), Advances in Math. **6** (1971) 389.

Antman P., and Keller, J.
[1] "Bifurcation theory and Nonlinear Eigenvalue Problems," Benjamin (1969).

Andronov, A., and Chaikin, C.
[1] "*Theory of Oscillations*" (S. Lefschetz ed.), Princeton University Press 1949.

Antonelli, P. L., Burghelea, D., and Kahn, P. J.
[1] *The Nonfinite type of some* $\text{Diff}_0(M^n)$, Bull. Am. Math. Soc. **76** (1970) 1246–1250 and Topology **11** (1972) 1–50.

Arnold, V.
[1] *Sur la geometrié differentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluids parfaits*, Ann. Inst. Grenoble **16**(1) (1966), 319–361.

Arnowitt, R., Deser, S., and Misner, C. W.,
[1] *The Dynamics of General Relativity*, in "Gravitation; an introduction to current research," ed. L. Witten, Wiley, New York, 1962.
[2] *Dynamical Structure and Definition of Energy in General Relativity*, Phys, Rev. **116** (1959) 1322–1330.

Avez, A.
[1] *Essais de géométric riemannienne hyperbolique global—applications à la relativité général.* Ann. Inst. Four. Grenoble **13** (1963) 105–190.
[2] *Le Problème des Conditions Initiales*, 163–167 in "Fluides et champs gravitationnels en relativité générale." Colloques internationaux du C.N.R.S. no. 170, Paris (1967).

Bardos, C., and Tartar L.
[1] *About Backward Uniqueness for Evolution Equations and Some Related Questions*, Arch. Rat. Mech. and Anal. (to appear).

Bass, J.
[1] *Les fonctions pseudo-aléatoire*, Mémorial des Sciences Mathématique, Fascicule CLIII, Gauthier-Villars, Paris (1962).
[2] *Fonctions stationnaires. Fonctions de Corrélation. Application à la représentatis, Spatio-temporelle de la turbulence*, Ann. Inst. H. Poincaré, **5** (1969) 135–193.

Batchelor, G. K.
[1] "The Theory of Homogeneous Turbulence," Cambridge University Press (1953).

Berger, M., and Ebin, D.
[1] *Some Decompositions of the Space of Symmetric Tensors on a Riemannian Manifold*, J. Diff. Geom. **3** (1969) 379–392.

Birkhoff, G.
[1] "Hydrodynamics; a Study in Logic, Fact and Similitude," Princeton University Press (1950).

Bishop, R., and Crittenden, R.
[1] "Geometry of Manifolds," Academic Press, N.Y. 1964.

Bowen, R.
[1] *Markov Partitions for Axiom* A *diffeomorphisms*, Am. Journ. Math. **92** (1970) 725–747, 907–918.

Brezis, H.
[1] *On flows of* $W^{k,p}$ *vector fields*, (private communication).

Brezis H., and Pazy, A.
[1] *Semi-groups of nonlinear contractions on convex sets*. J. Funct. An. **6** (1970) 232–281, **9** (1972) 63–74.

Brill, D., and Deser, S.
[1] *Variational Methods and Positive Energy in Relativity*, Ann. Phys. **50** (1968) 548–570.

Browder, F.
[1] *Existence theorems for non-linear partial differential equations*, Proc. Symp. Pure Math. Am. Math. Soc. **XVI** (1970) 1–60.

Bruslinskaya, N. N.
[1] *The origin of cells and rings at near-critical Reynolds numbers*, Uspekhi Mat. Nauk. **20** (1965) 259–260.
[2] *The Behavior of Solutions of the Equations of Hydrodynamics when the Reynolds number passes through a critical value*, Doklady **6**(4) (1965) 724–728.
[3] *Qualitative Integration of a system of n differential Equations in a region containing a singular Point and a Limit cycle*, Dokl. Akad. Nauk, SSSR **139** (1961) 9–12.

Cantor, M.
[1] *The group* $\mathcal{D}^s_I(\mathbb{R}^n)$ (preprint).
[2] Thesis, Berkeley (to appear).

Cerf, J.
[1] *Topology de certains Espaces de plongements*, Bul. Soc. Math. France **89** (1961) 227–380.
[2] "Sur les diffeomorphismes de la sphère de dimension trois ($\Gamma_4 = 0$)." Springer Lecture Notes #53, 1968.

Chafee, N.
[1] *The Bifurcation of One or More Closed Orbits from an Equilibrium Point of an Autonomous Differential System*, Journ. Diff. Equations **4** (1968) 661–679.

Chandresekar, S.
[1] "Hydrodynamic and Hydromagnetic Stability," Oxford University Press (1961).

Chernoff, P.
[1] *Note on Product Formulas for Operator Semi-Groups*, J. Funct. An. **2** (1968) 238–242.

Chernoff, P., and Marsden, J.
[1] "Hamiltonian Systems and Quantum Mechanics," (in preparation).
[2] *On Continuity and Smoothness of Group Actions*, Bull. Am. Math. Soc. **76** (1970) 1044.
[3] *Quantum Mechanics as a Lagrangian System*, (preprint).

Chillingworth, D., (ed.).
[1] "Proceedings of the Symposium on differential equations and dynamical systems." Springer lecture Notes in Mathematics 206 (1971).

Choquet-Bruhat, Y.
[1] *Espaces-temps einsteiniens généraux chocs gravitationels*, Ann. Inst. Henri Poincaré, **8** (1968) 327–338.
[2] *Solutions* C^∞ *d'équations hyperboliques non linéares*, C. R. Acad. Sc. Paris, **272** (1971) 386–388.
[3] *New Elliptic System and Global Solutions for the Constraints Equations in General Relativity*, Comm. Math. Phys. **21** (1971) 211–218.

Choquet-Bruhat, Y., and Geroch, R.
[1] *Global Aspects of the Cauchy Problem in General Relativity*, Comm. Math. Phys. **14** (1969) 329–335.

Chorin, A. J.
[1] *Turbulence, Spectra and Equilibrium*, Comm. Pure and Appl. Math. (to appear).
[2] *On the Convergence of Discrete Approximations to the Navier-Stokes Equations*, Math. of Comp. **23** (1969) 341–353.

Coddington, E., and Levinson, N.
[1] "Theory of Ordinary Differential Equations," McGraw-Hill, N.Y. (1955).

Coles, D.
[1] *Transition in Circular Couette Flow*, Journ. Fluid Mech. **21** (1965) 385–425.

Courant, R., and Hilbert, D.
[1] "Methods of Mathematical Physics," vol. I, II, Interscience, New York, 1953, 1962.

Cronin, J.
[1] *One-Sided Bifurcation Points*, J. Diff. Eq'ns. **9** (1971) 1–12.

Deser, S.
[1] *Covariant Decomposition of Symmetric Tensors and the Gravitational Cauchy Problem*, Ann. Inst. H. Poincaré VII (1967) 149–188.

DeWitt, B.
[1] *Quantum Theory of Gravity, I The Canonical Theory*, Phys. Rev. **160** (1113–1148), 1967.
[2] *Space time as a Sheaf of Geodesics in Superspace*, in "Relativity," ed. M. Carmeli, S. Fickler, and L. Witten, Plenum Press (1970).

Dieudonné, J. A.
[1] "Foundations of Modern Analysis," Academic Press, New York, 1960.

Dionne, P.
[1] *Sur les problèmes de Cauchy bien posés.* J. Analyse Math. **10** (1962-3) 1–90.
Dirac, P. A. M.
[1] *Generalized Hamiltonian Dynamics*, Can. Journ. Math. **2** (1950) 129–148.
Duff, G. F. D.
[1] *Differential forms in manifolds with boundary*, Ann. of Math. **56** (1952), 115–127.
[2] *On Turbulent Solutions of the Navier-Stokes Equations* (unpublished).
Duff, G., and Spencer, D.
[1] *Harmonic tensors on Riemannian manifolds with boundary*, Ann. of Math. **56** (1952), 128–156.
Duff, G., and Naylor, D.
[1] "Partial Differential Equations of Applied Mathematics," Wiley, N.Y. (1969).
Dunford, N. and Schwartz, J.
[1] "Linear Operators," Vol. I, Interscience (1958).
Eardley, D., Liang, E., and Sachs, R.
[1] *Velocity dominated singularities in irrotational dust cosmologies*, Journ. Math. Phys. **13** (1972) 99–107.
Earle, C. J., and Eells, J.
[1] *A fiber bundle description of Teichmüller theory*, J. Diff. Geom. **3** (1969) p. 19–43 (also Bull. Am. Math. Soc. **73** (1967) 557–559).
Earle, C. J., and Schatz, A.
[1] *Teichmüller theory for surfaces with boundary*, J. Diff. Geom. **4** (1970) 169–186.
Ebin, D.
[1] *The manifold of Riemannian metrics*, in Proc. Symp. Pure Math. xv, Amer. Math. Soc. 1970, 11–40 and Bull. Am. Math. Soc. **74** (1968) 1002–1004.
Ebin, D. G., and Marsden, J.
[1] *Groups of diffeomorphisms and the motion of an incompressible fluid*, Ann. of Math. **92** (1970), 102–163. (See also Bull. Am. Math. Soc. **75** (1969) 962–967.)
[2] *Groups of diffeomorphisms and the motion of an incompressible fluid*, II (in preparation).
Eells, J.
[1] *On the geometry of function spaces*, in Symposium de Topoloqia Algebrica, Mexico (1958) 303–307.
[2] *A setting for global analysis*, Bull. Amer. Math. Soc. **72** (1966) 751–807.
Eells, J., and Sampson, J.
[1] *Harmonic Maps of Riemannian Manifolds*, Am. J. Math. **86** (1964) 109–160.
Eliasson, H.
[1] *Geometry of Manifolds of Maps*, J. Diff. Geom. **1** (1967), 169–194.
Epstein, D. B. A.
[1] *The simplicity of certain groups of homeomorphisms*, Compositio Math. **22** (1970) 163–173.

Fabes, E. B., Jones, B. F., and Riviers, N. M.
[1] *The initial value problem for the Navier-Stokes equations with data in L^p*, Proc. Symp. Pure Math. **23** AMS (to appear).
Faddeev, L. D.
[1] *Symplectic structure and quantization of the Einstein gravitation theory*, Actes du Congrès Intern. Math. **3** (1970) 35–40.
Feynman, R. P., Leighton, R. B., and Sands, M.
[1] "The Feynman Lectures on Physics (II)," Addison-Wesley Co. Reading, Mass. (1964).
Finn, R.,
[1] *On Steady State Solutions of the Navier-Stokes Equations*, Arch. Rat. Mech. **3** (1959) 381–396, Acta Math. **105** (1961) 197–244.
[2] *Stationary Solutions of the Navier-Stokes Equations*, Proc. Symposia Appl. Math. **17** (1965) 121–153.
Finn, R., and Smith, D. R.
[1] *On the Stationary Solutions of the Navier-Stokes Equations in Two Dimensions*, Arch. Rat. Mech. Anal. **25** (1967) 26–39.
Fischer, A.
[1] *The Theory of Superspace*, in "Relativity," ed. M. Carmeli, S. Fickler and L. Witten, Plenum Press (1970).
Fischer, A., and Marsden, J.
[1] *The Einstein Equations of Evolution—A Geometric Approach*, J. Math. Phys. **13** (1972) 546–568.
[2] *The Einstein Evolution Equations as a First Order Symmetric Hyperbolic Quasi-Linear System* I, Comm. Math. Phys. (to appear). II (in preparation).
[3] *General Relativity, Partial Differential Equations and Dynamical Systems.* Proc. Symp. Pure Math. **23** AMS (to appear).
[4] *The Existence of Complete Asymptotically Flat Space times* (in preparation).
Flanders, H.
[1] "Differential Forms," Academic Press, New York (1963).
Foias, C.
[1] *Une remarque sur l'unicité des solutions des equations de Navier-Stokes en dimension* n. Bull. Soc. Math. France **89** (1961) 1–8.
Foias, C., and Prodi, G.
[1] *Sur le comportement global des équations de Navier-Stokes en dimension* 2. Rend. Sem. Mat. Padova XXXIX (1967) 1–34.
Fourès-Bruhat, Y.
[1] *Théorème d'existence pour certains systèmes d'équations aux dérivèes partielles non linéaires*, Acta Math., **88** (141–225), 1952.
Friedrichs, K. O.
[1] *Symmetric hyperbolic linear differential equations*, Comm. Pure and Appl. Math., **7** (1954) 345–392.
[2] "Special topics in fluid dynamics," Gordon and Breach, N.Y. 1966.
Friefeld, C.
[1] *One Parameter Subgroups Do Not Fill a Neighborhood of the Identity in an Infinite Dimensional Lie (Pseudo) Group*, in "Battelle Rencontes," ed. C. M. DeWitt and J. A. Wheeler, W. A. Benjamin (1968).

Fujita, H., and Kato, T.
[1] *On the Navier-Stokes Initial Value Problem, I*, Arch. Rat. Mech. Anal. **16** (1964) 269–315.
Gerlach, R.
[1] *The Derivation of the Ten Einstein Field Equations From the Semiclassical Approximation to Quantum Geometrodynamics*, Phys. Rev. **117**(I) (1969) 1929–1941.
Geroch, R.
[1] *What is a Singularity in General Relativity?* Ann. of Phys. **48** (1968) 526–540.
[2] *Some recent work on global properties of spacetimes*, Actes congrès Intern. Math. **3** (1970) 41–46.
Golovkin, K.
[1] *About Vanishing Viscosity in the Cauchy Problem for the Equations of Fluid Mechanics*, Memoire of Steklov Inst. **XCII**, Moscow (1966).
Gordon, W.
[1] *The Riemannian Structure of Certain Function Space Manifolds*, J. Diff. Geom. **4** (1970), 499–508.
Graff, D.
[1] Thesis, Princeton University (1971).
Guynter, N.
[1] *On the Basic Problem of Hydrodynamics*, Trudy Fiz-Mat. Inst. Steklov, **2** (1927) 1–168. See also Izv. Akad. Nauk (1926) 1326, 1503 (1927) 621, 735, 1139, and (1928) 9.
Hall, W. S.
[1] *The Bifurcation of Solutions in Banach Spaces*, Trans. Am. Math. Soc. **161** (1971) 207–218.
Hartman, P.
[1] "Ordinary Differential Equations," Wiley, N.Y. (1964).
[2] *The Swirling Flow Problem in Boundary Layer Theory*, Arch. Rat. Mech. Anal. **42** (1971) 137–156.
Hermann, R.
[1] "Differential Geometry and the Calculus of Variations," Academic Press, N.Y. (1968).
[2] "Lectures on Mathematical Physics (I, II)", W. A. Benjamin, N.Y. (1971).
Herman, M. R.
[1] *Simplicité du groupe de diffeomorphismes de class* C^∞, *isotopes a l'identité, du tore de dimension* n, C. R. Acad. Sc, Paris t. 273 (1971) 232–234.
Herman, M., et Sergeraert, F.
[1] *Sur un théorème d'Arnold et Kolmogorov*, C. R. Acad. Sc. Paris t. 273 (1971) 409–411.
Heywood, J.
[1] *On Stationary Solutions of the Navier-Stokes Equations as Limits of Nonstationary Solutions*, Arch. Rat. Mech. and Anal. **37**, No. 1 (1970), 48–60. See also Acta Math. **129** (1972) 11–34.
Hirsch, M., and Pugh, C.
[1] *Stable Manifolds and Hyperbolic Sets*, Proc. Symp. Pure Math. **XIV**, Am. Math. Soc. (1970) 133–163.

Hirsch, M., Pugh, C., and Schub, M.
[1] *Invariant Manifolds*, Bull. Am. Math. Soc. **76** (1970) 1015–1019 and (in preparation).
Hodge, V. W. D.
[1] "Theory and Applications of Harmonic Integrals," Sec. Ed. Cambridge, 1952.
Holt, M. (ed)
[1] "Proceedings of the Second International Conference on Numerical Methods in Fluid Mechanics," Springer Lecture Notes in Physics, Springer-Verlag (1971).
Hopf, E.
[1] *Abzweigung einer periodischen Losung von einer stationaren Losung eines Differential systems*, Ber. Math-Phys. Sachsische akademie der Wissenschaften Leipzig **94** (1942) 1–22.
[2] *A Mathematical Example Displaying the Features of Turbulence*, Comm. Pure Appl. Math. **1** (1948) 303–322.
[3] *Uber die Anfanswert-aufgabe fur die hydrodynamischen Grundgleichungen*, Math. Nachr. **4** (1951) 213–231.
[4] *Repeated branching through loss of stability, An Example*, Proc. Conf. on Diff. Equations, Univ. of Maryland (1955).
[5] *Remarks on the Functional-Analytic Approach to Turbulence*, Proc. Symp. Appl. Math., **XIII**, Amer. Math. Soc. (1962) 157–163.
[6] *On the right weak solution of the Cauchy Problem for a quasi-linear equations of first order*, J. Math. Mech. **19** (1969/70) 483–487.
Il'in, V. P. (ed).
[1] "Boundary Value Problems of Mathematical Physics and Related Aspects of Function Theory" (I), Seminars in Mathematics (Leningrad) (5) Consultants Bureau, N.Y. (1969).
Iooss, G.
[1] *Contribution à la theorie nonlineare de la stabilité des ecoulements laminaires*, Thèse, Faculté des Sciences, Paris VI (1971).
Irwin, M. C.
[1] *On the Stable Manifold Theorem*, Bull. London Math. Soc. **2** (1970) 196–198.
Itaya, N.
[1] *On the Cauchy Problem for the System of Fundamental Equations Describing the Movement of Compressible Viscous Fluid*, Kodai Math. Sem. Rep. **23** (1971) 60–120.
Joseph, D. D. and Sattinger, D. H.
[1] *Bifurcating Time Periodic Solutions and their Stability* (to appear, Trans. Am. Math. Soc.).
Jost, R., and Zehnder, E.
[1] *A generalization of the Hopf Bifurcation Theorem*, Helv. Phys. Acta, June (1972).
Judovich, V.
[1] *Non Stationary Flows of an Ideal Incompressible Fluid*, Z. Vycisl. Mat. i. Fiz. **3** (1963), 1032–1066.
[2] *Two-Dimensional Nonstationary Problem of the Flow of an Ideal Incompressible Fluid Through a Given Region*, Mat. Sb. N.S. **64** (1964) 562–588.

[3] *Example of the generation of a Secondary Stationary or Periodic Flow when there is a loss of Stability of the Laminar Flow of a Viscous Incompressible Fluid*, Prikl, Mat. Meh. **29** (1965) 453–467—J. Appl. Math. Mech., **29** (1965) 527–544.

Kahane, C.
[1] *On the Spatial Analyticity of Solutions of the Navier-Stokes Equations*, Arch. Rat. Mech. Anal. **33** (1969), 386–405.

Kaniel, S., and Shinbrot, M.
[1] *A Reproductive Property of the Navier-Stokes Equations*, Arch. Rat. Mech. Anal. **24** (1967) 302–324.
[2] *Smoothness of Weak Solutions of the Navier-Stokes Equations*, Arch. Rat. Mech. Anal. **24** (1967) 302–323.

Kato, T.
[1] *On Classical Solutions of the Two-Dimensional Non-Stationary Euler Equation*. Arch. Rat. Mech. and Anal. **25** (3) (1967) 188–200.
[2] *Nonstationary Flows of Viscous and Ideal Fluids in* $\mathbb{R}^3$. Journ. Funct. An. **9** (1972) 296–305.
[3] *Linear Evolution Equations of "Hyperbolic" Type*, Journ. Fac. Sci. Univ. of Tokyo, Sect. 1, **XVII** (1970) 241–258.

Kiselev, A. A., and Ladyzhenskaya, O. A.
[1] *On Existence and Uniqueness of the Solution of the Non Stationary Problem for a Viscous Incompressible Fluid*, Izvestiya Akad. Nauk, SSSR **21** (1957) 655–680.

Knops, R. J., and Payne, L. E.
[1] *On the Stability of Solutions of the Navier-Stokes Equations Backwards in Time*, Arch. Rat. Mech. Anal. **29** (1968) 331–335.

Kodaira, K.
[1] *Harmonic Fields in Riemannian Manifolds*, Ann. of Math **50** (1949), 587–665.

Kolmogorov, A. N.
[1] *The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds Numbers*, C. R. Acad. Sci. USSR **30** (1941) 301.
[2] *Dissipation of Energy in Locally Isotropic Turbulence*, C.R. Acad Sci. USSR **32** (1941) 16.

Kraichnan, R. H.
[1] *The Structure of Turbulence at Very High Reynolds Numbers*, Journ. Fluid Mech. **5** (1959) 497–543.
[2] *The Closure Problem of Turbulence Theory*, Proc. Symp. Appl. Math. **XIII**, Am. Math. Soc. (1962) 199–225.
[3] *Isotropic Turbulence and Inertial Range Structure*, Phys. Fluids **9** (1966) 1728–1752.

Krikorian, N.
[1] *Differentiable Structures on Function Spaces* (to appear, Trans. Am. Math. Soc.).

Kunzle, H.
[1] *Degenerate Lagrangian Systems*, Ann. Inst. H. Poincaré (A) **XI** (1969) 393–414.

Ladyzhenskaya, O. A.
[1] *Solution "in the large" of Non Stationary Boundary Value Problem for the Navier-Stokes System with two Space Variables*, Comm. Pure and Appl. Math. **12** (1959) 427–433.
[2] "The Mathematical Theory of Viscous Incompressible Flow" (2nd Edition), Gordon and Breach, N.Y. 1969.
[3] *Example of Nonuniqueness in the Hopf Class of Weak Solutions for the Navier-Stokes Equations*, Math. USSR-Izvestija **3** (1969) 229–236.
[4] (ed.) "Boundary Value Problems of Mathematical Physics and Related Aspects of Function Theory" (II, III), Seminars in Mathematics (Leningrad) (7, 11) Consultants Bureau, N.Y. (1970).

Lanczos, C.
[1] "The Variational Principles of Mechanics" (2nd Edition), University of Toronto Press (1962).

Landau, L. D., and Lifshitz, E. M.
[1] "Fluid Mechanics," Addison-Wesley, Reading, Mass., 1959.

Lang, S.
[1] "Introduction to Differentiable Manifolds," Interscience, New York 1962. (Second Edition 1972).
[2] "Analysis II," Addison-Wesley, Reading, Mass., 1969.

Lax, P.
[1] *Cauchy's Problem for Hyperbolic Equations and the Differentiability of Solutions of Elliptic Equations*, Comm. Pure and Appl. Math. **8** (1955) 615–633.

Leray, Jean
[1] *Étude de diverses equations integrales non-lineares et de quelques problèmes que pose l'hydrodynamique*, Journ. Math. Pures, Appl. **12** (1933) 1–82.
[2] *Essai sur les mouvements plans d'un liquide vesqueux que limitent des parois*, J. de Math. **13** (1934). 331–418.
[3] *Sue le mouvement d'un liquide visqueux emplissant l'espace*, Acta Math **63** (1934) 193–248.
[4] *Problemes non-linéaires*, Ensign, Math. **35** (1936) 139–151.
[5] *Lectures on hyperbolic equations with variable coefficients*, Inst. for Adv. Stud., Princeton, 1952.

Leslie, J.
[1] *On a differential structure for the group of diffeomorphisms*, *Topology* **6** (1967), 263–271.
[2] *Some Frobenius Theorems in Global Analysis* J. Diff. Geom. **2** (1968) 279–297.
[3] *On two classes of Classical Subgroups of Diff(M)*, J. Diff. Geom. **5** (1971) 427–436.

Lichnerowicz, A.
[1] "Relativistic Hydrodynamics and Magnetohydrodynamics," Benjamin, 1967.

Lichtenstein, L.
[1] "Grundlagen Der Hydromechanik," Verlag Von Julius Springer, Berlin, 1929.

[2] *Uber einige Existenzprobleme der Hydrodynamik homogener unzusammendrückbarer, reibungsloser Flüssigkeiter und die Helmholtzschen Wirbelsatze* Math. **2.23** (1925), 89–154, 309–316; **26** (1927), 196–323; **28** (1928), 387–415, 725; **32** (1930), 608.

Lin, C. C.
[1] "The Theory of Hydrodynamic Stability," Cambridge University Press (1955).

Lions, J. L.
[1] *Sur la régularité et l'unicité des solutions turbulentes des equations de Navier Stokes.* Rndiconti Seminario Math. Univ. Padova **30** (1960) 16–23.
[2] *Singular perturbations and singular layers in variational inequalities.* in Zarantonello [1].

Lions J. and Prodi G.
[1] *Un théorème d'existence et unicité dan les equations de Navier Stokes en dimension 2.* C. R. Acad. Sci. Paris **248** (1959) 3519–21.

Marcus, M., and Mizel, V.
[1] *Functional Composition of Sobolev Spaces*, Bull. Am. Math. Soc. **78** (1972) 38–42.

Marsden, J.
[1] *Hamiltonian one parameter groups*, Arch. Rat. Mech. and Anal. **28** 1968 (362–396).
[2] Publications du Département de Mathématiques de Lyon (1972).
[3] *The Hopf bifurcation for nonlinear semigroups* (to appear);
[4] Smoothness of nonlinear semigroups, Can. Math. Bull. (to appear).

Marsden, J. and Abraham, R.
[1] *Hamiltonian Mechanics on Lie Groups and Hydrodynamics*, Proc. Pure Math. XVI, Amer. Math. Soc. (1970) 237–243.

Martin, M. H.
[1] *The flow of a viscous fluid, I*, Arch. Rat. Mech. Anal. **41** (1971) 226–286.

Mather, J.
[1] Appendix of Smale **[2]**

McGrath, F. J.
[1] *Nonstationary plane flow of viscous and ideal fluids*, Arch. Rat. Mech. Anal. **27** (5) (1968), 329.

McLeod, J. B. and Serrin, J.
[1] *The Existence of Similar Solutions for some Laminar Boundary Layer Problems*, Arch. Rat. Mech. Anal. **31** (1969), 288–303.

Meyer, K. R.
[1] *Generic Bifurcation of Periodic Points*, Trans. Am. Math. Soc. **149** (1970) 95–107.

Milnor, J.
[1] "Morse Theory," Princeton University Press (1963).

Misner, C.
[1] *Feynman Quantization of General Relativity*, Rev. Mod. Phys. **29** (1957) 497–509.

Montgomery, D.
[1] *On Continuity in Topological Groups*, Bull. Am. Math. Soc. **42** (1936) 879.

Morrey, C. Jr., and Eells, A.
[1] *A Variational Method in the Theory of Harmonic Integrals* **63** (1956) 91–128.

Morrey, C. B., Jr.
[1] *A Variational Method in the Theory of Harmonic Integrals II*, Am. J. Math. **78** (1956) 137–170.
[2] "Multiple Integrals in the Calculus of Variations," Springer, 1966.

Moser, J.
[1] *On the Volume Elements on a Manifold*, Trans. of Amer. Math. Soc. **120** (1965), 286–294.

Naimark, J.
[1] Dokl Akad. Nauk SSSR **129** (1959) No. 4.
[2] *Motions Close to Doubly Asymptotic Motion*, Soviet. Math. Dokl. **8** (1967) 228–231.

Nash, J.
[1] Le problème de Cauchy pour les equations differentielles d'un fluide genéral, Bull Soc. Math. France **90** (1962), 487–497.

Nelson, E.
[1] "Topics in Dynamics I, Flows." Princeton University Press (1969).

VonNeumann, J.
[1] *Recent Theories of Turbulence*, Collected Works VI, Macmillan, N.Y. 1963, 437–472.

Nirenberg, L.
[1] *On Elliptic Partial Differential Equations*, Annali delli Scuola Norm. Sup., Pisa **13** (1959) 115–162.

Omori, H.
[1] *On the Group of Diffeomorphisms on a Compact Manifold*, in Proc. Symp. Pure Math. XV, Amer. Math. Soc. (1970) 167–184.
[2] *Regularity of Connections*, Notes, Univ. of Warwick, Coventry, Feb. 1971.
[3] *On I.L.H. Properties of Mapping Spaces* (to appear).

Orszag, S. A.
[1] *Analytical theories of turbulence.* J. Fluid Mech. 41 (1970), 363–386.

Palais, R.
[1] "Seminar on the Atiyah-Singer Index Theorem," Princeton, 1965.
[2] *On the Homotopy Type of Certain Groups of Operators*, Topology **3** (1965) 271–279.
[3] *Homotopy Theory of Infinite Dimensional Manifolds*, Topology 5 (1966) 1–16.
[4] "Foundations of Global Non-Linear Analysis," Benjamin, N.Y., 1968.
[5] *The Morse Lemma on Banach Spaces*, Bull. Am. Math. Soc. **75** (1969) 968–971.
[6] *Extending Diffeomorphisms*, Proc. Am. Math. Soc. **11** (1960) 274–277.

Payne, L. E., and Weinberger, H. F.
[1] *An Exact Stability Bound for Navier-Stokes Flow in a Sphere in* "Non-Linear Problems," ed. Langer, University of Wisconsin Press (1962).

Penot, J. P.
[1] *Variété différentiables d'applications et de chemins.* C. R. Acad. Sc. Paris **264** (1967) 1066–1068.

[2] *Une méthode pour construire des variétés d'applications au moyen d'un plangement*, C. R. Acad. Sc. Paris **266** (1968) 625–627.
[3] *Géométrie des variétés fonctionnelles*, Thèse, Paris (1970).
[4] *Sur la théorèm de Frobenius*, Bull. Math. Soc. France **98** (1970) 47–80.
[5] *Topologie faible sur des varietés de Banach.* C. R. Acad. Sc. Paris **274** (1972) 405–408.

Petrovskii, I.
[1] *Über das Cauchysche problem für lineare und nichtlineare hyperbolische partielle Differentialgleichungen*, Rec. Math. (Mat. Sbornik), N.S. 2, **44** (1937) 814–868.

Pirani, F.
[1] "Lectures on General Relativity," Brandeis Summer Institute of Physics, Volume One (1964).

Quinn, B.
[1] *Solutions with Shocks: An Example of an L_1-Contractive Semigroup*, Comm. Pure and Appl. Math. **XXIV** (1971) 125–132.

Rabinowitz, P. H.
[1] *Existence and Nonuniqueness of Rectangular Solutions of the Bénard Problem*, Arch. Rat. Mech. Anal. **29** (1967) 30–57.
[2] *Some Global Results for Non-linear Eigenvalues Problems*, J. Funct. Anal. 7 (1971) 487–513.

Reynolds, O.
[1] *On the Dynamical Theory of Incompressible Viscous Fluids and the Determination of the Criterion*, Phil. Trans. Roy. Soc. London **A186** (1895) 123–164.

Robbin, J.
[1] *On the Existence Theorem for Differential Equations*, Proc. Amer. Math. Soc. **19** (1968) 1005–1006.
[2] *Stable Manifolds of Hyperbolic Fixed Points*, Illinois Journ. Math. **15** (1971) 595–609.
[3] *A Structural Stability Theorem*, Ann. of Math. **94** (1971) 447–493.

Rosenblat, M., and Van Atta, C. (ed.).
[1] "Statistical Models and Turbulence," Springer Lecture Notes in Physics #12 (1972).

Ruelle, D.
[1] *Statistical Mechanics on a compact set with a Z^ν action satisfying expansiveness and specification* Publ. du dept. math. Univ. Lyon (1972).
[2] *Dissipative systems and differential Analysis*, Boulder Lectures (1971) (to appear).
[3] *Strange Attractors as a Mathematical Explanation of Turbulence*, pp. 292–299 of Rosenblatt-Van Atta [1].
[4] *Bifurcations with Symmetries* (in preparation).

Ruelle, D., and Takens, F.
[1] *On the Nature of Turbulence*, Comm. Math. Phys. **20** (1971) 167–192. **23** (1971) 343–344.

Saber, J.
[1] *On Manifolds of Maps*, Thesis, Brandeis University (1965).

Sattinger, D. H.
[1] *Bifurcation of Periodic Solutions of the Navier-Stokes Equations*, Arch. Rat. Mech. Anal. **41** (1971) 66–80.
[2] *On Global Solution of Nonlinear Hyperbolic Equations.* Arch. Rat. Mech. Anal. **30** (1968) 148–172.
[3] *The Mathematical Problem of Hydrodynamic Stability*, J. Math. and Mech. **19** (1971) 797–817.

Schauder, J.
[1] *Das Anfangswertproblem einer quasilinearen hyperbolischen Differentialgleichungen zweiter Ordnung in beliebiger Anzahl von unabhangigen Veranderlichen*, Fund. Math., **24** (1935) 213–246.

Schnute, J. and Shinbrot, M.
[1] *The Cauchy Problem for the Navier-Stokes Equations* (preprint).

Sel'kov, E.
[1] *Self-Oscillations in Glycose*, Europ. Journ. Biochem. **4** (1968) 79–86.

Serrin, J.
[1] *Mathematical Principles of Classical Fluid Dynamics*, Encyclopedia of Physics, Vol. 8/1, Springer (1959).
[2] *On the Stability of Viscous Fluid Motions*, Arch. Rat. Mech. Anal. **3** (1959) 1–13.
[3] *The Initial Value Problem for the Navier-Stokes Equations*, in "Non Linear Problems," ed. Langer, University of Wisconsin Press (1962).
[4] "Mathematical Aspects of Boundary Layer Theory," Lecture Notes, Univ. of Minnesota, 1962 (out of print).
[5] *On the Interior Regularity of Weak Solutions of the Navier-Stokes Equations*, Arch. Rat. Mech. Anal. **7** (1962) 187–195.
[6] *On the Mathematical Basis for Prandtl's Boundary Layer Theory: An Example*, Arch. Rat. Mech. Anal. **28** (1968) 217–225.

Shinbrot, M.
[1] *Fractional Derivatives of Solutions of the Navier-Stokes Equations*, Arch. Rat. Mech. Anal. **40** (1971) 139–154.

Simon, C.
[1] *The Onset of Turbulence* (notes).

Sinai, Ya G.
[1] *Mesures invariantes des Y-systèmes*, Actes Congrès intern. Math. tome **2** (1970) 929–940.

Smale, S.
[1] *Diffeomorphisms of the Two Sphere*, Proc. Amer. Math. Soc. **10** (1959) 621–626.
[2] *Differentiable Dynamical Systems*, Bull. Am. Math. Soc. **73** (1967) 747–817.

Sniatycki, J.
[1] *Lagrangian Systems, Canonical Systems and Legendre Transformations* (to appear in volume 2 of these proceedings).

Snyder, H. A.
[1] Phys. Fluids, **11** (1968) 728–734, 1599–1605.

Sobolev, S. L.
[1] "Applications of Funcational Analysis in Mathematical Physics," Translations of Math. Monographs, Vol. 7, Am. Math. Soc., Providence, R.I. (1963).

Sobolevskii, P. E.
[1] *An Investigation of the Navier-Stokes Equations by Means of the Theory of Parabolic Equations in Banach Spaces*, Sov. Math. Dokl. **5** (1964) 720–723.

Swann, H. S. G.
[1] *The Convergence with Vanishing Viscosity of Nonstationary Navier-Stokes Flow to Ideal Flow in* $\mathbb{R}^3$, Trans. A.M.S. **157** (1971) 373–397.

Synge, J. L.
[1] "Relativity, The General Theory," North-Holland (1960).

Taub, A. H.
[1] *On Hamilton's Principle for Perfect Compressible Fluids*, Proc. Symp. Appl. Math. **I**, Amer. Math. Soc. (1949) 148–157.

Taylor, E. F., and Wheeler, A.
[1] "Spacetime Physics," W. H. Freeman, San Francisco (1966).

Taylor, G. I.
[1] *Statistical Theory of Turbulence*, parts 1–4, Proc. Roy. Soc. **151** (1935) 421.

Temam, R.
[1] *Une méthode d'approximation de la solution des équation de Navier-Stokes Equations*, Bull. Soc. Math. France (1970).
[2] *Quelques méthodes de décomposition en analyse numerique*, Actes Congrès Intern. Math. (1970) tome **3**, 311–319.

Tomonaga, S.
[1] Prog. Theoret. Phys. **1** (1946) 27.

Ton, B. A.
[1] *Strongly Nonlinear Parabolic Equations*, J. Funct. Anal. **7** (1971) 147–155.

Trotter, H. F.
[1] *On the Product of Semi-groups of Operators*, Proc. Amer. Math. Soc. **10** (1959) 545–551.

Turing, A. M.
[1] *The Chemical Basis of Morphogenesis*, Phil. Trans. Roy. Soc. (B) **237** (1952) 37–72.

Velte, W.
[1] *Über ein Stabilitatskiterium der Hydrodynamik*, Arch. Rat. Mech. Anal. **9** (1962) 9–20.
[2] *Stabilitats verhalten und Verzweigung stationaret Losungen der Navier Stokesschen Gleichungen*, Arch. Rat. Mech. An. **16** (1964) 97–125.
[3] *Stabilitat und verzweigung stationaret Losungen der Navier-Stokesschen Gleichungen beim Taylorproblem*, Arch. Rat. Mech. Anal. **22** (1966) 1–14.

Warner, F.
[1] "Foundations of Differentiable Manifolds and Lie Groups," Scott, Foresman (1971).

Williams, R. F.
[1] *One Dimensional Non-Wandering Sets*, Topology **6** (1967) 473–487.

Weyl, H.
[1] *The Method of Orthogonal Projection in Potential Theory*, Duke Math. J. **7** (1940) 411–444.

Weinstein, A.
[1] *Symplectic Manifolds and Their Lagrangian Submanifolds*, Advances in Math. **6** (1971) 329–346.

Wheeler, J. A.
[1] *Geometrodynamics and the Issue of the Final State*, in "Relativity, Groups and Topology", ed. DeWitt and DeWitt, Gordon and Breach, New York, 1964.

Wolibner, W.
[1] *Un théorème sur l'existence du mouvement plan d'un fluide parfait homogène, incompressible, pendant un temps infiniment longue*, Math. Z, **37** (1933) 698–726.

Yano, K.
[1] "Integral Formulas in Riemannian Geometry," Marcel Dekker, N.Y. (1971).

Yosida, K.
[1] "Functional Analysis," Springer, 1965.

Zarantonello, E. H. (ed.)
[1] "Contributions to Non-linear Functional Analysis" Academic Press, N.Y. (1971).

Remarks added in proof: (1) The simplicity of $\mathscr{D}(M)$ mentioned on page 166 has just been announced in the general case by B. Thurston (Berkeley graduate student).

(2) Marsden [3] shows that the ideas in Ruelle-Takens [1] can be applied to the Navier-Stokes equations. This provides an alternative to the power series methods used in Joseph and Sattinger [1].

(3) A useful general reference for bifurcation problems in Banach spaces is M. Crandall and P. Rabinowitz, *Bifurcation from Simple Eigenvalues*, Journ. Funct. An. 8 (1971) 321–340. Their theorem 1.7 seems to be more general than theorem 3.3.1, p. 175.

(4) Some recent work on bifurcations in biological and chemical systems has been done by S. Smale and N. Kopell. Smale has an example of an attracting closed orbit for a coupled system which have, separately, only an attracting fixed point. Kopell has used the Hopf bifurcation to study oscillations occurring in certain chemical systems related to the now celebrated "Jabotinsky reaction"; cf. Sel'kov [1].

(5) A proof of 1.37, p. 189 using similar ideas, but without the use of infinite dimensional manifolds has been given by H. Brezis and J. Bourgilignon (preprint).

(6) Results on non-singular asymptotically flat spacetimes, as mentioned on p. 248 have recently been discussed, independently, by L. Marder, Nature, **235** (1972) 379.