EDWARD NELSON

Topics in Dynamics

PRINCETON LEGACY LIBRARY

TOPICS IN DYNAMICS

I: FLOWS

BY

EDWARD NELSON

PRINCETON UNIVERSITY PRESS

AND THE

UNIVERSITY OF TOKYO PRESS

PRINCETON, NEW JERSEY

1969

L.C. Card: 79-108265

S.B.N.: 691-08080-1

A.M.S. 1968: 3404, 4615

Published in Japan exclusively
by the University of Tokyo Press;
in other parts of the world by
Princeton University Press

Printed in the United States of America

I. FLOWS

These are the lecture notes for the first term of a course on differential equations, given in Fine Hall the autumn of 1968.

It is a pleasure again to thank Miss Elizabeth Epstein for her typing.

I. FLOWS

In classical mechanics the state of a physical system is represented by a point in a differentiable manifold M and the dynamical variables by real functions on M . In quantum mechanics the states are given by rays in a Hilbert space $\mathcal{H}$ and the dynamical variables by self-adjoint operators on $\mathcal{H}$. In both cases motion is represented by a flow; that is, a one-parameter group of automorphisms of the underlying structure (diffeomorphisms or unitary operators).

The infinitesimal description of motion is in the classical case by means of a vector field and in quantum mechanics by means of a self-adjoint operator One of the central problems of dynamics is the integration of the equations of motion to obtain the flow, given the infinitesimal description of the flow.

1. Differential calculus

In recent years there has been an upsurge of interest in infinite dimensional manifolds. The theory has had important applications to Morse theory, transversality theory, and in other areas. It might be thought that an infinite dimensional manifold with a smooth vector field on it is a suitable framework for discussing classical dynamical systems with infinitely many degrees of freedom. However, classical dynamical systems of infinitely many degrees of freedom are usually described in terms of partial differential operators, and partial differential operators cannot be formulated as everywhere-defined operators on a

Banach space. We will be concerned only with finite dimensional manifolds. Despite this, I will begin by discussing the general case. I do this for two reasons: because the theory is useful in other branches of mathematics and because the fundamental concepts are clearer in the general context.

Let E be a real Banach space. That is, E is a real vector space with a function $x \rightsquigarrow \|x\|$ mapping E into the real numbers $\mathbb{R}$ such that $\|x\| \geq 0$, $\|x\| = 0$ only if $x = 0$, $\|ax\| = |a|\|x\|$, $\|x+y\| \leq \|x\| + \|y\|$, and E is complete: if $\|x_n - x_m\| \longrightarrow 0$ there is an x in E with $\|x_n - x\| \longrightarrow 0$. For example, E may be an s-dimensional Euclidean space $\mathbb{R}^s$ in the norm $\|x\| = (x_1^2 + \ldots + x_s^2)^{\frac{1}{2}}$. If F is another Banach space we denote by $L(E,F)$ the Banach space of all continuous linear mappings of E into F in the norm $\|A\| = \sup\{\|Ax\| : \|x\| \leq 1\}$. We abbreviate $L(E,E)$ by $L(E)$.

Let U be an open subset of the Banach space E, and let x be in U A function $f: U \longrightarrow F$ (where F is a Banach space) is said to be (Fréchet) <u>differentiable at</u> x in case there is an element $Df(x)$ of $L(E,F)$ such that

$$f(x+y) = f(x) + Df(x)y + o(y) ,$$

where $o(y)$ is a function defined in a neighborhood of 0 such that $\|o(y)\|/\|y\| \longrightarrow 0$ as $y \longrightarrow 0$ with $y \neq 0$. It is clear that $Df(x)$ is unique if it exists. It is called the (Fréchet) <u>derivative of</u> f <u>at</u> x. The function $f: U \longrightarrow F$ is called <u>differentiable</u> in case it is differentiable at all points x in U, and it is called C^1 in case it is differentiable and $x \rightsquigarrow Df(x)$ is continuous from U to $L(E,F)$. If f is C^1 then Df is a function from U into the Banach space $L(E,F)$, so it makes sense to ask whether Df is C^1.

The function f is said to be C^2 in case f is C^1 and Df is C^1 and, by recursion, f is said to be C^k in case f is C^1 and Df is C^{k-1}. (A trivially equivalent definition is that f is C^k in case f is C^{k-1} and $D^{k-1}f$ is C^1. Sometimes one definition and sometimes the other suggests the more convenient way to organize an induction proof to show that f is C^k.) Similarly, we define f to be k times differentiable in case it is differentiable and Df is $k-1$ times differentiable (or equivalently, in case it is $k-1$ times differentiable and $D^{k-1}f$ is differentiable). Notice that if f is differentiable at x it is continuous at x. Consequently a differentiable function is continuous, and a k times differentiable function is C^{k-1}.

Let $E_1,\ldots,E_n$ be Banach spaces, and consider their Cartesian product $E_1 \times\ldots\times E_n$. It is possible to give this a Banach space structure by defining addition and scalar multiplication componentwise and giving an element the norm which is the sum of the norms of its components. This Banach space is denoted by $E_1 \oplus\ldots\oplus E_n$ and called the direct sum of the Banach spaces $E_1,\ldots,E_n$. Elements of it are denoted by $x_1 \oplus\ldots\oplus x_n$, where x_i is in E_i. Frequently we wish to consider multilinear forms on $E_1 \times\ldots\times E_n$; that is, functions on $E_1 \times\ldots\times E_n$ which are linear in each variable separately. If f is also a Banach space, we let $L(E_1 \times\ldots\times E_n, F)$ be the Banach space of all continuous multilinear forms on $E_1 \times\ldots\times E_n$ with values in F, with the norm

$$\|A\| = \sup\{\|A(y_1,\ldots,y_n): \|y_1\|,\ldots,\|y_n\| \leq 1\} .$$

This Banach space may be identified with the Banach space

$$L(E_1,\ldots,L(E_{n-1},L(E_n,F))\ldots)$$

under the identification which takes an element A of the latter into the form given by

$$A(y_1,\ldots,y_n) = ((\ldots(Ay_1)\ldots)y_{n-1})y_n \ .$$

If $E_1 = \ldots = E_n = E$, we abbreviate $L(E_1 \times\ldots\times E_n,F)$ by $L^n(E,F)$. The set of symmetric elements of it is denoted by $L^n_{sym}(E,F)$. If A is in $L(E_1 \times\ldots\times E_n,F)$ we denote the value $A(y_1,\ldots,y_n)$ by $Ay_1 \ldots y_n$. Also, if y is in E then y^n means $(y,\ldots,y)$ n times, so that Ay^n is defined if A is in $L^n(E,F)$. If $f\colon U \longrightarrow F$ (with U open in E) is k times differentiable then D^kf takes values in $L^n(E,F)$.

<u>Theorem</u> 1 (<u>product rule</u>). <u>Let</u> E, F_1, F_2, <u>and</u> G <u>be Banach spaces, let</u> U <u>be open in</u> E, <u>let</u> $f\colon U \longrightarrow F_1$ <u>and</u> $g\colon U \longrightarrow F_2$ <u>be</u> C^k, <u>let</u> $(z_1,z_2) \rightsquigarrow z_1 \cdot z_2$ <u>be in</u> $L(F_1 \times F_2,G)$ <u>and define</u> $f\cdot g$ <u>by</u> $(f\cdot g)(x) = f(x)\cdot g(x)$. <u>Then</u> $f\cdot g\colon U \longrightarrow G$ is C^k <u>and</u>

$$D(f\cdot g)(x)y = Df(x)y\cdot g(x) + f(x)\cdot Dg(x)y \ . \tag{1}$$

<u>Proof</u>. Suppose f and g are C^1. Then

$$f(x+y) = f(x) + Df(x)y + o(y)$$

$$g(x+y) = g(x) + Dg(x)y + o(y)$$

so that

$$f(x+y)\cdot g(x+y) = f(x)\cdot g(x) + Df(x)y\cdot g(x) + f(x)\cdot Dg(x)y + o(y) \ .$$

Thus $f\cdot g$ is C^1 and (1) holds, so that the theorem is proved for $k = 1$ Suppose the theorem to be true for $k-1$, and let f and g be C^{k-1}.

Then (1) holds. The mapping $\mu\colon L(E,F_1) \times F_2 \longrightarrow L(E,G)$ given by $(A,z) \rightsquigarrow B$, where $By = Ay\cdot z$, is continuous and bilinear. Now Df and g are C^{k-1}, so by the theorem for $k-1$, $x \rightsquigarrow \mu(Df(x),g(x))$ is C^{k-1}, and similarly for the other term. Therefore $D(f\cdot g)$ is C^{k-1}, so $f\cdot g$ is C^k. This concludes the proof.

The same proof shows that the theorem with "C^k" replaced by "k times differentiable" is true.

<u>Theorem</u> 2 (<u>chain rule</u>). <u>Let</u> E, F, <u>and</u> G <u>be Banach spaces</u>, <u>let</u> U <u>be open in</u> E, <u>let</u> V <u>be open in</u> F, <u>and let</u> $f\colon U \longrightarrow V$ <u>and</u> $g\colon V \longrightarrow G$ <u>be</u> C^k. <u>Then</u> $g\circ f$ <u>is</u> C^k <u>and</u>

(2) $$D(g\circ f)(x) = Dg(f(x))Df(x) .$$

<u>Proof.</u> Suppose f and g are C^1. Then

$$f(x+y) = f(x) + Df(x)y + o(y) ,$$

$$\begin{aligned}(g\circ f)(x+y) &= g(f(x+y)) \\ &= g(f(x)) + Dg(f(x))(Df(x)y + o(y)) + o(Df(x)y + o(y)) \\ &= g(f(x)) + Dg(f(x))Df(x)y + o(y) .\end{aligned}$$

Hence $g\circ f$ is C^1 and (2) holds. Thus the theorem holds for $k = 1$. Suppose the theorem to be true for $k-1$, and let f and g be C^k. Then Dg and f are C^{k-1}, so $Dg\circ f$ is C^{k-1}. Also Df is C^{k-1}. The mapping of $L(F,G) \times L(E,F)$ into $L(E,G)$ which takes two linear operators into their product is continuous and bilinear, so by Theorem 1, $(Dg\circ f)(Df)$ is C^{k-1}. By (2), therefore, $D(g\circ f)$ is C^{k-1} and $g\circ f$ is C^k, which completes the proof.

The following formulas are easily proved by induction, for C^k

functions f and g :

$$D^k(f\cdot g)(x)y_1 \ldots y_k = \sum_{q=0}^{k} \Sigma\, D^q f(x)y_{i_1} \ldots y_{i_q} \cdot D^{k-q}g(x)y_{j_1} \ldots y_{j_{k-q}} ,$$

where the inner sum is over all $\binom{k}{q}$ partitions of $y_1,\ldots,y_k$ into two sets with $i_1 <\ldots< i_q$ and $j_1 <\ldots< j_{k-q}$, and

$$D^k(g\circ f)(x)y_1\ldots y_k =$$

$$\sum_{q=1}^{k} \Sigma\, D^q g(f(x))D^{r_1}g(x)y_1^{(1)}\ldots y_{r_1}^{(1)}D^{r_2}g(x)y_1^{(2)}\ldots y_{r_2}^{(2)}\ldots D^{r_q}g(x)y_1^{(q)}\ldots y_{r_q}^{(q)},$$

where the inner sum is over all $k!/r_1!\ldots r_q!$ partitions of $y_1,\ldots,y_k$ into q sets with $r_1,r_2,\ldots,r_q$ elements and the natural ordering in each set.

Let us define

$$(D^q f\cdot D^{k-q}g)(x)y_1\ldots y_k = D^q f(x)y_1\ldots y_q \cdot D^{k-q}g(x)y_{q+1}\ldots y_k$$

and

$$(D^{r_1}g\cdot D^{r_2}g\ldots D^{r_q}g)(x)y_1\ldots y_k =$$

$$D^{r_1}g(x)y_1\ldots y_{r_1}D^{r_2}g(x)y_{r_1+1}\ldots y_{r_1+r_2}\ldots D^{r_q}g(x)y_{r_1+\ldots r_{q-1}+1}\ldots y_k .$$

We shall see later that if f is C^k then $D^k f$ is symmetric. Let us denote by Sym the symmetrizing operator; that is, if $\varphi \in L^k(E,F)$ then Sym φ in $L^k_{sym}(E,F)$ is defined by

$$(\mathrm{Sym}\ \varphi)(y_1,\ldots,y_k) = \frac{1}{k!}\sum_{\pi} \varphi(y_{\pi(1)},\ldots,y_{\pi(k)}) ,$$

where the summation is over all permutations π of $1,\ldots,k$. Then we may write

$$D^k(f\cdot g) = \operatorname{Sym} \sum_{q=0}^{k} \binom{k}{q} D^q f \cdot D^{k-q} g ,$$

$$D^k(g\circ f) = \operatorname{Sym} \sum_{q=1}^{k} \sum_{r_1+\dots+r_q=k} \frac{k!}{r_1!\dots r_q!} (D^q f)\circ g \cdot D^{r_1} g\cdot D^{r_2} g\dots D^{r_q} g .$$

(The formulas on p.3 of [6] should be corrected to take symmetrization into account.)

The following is another proof of a theorem of Abraham [6, p.6]. By $o(y^k)$ we mean a function such that $o(y^k)/\|y\|^k \longrightarrow 0$ as $y \longrightarrow 0$ with $y \neq 0$.

<u>Theorem</u> 3 (<u>converge of Taylor's theorem</u>). <u>Let</u> E <u>and</u> F <u>be Banach spaces, let</u> U <u>be open in</u> E , <u>and suppose that</u> $f: U \longrightarrow F$ <u>satisfies</u>

$$(3) \qquad f(x+y) = a_0(x) + a_1(x)y + \frac{a_2(x)}{2!} y^2 + \dots + \frac{a_k(x)}{k!} y^k + o(y^k)$$

<u>where the</u> $a_j(x)$ <u>are in</u> $L^j_{sym}(E,F)$ <u>and each</u> a_j <u>is continuous. Then</u> f <u>is</u> C^k <u>and</u> $a_j = D^j f$ <u>for</u> $j = 0,1,\dots,k$.

<u>Proof.</u> For $k = 1$ this is the definition. Suppose the theorem is true for $k-1$. Then in (3), since $(a_k(x)/k!)y^k = o(y^{k-1})$, we know that $a_j = D^j f$ for $j = 0,1,\dots,k-1$. Now let us expand $f(x+y+z)$ in two different ways:

$$f(x+y+z) = f(x+y) + Df(x+y)z + \dots + \frac{1}{(k-1)!} D^{k-1}f(x+y)z^{k-1} + \frac{a_k(x+y)}{k!} z^k + o(z^k) ,$$

$$f(x+y+z) = f(x) + Df(x)(y+z) + \dots + \frac{1}{(k-1)!} D^{k-1}f(x)(y+z)^{k-1} + \frac{a_k(x)}{k!} (y+z)^k + o((y+z)^k) .$$

Fix x and restrict z so that $\frac{1}{4}\|y\| \le \|z\| \le \frac{1}{2}\|y\|$. Then it does not matter whether we write $o(z^k)$, $o((y+z)^k)$, or $o(y^k)$. Subtract the two equations, collecting coefficients of z and denoting the coefficient of z^j by $g_j(y)$. Then

$$(4) \qquad g_0(y) + g_1(y)z + \ldots + g_{k-1}(y)z^{k-1} + g_k(y)z^k = o(y^k) .$$

Now

$$g_k(y)z^k = \frac{1}{k!}[a_k(x+y) - a_k(x)]z^k ,$$

and by the continuity of a_k this is $o(y^k)$, so we may drop this term. We claim that each term separately in (4) is $o(y^k)$. To see this, let $\lambda_1, \ldots, \lambda_k$ be distinct numbers, and replace z by $\lambda_i z$ for $i = 1, \ldots k$. In this way we obtain k equations which we write as

$$\begin{pmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{k-1} \end{pmatrix} \begin{pmatrix} g_0(y) \\ g_1(y)z \\ \vdots \\ g_{k-1}(y)z^{k-1} \end{pmatrix} = \begin{pmatrix} o(y^k) \\ o(y^k) \\ \vdots \\ o(y^k) \end{pmatrix}$$

Since the λ_i are distinct, the matrix is invertible (it is the Vandermonde matrix with determinant $\prod_{i<j} (\lambda_j - \lambda_i))$. Therefore each $g_j(y)z^i$ is $o(y^k)$. In particular, this is true for $j = k-1$. But (and here we use the symmetry of a_k)

$$g_{k-1}(y)z^{k-1} = \left[\frac{D^{k-1}f(x+y)}{(k-1)!} - \frac{D^{k-1}f(x)}{(k-1)!} - \frac{ka_k(x)y}{k!}\right]z^{k-1} .$$

Therefore the term in brackets is $o(y)$. By definition of the derivative, this means that $D^k f(x) = a_k(x)$, and since a_k is continuous, f is C^k . This completes the proof.