GEOMETRICAL PHYSICS

Part I: Manifolds

(Lecture Notes for Math. 259A)

by

Theodore Frankel

Copyright: September 23, 1986.

Preface

These notes are for the first quarter of a 3 quarter course (Ma 259 ABC) in geometrical tools and notions useful for theoretical physics and engineering.

My aim in Ma. 259A has been to give the reader a working knowledge of manifolds (including their elementary topology), vector fields and exterior differential calculus, so that further reading in texts (e.g. Arnold's book on mechanics) and in research papers would be easier. To achieve this, it is absolutely necessary that the reader do the assigned problems. Almost all of the problems are simple when attacked by the methods presented; if you are getting into horrible computations, you probably are not taking advantage of the machinery that has been developed in the lectures.

Each written lecture will be covered in a single class lecture. Since the written lecture usually contains more material than can be covered orally in 50 minutes, I'll try to give enough of the important ideas in class so that the reader can easily finish reading the lecture afterwards.

A glance at the table of contents will show how physical applications (mechanics, electromagnetism, thermodynamics) occur during the course. I ask that the physics and engineering students have patience; I introduce applications as early as I feel is possible subject to the constraint that the geometric machinery and appropriate ideas should be available to give the best understanding of the physical notions and problems.

Ma 259 B will cover Riemannian geometry (i.e curvature) with applications in particular to general relativity.

Ma 259C will concern itself with bundles, connections and characteristic classes, with applications including Yang Mills fields.

I would like to thank Poul Hjorth for many discussions about the course.

I am indebted to Prof. Hidenori Murakami for his encouragement and belief that such a set of class notes would be useful.

Finally I wish to thank Neda Crimmins for creating such an attractive manuscript.

Ted Frankel
La Jolla, Sept 22, 1986.

Books Referred to in the Notes for Ma. 259A

[A] : Arnold, V, Mathematical methods of classical mechanics Springer, 1978.

[A,M,R] : Abraham, R. Marsden, J, and Ratiu, T, Manifolds, Tensor analysis and applications, Addison Wesley, 1983.

[C] : Chandrasekhar, S, An introduction to the study of Stellar Structure, Dover, 1939

[Fe] : Feynman, R, Leighton, R, and Sands M, The Feynman Lectures on Physis, II, Addison Wesley, 1964.

[Fr] : Frankel, T., Gravitational Curvature, Freeman, 1979.

[G,P] : Guillemin, V and Pollack, A, Differential Topology, Prentice-Hall, 1974.

Table of Contents

Lecture 1. Submanifolds of Euclidean Space 1

1. $\mathbb{R}^n$ = Euclidean n space 1
2. Submanifolds of $\mathbb{R}^{n+r}$ 1
   Problem 1 2
3. Geometry of Jacobian Matrices; the "Differential" 2
4. The Main Theorem on Submanifolds of $\mathbb{R}^{n+r}$ 4
5. A nontrivial example; the configuration space of a rigid body 4
   Problem 2 6
   Problem 3 6

Lecture 2. Manifolds 7

1. Idea of a Manifold 7
2. Examples 7
   Problem 4 10
   Problem 5 12
3. A Rigorous Definition of Manifold 13
   Problem 6 15

Appendix on Point Set Topology 14

Lecture 3. Vectors and Mappings 16

1. Tangent, or "Contravariant" vectors 16
   Problem 7 19
2. Submanifolds of manifolds 19
3. The Inverse Function Theorem 22

Lecture 4. Vector Fields and Flows 23

1. Vector fields and flows 23
   Problem 8 25
2. Vector Fields on Manifolds 26

Appendix on "Fundamental Theorems" 27

Lecture 5. Covectors and Riemannian Metrics 31

1. Some Linear Algebra; the Dual Space 31
2. The Differential of a Function 33
   Problem 9 35
3. Riemannian Manifolds; the Gradient Vector 35
   Problem 10 40
4. Morse deformations 40
   Problem 11 41

Lecture 6. Tangent and Cotangent Bundles. Phase Space 42

1. The tangent bundle 42
2. The Cotangent Bundle 45
3. Phase space of Mechanics 46
4. Induced mappings on vectors and forms 48
   Problem 12 50
5. The Poincaré 1-form on phase space T^*M^n 50
   Problem 13 51

Lecture 7. Tensors 52

1. Linear algebra again 52
   Problem 14 53
2. Transformation Properties 56
3. Einstein Summation Convention 56
4. Tensor Fields on Manifolds 56
   Problem 15 57
5. Pull-backs, and a digression into elasticity 57

Lecture 8. Exterior or Grassmann Algebra 60

1. The Tensor Product 60
2. The Grassmann or Exterior Algebra 60
   Problem 16 65
   Problem 17 66
3. Geometric meaning of forms 66
4. Special cases of the exterior product 67
5. Computations and relation with vector analysis 68
   Problem 18 69

Lecture 9. Exterior Differentiation 70

1. The Exterior Differential d 70
2. Examples in $\mathbb{R}^3$ 73
3. A concise coordinate expression for d 74

Lecture 10. Pull-Backs 75

   Problem 19 76

Lecture 11. Orientation and Pseudo-forms 80

1. Orientation of a manifold 80
   Problem 20 81
   Problem 21 83
2. 1-sided, 2-sided 83
   Problem 22 83
3. Pseudo-forms $\equiv$ twisted forms $\equiv$ forms of odd kind 84
4. The Volume Pseudo n-form on a Riemannian Manifold 85
   Problem 23 86

28-34

Lecture 12. Interior Products and Vector Analysis in $\mathbb{R}^3$ 87

1. Interior product (contraction) 87
Problem 24 88
2. Vector analysis 88
Problem 25 93

Lecture 13. Integration of Forms Over Parameterized Subsets 94

1. Integration of a (true) p-form in $\mathbb{R}^p$ 94
2. Integratiuon over parameterized subsets 94
3. The integrals are independent of the paramerization 100
Problem 26 101
Problem 27 102

Lecture 14. Integration and Partitions of Unity 103

1. Manifolds with boundary 103
2. Partition of Unity 105
3. Integration of forms over compact oriented submanifolds of M^n 107
4. Manifolds admit Riemannian metrics 108

Lecture 15. Stokes theorem for True Forms 109

1. Orienting the boundary of an oriented manifold 109
2. Stokes' Theorem 110

Lecture 16. Integration of Pseudo-forms. Maxwell's Equations 114

1. Integration of pseudo-forms 114
2. Stokes' theorem for pseudo-forms 117
3. Electromagnetism and Maxwells equations 118
4. Maxwell's equations in "differential form" 121

Lecture 17. The Lie Derivative of a Vector Field 124

1. The Lie bracket 124
Problem 28 126
2. Invariant vector fields; Jacobi's variational equation 126
3. The Flow Associated to $\mathcal{L}_X Y = [X,Y]$ 128

Lecture 18. Lie Derivatives of Exterior Forms 131

1. Lie derivatives of forms 131
Problem 29 134
2. Cartan's formula 134
Problem 30 135
Problem 31 135
Problem 32 137
3. Vector analysis again 137

Lecture 19. Differentiation of Integrals. The Divergence Theorem 140

1. Time independent case 140
2. The Divergence Theorem 141
3. The time dependent case 142
Problem 33 146
4. Euler's equations for a perfect fluid 146
Problem 34 46

Lecture 20. A Problem Set on Hamiltonian Mechanics 148

1. Time-independent Hamiltonians 148
2. Time dependent Hamiltonians 155
Problems M.1 to M.12

Lecture 21. Volterra's Theorem 159

1. Stokes' theorem again 159
2. Closed forms and exact forms 159
Problem 35 159
3. The converse of the Poincaré lemma 163
Problem 36 167
4. Finding potentials without Volterra's formula 167
Problem 37 168

Lecture 22. Topology, deRham, and Electromagnetism 169

Lecture 23. Topological Aspects of Maxwell's Equations 176

Problem 38 176
1. Compact spaces 176
Problem 39 178
2. Magnetic fields in a wormhole topology 179
3. Currents in a wormhole topology 182
Problem 40 184

Lecture 24. Brouwer Degree and Kronecker Index 188

1. The Brouwer degree of a map 188
Problem 41 191
2. Kronecker index of a vector field 191
Problem 42 192
Problem 43 192
Problem 44 193
Problem 45 193
3. Gauss looping integral 194
Problem 46 195
Problem 47 195

Lecture 25. The Frobenius Integrability Theorem 196

1. Planes in $\mathbb{R}^3$ 196
2. Distributions and Vector Fields 198
3. Distributions and Pfaffians (1-forms) 198
4. The Frobenius Theorem 201
   Problem 48 202

Lecture 26. Integrability and Constraints 205

1. Foliations and the maximal leaves 205
2. Systems of Mayer-Lie 207
   Problem 49 208
3. Holonomic and nonholonomic constraints 209

Lecture 27. Thermodynamics à la Caratheodory 213

1. The first law of thermodynamics 213
2. Elementary changes of state 214
3. The second law 215
4. Behavior of brackets under mapping 216
5. Caratheodory's Theorem 218
6. Integrating factors and the second law 220

Lecture 28. Poincaré's Theorem: A Preview of characteristics Classes 222

1. The sum of the indices of a vector field 222
2. $\chi(M)$ and the Stiefel vector field 227
3. Concluding remarks 231

GEOMETRICAL PHYSICS

Lecture 1. Submanifolds of Euclidean Space

(1) $\mathbb{R}^n$ = Euclidean n space, is not necessarily considered as a vector space. There are many global coordinates systems $x = (x^1, \ldots, x^n)$. This is the most important n dimensional manifold (defined later).

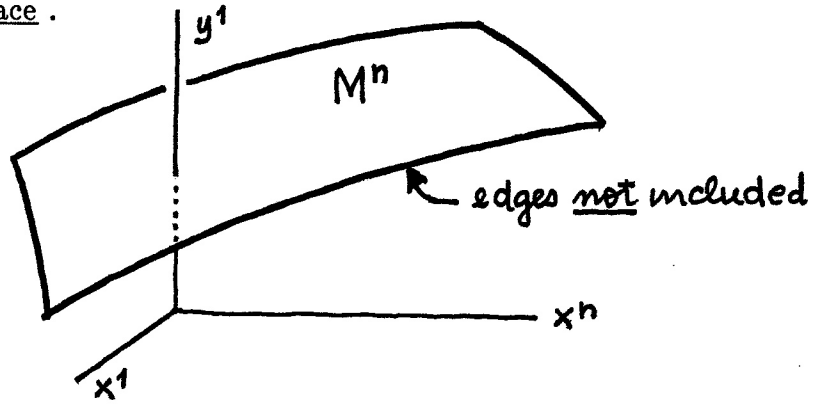
(2) Submanifolds of $\mathbb{R}^{m+r}$. $M^n \subset \mathbb{R}^{n+r}$ is an n-dimensional submanifold of $\mathbb{R}^{n+r}$ if locally M^n can be described in some coordinate system $(x^1, \ldots, x^n, y^1, \ldots, y^r) = (x,y)$ of $\mathbb{R}^{n+r}$ by r smooth (or at least continuously differentiable) functions

$$y^\alpha = f^\alpha(x^1, \ldots, x^n), \quad \text{i.e.,} \quad y = f(x) \quad .$$

$(x^1, \ldots, x^n)$ then forms a local (curvilinear) coordinate system on M^n.

Examples.

(a) $y^1 = f(x^1, \ldots, x^n)$ describes an n dimensional submanifold in $\mathbb{R}^{n+1}$. In this case $(x^1, \ldots, x^n)$ forms a global curvilinear coordinate system on M^n. When $n = 2$, M is a surface.



(b) Given r functions $F^\alpha(x^1, \ldots, x^n, y^1, \ldots, y^r)$ of (n+r) variables we may consider the locus $M^n \subset \mathbb{R}^{n+r}$ defined by the equations

$$F^\alpha(x,y) = c^\alpha, \quad (c^1, \ldots, c^r \text{ constants}) \quad .$$

If at $(x_0,y_0) \in M^n$ the Jacobian determinant is nonzero,

$$\frac{\partial(F^1,F^2, \ldots, F^r)}{\partial(y^1,y^2, \ldots, y^r)} (x_0,y_0) \neq 0 \quad ,$$

the implicit function theorem assures us that locally , near (x_0,y_0) , we may solve $F^\alpha(x,y) = c^\alpha$, $\alpha = 1, \ldots, r$, for y's in terms of x's

$$y^\alpha = f^\alpha(x^1, \ldots, x^n) \quad .$$

Thus near (x_0,y_0) M^n is a submanifold of $\mathbb{R}^{n+r}$. If the Jacobian is $\neq 0$ at all $(x_0,y_0) \in M^n$, then the entire M^n is a submanifold of $\mathbb{R}^{n+r}$.

More generally, suppose we have r functions of n+r variables, $F^\alpha(x^1, \ldots, x^{n+r})$. Suppose that at each $x_0 = (x_0^1,\ldots,x_0^{n+r})$ for which $F^\alpha(x_0) = c^\alpha$, $\alpha = 1, \ldots, r$, we have that the matrix

$$\frac{\partial F^\alpha}{\partial x^i}(x_0) \quad \begin{matrix} \alpha = 1,\ldots,r \\ i = 1, \ldots, n+r \end{matrix}$$

has rank r . Then the locus $F^\alpha = c^\alpha$ defines an n dimensional submanifold of $\mathbb{R}^{n+r}$ (since locally we may solve for r of the coordinates in terms of the remaining n).

- **Problem 1.** Investigate the locus

$$x^2 + y^2 - z^2 = c$$

in $\mathbb{R}^3$, for $c > 0$, $c = 0$ and $c < 0$. Are they submanifolds? What if the origin is omitted from a locus? Draw all three loci, for $c = +1 , 0 , -1$, in one picture.

(3) **Geometry of Jacobian Matrices; the "Differential."** The tangent space to $\mathbb{R}^n$ at $x_0 \in \mathbb{R}^n$ is by definition the vector ($\equiv$ linear) space $\mathbb{R}^n_{x_0}$ of all tangent vectors in $\mathbb{R}^n$ based at x_0 .

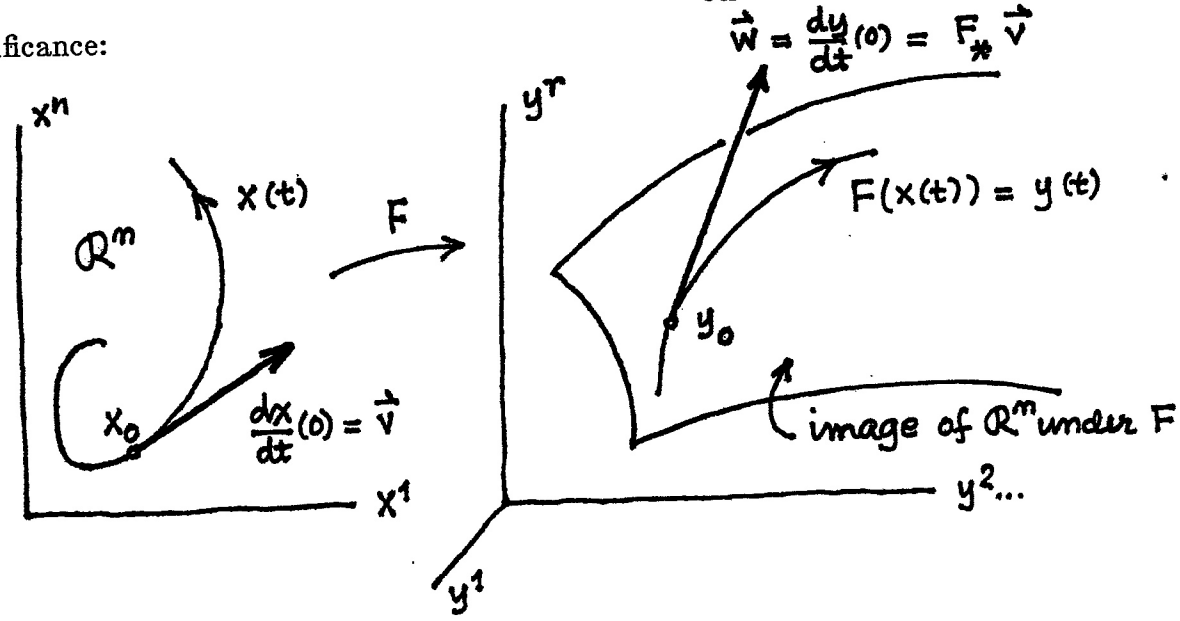
Let $x = (x^1, \ldots, x^n)$ be coordinates for $\mathbb{R}^n$, let $y = (y^1, \ldots, y^r)$ be coordinates for $\mathbb{R}^r$. Let $F : \mathbb{R}^n \to \mathbb{R}^r$ be a smooth map ($\equiv$ function). In coordinates, F is described by

$$y^\alpha = F^\alpha(x) \qquad \alpha = 1,\ldots,r \quad .$$

or

$$y = F(x) \quad .$$

Suppose $y_0 = F(x_0)$. The Jacobian matrix $\frac{\partial y^\alpha}{\partial x^i}(x_0)$ has the following significance:



Let $v \in \mathbb{R}^n_{x_0}$ be a vector at $x_0 \in \mathbb{R}^n$. Take any smooth curve $x(t)$ such that $x(0) = x_0$ and $\frac{dx}{dt}(0) = v$. The image of this curve

$$y(t) = F(x(t))$$

has tangent vector w at y_0 given by the chain rule

$$w^\alpha = \frac{dy^\alpha}{dt}(0) = \sum_{i=1}^{n} \frac{\partial y^\alpha}{\partial x^i}(x_0)\,\frac{dx^i}{dt}(0) = \sum_{i=1}^{n} \frac{\partial y^\alpha}{\partial x^i}(x_0) v^i \quad .$$

The assignment $v \to w$ is a linear transformation called the differential of F, written F_* or DF

$$DF \equiv F_* : \mathbb{R}^n_{x_0} \to \mathbb{R}^r_{y_0}; \qquad F_*(v) = w$$

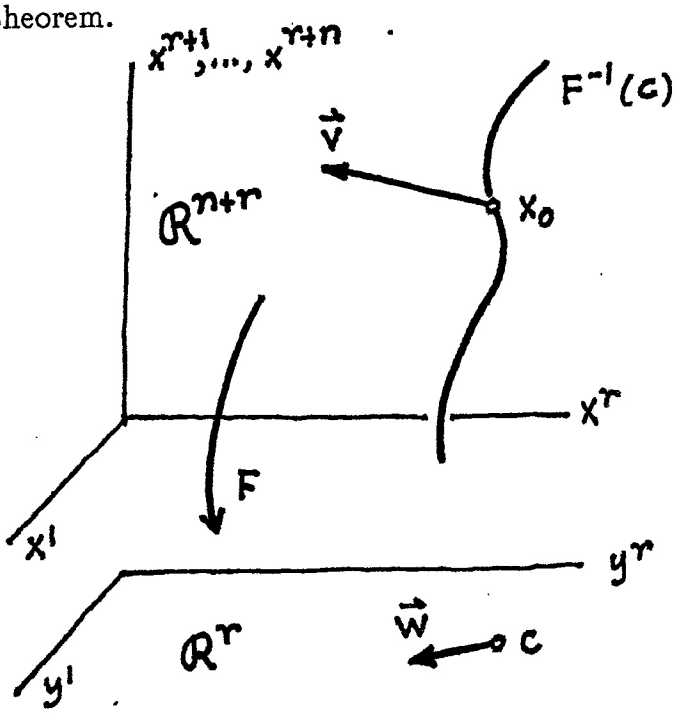
and the Jacobian $\frac{\partial y^\alpha}{\partial x^i}(x_0)$ is its matrix.

> The geometric definition of F_* given above, as illustrated in the figure, is not only more important conceptually than its matrix expression $\partial y^\alpha / \partial x^i$, it is also frequently easier to compute with.

This philosophy will be nicely illustrated in §5 below.

Note that the statement "F has rank r at x_0," i.e., $\frac{\partial y^\alpha}{\partial x^i}(x_0)$ has rank r, is equivalent to the statement that the differential $F_* : \mathbb{R}^n_{x_0} \to \mathbb{R}^r_{y_0=F(x_0)}$ is <u>onto</u> ("surjective") i.e., given any vector w at y_0 there is at least one vector v at x_0 such that $F_*(v) = w$.

(4) **The Main Theorem on Submanifolds of $\mathbb{R}^{n+r}$** is merely a restatement of the implicit function theorem.



Let $F : \mathbb{R}^{n+r} \to \mathbb{R}^r$ and suppose that $F^{-1}(c) = \{x \in \mathbb{R}^{n+r} \mid F(x) = c\}$ is not empty. Suppose further that for every $x_0 \in F^{-1}(c)$

$$F_* : \mathbb{R}^{n+r}_{x_0} \to \mathbb{R}^r_c$$

is <u>onto</u> . Then $F^{-1}(c)$ is an n dimensional submanifold of $\mathbb{R}^{n+r}$. (The best example to keep in mind is the linear map $\mathbb{R}^{n+r} \to \mathbb{R}^r$ $(x^1, \ldots, x^r, x^{r+1}, \ldots, x^{r+n}) \xrightarrow{F} (x^1, \ldots, x^r)$. What are the coordinates for $F^{-1}(c)$ here?)

(5) **A nontrivial example; the configuration space of a rigid body.** If the body has one point fixed, its configuration space is the group SO(3) of <u>rotations</u> in $\mathbb{R}^3$,

i.e., matrices $x = (x_{ij})$ such that ${}^t x = x^{-1}$ (t denoting transpose) and $\det x > 0$. (${}^t x = x^{-1}$ defines the orthogonal group $0(n)$). By assigning the coordinates $x_{11}, x_{12}, x_{13}, x_{21}, \ldots, x_{33}$ to any matrix x we see that the space of all 3×3 matrices may be represented by $\mathbb{R}^9$, nine dimensional Euclidean space. $0(3)$ is the locus in $\mathbb{R}^9$ defined by the equation ${}^t x x = I$, i.e., the quadratic equations (i,k)

$$\sum_{j=1}^{3} x_{ji}\, x_{jk} = \delta_{ij} \qquad \begin{matrix} i = 1,2,3 \\ k = 1,2,3 \end{matrix} .$$

We claim that $0(3)$ is a 3-dimensional submanifold of $\mathbb{R}^9$. First note that since ${}^t x x$ is a symmetric matrix there are really only six equations, not nine; equation (i,k) is the same as equation (k,i). This suggests that we do the following. Let

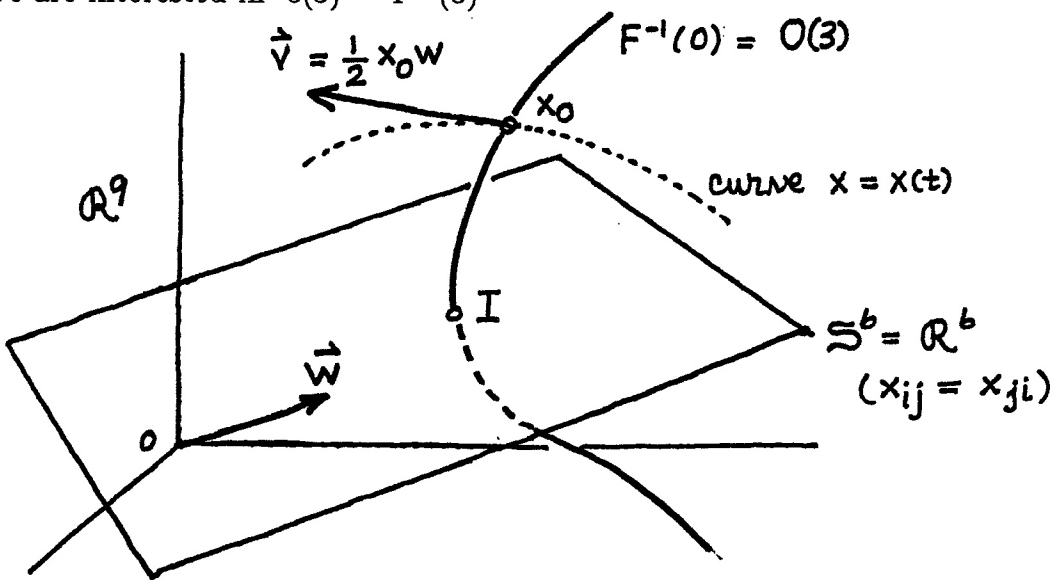
$\mathcal{S}^6 =$ the set of all symmetric 3×3 matrices. Note that since $\mathcal{S}^6$ is defined by the linear equations

$$x_{ij} - x_{ji} = 0 \quad .$$

$\mathcal{S}^6$ is actually a vector subspace $\mathbb{R}^6 \subset \mathbb{R}^9$. We now define a map F from all of $\mathbb{R}^9$ to $\mathcal{S}^6 = \mathbb{R}^6$ by

$$F(x) = {}^t x x - I \qquad F : \mathbb{R}^9 \to \mathbb{R}^6$$

and we are interested in $0(3) = F^{-1}(0)$



Let $x_0 \in F^{-1}(0) = 0(3)$. We must show that $F_* : \mathbb{R}^9_{x_0} \to \mathbb{R}^6_0 = \mathcal{S}^6$ is onto. Let w be tangent to $\mathcal{S}$ at 0; then w is itself a symmetric matrix. We must find v, a vector to $\mathbb{R}^9$ at x_0 such that $F_* v = w$. Take a curve $x(t)$ of matrices such

that $x(0) = x_0$; its tangent vector at x_0 is $\dot{x}(0)$. The image curve $F(x(t)) = {}^t x(t)x(t) - I$ has tangent vector at $t = 0$

$$\frac{d}{dt} F(x(t))|_{t=0} = {}^t\dot{x}(0)x_0 + {}^t x_0\dot{x}(0)$$

and we wish this to be w. It is sufficient to solve ${}^t x_0\dot{x}(0) = \frac{1}{2} w$. Since $x_0 \in 0(3)$, ${}^t x = x^{-1}$ and we have solution $v = \dot{x}(0) = \frac{1}{2} x_0 w$ (which, e.g. is the tangent vector to the curve $x(t) = x_0 + t \cdot \frac{1}{2} x_0 w$). $\therefore F_*$ is onto and $0(3) = F^{-1}(0)$ is a submanifold of $\mathbb{R}^9$ of dimension $9 - 6 = 3$.

Determinant $\det : \mathbb{R}^9 \to \mathbb{R}$ is continuous and $\det = \pm 1$ on $0(3)$. $0(3) = SO(3) \cup \{x \in 0(3) \mid \det x = -1\}$. These two subsets are not connected since det is continuous. $\therefore SO(3)$ is also a submanifold. □

- **Problem 2.** What can we say about $SO(n)$?

- **Problem 3.** Is $Sl(n) \equiv \{n \times n$ real matrices $x \mid \det x = 1\}$ a submanifold of $\mathbb{R}^{n^2}$? (Hint: You will need to know something about $\frac{\partial}{\partial x_{ij}} \det x$. This is an example where the analytical expression of F_* is perhaps easier, i.e., use §2(b).)

Lecture 2. Manifolds

We first give an intuitive treatment of manifolds. Rigorous definitions will be given in (3). Some topological notions are defined in the appendix to this chapter but an intuitive "feeling" is more important at this stage.

(1) **Idea of a Manifold.** An **n-dimensional** (differentiable) manifold M^n is a space that is locally $\mathbb{R}^n$ in the following sense; it is covered by a family of (curvilinear) coordinate systems $\{U_\alpha ; x_\alpha^1, x_\alpha^2, \ldots, x_\alpha^n\}$ such that a point $p \in U_\alpha \cap U_\beta$ that lies in two coordinate patches will have its two sets of coordinates related differentiably

$$x_\beta^i(p) = f_{\beta\alpha}^i(x_\alpha^1(p), \ldots, x_\alpha^n(p)) \qquad i = 1, \ldots, n$$

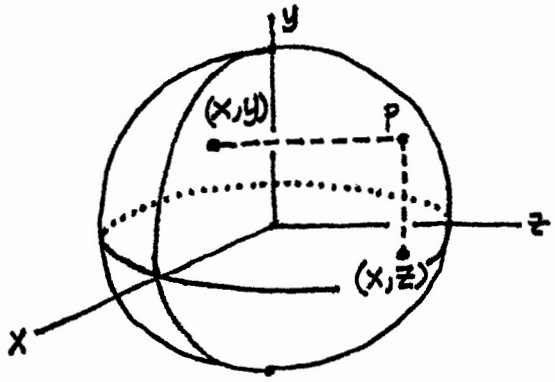
(if the functions $f_{\beta\alpha}^i$ are C^∞, or real analytic, we say that M is C^∞, or real analytic, etc.).

(2) **Examples.**

(a) $M^n = \mathbb{R}^n$, covered by a single coordinate system $x^1 \cdots x^n$.

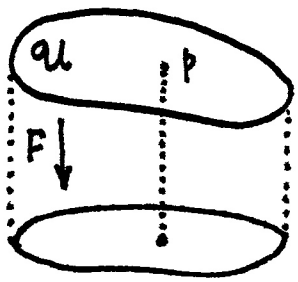
(b) $M^n = S^n$, the unit sphere in $\mathbb{R}^{n+1}$. We illustrate with $n = 2$. Cover S^2 with six "open" subsets

$$U_x^+ = \{p \in S^2 \mid x(p) > 0\}$$
$$U_x^- = \{p \in S^2 \mid x(p) < 0\}$$
$$U_y^+ = \{p \in S^2 \mid y(p) > 0\}$$
$$\vdots \quad \text{etc., etc.}$$



The p illustrated sits in $U_x^+ \cap U_y^+ \cap U_x^+$. Project U_x^+ on $y - z$ plane; this introduces y,z as curvilinear coordinates in U_x^+, etc., etc. For $p \in U_y^+ \cap U_z^+$ we have two sets of coordinates $u_1 = x$, $u_2 = z$ and $v_1 = x$, $v_2 = y$, and these are related by $v_1 = u_1$, $v_2 = \sqrt{1-(x^2+z^2)} = \sqrt{1-u_1^2-u_2^2}$ and these are infinitely differentiable in $U_y^+ \cap U_z^+$ (why?).

S^2 is " locally $\Re^2$" in this precise sense; the typical point p (shown above) has a neighborhood (in the topology of S^2 induced as a subset of $\Re^3$) that is homeomorphic (via projection F) to an open "ball" in $\Re^2$.



If two sets of coordinates are related differentiably in an overlap we shall say that they are compatible.

On S^2 we could introduce (instead of the coordinates above) the usual spherical coordinates θ,ϕ. They do not work for the entire S^2 (e.g. at the poles), but where they do work they are compatible with our first coordinates.

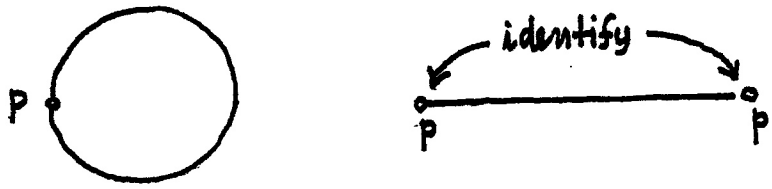
On S^2 one could introduce coordinates via stereographic projection onto the planes $z = 1$ and $z = -1$ again failing at South and North poles respectively). Where they are good coordinates they are again compatible with our first set. On a manifold we should include all coordinate systems which are compatible with those that originally defined the manifold. Such a collection of coordinate systems forms a " maximal atlas ."

(c) If M^n is a manifold with coordinates $\{U_\alpha ; x_\alpha^1, \ldots, x_\alpha^n\}$ and W^r is a manifold with coordinates $\{V_\beta ; y_\beta^1, \ldots, x_\beta^r\}$ we can form the product manifold

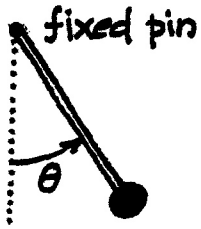
$$L^{n+r} = M^n \times W^r = \{(p,q) \mid p \in M \quad \text{and} \quad q \in W\}$$

by using as local coordinates $x_\alpha^1, \ldots, x_\alpha^n, y_\beta^1, \ldots, y_\beta^r$ in $U_\alpha \times V_\beta$.

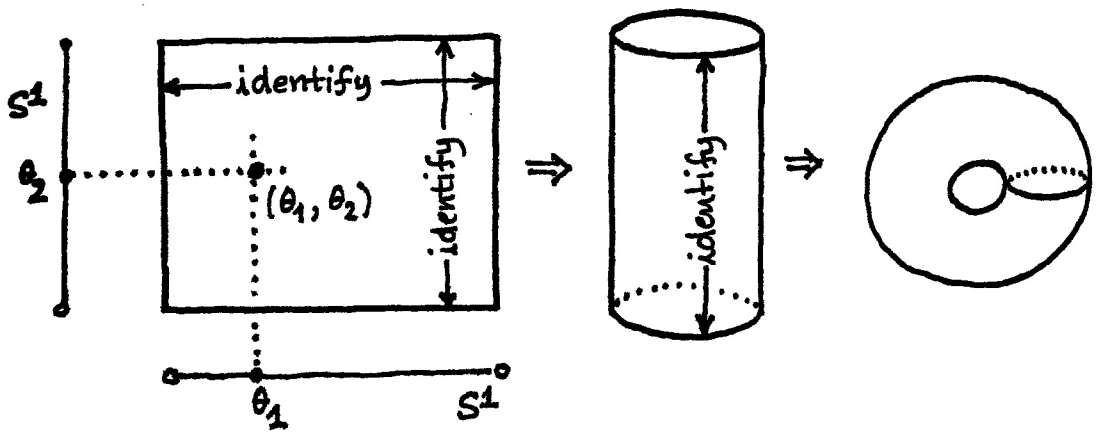
Note that S^1 is simply the unit circle in $\Re^2$ (traditionally we use $\theta = \tan^{-1}(y/x)$ as coordinate, knowing full well the dangers; we should use at least two such angular coordinates to cover all of S^1). Topologically S^1 is conveniently represented by an interval on the real line with endpoints identified



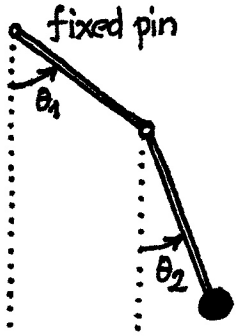
S^1 is the configuration space for a rigid pendulum constrained to oscillate in the plane



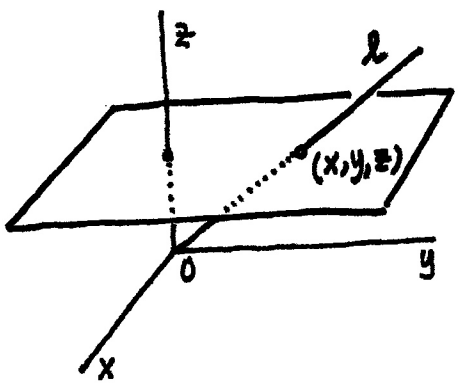
The n- dimensional torus $T^n \equiv S^1 \times S^1 \times \cdots \times S^1$ has local coordinates given by n angular parameters $\theta_1, \ldots, \theta_n$. Topologically it is the n-cube (product of n intervals) with identification. For $n = 2$



T^2 is the configuration space of a planar double pendulum .



(d) **Real Projective n-space** $P^n(\mathbb{R})$ is the space of all unoriented lines l passing through the origin of $\mathbb{R}^{n+1}$; we illustrate with the real projective plane $P^2(\mathbb{R})$ of lines through 0 of $\mathbb{R}^3$.



Note that such a line l is completely determined by giving a point $[x,y,z]$ (other than the origin) on it. Note also that if $\lambda \neq 0$, then $[\lambda x,\lambda y,\lambda z]$ describes the same line l, that is, the same point $l \in P^2(\mathbb{R})$. $[x,y,z] \neq [0,0,0]$ are called the homogeneous coordinates of the point $l \in P^2(\mathbb{R})$.

(**Warning:** Our lines are not directed!)

Cover $P^2(\mathbb{R})$ by three sets

$U_x = \{$those lines l not lying in the yz plane$\}$ etc., etc.

Introduce coordinates in U_x, U_y, U_z as follows. In U_z, given l, pick point $[x,y,z]$ on l. Define coordinates

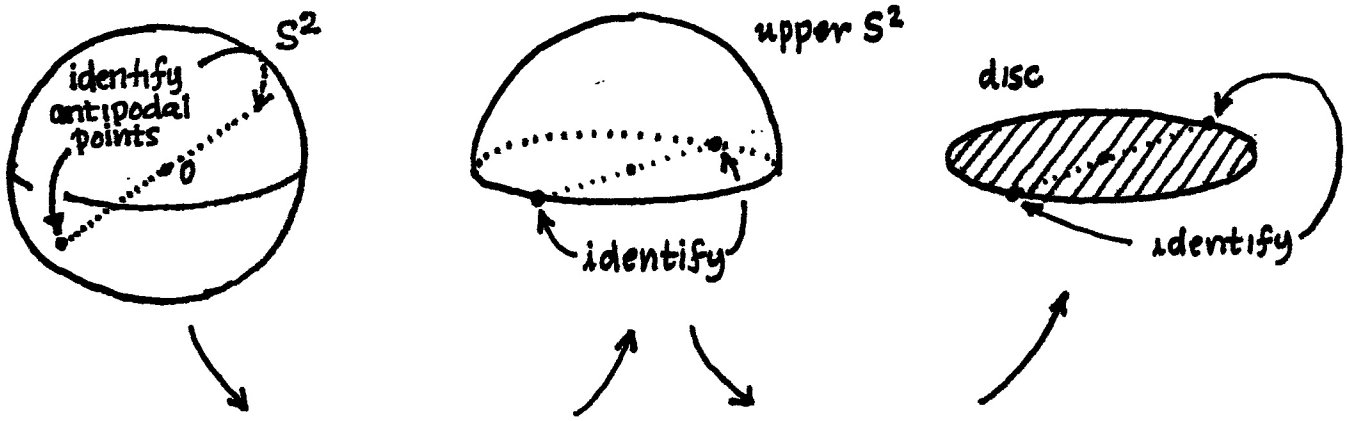
$$u_1 = \frac{x}{z}, \quad u_2 = \frac{y}{z} .$$

Note that in U_z we are always free, for example, to choose $[x,y,z]$ with $z = 1$. Then the coordinates $u_1 = x$, $u_2 = y$ are merely the x,y coordinates of the point where the line l pierces the plane $z = 1$.

Similarly, pick coordinates in U_x, U_y.

Problem 4. Show that $P^2(\mathbb{R})$ is a C^∞ 2-dimensional manifold.

Note that if we had considered directed lines $\vec{l}$ through the origin of $\mathbb{R}^{n+1}$ we would have ended up with the manifold S^n, since each such $\vec{l}$ pierces the unit S^n in exactly 1 " forward" point. Since we are using un- directed lines, each l intersects S^n in a pair of antipodal points; $P^n(\mathbb{R})$ is topologically S^n with antipodal points identified! Topologically

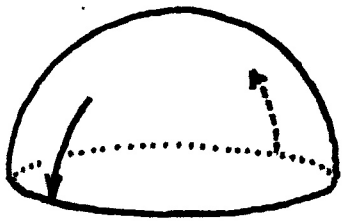


In this step we throw away redundant points in lower hemisphere.

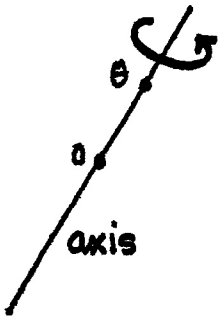
In this step we flatten the hemisphere to a disc, still identifying points on the boundary.

and so topologically $P^2(\mathbb{R})$ is the 2-dimensional disc $\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ with antipodal points on the boundary circle $x^2 + y^2 = 1$ identified. Similarly for $P^n(\mathbb{R})$.

Note that there is no "edge" to $P^2(\mathbb{R})$; just as one leaves the northern hemisphere one re-enters this hemisphere at the antipodal point.

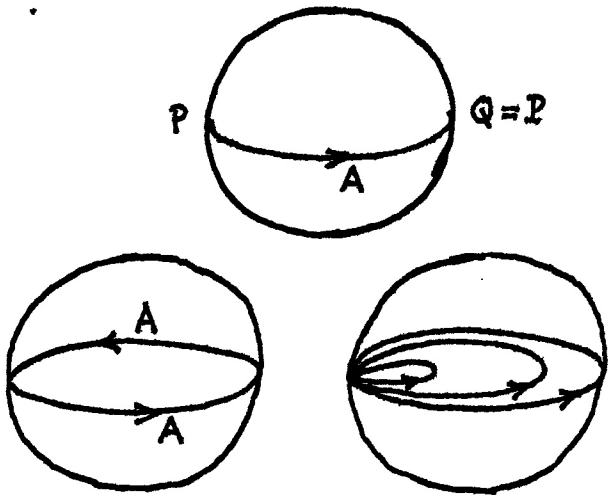


(e) It is a fact that <u>every submanifold of</u> $\mathbb{R}^n$ <u>is a manifold</u>. (We verified this in the case S^n in part (b).) In Lecture 1, §5 we showed that SO(3) is a 3-dimensional submanifold of $\mathbb{R}^9$; hence it is a 3-manifold. But any rotation in $\mathbb{R}^3$ is determined by giving its axis and the angle θ of the rotation. Use the right hand rule to associate with a rotation a point at distance θ radians from the origin along the axis. Of course the



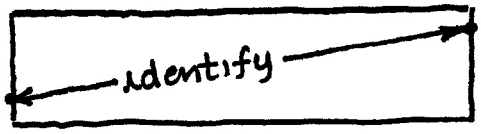
points $\theta = \pi$ and $\theta = -\pi$ represent exactly the same rotation! Thus the points of SO(3) can be represented by points of the ball of radius π, center at 0, with antipodal points on the boundary 2-sphere identified. <u>Topologically</u> SO(3) <u>is</u> $P^3(\mathbb{R})$, <u>the real projective</u> 3-<u>space</u> of <u>lines through</u> 0 of $\mathbb{R}^4$.

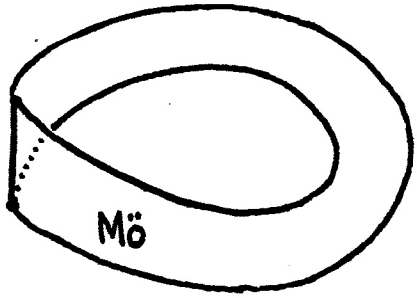
Note that if we use as model of $SO(3) = P^3(\mathbb{R})$ the solid ball of radius π with antipodal points on the boundary 2-sphere identified, then the great half circle A is a closed curve in $P^3(\mathbb{R})$ since the endpoints P and Q are really the same point.



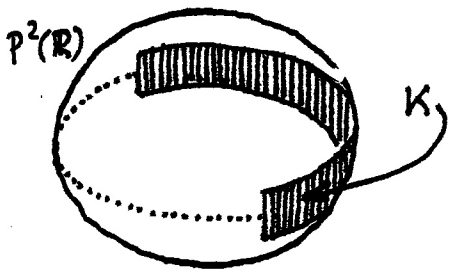
This curve A cannot be continuously shrunk to a point in $P^3(\mathbb{R})$ (go ahead, try it, but remember that end- points must always be points that are identified). However, if we go around A twice we can shrink the resulting closed curve to a point. This simple fact is the reason why spinors can exist ! We shall discuss this more when we come to Lie groups.

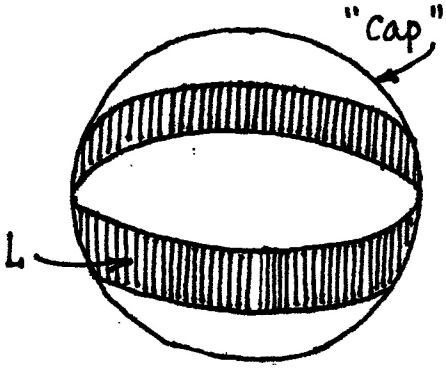
- **Problem** 5. The Möbius band Mö is the familiar strip of paper with two edges identified after a half twist.



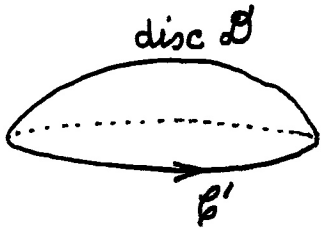


"Show" that the indicated subset K of $P^2(\mathbb{R})$ is a Möbius band. So is the subset L (why?).



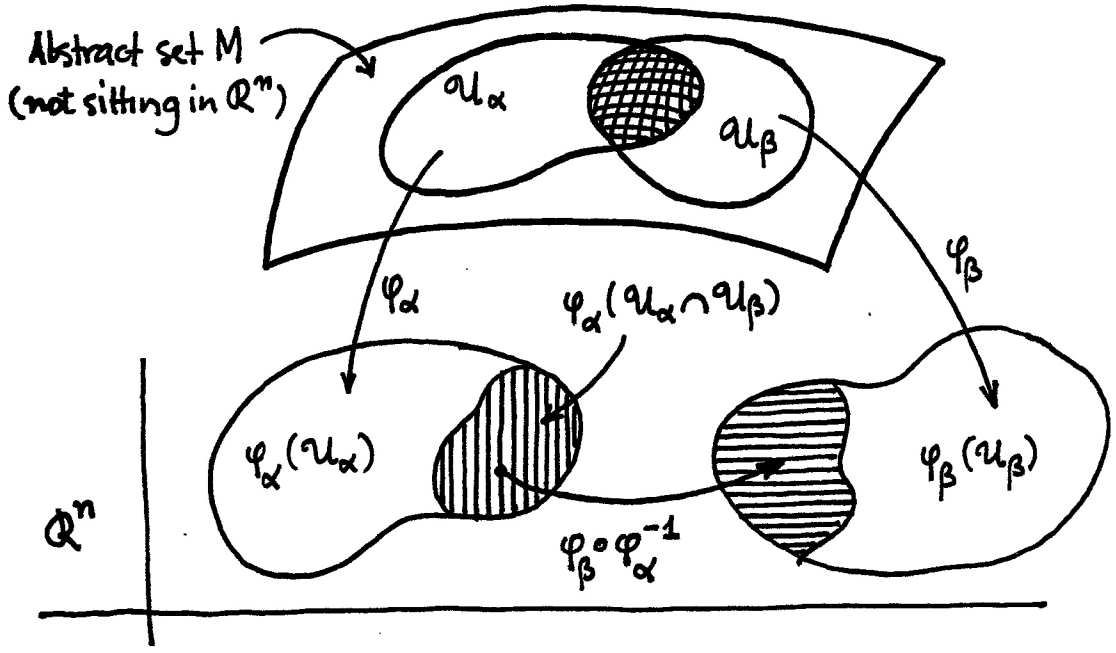


This shows that if we remove a "cap" from $P^2(\mathbb{R})$ the remaining space is a Möbius band L . This in turn shows us that if we take a Möbius band L (which has a single edge C) and take a disc D



(which has a single edge C') and if we sew these two edges together $C \leftrightarrow C'$, the resulting surface is $P^2(\mathbb{R})$! (My vast sewing experiences indicate that this really cannot be done in $\mathbb{R}^3$; L and $\mathcal{D}$ try to interpenetrate each other. However sewing is much easier in $\mathbb{R}^4$; there it can be done.)

(3) A Rigorous Definition of Manifold (see appendix for topological terms). Let M be any <u>set</u> (without a topology) that has a covering by subsets $M = \bigcup_\alpha u_\alpha$, with each u_α in 1:1 correspondence $\phi_\alpha : u_\alpha \to \mathbb{R}^n$ with an <u>open</u> subset $\phi_\alpha(u_\alpha)$ of $\mathbb{R}^n$.



We require that each $\phi_\alpha(u_\alpha \cap u_\beta) \subset \mathbb{R}^n$ be an <u>open</u> subset of $\mathbb{R}^n$. We also require that the overlap maps $\phi_\beta \circ \phi_\alpha^{-1}$

$$\phi_\alpha(u_\alpha \cap u_\beta) \xrightarrow{\phi_\alpha^{-1}} M \xrightarrow{\phi_\beta} \mathbb{R}^n$$

be differentiable (we know what this means since $\phi_\beta \circ \phi_\alpha^{-1}$ are maps from subsets of $\mathbb{R}^n$ to subsets of $\mathbb{R}^n$). Take a maximal atlas (see (b)) of such coordinate systems u_α , ϕ_α . Now define a <u>topology</u> in M by declaring $V \subset M$ to be <u>open</u> provided given any $p \in V$ there is a coordinate chart u_α such that $p \in u_\alpha \subset V$. If the resulting topology is Hausdorff with a countable base, we say that M is an n-dimensional differentiable manifold.

Appendix on Point Set Topology

The <u>open</u> ball in $\mathbb{R}^n$, radius ϵ, center at $x_0 \in \mathbb{R}^n$ is

$$B_{x_0}(\epsilon) = \{x \in \mathbb{R}^n \mid ||x - x_0|| < \epsilon\} \quad .$$

where $||x - x_0||^2 = \sum_{i=1}^{n} |x^i - x_0^i|^2$.

$$\bar{B}_{x_0}(\epsilon) = \{x \in \mathbb{R}^n \mid ||x - x_0|| \leq \epsilon\}$$

defines the <u>closed ball</u>.

A set $U \subset \mathbb{R}^n$ is declared <u>open</u> if given $x_0 \in U$ there is an open ball $B_{x_0}(\epsilon) \subset U$. For example $B_0(1)$ is open, $\bar{B}_0(1)$ is not (<u>try</u> to find an open ball in $\bar{B}_0(1)$ centered at $(1,0, \ldots, 0) \in \bar{B}_0(1)$.) Note that the <u>complement</u> $\mathbb{R}^n \backslash \bar{B}_0(1)$ of the closed ball

$$\{x \in \mathbb{R}^n \mid x \notin \bar{B}_0(1)\}$$

<u>is</u> an open subset of $\mathbb{R}^n$. It is not difficult to see that if U_1 and U_2 are open in $\mathbb{R}^n$ then so is $U_1 \cap U_2$ (in fact any <u>finite</u> intersections of open sets are open) and if $\{U_\alpha\}$ is <u>any</u> collection of open sets (even uncountable) then their union $\bigcup_\alpha U_\alpha$ is trivially open.

<u>Definition:</u> A <u>topological space</u> is a set M with a distinguished collection of subsets called <u>open</u> subsets, satisfying

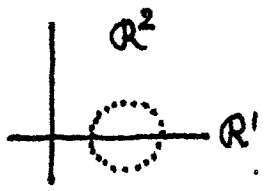
(i) M is open and the empty set ϕ is open.

(ii) U open and V open $\Rightarrow U \cap V$ open.

(iii) The union of any collection of open subsets is open.

These open subsets "define" the topology of M. $\mathbb{R}^n$ is a topological space with open as defined above.

Let A be any subset of a topological space M. Define a topology for A (the " <u>induced</u> " or " <u>subspace</u> " <u>topology</u>) by declaring $U_A \subset A$ is open if there is some open subset U of M such that $U_A = U \cap A$. One checks easily that these sets do indeed define a topology for A.

For example, look at $A = \mathbb{R}^1 \subset \mathbb{R}^2$ 

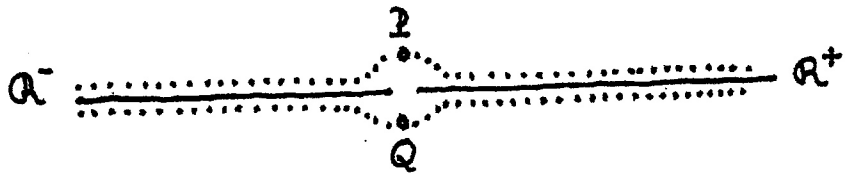
. Check that the open subsets of A are exactly those that are open in the original topology of $\mathbb{R}^1$.

Definition. A neighborhood (nbhd) of $x \in M$ is any open subset of M containing x.

Definition. If $f : M \to N$ is a map of a topological space M into a topological space N we say F is continuous if whenever $V \subset N$ is open in N then the inverse image $F^{-1}(V) = \{x \in M \mid F(x) \in V\}$ is open in M. (This reduces to the usual ϵ,δ definition when $M = \Re^m$ and $N = \Re^n$).

Definition. $F : M \to N$ is a homeomorphism if F is continuous, F is 1:1 and onto (and thus $F^{-1} : N \to M$ exists) and F^{-1} is continuous. We also say that "M is topologically N," or M is "homeomorphic to N."

Definition. A topological space M is Hausdorff if every pair of distinct points has at least one pair of disjoint nbhds. Most common spaces are Hausdorff. We construct a non-Hausdorff example as follows. Let $\tilde{M}$ consist of ($\Re$-origin) $\cup$ two other points P and Q.



Declare as open subsets $\tilde{M}$, ϕ , $\Re^+ \cup \Re^-$, $\Re^+ \cup \Re^- \cup \{P\}$ and $\Re^+ \cup \Re^- \cup \{Q\}$.

Definition. A topological space M has a countable base provided that there is a countable collection of open sets $\{V_1, V_2, \ldots, \}$, such that any open $U \subset M$ is a union of sets V_α .

- **Problem** 6. Check that example $\tilde{M}$ above satisfies all the criteria for being a non-Hausdorf topological space.

Lecture 3. Vectors and Mappings

(1) **Tangent, or "Contravariant" vectors .** Let $p = p(t)$ be a curve lying on a manifold M^n. In a coordinate system (u,x) about the point $p_0 = p(0)$ the curve will be described by n functions $x^i = x^i(t)$. The velocity vector $\frac{dp}{dt}(0)$ at $t = 0$ is described by the n-tuple of real numbers $\frac{dx^1}{dt}(0), \ldots, \frac{dx^n}{dt}(0)$. If p_0 also lies in a second coordinate system $(\bar{u},\bar{x})$ then this same velocity vector $\frac{dp}{dt}(0)$ would be described by another n-tuple $\frac{d\bar{x}^1}{dt}(0), \ldots, \frac{d\bar{x}^n}{dt}(0)$ related to the first by the chain rule

$$\frac{d\bar{x}^i}{dt}(0) = \sum_{j=1}^{n} \frac{\partial \bar{x}^i}{\partial x^j}(p_0)\, \frac{dx^j}{dt}(0) \quad .$$

Definition. A tangent, or contravariant, vector or simply a vector at $p_0 \in M^n$, call it v, assigns to each coordinate system (u_α,x_α) holding p_0, an n-tuple

$$(v_\alpha^i);\ v_\alpha^1,\ v_\alpha^2, \ldots, v_\alpha^n$$

such that if $p_0 \in u_\alpha \cap u_\beta$, we have

$$\boxed{v_\beta^j = \sum_{j=1}^{n} \frac{\partial x_\beta^j}{\partial x_\alpha^j}(p_0)\, v_\alpha^j \,.} \tag{3.1}$$

It is not hard to see that such a vector is indeed the velocity vector to some (in fact infinitely many) curves through p_0.

Associated to any such vector v there is a first order differential operator D_v; it acts on functions defined near p_0 by

$$D_v(f) \equiv \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(p_0) v^i \quad .$$

As it stands, this operator seems to depend on the coordinate system x used. However, in an overlap, by the chain rule

$$D_v^{(x)}(f) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial f}{\partial \bar{x}^j} \frac{\partial \bar{x}^j}{\partial x^i} \right) v^i = \sum_{j=1}^{n} \frac{\partial f}{\partial \bar{x}^j} \left(\sum_{i=1}^{n} \frac{\partial \bar{x}^j}{\partial x^i} v^i \right)$$

$$= \sum_{j=1}^{n} \frac{\partial f}{\partial \bar{x}^j} \bar{v}^j = D_v^{(\bar{x})}(f) \quad .$$

Thus

$$\boxed{v_{p_0}(f) \equiv D_v(f) \equiv \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}(p_0) v^i} \tag{3.2}$$

is, in fact, independent of the coordinate system used!

A tangent vector at p_0 can thus be thought of either as the velocity vector to a curve through p_0 or as a first order differential operator that takes the special form

$$\boxed{v_{p_0} = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}\Big|_{p_0}} \tag{3.2′}$$

in a local coordinate system.

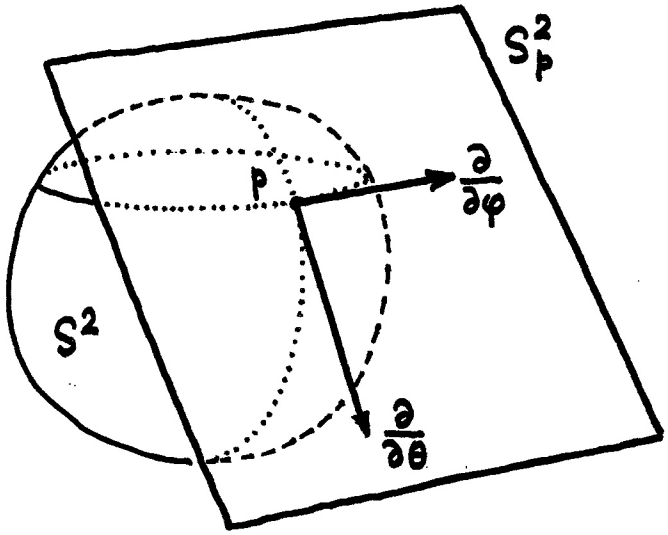
Definition. The <u>tangent space</u> to M^n at the point $p \in M^n$ is the vector ($\equiv$ linear) space M_p^n consisting of all tangent vectors to M^n at p. If (u,x) is a coordinate system for M^n holding p then the n vectors

$$e_1 = \frac{\partial}{\partial x^1}\Big|_p, \ldots, e_n = \frac{\partial}{\partial x^n}\Big|_p$$

form, from (3.2′), a <u>basis</u> for M_p^n, called a <u>coordinate basis</u>; M_p^n is an n dimensional vector space.

If $M^n \subset \mathbb{R}^N$ is a submanifold of $\mathbb{R}^N$, then M^n_p is the usual tangent n-space to M^n at p.

Example: We have drawn the unit 2-sphere $M^2 = S^2$ in $\mathbb{R}^3$ with the usual geographical coordinates.* The equations defining S^2 are $x = \sin\theta\cos\phi$, $y = \sin\theta\sin\phi$, $z = \cos\theta$. The tangent vector $\frac{\partial}{\partial\theta}$ is the velocity vector to a line of longitude, i.e., keep longitude ϕ constant and parameterize this curve by co-latitude θ, etc. Note that the vectors $\frac{\partial}{\partial\theta}$ and $\frac{\partial}{\partial\phi}$ do not live in S^2; they live in a vector space S^2_p attached to S^2 at p. Vectors at $q \neq p$ live in a different vector space S^2_q.



~~ **WARNING:** Because S^2 is a submanifold of $\mathbb{R}^3$, and because $\mathbb{R}^3$ carries a "standard" metric, it makes some sense to talk about the length of tangent vectors to S^2. For example we would say $\|\frac{\partial}{\partial\theta}\| = 1$ and $\|\frac{\partial}{\partial\phi}\| = \sin\theta$. However, our definition of manifold M^n (Lecture 1(3)) does not require that M^n be a subset of some $\mathbb{R}^N$; we do not have the notion of length of a tangent vector to a general M^n. For example the configuration space of a thermodynamical system is described by coordinates p (pressure), v (volume), T (temperature),--and the notions

$$\|\frac{\partial}{\partial p}\|, \quad \|\frac{\partial}{\partial v}\|, \quad \|\frac{\partial}{\partial T}\|, \quad \text{etc.}$$

*except we use co-latitude instead of latitude!

seem to have little (if any) physical significance. If we wish to talk about the "length" of a vector we shall be forced to introduce an additional structure on the manifold in question. The most common structure used is called a Riemannian structure, or metric, to be introduced in our next lecture.

• **Problem 7.** Why not define $\|v_{p_0}\|$ from (3.2′) by

$$\|v_{p_0}\|^2 = \sum_{i=1}^{n} (v^i)^2 \quad ?$$

(2) **Submanifolds of manifolds.** If M^n and V^r are manifolds and $F : M^n \to V^r$ is a map, then it is described in terms of local coordinates x near $p \in M^n$ and y near $F(p) \in V^r$ by r functions of n variables

$$y^\alpha = F^\alpha(x^1, \ldots, x^n) \quad .$$

We also shall write this briefly $y = F(x)$, or more dangerously as $y = y(x)$. The map F is continuous, differentiable,... provided the functions F^α are continuous, differentiable,... functions of $x^1, \ldots, x^n$.

$F : M^n \to V^n$ is a diffeomorphism provided F is 1:1 and onto, and both F and F^{-1} are differentiable.

Example. $F : \mathbb{R} \to \mathbb{R}$ defined by $y = x^3$ is a homeomorphism (see appendix to lecture 2) but not a diffeomorphism because F^{-1} (i.e., $x = y^{1/3}$) is not differentiable at $y = 0$.

Definition. $W^r \subset M^n$ is an (embedded) submanifold of M^n provided W^r is locally described as the common locus $F^1(x^1, \ldots, x^n) = 0, \ldots, F^{n-r}(x^1, \ldots, x^n) = 0$ of $(n-r)$ differentiable functions that are independent in the sense $\left[\dfrac{\partial F^\alpha}{\partial x^i}\right]$ has rank $n-r$.

The implicit function theorem assures us that W^r can be described in some coordinate system as the points where

$$x^{r+1} = f^{r+1}(x^1, \ldots, x^r), \ldots, x^n = f^n(x^1, \ldots, x^r) \quad .$$

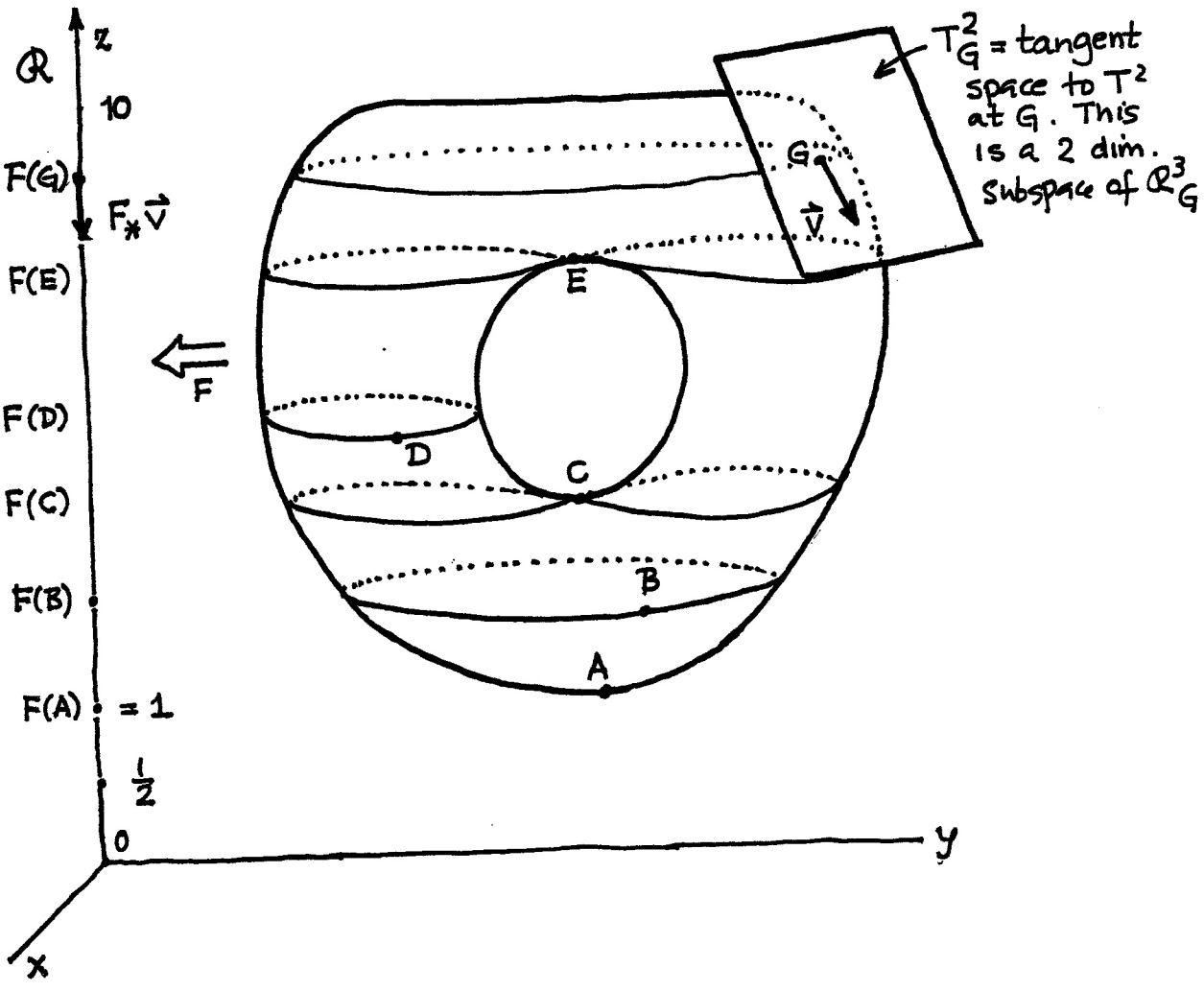
Every embedded submanifold of M^n is itself a manifold ! Note: There are other notions of submanifold (they are not "embedded"); we shall talk of them later.

The differential $F_* = DF$ of a differentiable map $F : M^n \to V^r$ has the same geometrical significance as in the case $\mathbb{R}^n \to \mathbb{R}^r$ of Lecture 1(3). In local coordinates it has matrix at $p \in M^n$ given again by

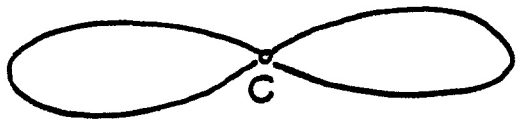
$$(F_*)_i^{\alpha} = (DF)_i^{\alpha} \equiv \frac{\partial F^{\alpha}}{\partial x^i}(p) \quad .$$

The main theorem on submanifolds of a manifold is exactly as in Lecture 1(4) with $\mathbb{R}^{n-r}$ and $\mathbb{R}^r$ replaced by M^{n+r} and V^r. Let's look at a good example to illustrate everything.

Example. Consider a 2-dimensional torus T^2 embedded in $\mathbb{R}^3$ (before we became so sophisticated we would have said "surface of a doughnut in space"). We have drawn it flat on top. Define a differentiable map $F : T^2 \to \mathbb{R}$ by $F(p) =$ z coordinate of p , i.e., $F(p) =$ height of p above the xy plane.



Note that $F : T^2 \to \mathbb{R}$ is simply the restriction to the submanifold T^2 of the (linear) map $\hat{F} : \mathbb{R}^3 \to \mathbb{R}$ given by $\hat{F}(x,y,z) = z$. The differential $\hat{F}_* : \mathbb{R}^3_G \to \mathbb{R}_{F(G)}$ is from the geometric meaning (Lecture 1(3)) simply the projection of vectors v at $G \in \mathbb{R}^3$ onto the $\mathbb{R}$ $(= z)$ axis. For the submanifold $T^2 \subset \mathbb{R}^3$, $F_* : T^2_G \to \mathbb{R}_{F(G)}$ is simply the restriction of this map to tangent vectors to the submanifold. Note that $F_* : T_u \to \mathbb{R}_{F(u)}$ maps T_u <u>onto</u> the 1-dimensional vector space $\mathbb{R}_{F(u)}$ at all $u \in T^2$ where the tangent plane T_u is not horizontal and that except for such points $F^{-1}(u)$ appears to be an embedded 1-dimensional submanifold of T^2. Note that at the point C where T^2_C is horizontal $F^{-1}(F(c))$ is a figure 8.



This is <u>not</u> a 1-dimensional manifold (nor submanifold); no neighborhood of C looks like a 1-ball (i.e., interval) ——— .

Definition. If $F : M^n \to V^r$ is a differentiable map we say

(i) $x \in M^n$ is a <u>regular point</u> if $F_* : M^n_x \to V^r_{F(x)}$ is <u>onto</u>, (otherwise x is a <u>critical point</u>).

(ii) $y \in V^r$ is a <u>regular value</u> provided either $F^{-1}(y)$ is empty, or $F^{-1}(y)$ consists entirely of regular points.

Theorem. If $F : M^n \to V^r$ is differentiable and if $y \in V^r$ is a regular <u>value</u> then either $F^{-1}(y)$ is empty or it is an embedded submanifold of M^n of dimension $n-r$.

This is simply a rewording of our main theorem on submanifolds. It is illustrated in our example by the values $\frac{1}{2}$ (where $F^{-1}\left(\frac{1}{2}\right) = \phi$), F(B), F(D), F(G). Note that $A = F^{-1}(1)$ <u>is</u> a submanifold in spite of 1 being a critical value; however its dimension is wrong.

Theorem. (A.P Morse, Sard). Almost all <u>values</u> of F are regular values, provided $F : M^n \to V^r$ is of differentiability class C^k with ~~$k \geq n-r+1$~~ $k-1 \geq \max(n-r, 0)$.

Thus for all points $y \in F(M)$ except perhaps for some set of measure 0 (to be defined later), $F^{-1}(y)$ is an embedded submanifold of dimension n–r.

In our example note that the flat top of T^2 consists of a huge set of critical <u>points</u>, but the critical <u>values</u> consist only of the points $z = 1$, $F(C)$, $F(E)$, 10. For a proof of "Sard's theorem," see (A,M,R, p. 551).

(3) **The Inverse Function Theorem** states that if $F : M^n \to V^n$ is such that $F_* : M^n_x \to V^n_{F(x)}$ is an isomorphism then F is locally a diffeomorphism near x, i.e., there is an open subset U of M^n holding x such that F(U) is open in V^n and $F : U \to F(U)$ is a diffeomorphism.

The inverse function theorem is essentially equivalent to the implicit function theorem; each is easily proved from the other.

It is important to realize that the theorem is local. For example, if $F : M^n \to V^n$ is such that F_* is an isomorphism at each $x \in M^n$, F need not be globally 1:1 (of course it is locally so). $F : \mathbb{R}^1 \to S^1$ defined by $F(t) = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}$ is a simple example to keep in mind.

The inverse function theorem is useful for introducing new coordinates; if $y^\alpha = y^\alpha(x^1, \ldots, x^n)$ are n functions such that

$$\frac{\partial(y^1, \ldots, y^n)}{\partial(x^1, \ldots, x^n)}(x_0) \neq 0$$

then the y's form a local coordinate system in some neighborhood of x_0.

Lecture 4. Vector Fields and Flows

(1) **Vector fields and flows.** By definition, a vector field on $\mathcal{U} \subset M^n$ assigns, in a differentiable manner a vector v_p to each $p \in \mathcal{U}$.

In a coordinate patch $(\mathcal{U},x)$ we have n linearly independent "coordinate" vector fields $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$. Any vector field on this patch has a local expression

$$v = \sum_{i=1}^{n} v^i(x) \frac{\partial}{\partial x^i}$$

where $v^i(x)$ are differentiable functions of $(x^1, \ldots, x^n)$, $i = 1, \ldots, n$.

Given a "stationary" (i.e., time independent) flow of water in $\mathbb{R}^3$ we can construct a 1-parameter family of maps

$$\phi_t : \mathbb{R}^3 \to \mathbb{R}^3 \quad ;$$

ϕ_t takes the molecule that is located at p when $t = 0$ to the position of the same molecule t seconds later. Because the flow is time-independent we have

$$\begin{cases} \phi_s(\phi_t(p)) = \phi_{s+t}(p) = \phi_t(\phi_s(p)) \\ \phi_{-t}(\phi_t(p)) = p\,, \qquad \text{i.e.,} \quad \phi_{-t} = (\phi_t)^{-1} \end{cases}.$$

and so if the flow is differentiable we have a 1-parameter "group" of diffeomorphisms. Associated with such a flow, i.e., associated with a differentiable 1-parameter group of diffeomorphisms ϕ_t, is a time-independent velocity vector field

$$v_p = \left(\frac{d}{dt} \phi_t p \right)_{t=0} .$$

(Thought of as a differential operator

$$v_p(f) = \frac{d}{dt} f(\phi_t p) \Big|_{t=0} .)$$

We thus have the almost trivial observation that to each "flow" $\{\phi_t\}$ we can associate the "velocity vector field" $v_p = \frac{d}{dt} (\phi_t p) \Big|_{t=0}$. The converse result, perhaps the most important theorem relating calculus to science, states, roughly speaking, that to each vector field v one may associate a flow ϕ_t such that v is its velocity field and that ϕ_t is found by solving the system of differential equations

$$\frac{dx^i}{dt} = v^i(x^1(t), \ldots, x^n(t)) \quad .$$

That is, one finds the integral curves of the above system and $\phi_t p$ says "move along the integral curve through p for t seconds." For future reference we shall give a precise statement of this fundamental theorem on solutions of ordinary differential equations, but definitions will be lacking. We refer the reader to (A,M,R, Chapter 4) for details. We shall also give an extremely rough idea of a proof in the appendix of this lecture.

A Banach space E is a "complete" normed linear space; it may be infinite dimensional and is a vast generalization of $\mathbb{R}^n$; you may use $\mathbb{R}^n$ rather than E if you wish in the following. "Completeness" means that every Cauchy sequence converges.

The Fundamental Theorem on Vector Fields. Let v be a C^k vector field, $k \geq 1$, on an open subset $\mathcal{U}$ of a Banach space E

$$v : \mathcal{U} \to E \quad .$$

Then for each $p_0 \in \mathcal{U}$ there is a curve γ mapping an interval $(-b,b)$ of the real line $\mathbb{R}$

$$\gamma : (-b,b) \to \mathcal{U}$$

such that

$$\left.\begin{aligned} &\gamma(0) = p_0 \\ &\frac{d\gamma}{dt}(t) = v_{\gamma(t)} \end{aligned}\right\} \quad \text{this says that } \gamma \text{ is an integral curve of } v$$

for all $t \in (-b,b)$. Any two such curves are equal on the intersection of their domains ("uniqueness"). Moreover, there is a nbhd $\mathcal{U}_0$ of $p_0 \in \mathcal{U}$ and a real number $\epsilon > 0$ and a C^k map

$$\Phi : \mathcal{U}_0 \times (-\epsilon,\epsilon) \to E$$

such that the curve

$$t \in (-\epsilon,\epsilon) \longmapsto \phi_t(p) = \Phi(p,t)$$

(this means t gets sent into $\phi_t(p)$) is a curve through $p \in \mathcal{U}_0$ satisfying the differential equation

$$\frac{\partial}{\partial t}\phi_t(p) = v_{\phi_t(p)}$$

for all $t \in (-\epsilon,\epsilon)$. Moreover, if t, s, and t+s are all in $(-\epsilon,\epsilon)$, then

$$\phi_t \circ \phi_s = \phi_{t+s} = \phi_s \circ \phi_t$$

and $\{\phi_t\}$ defines a local flow of diffeomorphisms.

Example. $E = \mathbb{R} = \mathbb{R}^1$, the real line, v is the vector field $v = x\frac{\partial}{\partial x}$, i.e., v has the single component x at the point in $\mathbb{R}$ with coordinate x. To find ϕ_t we simply solve

$$\begin{cases} \dfrac{dx}{dt} = x \\ x(0) = x_0 \end{cases}$$

and get $x(t) = e^t x_0$, i.e., $\phi_t(x_0) = e^t x_0$. In this example (as for all linear vector fields) ϕ_t is defined for all t, i.e., the interval $(-\epsilon,\epsilon)$ can be chosen to be $(-\infty,\infty)$.

- **Problem 8.** As in the above example, consider the vector field $v = x^2\frac{\partial}{\partial x}$, i.e., the differential equation

$$\begin{cases} \dfrac{dx}{dt} = x^2 \\ x(0) = x_0 \end{cases}$$

Are solutions defined for all t? Let $x_0 = 1$ and let U_0 be $\frac{1}{2} < x < \frac{3}{2}$. Find the largest ϵ of the fundamental theorem for which $\Phi : U_0 \times (-\epsilon,\epsilon) \to \mathbb{R}$ is defined.

Note that we have required that the vector field v be differentiable, i.e., of class C^k for $k \geq 1$. One can lose uniqueness, for example, if the vector field is only continuous. On the real line, consider the differential equation $\frac{dx}{dt} = 3x^{\frac{2}{3}}$, i.e., the vector field $v = 3x^{\frac{2}{3}}\frac{\partial}{\partial x}$. The usual solutions are of the form $x(t) = (t-c)^3$ but there is also the "singular solution" $x(t) \equiv 0$.

(2) **Vector Fields on Manifolds.** If v is a C^k vector field, $k \geq 1$, on an open subset $\mathcal{U}$ of a manifold M^n then we can again recover a 1-parameter local group of diffeomorphisms $\phi_t : \mathcal{U} \to \mathcal{U}$, for the following reason. If $\mathcal{U}$ is contained in a coordinate patch $(\mathcal{U}_\alpha\ x_\alpha)$ we proceed just as above; $\mathcal{U}$ basically is an open subset of $\mathbb{R}^n$. If $\mathcal{U}$ is not contained in a single coordinate patch, consider a point $p \in \mathcal{U}$, $p \in \mathcal{U}_\alpha \cap \mathcal{U}_\beta$. In $\mathcal{U}_\alpha$ we solve the differential equations

$$\frac{dx_\alpha^i}{dt} = v_\alpha^i(x_\alpha^1, \ldots, x_\alpha^n)$$

as before. In $\mathcal{U}_\beta$ we solve the differential equations .

$$\frac{dx_\beta}{dt} = v_\beta^i(x_\beta^1, \ldots, x_\beta^n)$$

Because of the transformation rule (3.1), these differential equations in $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ say exactly the same thing ! Uniqueness of solutions allows us to construct unique diffeomorphisms ϕ_t.

~~ **WARNING.** Let $f : M^n \to \mathbb{R}$ be a differentiable real valued function on M^n. In elementary mathematics it is often said that the n-tuple $\frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \ldots, \frac{\partial f}{\partial x^n}$ in the coordinate system x forms a vector field "grad f ." But if one looks at the transformation properties

$$a_\beta^i \equiv \frac{\partial f}{\partial x_\beta^i} = \sum_{j=1}^n \frac{\partial f}{\partial x_\alpha^j} \frac{\partial x_\alpha^j}{\partial x_\beta^i} = \sum_{j=1}^n \frac{\partial x_\alpha^j}{\partial x_\beta^i} a_\alpha^j$$

and compares this with (3.1) we see that this "grad f" is not a contravariant vector field. Consequently the two systems of differential equations (in $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$)

$$\frac{dx_\alpha^i}{dt} = \frac{\partial f}{\partial x_\alpha^i} \quad \text{and} \quad \frac{dx_\beta^j}{dt} = \frac{\partial f}{\partial x_\beta^j}$$

do not say the same thing and $\therefore$ do not yield a flow ϕ ! What do $\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}$ form? See Lecture 5.

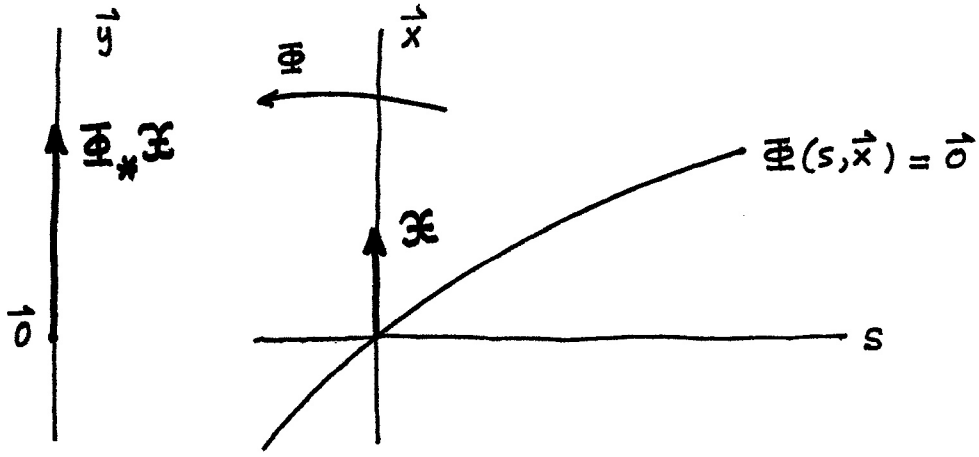
Appendix on "Fundamental Theorems"

Consider a curve $\tilde{\gamma}$ as a point in a huge function space. For any real number s (not to be confused with t) consider the differential equation

$$\frac{d\tilde{\gamma}}{dt} - s\, v(\tilde{\gamma}) = 0$$

as an equation $\Phi(s,\tilde{\gamma}) = 0$; we wish to solve this equation for $\tilde{\gamma}$ as a function of s. Then $\gamma(t) \equiv \tilde{\gamma}\left(\frac{t}{s}\right)$ will solve the equation $\frac{d\gamma}{dt}(t) = v(\gamma(t))$. For details of the following, see A.P.M., page 202.

Consider a special case of the implicit function theorem. Let $\Phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, i.e., $\vec{y} = \Phi(s,\vec{x})$, be a system of n functions $y^i = \Phi^i(s,x^1, \ldots, x^n)$ of (n+1)-variables. Suppose $\Phi(0,\vec{0}) = \vec{0}$. We wish to solve $\Phi(s,\vec{x}) = \vec{0}$ for $\vec{x} = f(s)$ near $s = 0$, $\vec{x} = \vec{0}$.



The implicit function theorem says that we can do this provided the "partial" differential (keeping $s = 0$) is a matrix $\frac{\partial y^i}{\partial x^j}(0,\vec{0})$ that represents an isomorphism, i.e., a linear transformation that is 1:1 and onto. One can prove a very general version of this, considering a map $\Phi : \mathbb{R} \times E \to F$ where E and F are infinite dimensional Banach spaces!

We are interested in solving the differential equations $\frac{\partial\gamma(t)}{dt} = v(\gamma(t))$, $\gamma(0) = \vec{0}$; here $\gamma(t)$ is to be a differentiable curve in $\mathbb{R}^n$ and $v : \mathbb{R}^n \to \mathbb{R}^n$ is a given differentiable vector field. For technical reasons we shall consider instead the system

$$\begin{cases} \dfrac{d\gamma}{dt}(t) - s\, v(\gamma(t)) = 0 \\ \gamma(0) = 0 \end{cases}$$

where s is a real number. Since $\gamma(t)$ is to be a differentiable curve in $\mathbb{R}^n$ we think of γ as a map of a closed interval $I = [-1,+1] = \{t \mid -1 \leq t \leq +1\} \subset \mathbb{R}$,

$$\gamma : I \longrightarrow \mathbb{R}^n \quad .$$

Let $C_0^1 =$ the space of all C^1 functions $\gamma : I \to \mathbb{R}^n$ such that $\gamma(0) = \vec{0}$. It turns out that C_0^1 is a Banach space (with suitable norm). Now for $\gamma \in C_0^1$, $\dfrac{d\gamma}{dt}(t)$ defines a new curve in $\mathbb{R}^n$, but it need only be continuous and $\dfrac{d\gamma}{dt}(0)$ need not be $\vec{0}$. Put

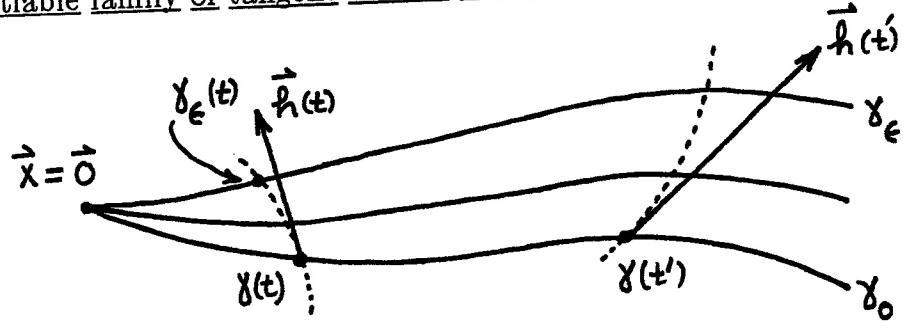
$$C^0 = \{\gamma : I \to \mathbb{R}^n \mid \gamma \text{ is continuous}\} \quad .$$

Then $\Phi(s,\gamma)(t) = \dfrac{d\gamma}{dt}(t) - s\, v(\gamma(t))$ defines a map

$$\Phi : \mathbb{R} \times C_0^1 \longrightarrow C^0 \quad .$$

Note that $\gamma(t) \equiv 0$ is a solution to $\Phi(0,\gamma) = 0$. Let us compute the "partial" differential of Φ, keeping $s = 0 \in \mathbb{R}$.

First, what is a tangent vector to the linear space C_0^1 at the "point" γ? Well, we have to consider a 1-parameter family γ_ϵ of such γ (i.e., a curve of such γ's). We then have $\gamma_\epsilon(t)$ and we consider $\dfrac{\partial}{\partial \epsilon}\gamma_\epsilon(t)\Big|_{\epsilon=0} \equiv h(t)$. <u>This is simply a differentiable family of tangent vectors along the curve</u> γ <u>in</u> $\mathbb{R}^n$



Of course $h(0) = \vec{0}$. The tangent space to C_0^1 (i.e., the collection of all such h) is simply another copy of C_0^1! Likewise, a tangent vector to $\{k(t)\}$ in C^0 is a continuous vector field defined in $\mathbb{R}^n$ along the curve $k(t)$. For "partial" differential

of Φ at $(s = 0\,,\ \gamma \equiv 0)$ we will have a linear map

$$\Phi_*^{s=0} = \overset{s=0}{D\Phi} : C_0^1 \longrightarrow C^0 \quad .$$

Evaluate it on the tangent vector h at $s = 0\,,\ \gamma = 0$ and so $\gamma(t) \equiv 0)$

$$\begin{aligned}
\overset{s=0}{D\Phi}(h) &\equiv \frac{\partial}{\partial \epsilon}\, \Phi(0,\vec{0} + \epsilon h)\,\Big|_{\epsilon=0,\ s=0}(t) \\
&= \frac{\partial}{\partial \epsilon}\left[\frac{\partial(\vec{0} + \epsilon\, h(t))}{\partial t} - s\, v(\vec{0} + \epsilon\, h(t))\right]_{\epsilon=0,\ s=0} \\
&= \left[\frac{\partial h(t)}{\partial t} - s\,\frac{\partial}{\partial \epsilon}\, v(\epsilon\, h(t))\,\Big|_{\epsilon=0}\right]_{s=0} = \frac{\partial h}{\partial t}\,(t)
\end{aligned}$$

and so $\overset{s=0}{D\Phi}(h)(t) = \dfrac{dh(t)}{dt}$. We see first that this differential is 1:1, for if $\dfrac{dh}{dt} = 0 \in C^0$ then $\dfrac{dh}{dt}(t) = 0$ for all $-1 \le t \le +1$ and $h(0) = 0$, thus $h(t) \equiv 0$. $\overset{s=0}{D\Phi}$ is 1:1. Also $\overset{s=0}{D\Phi}$ is onto, for if $k(t)$ is a given continuous curve, i.e., $k \in C^0$, then we can find h such that $D\Phi(h) = k$, namely

$$h(t) = \int_0^t k(\tau)d\tau \quad .$$

It can also be shown (using the norms of C_0^1 and C^0, that $\overset{s=0}{D\Phi}$ is continuous (i.e., bounded). Thus $\overset{s=0}{D\Phi}$ is an isomorphism at $s = 0\,,\ \gamma = 0$. By the generalized implicit function theorem we can solve

$$\Phi(s,\tilde{g}) = \frac{d\tilde{g}}{dt}\,(t) - s\, v(\tilde{g}(t)) = 0$$

for $\tilde{g} = \gamma_s$ provided s is sufficiently near 0. But then if we define, for such an $s \neq 0$

$$\gamma(t) \equiv \gamma_s\left(\frac{t}{s}\right)$$

one easily gets $\dfrac{d\gamma}{dt}\,(t) = v(\gamma(t))$, as desired.

A slight modification of the above shows in fact that the solutions depend differentiably on the initial point $\gamma(0)$ chosen. The full force of the implicit

function theorem will also show that the solution γ depends differentiably on the vector field v.

I hope this sketch convinces you of a need to have an infinite dimensional version of the implicit function theorem.

Lecture 5. Covectors and Riemannian Metrics

(1) **Some Linear Algebra; the Dual Space.** Let E be a real vector space (E need not be equipped with a basis and while E may be ∞ dimensional our main concern at present is with finite dimensional E). If E is n-dimensional, let $e_1 \cdots e_n$ be a basis, i.e., a linearly independent set of vectors in E that span E. Thus if $v \in E$, v has a unique expansion

$$v = \sum_{i=1}^{n} e_i v^i \qquad \left(= \sum_{i=1}^{n} v^i e_i \right)$$

for real numbers $v^1, \ldots, v^n$, the "components" of v. We shall usually write the components v^i to the <u>right</u> of the basis vectors. This is not done in most treatments of linear algebra but it does have many advantages. The <u>dis</u>advantage is that equation (3.2′) should be written $v_{p_0} = \sum_{i=1}^{n} \frac{\partial}{\partial x^i}\Big|_{p_0} v^i$, making it <u>appear</u> that $\frac{\partial}{\partial x^i}$ is operating on v^i, which is not the case.

Symbolically, we write matrices $e = (e_1, \ldots, e_n)$, $v = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$; the first matrix is symbolic because the entries are <u>vectors</u> rather than numbers. We then have a symbolic matrix product

$$v = \sum_{i=1}^{n} e_i v^i = (e_1 \cdots e_n) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \equiv ev \quad .$$

We see that an n-dimensional vector space E <u>with a basis</u> is isomorphic to $\mathbb{R}^n$, $v \in E \leftrightarrow \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \in \mathbb{R}^n$. This isomorphism $E \approx \mathbb{R}^n$ however is <u>unnatural</u> since it requires a basis to be given in E and there are many bases in E, each giving a different isomorphism with $\mathbb{R}^n$.

Definition. A (real) <u>linear functional</u> α on E is a linear transformation $\alpha : E \to \mathbb{R}$ to the 1-dimensional vector space $\mathbb{R} = \mathbb{R}^1$.

Thus α is a real valued function on E such that

$$\alpha(c_1 v + c_2 w) = c_1\alpha(v) + c_2\alpha(w)$$

for all vectors v,w in E and real numbers c_1 , c_2 .

Example. If E has basis $e_1 \cdots e_n$ and if $a_1, \ldots, a_n$ are real numbers, then

$$\alpha(v) = \alpha\left(\sum e_i v^i\right) \equiv \sum a_i v^i$$

defines a linear functional (check this). Moreover, we claim that every linear functional is of this form, for if β is a linear functional $\beta(v) = \beta\left(\sum e_i v^i\right) \underset{\substack{\uparrow \\ \text{(why?)}}}{=} \sum \beta(e_i)v^i$, $\therefore$ put

$a_i \equiv \beta(e_i)$.

Example. E = space of continuous real valued functions $f : \mathbb{R} \to \mathbb{R}$ (an ∞ dimensional space). For each $a \in \mathbb{R}$ we have the Dirac functional δ_a defined by $\delta_a(f) \equiv f(a)$

Definition. If E is a vector space, the collection of all linear functionals α on E defines a new vector space E*, the dual space to E, under operations

$$\begin{cases} (\alpha+\beta)(v) \equiv \alpha(v) + \beta(v) & \alpha,\beta \in E^* \\ (c\alpha)(v) \equiv c\,\alpha(v) & c \in \mathbb{R}\,,\ v \in E \end{cases}$$

If $e_1, \ldots, e_n$ is a basis for E, define the dual basis $\sigma^1, \ldots, \sigma^n$ of E* by

$$\sigma^i(e_j) = \delta_j^{\,i} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}\,, \quad \text{the Kronecker delta} \quad .$$

(note that from our example above $\beta\left(\sum e_i v^i\right) = \sum \beta(e_i)v^i$ we see that a linear functional β on E is determined when we know its values $b_i \equiv \beta(e_i)$ on any basis e). Thus $\sigma^i(v) = \sigma^i\left(\sum_j e_j v^j\right) = \sum_j \sigma^i(e_j)v^j = \sum_j \delta_j^{\,i} \cdot v^j = v^i$, and so σ^i is the linear functional that reads off the i^{th} component of any vector (with respect to basis $e_1, \ldots, e_n$). (Note, e.g., $(v+w)^i = v^i + w^i$, etc.) It is easy to show in this finite dimensional case that the $\sigma^1, \ldots, \sigma^n$ form a basis, for

(i) $0 = \sum a_i\sigma^i \Rightarrow \left(\sum a_i\sigma^i\right)(e_j) = 0 \Rightarrow \sum a_i\sigma^i(e_j) = 0$
$\Rightarrow \sum_i a_i\delta_j^{\,i} = 0 \Rightarrow a_j = 0$, all j . $\therefore$ linearly independent.

(ii) If $\beta \in E^*$, $\beta(v) = \sum \beta(e_j)v^j$. But then

$$\boxed{\beta = \sum_j \beta(e_j)\sigma^j \,,} \tag{5.1}$$

i.e., $\sigma^1, \ldots, \sigma^n$ span E^* because β and $\sum \beta(e_j)\sigma^j$ are two linear functionals that take the same value on every vector v.

We write symbolically $\sigma = \begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^n \end{pmatrix}$. Then for $\beta \in E^*$, $\beta = \sum b_i\sigma^i$, $b_i \equiv \beta(e_i)$, can be written

$$\beta = \sum_i b_i\sigma^i = (b_1 \cdots b_n)\begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^n \end{pmatrix} \equiv b\sigma \quad .$$

Thus when vectors $v \in E$ are identified with column vectors $v = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$ linear functionals $\beta \in E^*$ are identified with row vectors $b = (b_1 \cdots b_n)$. Then

$$\beta(v) = (b_1 \cdots b_n)\begin{pmatrix} \sigma^1 \\ \vdots \\ \sigma^n \end{pmatrix}(e_1 \cdots e_n)\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} = (b_1 \cdots b_n)\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}$$

is $b\sigma ev = b(\sigma e)v = bIv = bv$, $I =$ identity .

$\sim\sim$ Note that E and E^* are <u>different spaces</u> . In the example of the Dirac δ-functional, one would never confuse δ_a with a continuous function!

(2) **The Differential of a Function.** If $f : M^n \to \mathbb{R}$, its "differential" df is <u>by definition</u> the linear functional on tangent spaces $df : M_p^n \to \mathbb{R}$ given by (see (3.2)):

$$\boxed{df(v_p) = v_p(f)} \tag{5.2}$$

In local coordinates, $e_1 = \dfrac{\partial}{\partial x^1}\Big|_p, \ldots, e_n = \dfrac{\partial}{\partial x^n}\Big|_p$ form a basis for M_p^n

$$df\left(\sum v^i(p)\,\frac{\partial}{\partial x^i}\Big|_p\right) = \sum v^i(p)\,\frac{\partial f}{\partial x^i}\Big|_p \quad .$$

In particular we may consider the differential of a coordinate function

$$\begin{cases} dx^i(e_j) = dx^i\left(\dfrac{\partial}{\partial x^j}\right) \equiv \dfrac{\partial x^i}{\partial x^j} = \delta_j{}^i \\ dx^i(v) = dx^i\left(\sum\limits_j v^j \dfrac{\partial}{\partial x^j}\right) = \sum\limits_j v^j\, dx^i\left(\dfrac{\partial}{\partial x^j}\right) = v^i \end{cases}$$

Thus the linear functional dx^i reads off the i^{th} component of any vector , when expressed in terms of the coordinate basis $e_1 = \dfrac{\partial}{\partial x^1}, \ldots, e_n = \dfrac{\partial}{\partial x^n}$, and furthermore $dx^1, \ldots, dx^n$ form the dual basis to the $\dfrac{\partial}{\partial x}$!

The $dx^1, \ldots, dx^n$ give a basis for $(M_p^n)^*$, not M_p^n . The most general linear functional to M_p^n is then of the form

$$\begin{cases} \alpha = \sum\limits_i a_i\, dx^i \\ a_i \equiv \alpha\left(\dfrac{\partial}{\partial x^i}\right) \end{cases} \tag{5.3}$$

(see (5.1)). A linear functional $\alpha : M_p^n \to \mathbb{R}$ is called a covariant vector, or covector , or 1-form , or Pfaffian . A differentiable assignment of a covector is locally of the form

$$\alpha = \sum a_i(x) dx^i$$

where $a_i(x)$ are differentiable functions. This is called a covector field, etc. The components $a_i(x)$ are simply $\alpha\left(\dfrac{\partial}{\partial x^i}\right)$. Then for any differentiable function f

$$df = \sum_i df\left(\frac{\partial}{\partial x^i}\right) dx^i = \sum_i \frac{\partial f}{\partial x^i}\, dx^i \,, \quad \text{i.e.,}$$

$$\boxed{df = \sum_i \frac{\partial f}{\partial x^i}\, dx^i} \tag{5.4}$$

Thus the numbers $\frac{\partial f}{\partial x^1}, \ldots, \frac{\partial f}{\partial x^n}$ (considered in Lecture 4, §4), are the components not of a vector but rather of a covector, the differential of f. Thus in spherical coordinates r,θ,ϕ in $\mathbb{R}^3$

$$df = \frac{\partial f}{\partial r}\, dr + \frac{\partial f}{\partial \theta}\, d\theta + \frac{\partial f}{\partial \phi}\, d\phi$$

and these three components are nothing like the components of what one calls the "gradient vector" of f. To discuss the gradient vector we must introduce a Riemannian structure, as we will, shortly.

Under a change of local coordinates

$$d\bar{x}^i = \sum_j d\bar{x}^i\left(\frac{\partial}{\partial x^j}\right)dx^j = \sum_j \frac{\partial \bar{x}^i}{\partial x^j}\, dx^j \quad .$$

(Note that this is simply a statement of the chain rule.) Then for a general 1-form

$$\sum_j a_j dx^j = \sum_j a_j \sum_i \frac{\partial x^j}{\partial \bar{x}^i}\, d\bar{x}^i \equiv \sum_i \bar{a}_i\, d\bar{x}^i$$

says

$$\bar{a}_i = \sum_j a_j \frac{\partial x^j}{\partial \bar{x}^i}$$

i.e.,

$$\boxed{a_i{}^\beta = \sum_{j=1}^{n} a_j{}^\alpha \frac{\partial x_\alpha^j}{\partial x_\beta^j}} \tag{5.5}$$

is the transformation law for the components of a covariant vector. Compare with (3.1) for contravariant vector components.

- **Problem 9.** If v is a vector and α is a covector, compute directly that $\sum \bar{a}_i \bar{v}^i = \sum a_i v^i$. What happens if w is another vector and one considers $\sum v^i w^i$?

(3) **Riemannian Manifolds; the Gradient Vector.** Let E be an n-dimensional real vector space with an inner (≡ scalar) product $\langle\ ,\ \rangle$. Thus for each pair of vectors v, w of E, $\langle v,w \rangle$ is a real number and $\langle v,w \rangle$ is linear in each entry when

the other is held fixed and $\langle v,w\rangle = \langle w,v\rangle$. We further demand that $\langle\ ,\ \rangle$ be <u>non-singular</u> in the sense that $\langle v,w\rangle = 0$ for all $w \in E$ implies that $v = 0 \in E$. To accommodate relativity we shall not always demand that $\|v\|^2 \equiv \langle v,v\rangle$ be positive or nonzero for $v \neq 0$. If $\langle v,v\rangle > 0$, all $v \neq 0$, we say $\langle\ ,\ \rangle$ is <u>positive definite.</u> If $e_1, \ldots, e_n$ is a basis of E, then the matrix

$$\boxed{g_{ij} \equiv \langle e_i,e_j\rangle} \tag{5.6}$$

is called the (matrix of the) <u>metric tensor</u>. It is symmetric and positive definite if $\langle\ ,\ \rangle$ is. Then if $v = \sum e_i v^i$, $w = \sum e_i w^i$

$$\langle v,w\rangle = \Big\langle \sum_i e_i v^i , \sum_j e_j w^j \Big\rangle = \sum_{i,j} v^i w^j \langle e_i,e_j\rangle$$
$$= \sum_{i,j} v^i w^j g_{ij} = \sum_{i,j} v^i g_{ij} w^j$$

$$\boxed{\langle v,w\rangle = \sum_{i,j} g_{ij} v^i w^j} \tag{5.7}$$

Note that if $e_1 \cdots e_n$ forms an <u>orthonormal</u> (o.n.) basis then $g_{ij} = \delta_{ij}$ is simply the identity matrix and $\langle v,w\rangle = \sum_i v^i w^i$.

By hypothesis $\langle v,w\rangle$ is a <u>linear function</u> of w when v is held fixed. Thus if $v \in E$, the function f_v defined by

$$f_v(w) = \langle v,w\rangle$$

is a linear functional, $f_v \in E^*$. Thus, putting $\nu = f_v$, $\nu = \sum \nu_i \sigma^i$, we have (see (5.1)):

$$\nu_j = \nu(e_j) = \langle v,e_j\rangle = \Big\langle \sum_i v^i e_i, e_j \Big\rangle$$
$$= \sum_i v^i \langle e_i,e_j\rangle = \sum_j v^i g_{ij} \quad .$$

Thus if E has an inner product, to each <u>vector</u> $v \in E$ with components v^i we can associate a <u>covector</u> ν, with components $\nu_j = \sum_i v^i g_{ij}$, having the property

$$\boxed{\nu(w) = \langle v,w \rangle} \tag{5.8}$$

all $w \in E$. Note that if $e_1 \cdots e_n$ is an orthonormal basis then $\nu_i = v^i$. $\nu_j = \sum_i v^i g_{ij}$ are called the "covariant components" of the vector v. Note that in our finite dimensional case, every covector $\alpha \in E^*$ is of the form $\alpha(w) = \langle v,s \rangle$ for some vector v. This is true because if $\alpha = \sum a_i \sigma^i$, then $\alpha(w) = \sum_j a_j w^j$ and we can write $\alpha(w) = \langle v,w \rangle = \sum_{i,j} v^i g_{ij} w^j$ merely by solving

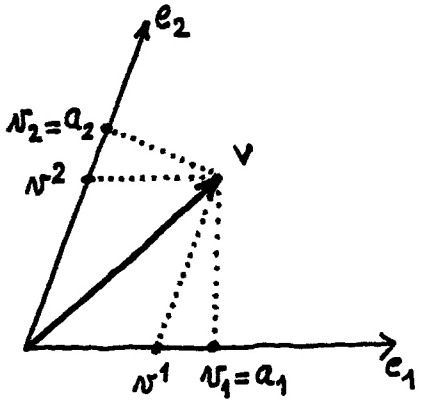
$$a_j = \sum_i v^i g_{ij}$$

for v^i. The solution is

$$\boxed{v^i = \sum_j g^{ij} a_j} \tag{5.9a}$$

$$(g^{ij}) \equiv \text{inverse matrix to } (g_{ij}) \ .$$

In the figure below e_1 and e_2 are unit vectors, but not orthogonal. Note that in general



$$a_j = \sum_i v^i g_{ij} = \sum_i v^i \langle e_i,e_j \rangle$$
$$= \left\langle \sum_i v^i e_i, e_j \right\rangle = \langle v,e_j \rangle$$

$$\boxed{a_j = \langle v,e_j \rangle} \tag{5.10}$$

Notation. In tensor analysis we use the same letter for the components of the covector ν associated with the vector v; we merely lower the indices. Thus if $v = \sum_i e_i v^i$, then $\nu = \sum v_j \sigma^j$ where

$$\boxed{v_j = \sum_i v^i g_{ij}} \tag{5.9b}$$

Likewise, if $\alpha = \sum a_i \sigma^i$ is a covector, we can associate a vector A with it having the property $\alpha(w) = \langle A, w \rangle$, all $w \in E$. Then we write $a^i \equiv A^i = \sum_j g^{ij} a_j$, where (g^{ij}) is the inverse matrix to (g_{ij}). In finite dimensional spaces E, an inner product $\langle \, , \rangle$ allows us to "identity" E with E*. In terms of a basis $v_1, \ldots, v_n \leftrightarrow v^1, \ldots, v^n$.

Definition. A Riemannian metric on a manifold M^n assigns in a differentiable fashion a positive definite inner product $\langle \, , \rangle$ in each tangent space M_p^n.

(If $\langle \, , \rangle$ is not positive definite it is called indefinite and the resulting structure on M^n is a "pseudo-Riemannian" metric). Then in terms of the coordinate basis $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ we have the differentiable matrices ("metric tensor")

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle .$$

In an overlap, with new coordinates $\bar{x}$

$$\bar{g}_{ij} = \left\langle \frac{\partial}{\partial \bar{x}^i}, \frac{\partial}{\partial \bar{x}^j} \right\rangle = \left\langle \sum_l \frac{\partial x^l}{\partial \bar{x}^i} \frac{\partial}{\partial x^l}, \sum_m \frac{\partial x^m}{\partial \bar{x}^j} \frac{\partial}{\partial x^m} \right\rangle$$

$$= \sum_{l,m} \frac{\partial x^l}{\partial \bar{x}^i} \frac{\partial x^m}{\partial \bar{x}^j} \left\langle \frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^m} \right\rangle ,$$

i.e.,

$$\bar{g}_{ij} = \sum_{l,m} \frac{\partial x^l}{\partial \bar{x}^i} \frac{\partial x^m}{\partial \bar{x}^j} g_{lm} \tag{5.11}$$

Definition. If M^n is a (pseudo-) Riemannian metric and f is a differentiable function $f : M^n \to \mathbb{R}$, the gradient vector

$$\text{grad } f = \nabla f$$

is the contravariant vector associated to the covector df

$$df(w) = \langle \nabla f, w \rangle . \tag{5.12}$$

In coordinates, from (5.9a)

$$\boxed{(\nabla f)^i = \sum_j g^{ij} \frac{\partial f}{\partial x^j}} \tag{5.13}$$

Example. In relativity (special) $M^4 = \mathbb{R}^4$ with metric given in inertial coordinates $t = x^0$, $x = x^1$, $y = x^2$, $z = x^3$, by

$$g_{ij} = \left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle = \begin{cases} 1 \text{ if } i = j = 1, 2, \text{ or } 3 \\ -1 \text{ if } i = j = 0 \\ 0 \text{ otherwise} \end{cases},$$

i.e., $(g_{ij}) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Then

$$df = \frac{\partial f}{\partial t}\, dt + \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz$$

i.e.,

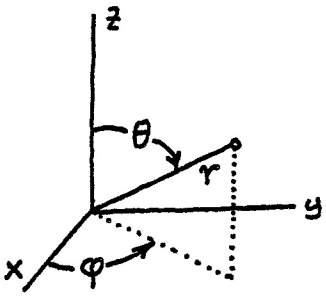
$$df \text{ "}=\text{"} \left(\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

but

$$\nabla f \text{ "}=\text{"} \begin{pmatrix} -\dfrac{\partial f}{\partial t} \\ \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \\ \dfrac{\partial f}{\partial z} \end{pmatrix}, \text{ i.e., } \nabla f = -\frac{\partial f}{\partial t}\frac{\partial}{\partial t} + \frac{\partial f}{\partial x}\frac{\partial}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial}{\partial y} + \frac{\partial f}{\partial z}\frac{\partial}{\partial z}.$$

Example. In $\mathbb{R}^3$ with Cartesian coordinates x^1, x^2, x^3 we have standard metric $g_{ij}(x) = \delta_{ij}$. Hence df and ∇f will have exactly the same components $(df)_i = \dfrac{\partial f}{\partial x^i} = (\nabla f)^i$.

• **Problem 10.** In $\mathbb{R}^3$ let $x^1 = x$, $x^2 = y$, $x^3 = z$ be usual Cartesian coordinates and let $\bar{x}^1 = r$, $\bar{x}^2 = \theta$, $\bar{x}^3 = \phi$ be usual spherical coordinates. Compute the metric tensor $\bar{g}_{ij}$ in the coordinates $\bar{x}$, i.e., in spherical coordinates



$$\bar{g}_{r\theta} \equiv \bar{g}_{12} = \left\langle \frac{\partial}{\partial r}, \frac{\partial}{\partial \theta} \right\rangle \text{ etc.}$$

(Note; don't fiddle with matrices, just use chain rule $\frac{\partial}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \cdots$). Compute the coefficients $\overline{(\nabla f)}^i$ in

$$\nabla f = (\overline{\nabla f})^r \frac{\partial}{\partial r} + (\overline{\nabla f})^\theta \frac{\partial}{\partial \theta} + (\overline{\nabla f})^\phi \frac{\partial}{\partial \phi} \quad .$$

Verify that $\frac{\partial}{\partial r}$, $\frac{\partial}{\partial \theta}$, $\frac{\partial}{\partial \phi}$ are orthogonal, but not all unit vectors. Put

$$\hat{e}_r = \frac{\frac{\partial}{\partial r}}{\left\|\frac{\partial}{\partial r}\right\|} \quad \hat{e}_\theta = \frac{\frac{\partial}{\partial \theta}}{\left\|\frac{\partial}{\partial \theta}\right\|} \quad \hat{e}_\phi = \frac{\frac{\partial}{\partial \phi}}{\left\|\frac{\partial}{\partial \phi}\right\|}$$

and write ∇f in terms of this orthonormal set

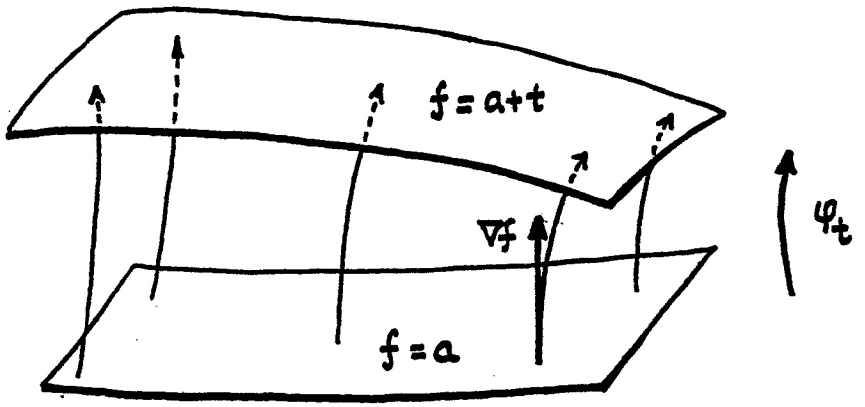
$$\nabla f = (\nabla f)^r \hat{e}_r + (\nabla f)^\theta \hat{e}_\theta + (\nabla f)^\phi \hat{e}_\phi \quad .$$

Check that you get the usual expressions in physics books for the components of grad f in this orthonormal system. (You may then wish to ask yourself; why do we need ∇f in spherical coordinates when df is so simple?)

(4) **Morse deformations.** Let $f : M^n \to \mathbb{R}$ be a smooth function defined on Riemmanian M^n. If $f(x_0) = a$ then the "level set" of f through x_0 is the set $M^{n-1}(a) = \{x \mid f(x) = a\}$. If $df = \sum_i \frac{\partial f}{\partial x^i} dx^i$ does not vanish at x_0 then, near x_0, $M^{n-1}(a)$ is an (n-1)-dimensional submanifold of M^n (keep the toroidal picture of Lecture 3, §2 in mind!). If x(t) is a curve in $M^{n-1}(a)$ through x_0 then its velocity vector dx/dt is "annihilated" by df; $df\left(\frac{dx}{dt}\right) = \sum \frac{\partial f}{\partial x^i}\frac{dx^i}{dt} = 0$ since f(x(t)) is constant. We would like to say df is "orthogonal" to the tangent space $M^{n-1}_{x_0}(a)$ to $M^{n-1}(a)$ but this makes no sense since df is not a contravariant vector. Its contravariant version ∇f is however orthogonal to $M^{n-1}(a)$ in the sense that $\left\langle \nabla f, \frac{dx}{dt} \right\rangle = df\left(\frac{dx}{dt}\right) = 0$ for all tangents to $M^{n-1}(a)$ at x_0.

Recall that one does not get a well defined flow by considering $\frac{dx^i}{dt} = \frac{\partial f}{\partial x^i}$. If, however, $\lambda(x)$ is any positive smooth function, $\lambda(x) > 0$, defined on M^n in a neighborhood of x_0, then $\frac{dx}{dt} = \lambda(x)\,\nabla f$, i.e., $\frac{dx^i}{dt} = \lambda(x) \sum_j g^{ij} \frac{\partial f}{\partial x^j}$, does have a well-defined flow; its integral curves are tangent to ∇f, orthogonal to the level hypersurfaces $f =$ constant (any (n-1)-dimensional submanifold of an M^n is called a hypersurface) and are called "gradient curves or lines" or curves of "steepest ascent".

- **Problem 11.** Since $\nabla f \neq 0$ at x_0 it is also nonvanishing in some neighborhood of x_0. Show that one can choose λ so that one has $\frac{df}{dt} \equiv 1$ along the gradient curves. This shows that the flow ϕ_t generated by this vector field $\lambda(x)\nabla f$ sends the level hypersurface $M^{n-1}(a)$ locally into the level hypersurface $M^{n-1}(a+t)$!



Note that these flows ϕ_t are only locally defined and then for only short times perhaps; that's all the fundamental theorem of Lecture 4 guarantees. There are theorems that tell us when the flow ϕ_t generated by a differentiable vector field v is gloablly defined, for all t; for example if M^n is a compact manifold this is so. We shall discuss such things later, in Lecture ?, but the interested reader can see A,M, and R, 4.1.20, page 196.

Lecture 6. Tangent and Cotangent Bundles. Phase Space

(1) **The tangent bundle** to a differentiable manifold M^n is the space of all tangent vectors to M^n. Thus a "point" in this new space consists of a pair (x,v) where x is a point of M^n and v is a tangent vector to M^n at x, i.e., $v \in M^n_x$. Denote the tangent bundle to M^n by TM^n. We can introduce local coordinates in TM as follows. Let $(x,v) \in TM^n$. x lies in some local coordinate system $U, x^1, \ldots, x^n$ of M^n. At x we have a (coordinate) basis $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ for M^n_x. We can then express $v = \sum_{i=1}^{n} v^i \frac{\partial}{\partial x^i}$. Then (x,v) is completely described by the 2n-tuple of real numbers.

$$x^1, \ldots, x^n, v^1(x), \ldots, v^n(x) \quad \begin{cases} \text{here } x = (x^1, \ldots, x^n) \in u \\ v = (v^1, \ldots, v^n) \in \mathbb{R}^n \end{cases}$$

If x also lies in the coordinate patch $\bar{u}, \bar{x}^1, \ldots, \bar{x}^n$, the same $(x,v) \in TM^n$ would be described by a new 2n-tuple

$$\bar{x}^1, \ldots, \bar{x}^n, \quad \bar{v}^1(x), \ldots, \bar{v}^n(x)$$

where

$$\begin{cases} \bar{x}^i = \bar{x}^i(x^1, \ldots, x^n) \\ \bar{v}^i = \sum_j \frac{\partial \bar{x}^i}{\partial x^j}(x)\, v^j(x) \end{cases} \tag{6.1}$$

We see that TM is a 2n- dimensional differentiable manifold!

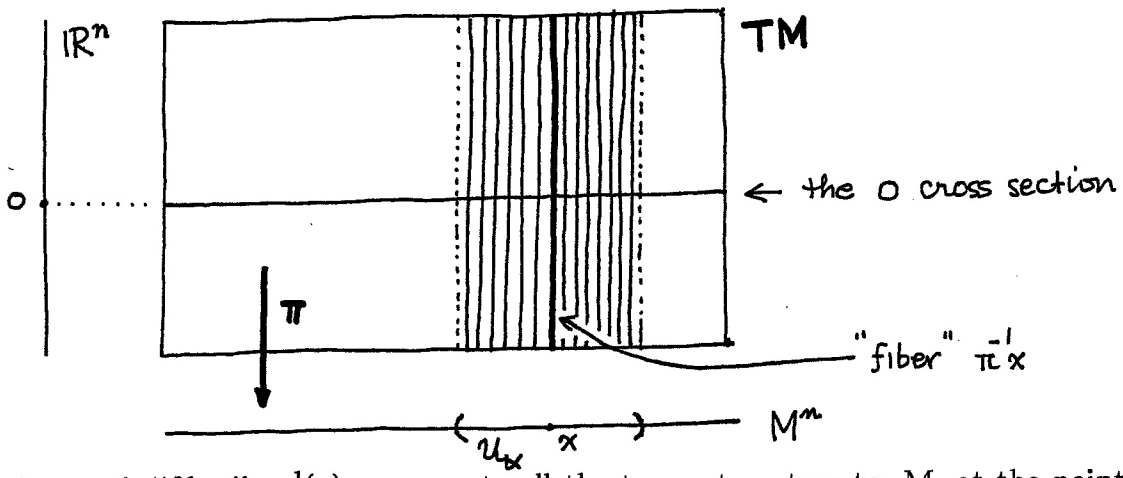
Note that we have a natural mapping

$$\pi : TM \to M \quad .$$

This merely takes a point of TM, i.e., a pair consisting of a vector v at a point $x \in M$ and gives back the point x at which the vector is located. Over the coordinate patch $u_\alpha, x^1_\alpha, \ldots, x^n_\alpha$ on M we have a coordinate patch

$$(x^1_\alpha, \ldots, x^n_\alpha, v^1_\alpha, \ldots, v^n_\alpha)$$

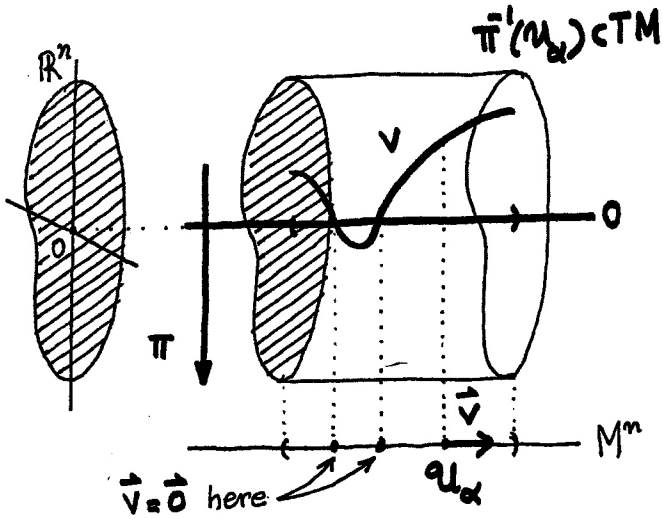
of TM and π is simply the map $\pi(x,v) = x$. Schematically we draw the following picture.



Note that each "fiber" $\pi^{-1}(x)$ represents all the tangent vectors to M at the point x; this is simply M^n_x, a vector space isomorphic to $\mathbb{R}^n$. $\pi^{-1}(U_\alpha)$ is in a natural way simply

$$\pi^{-1}(U_\alpha) = U_\alpha \times \mathbb{R}^n \quad .$$

~~ **Warning.** TM is not usually globally $M \times \mathbb{R}^n$, even though locally it is a product $\pi^{-1}(\mathcal{U}_\alpha) = \mathcal{U}_\alpha \times \mathbb{R}$; it is usually twisted. Locally "over $\mathcal{U}_\alpha$," we always draw a product picture. Note that we always have a copy of M^n contained in TM, namely the 0 vector field to M^n.



Given another vector field v on M we get another copy of M^n by means of the mapping $v : M \to TM$ locally given by $(x^1, \ldots, x^n) \to (x^1, \ldots, x^n, v^1(x), \ldots, v^n(x)$ where $v = \sum v^i(x) \dfrac{\partial}{\partial x^i}$ is the local expression of v. Such a map $v : M \to TM$ satisfies $\pi \circ v =$ identity map $M \to M$ and is called a "(cross) section" of TM.

Example. In mechanics the configuration of a dynamical system is usually described as a point in an n-dimensional manifold M^n, the configuration space. The coordinates x are usually denoted by $q^1, \ldots, q^n$, the "generalized coordinates." For example if we are considering two mass points moving on the real line, $M^2 = \mathbb{R} \times \mathbb{R}$ with coordinates q^1, q^2 (one for each particle). A tangent vector to the configuration space M^n

is thought of, in mechanics, as a velocity vector ; its components with respect to the coordinates $q^1, \ldots, q^n$ are written $\dot{q}^1, \ldots, \dot{q}^n$, rather than $v^1, \ldots, v^n$. These are the generalized velocities . Thus TM^n is the generalized velocity space to the configuration space, but there is no standard name for this space in mechanics (it is not the phase space, see §3). The configuration space need not be Euclidean space. For the planar double pendulum, Lecture 2(§2c). $M^2 = S^1 \times S^1 = T^2$. For the spatial (single) pendulum M^2 is the 2-sphere (center at pin).

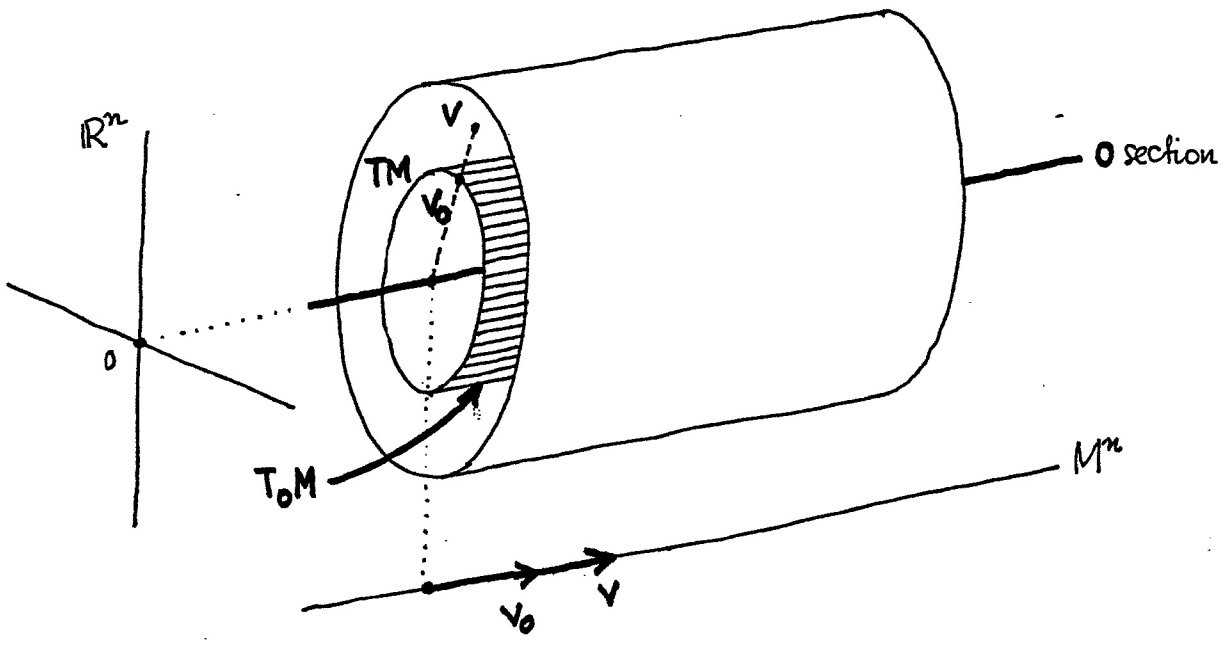
Example. If M^n is a Riemannian manifold, i.e., a manifold endowed with a Riemanian metric (Lecture 5, §3), then instead of considering all tangent vectors to M^n we may restrict ourselves to all unit tangent vectors. Thus if we are in the coordinate patch $(x^1, \ldots, x^n, v^1, \ldots, v^n)$ of M^n we may restrict ourselves to the subset defined by $\| v \|^2 = 1$, i.e., the unit tangent bundle T_0M^n

$$T_0M^n : \sum_{i,j} g_{ij}(x)v^iv^j = 1 \ .$$

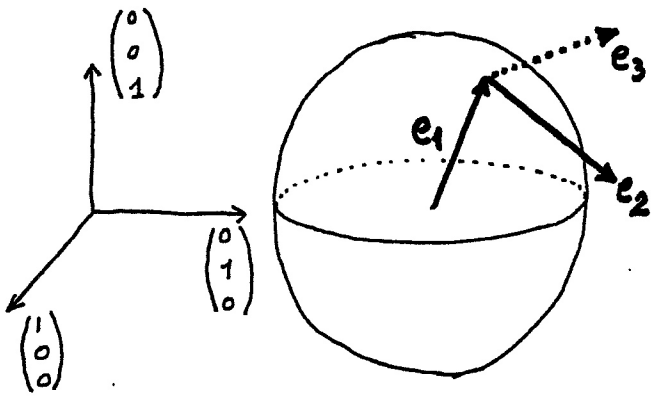
On TM^n we are putting $f(x,v) = \sum g_{ij}(x)v^iv^j = 1$. Note that $(x^1, \ldots, x^n, v^1, \ldots, v^n)$ are the local coordinates for TM and

$$\frac{\partial f}{\partial v^i} = 2 \sum_{j=1}^{n} g_{ij}(x)v^j \ .$$

Since $\det (g_{ij}) \neq 0$, we conclude that not all $\frac{\partial f}{\partial v^i}$ can be zero for $v \neq 0$. Thus we can solve for some coordinate v^α in terms of the remaining v's and x's. We conclude (see Lecture 2, §2), that T_0M^n is a (2n-1)- dimensional submanifold of the manifold TM^n. In particular T_0M^n is itself a (2n-1)-dimensional manifold.



Example. Let $M^2 = S^2$ be the unit 2-sphere in $\mathbb{R}^3$. The tangent bundle $TM^2 = TS^2$ is 4-dimensional. We shall consider instead the 3-dimensional T_0S^2.



A unit tangent vector e_2 to S^2 is located at the point described by the unit vector e_1 from the origin. Thus a point in T_0S^2 is uniquely described by e_1 and e_2 together. For free we may throw in $e_3 \equiv e_1 \times e_2$. Thus a point in T_0S^2 is uniquely described by the orthonormal triple e_1, e_2, e_3 of vectors in $\mathbb{R}^3$. By parallel translating e_2 and e_3 to the origin we have an orthonormal basis of $\mathbb{R}^3$, and comparing with the standard basis, $\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}$ we see that the points of T_0S^2 are in 1:1 correspondence with the points of the rotation group SO(3). It should also be clear that the topology of T_0S^2 is the same as that of SO(3) (nearby tangent vectors e_2 correspond to nearby orthogonal matrices). We have seen in Lecture 2, §e, that SO(3) is topologically $P^3(\mathbb{R})$.

The unit tangent bundle T_0S^2 to the 2-sphere is topologically the 3-dimensional real projective 3-space $P^3(\mathbb{R}) \approx SO(3)$.

(2) **The Cotangent Bundle** to M^n is the space T^*M^n of all covectors on M^n. A point in T^*M^n is a pair (x,α) where α is a covector at the point x. If x is in the coordinate patch U, $x^1, \ldots, x^n$ of M^n, we have basis $dx^1, \ldots, dx^n$ for $M_x^n)^*$, and α can be expressed as $\alpha = \sum a_i\, dx^i$. (x,α) is completely described by the $2n$ tuple

$$x^1, \ldots, x^n, a_1(x), \ldots, a_n(x)$$

and we have, if p is also in $\overline{U}$, $\overline{x}^1, \ldots, \overline{x}^n$,

$$\begin{cases} \overline{x}^i = \overline{x}^i(x^1, \ldots, x^n) \\ \overline{a}_i = \sum\limits_j \dfrac{\partial x^j}{\partial \overline{x}^i}\, a_j \end{cases} \tag{6.2}$$

T^*M^n is again a $2n$-dimensional manifold.

(3) **Phase space of Mechanics.** Let M^n be the configuration space of a dynamical system, with local generalized coordinates $q^1, \ldots, q^n$. The velocity space is simply TM^n, the tangent bundle to M^n, with local coordinates $q^1, \ldots, q^n$, $\dot{q}^1, \ldots, \dot{q}^n$. For simplicity we shall consider a time-independent Lagrangian $L = L(q,\dot{q})$, thus $L : TM^n \to \mathbb{R}$; this is merely a function on the tangent bundle to M^n. (If the Lagrangian is time dependent we would consider $L : TM^n \times \mathbb{R} \to \mathbb{R}$, where t is the coordinate in $\mathbb{R}$ on the left hand side.) To go from the Euler-Lagrange form of least action principle $\delta \int L(q,\dot{q})dt = 0$ to that of Hamilton one is led to make the definition

$$p_i(q,\dot{q}) \equiv \frac{\partial L}{\partial \dot{q}^i} \quad .$$

In many books on mechanics this is looked upon merely as a change of coordinates in TM, i.e., one switches from coordinates $q^1, \ldots, q^n$, $\dot{q}^1, \ldots, \dot{q}^n$ to coordinates $q^1, \ldots, q^n$, $p_1, \ldots, p_n$. This is not so ! The p_i's do not represent tangent vector components because they transform not like the v's in (3.1) but rather

$$p_i = \frac{\partial L}{\partial \dot{q}^i} = \sum_j \frac{\partial L}{\partial \overline{q}^j} \frac{\partial \overline{q}^j}{\partial \dot{q}^i} + \frac{\partial L}{\partial \dot{\overline{q}}^j} \frac{\partial \dot{\overline{q}}^j}{\partial \dot{q}^i} \quad .$$

Now $\overline{q}^j = \overline{q}^j(q^1, \ldots, q^n)$ alone, no $\dot{q}$, $\therefore \frac{\partial \overline{q}^j}{\partial \dot{q}^i} = 0$. Furthermore, again from (3.1)

$$\frac{\partial \dot{\overline{q}}^j}{\partial \dot{q}^i} = \frac{\partial \overline{q}^j}{\partial q^i} \quad .$$

Consequently

$$p_i = \sum_j \frac{\partial L}{\partial \dot{\overline{q}}^j} \frac{\partial \dot{\overline{q}}^j}{\partial \dot{q}^i} = \sum_j \overline{p}_j \frac{\partial \overline{q}^j}{\partial q^i}$$

or

$$\overline{p}_j = \sum_i \frac{\partial q^i}{\partial \overline{q}^j} p_i$$

and so the p's represent covectors on M^n, not vectors (see (6.2))! Thus the p's represent coordinates in the cotangent bundle to M^n and the formulas (where from now on we use p_i rather than a_i for coordinates in T*M)

$$\begin{cases} q^i = q^i \\ p_i = \dfrac{\partial L}{\partial \dot{q}^i} \end{cases} \tag{6.3}$$

really are defining a differentiable map

$$TM^n \longrightarrow T^*M^n \quad .$$

In mechanics the space with local coordinates (q,p) is called " phase space" ; we see that phase space is simply the cotangent bundle to the configuration space M^n .

While the $\dot{q}^i$ are called (generalized) velocities , the p_i are called generalized momenta . This is suggested by the following situation. The Lagrangian is frequently of the form

$$L(q,\dot{q}) = T(q,\dot{q}) - V(q) \qquad \begin{matrix} T = \text{kinetic energy} \\ V = \text{potential energy} \end{matrix}$$

where V is usually independent of $\dot{q}$ and T frequently is a positive definite quadratic form in the velocities

$$T(q,\dot{q}) = \tfrac{1}{2} \sum g_{ij}(q)\dot{q}^i\dot{q}^j$$

(for example, for two mass points moving on the line we have $T = \frac{1}{2} m_1(\dot{q}^1)^2 + \frac{1}{2} m_2(\dot{q}^2)^2$, where m_1 and m_2 are the masses). Then

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i} = \frac{\partial}{\partial \dot{q}^i} \frac{1}{2} \sum_{i,j} g_{ij}(q)\dot{q}^i\dot{q}^j$$

$$= \sum_j g_{ij}\dot{q}^j \quad .$$

Thus 2T defines a Riemannian metric on M^n and then the p_i are merely the covariant components (see 5.9b) of the "velocity", (again exhibiting p's as covariant quantities). In the special example of two mass points on the line

$$p_1 = m_1\dot{q}_1 \qquad p_2 = m_2\dot{q}_2$$

are indeed what everyone calls the momenta of the two particles.

The tangent and cotangent bundles TM^n and T^*M^n are distinct geometric objects. However, if M^n carries a Riemannian metric, we can define a diffeomorphism $TM^n \to T^*M^n$. Locally it maps the coordinate patch $(q,\dot{q}) \to$ the coordinate path (q,p) by

$$\begin{cases} q^i = q^i & \qquad q^i = q^i \\ & \text{with inverse} \\ p_i = \sum\limits_j g_{ij}(q)\dot{q}^j & \qquad \dot{q}^i = \sum\limits_j g^{ij}p_j \end{cases} .$$

If instead, we are given a Lagrangian function $L(q,\dot{q})$ on TM we can define a local map $TM \to T^*M$ by

$$\begin{cases} q^i = q^i \\ p_i = \dfrac{\partial L}{\partial \dot{q}^i} \end{cases} .$$

If this map is invertible we then have a diffeomorphism. Is there no real difference between TM and T^*M? To answer this we first need to digress on mappings.

(4) **Induced mappings on vectors and forms.** Let $F: M^n \to W^r$ be a differentiable mapping. Locally, if $x^1, \ldots, x^n$ are coordinates for M^n near p and $y^1, \ldots, y^r$ are coordinates for W^r near $F(p)$, then near p, F is described by

$$y^\beta = y^\beta(x^1, \ldots, x^n) = y^\beta(x^i) \qquad \begin{matrix} i = 1, 2, \ldots, n \\ \beta = 1, 2, \ldots, r \end{matrix} .$$

We emphasize that we needn't have $r = n$, and even when $r = n$ we don't ask that F be 1:1, etc.; we simply have a differentiable map between manifolds! Since M and W are locally Euclidean, we have an induced map, the differential of F, as in Lecture 3,

$$F_* : M^n_p \to W^r_{F(p)} .$$

Geometrically, it again takes tangents (i.e., velocity vectors to curves through p) onto the corresponding velocity vectors to the image curves through $F(p)$. It is a linear transformation with matrix the Jacobian

$$\boxed{F_*\left(\frac{\partial}{\partial x^i}\right) = \sum_{\beta=1}^{r} \left(\frac{\partial y^\beta}{\partial x^i}\right) \frac{\partial}{\partial y^\beta}} . \tag{6.4}$$

Thus F_* is just an expression of the chain rule.

~~ **WARNING.** F_* does not take vector fields on M^n into vector fields on W^r. This is because F may not be 1:1; it may be that $F(p) = F(p')$ and there is no reason to believe that if v is a vector field on M then $F_*(v_p) = F_*(v_{p'})$, (imagine trying to project a general vector field in $\mathbb{R}^3$ into $\mathbb{R}^2$).

Returning again to $F : M^n \to W^r$, from $F_* M^n_p \to W^r_{F(p)}$ we may consider the dual linear transformation on covectors, i.e., on 1-forms

$$\left.\begin{array}{rl} & F^* : W_y^{r*} \to M^{n*}_{F^{-1}(y)} \\ \text{by} & \\ & (F^*\beta)(v) = \beta(F_* v) \end{array}\right\} \tag{6.5}$$

i.e., we "pull-back" 1-forms at $y \in W$ to give 1-forms at each $F^{-1}(y)$ Thus if $\beta = \sum b_\alpha dy^\alpha$ is a 1-form at $y \in W$, then (from (5.1))

$$F^*\beta = \sum_{i=1}^{n} (F^*\beta)\left(\frac{\partial}{\partial x^i}\right) dx^i = \sum_{i=1}^{n} \beta\left(F_* \frac{\partial}{\partial x^i}\right) dx^i = \text{(from (6.4))}$$

$$= \sum_{i=1}^{n} \beta\left(\sum_{\alpha=1}^{r} \frac{\partial y^\alpha}{\partial x^i} \frac{\partial}{\partial y^\alpha}\right) dx^i = \sum_{i,\alpha} \frac{\partial y^\alpha}{\partial x^i} \beta\left(\frac{\partial}{\partial y^\alpha}\right) dx^i$$

$$= \sum_{i\alpha} \frac{\partial y^\alpha}{\partial x^i} b_\alpha \, dx^i$$

i.e.,

$$F^*\left(\sum_{\alpha=1}^{r} b_\alpha(y) dy^\alpha\right) = \sum_{i=1}^{n} \left(\sum_{\alpha} b_\alpha(y(x)) \frac{\partial y^\alpha}{\partial x^i}\right) dx^i \tag{6.6}$$

just expresses the chain rule again ! In particular the pull-back F^* gives

$$\boxed{F^*(dy^\alpha) = \sum_{i=1}^{n} \frac{\partial y^\alpha}{\partial x^i} \, dx^i} \tag{6.7}$$

In summary we have, for each $F : M \to W$ induced maps

$$F_* : M_x \to W_{f(x)} \qquad \text{(the "differential" of } F\text{)}$$

and

$$F^* : W_y^* \to M^*_{F^{-1}(y)} \qquad \text{(the "pull–back" of } F\text{)} \quad .$$

There is, however, an important difference between F_* and F^*. Note that if β is a covector field on W^r, then $F^*\beta$ is again a covector field on M, as is evident from (6.6); it matters not that we may have $F(p) = F(p')$: a form β at $y \in W$ pulls back to a 1-form at each point of the inverse image $F^{-1}(w)$. F^* pulls back covector fields on W to give covector fields on M! This will turn out to yield a very important advantage of forms over contravariant vector fields.

- **Problem 12.** Let $\underset{(x^1,\ldots,x^n)}{M^n} \xrightarrow{F} \underset{(y^1,\ldots,y^r)}{W^r} \xrightarrow{G} \underset{(z^1,\ldots,z^s)}{V^s}$ and so $G \circ F : M^n \to V^s$ Show (using chain rule) that for induced linear maps:

$$(G \circ F)_* = G_* \circ F_* : M^n_x \longrightarrow V^s_{G(F(x))}$$

$$(G \circ F)^* = F^* \circ G^* : V_z^* \longrightarrow M^*_{F^{-1}(G^{-1}v)}$$

(5) **The Poincaré 1-form on phase space** T^*M^n. There are natural maps ("projections") taking the tangent and the cotangent bundles back to the space M^n (by abuse of language we shall call both maps π)

$$\pi : TM \to M \quad \text{takes} \quad (q,\dot{q}) \to (q)$$
$$\pi : T^*M \to M \quad \text{takes} \quad (q,p) \to (q)$$

In words, the point $(q,\dot{q}) \in TM$ represents a tangent vector to M based at the point q. π takes this vector into its base point on M. Likewise for $\pi : T^*M \to M$.

Theorem.	There is a globally defined 1-form on every cotangent bundle T^*M, the Poincare 1-form λ. In local coordinates it is defined by $$\lambda = \sum_{i=1}^{n} p_i dq^i$$

(Note that such expressions make sense on T^*M, not M, because the q's and p's are coordinates on T^*M.)

Proof. We only need show that λ is well defined on an overlap of coordinates. But if $\bar{q}, \bar{p}$ are other local coordinates for T^*M we have

$$d\bar{q}^i = \sum_j \frac{\partial \bar{q}^i}{\partial q^j}\, dq^j + \frac{\partial \bar{q}^i}{\partial p_j}\, dp_j \quad .$$

But $\bar{q}^i$ are independent of p_j, so $\frac{\partial \bar{q}^i}{\partial p_j}\, dp_j = 0$ and we have

$$\sum_i \bar{p}_i\, d\bar{q}^i = \sum_i \bar{p}_i \sum_j \frac{\partial \bar{q}^i}{\partial q^j}\, dq^j = \sum_j p_j\, dq^j$$

as desired. □

There is a simple geometric explanation for the existence of the form λ. At the point $(q,p) \in T^*M$ the form λ is the following. (q,p) represents the 1-form $\sum p_i dq^i$ at the point $\pi(q,p) = q$. We may pull back this form to give a 1-form $\pi^*\left(\sum p_i dq^i\right)$ at (q,p). This is the form λ. (Note we are using q for coordinates in both M and in T^*M. In terms of coordinates (q,p) in T^*M and (q) in M the map $\pi : T^*M \to M$ is given simply by $q^i = q^i$ (independent of p). Thus for pull-back of the form $\sum p_i\, dq^i$ on M we have

$$\pi^*\left(\sum p_i\, dq^i\right) = \sum p_i \cdot \pi^* dq^i = \sum_{i,j} p_i \left(\frac{\partial q^i}{\partial q_j}\, dq^j + \frac{\partial q^i}{\partial p_j}\, dp_j\right)$$

$$= \sum_{i,j} p_i \delta_j^{\,i}\, dq^j = \sum p_i dq^i \quad ,$$

as claimed.

This is a great advantage for T^*M over TM. There is no naturally defined 1-form on the tangent bundle.

- **Problem 13.** Is $\sum \dot{q}^i \frac{\partial}{\partial q^i}$ a well-defined vector field on TM? Is $\sum \dot{q}^i \frac{\partial}{\partial \dot{q}^i}$? (Hint: $\frac{\partial}{\partial \bar{x}^i}\left(\frac{\partial \bar{x}^j}{\partial x^k}\right) = \sum_r \frac{\partial^2 \bar{x}^j}{\partial x^r \partial x^k} \frac{\partial x^r}{\partial \bar{x}^i}$.) If any of these are well defined, give the geometric (i.e., non-coordinate) meaning.

Lecture 7. Tensors

(1) **Linear algebra again.** Covectors are linear functionals on vectors (see Lecture 5, §2). Note that we may then think of vectors as linear functionals of covectors by the trivial definition $v(\alpha) \equiv \alpha(v)$ $(= \sum a_i v^i$ in coordinates). Although we shall define tensors over a general finite dimensional vector space E we shall usually be concerned with the case $E = M_p^n$, the tangent space to a manifold at a point p. Consequently, instead of using $e_1, \ldots, e_n$ for basis of E and $\sigma^1, \ldots, \sigma^n$ for dual basis of E^* we shall simply write $e_i = \dfrac{\partial}{\partial x^i}$ and $\sigma^i = dx^i$, but remember that our constructions are really just linear algebra.

Definition. A covariant tensor of rank r is a multilinear real valued function

$$Q : \underbrace{E \times \cdots \times E}_{r \text{ factors}} \longrightarrow \mathbb{R} \quad .$$

Thus Q is a function of r-tuples of vectors, $Q(v_{(1)}, \ldots, v_{(r)})$ that is linear in each $v_{(i)}$ when remaining entries are kept fixed (note: $v_{(i)}$ is a vector, not a component).

Example. The metric tensor $\mathcal{G}: E \times E \to \mathbb{R}$ defined by a scalar product (see Lecture 5, §3)

$$\mathcal{G}(v,w) = \langle v,w \rangle = \sum g_{ij} v^i w^j$$

is clearly linear in v and linear in w.

In components we have, from multilinearity

$$\begin{aligned} Q(v_{(1)}, \ldots, v_{(r)}) &= Q\left(\sum v_{(1)}^{i_1} \frac{\partial}{\partial x^{i_1}}, \sum v_{(2)}^{i_2} \frac{\partial}{\partial x^{i_2}}, \ldots, \sum v_{(r)}^{i_r} \frac{\partial}{\partial x^{i_r}} \right) \\ &= \sum_{i_1} v_{(1)}^{i_1} Q\left(\frac{\partial}{\partial x^{i_1}}, \sum v_{(2)}^{i_2} \frac{\partial}{\partial x^{i_2}}, \ldots, \sum v_{(r)}^{i_r} \frac{\partial}{\partial x^{i_r}} \right) \end{aligned}$$

$$= \sum_{i_1,i_2} v_{(1)}^{i_1} v_{(2)}^{i_2} Q\left(\frac{\partial}{\partial x^{i_1}} \frac{\partial}{\partial x^{i_2}}, \ldots, \sum v_{(r)}^{i_r} \frac{\partial}{\partial x^{i_r}}\right)$$

$$= \cdots = \sum_{i_1,i_2,\ldots,i_r} v_{(1)}^{i_1} v_{(2)}^{i_2} \cdots v_{(r)}^{i_r} Q\left(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_r}}\right)$$

i.e.,

$$Q = \sum_{i_1,\ldots,i_r} Q_{i_1 \cdots i_r} v_{(1)}^{i_1} v_{(2)}^{i_2} \cdots v_{(r)}^{i_r}, \quad Q_{i_1 \cdots i_r} \equiv Q\left(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_r}}\right) \tag{7.1}$$

The space of all covariant tensors of rank r, i.e., the space of all such Q's, forms a vector space (we know how to add functions and multiply them by numbers) which we write as

$$E^* \otimes E^* \otimes \cdots \otimes E^* \equiv \otimes^r E^* \quad .$$

It has dimension n^r (count the number of "components" $Q_{i_1 \cdots i_r}$).

If α and β are covectors, $\alpha \in E^*$, $\beta \in E^*$ we can form the 2nd rank covariant tensor $\alpha \otimes \beta$ defined as follows. We must tell how $\alpha \otimes \beta : E \times E \to \mathbb{R}$.

$$\alpha \otimes \beta\,(v,w) \equiv \alpha(v)\beta(w) \quad .$$

In components $\alpha = \sum a_i dx^i$, $\beta = \sum b_i dx^i$, and from (7.1)

$$(\alpha \otimes \beta)_{ij} = \alpha \otimes \beta \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) = \alpha\left(\frac{\partial}{\partial x^i}\right) \beta\left(\frac{\partial}{\partial x^j}\right) = a_i b_j \quad .$$

$\{a_i b_j\}$, $i,j = 1, \ldots, n$, form the components of the tensor product of the covectors α and β (the α and β are of course covariant tensors of rank 1; scalars will be considered tensors of rank 0).

- **Problem 14.** Show that the 2nd rank covariant tensor $\sum_{i,j} a_i b_j\, dx^i \otimes dx^j$ has the same values as $\alpha \otimes \beta$ on any pair of vectors, and so

$$\alpha \otimes \beta = \sum_{i,j} a_i b_j\, dx^i \otimes dx^j \quad .$$

Now consider contravariant tensors.

Definition. A contravariant tensor of rank s is an s-linear (i.e., multilinear) function of covectors

$$T : \underbrace{E^* \times \cdots \times E^*}_{\text{s factors}} \to \mathbb{R} \quad .$$

As in (7.1) we show immediately that

$$T(\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(s)}) = \sum_{j_1, \ldots, j_s} \alpha_{(1)j_1} \alpha_{(2)j_2} \cdots \alpha_{(s)j_s} T^{j_1 \cdots j_s} \tag{7.2}$$

$$T^{j_1 \cdots j_s} \equiv T(dx^{j_1}, \ldots, dx^{j_s}) \quad .$$

Again we shall write

$$T \in E \otimes \cdots \otimes E = \otimes^s E \quad .$$

Examples. The inverse of the metric tensor $\mathcal{G}^{-1}$ measures the scalar product of any pair of covectors

$$\mathcal{G}^{-1}(\alpha,\beta) = \sum g^{ij} a_i b_j \quad .$$

Note that for a pair of (contravariant) vectors $v \otimes w = \sum v^i \frac{\partial}{\partial x^i} \otimes \sum w^j \frac{\partial}{\partial x^j}$,

$$v \otimes w = \sum_{i,j} v^i w^j \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \qquad \text{(see Problem 14)} \quad .$$

You should verify that we are justified in writing

$$\mathcal{G} = \sum g_{ij}\, dx^i \otimes dx^j \qquad \text{and} \qquad \mathcal{G}^{-1} = \sum g^{ij} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j} \quad .$$

Finally we consider a third type of tensor, which actually generalizes those above.

Definition. A <u>mixed tensor</u>, r times covariant, s times contravariant is a multilinear function $W : \underbrace{E^* \times \cdots \times E^*}_{s} \times \underbrace{E \times \cdots \times E}_{r} \longrightarrow \mathbb{R}$

$$W(\alpha_{(1)}, \ldots, \alpha_{(s)}, v_{(1)}, \ldots, v_{(r)})$$

$$= \sum_{\substack{i_1 \ldots i_r \\ j_1 \ldots j_s}} \alpha_{(1)j_1} \cdots \alpha_{(s)j_s} W^{j_1 \ldots j_s}{}_{i_1 \ldots i_r} v_{(1)}^{i_1} \ldots v_{(r)}^{i_r}$$

$$W^{j_1 \ldots j_s}{}_{i_1 \ldots i_r} \equiv W\left(dx^{j_1}, \ldots, dx^{j_s}, \frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_r}}\right) .$$

Example. Let $A : E \to E$ be a <u>linear transformation</u>. In terms of a basis, $Ae_i = \sum_j e_j A^j{}_i$ and $A\left(\sum_i e_i v^i\right) = \sum_{i,j} e_j A^j{}_i v^i$, or briefly $(Av)^i = \sum_j A^i{}_j v^j$. Then A defines a mixed tensor $\tilde{A}$, once covariant once contravariant, by

$$\tilde{A}(\alpha, v) = \alpha(Av) = \sum_{i,j} a_j A^j{}_i v^i .$$

Conversely, such a mixed tensor $\tilde{B}$ defines a linear transformation B, by, Bv is that unique vector such that $\alpha(Bv) = \tilde{B}(\alpha, v)$, all $\alpha \in E^*$ There is no difference, really, between a linear transformation $A : E \to E$ and its associated mixed tensor $\tilde{A}$; we shall therefore drop the wiggle $\tilde{}$. Of course the identity transformation I is represented by the mixed tensor whose components form the Kronecker δ

$$\delta^i_j .$$

In conclusion, there are three types of matrices (A_{ij}), (A^{ij}), $(A_j{}^i)$. The first and second are the matrices of bilinear forms (on E or E^* respectively), while only the last represents a linear transformation from E to E.

If E has an inner product we may associate to any linear transformation $\mathcal{G} : E \to E$ a covariant bilinear form A as follows

$$A(v,w) \equiv \langle v, \mathcal{G}w \rangle ,$$

i.e.,

$$\sum_{i,j} A_{ij} v^i w^j \equiv \sum_{i,j} v^i g_{ij} \sum_k \mathcal{G}^j{}_k w^k$$

i.e.,

$$A_{ik} = \sum_k g_{ij}\, \mathcal{A}^j{}_k \tag{7.3}$$

i.e., A_{ij} is simply $\mathcal{A}^j{}_k$ with the index j lowered by means of the metric tensor. As in Lecture 5, §3, we would use the same letter (A or $\mathcal{A}$) for both of these tensors; the placement of the indices would explain which is which. In terms of an orthonormal (o.n) basis of E, $g_{ij} = \delta_{ij}$ and so $A_{ij} = \mathcal{A}^i{}_j$, etc. It is for this reason that the difference between a linear transformation and its associated bilinear form is not stressed in elementary linear algebra, but we shall always want to know which is which!

(2) **Transformation Properties.** A mixed tensor W has components (with respect to basis $\frac{\partial}{\partial x}$ of E and dx of E*) given by

$$W^{j_1 \cdots j_s}{}_{i_1 \cdots i_r} = W\left(dx^{j_1}, \ldots, dx^{j_s}, \frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_r}}\right) .$$

Under a change of basis $\frac{\partial}{\partial \bar{x}^i} = \sum_k \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial}{\partial x^k}$, $d\bar{x}^j = \sum_l \frac{\partial \bar{x}^j}{\partial x^l} dx^l$ we get by multilinearity

$$\begin{aligned} \bar{W}^{j_1 \cdots j_s}{}_{i_1 \cdots i_r} &= W\left(d\bar{x}^{j_1}, \ldots, d\bar{x}^{j_s}, \frac{\partial}{\partial \bar{x}^{i_1}}, \ldots, \frac{\partial}{\partial \bar{x}^{i_r}}\right) \\ &= \sum_{\substack{k_1 \cdots k_r \\ l_1 \cdots l_s}} \frac{\partial \bar{x}^{j_1}}{\partial x^{l_1}} \cdots \frac{\partial \bar{x}^{j_s}}{\partial x^{l_s}} W^{l_1 \cdots l_s}{}_{k_1 \cdots k_r} \frac{\partial x^{k_1}}{\partial \bar{x}^{i_1}} \cdots \frac{\partial x^{k_r}}{\partial \bar{x}^{i_r}} . \end{aligned} \tag{7.4}$$

(3) **Einstein Summation Convention.** A repeated index, once contravariant (i.e., upper) and once covariant (lower) implies a summation over that index, $W^{\cdots j \cdots}{}_{\cdots j \cdots} \equiv \sum_{j=1}^{n} W^{\cdots j \cdots}{}_{\cdots j \cdots}$. For example $A^i{}_i = \mathrm{tr}(A)$ is simply the trace of the linear transformation A.

(4) **Tensor Fields on Manifolds** of course have components varying on the manifold. A Riemannian metric $\langle v,w \rangle = g_{ij} v^i w^j$ is the most important 2nd rank covariant tensor. Let A be a covariant vector field on M^n (a "vector potential") $A = \sum A_i(x) dx^i$. Introduce the usual notation

$$\partial_i(f) \equiv \frac{\partial f}{\partial x^i} .$$

It will be clear from the calculation below that the matrix $\partial_j A_i = \dfrac{\partial A_i}{\partial x^j}$ does not constitute the matrix of a tensor because of its transformation properties. However, one can form the "field strength" matrix

$$F_{ij} = \partial_i A_j - \partial_j A_i \quad .$$

Proposition. $F_{ij} = \partial A_j - \partial_j A_i$ are the components of a covariant 2nd rank tensor.

Proof. We merely need to check the transformation law. Since A is a covector $\bar{A}_j = \dfrac{\partial x^l}{\partial \bar{x}^j} A_l$, and so

$$\bar{F}_{ij} = \frac{\partial \bar{A}_j}{\partial \bar{x}^i} - \frac{\partial \bar{A}_i}{\partial \bar{x}^j} = \frac{\partial}{\partial \bar{x}^i}\left(\frac{\partial x^l}{\partial \bar{x}^j} A_l\right) - \frac{\partial}{\partial \bar{x}^j}\left(\frac{\partial x^l}{\partial \bar{x}^i} A_l\right)$$

$$= \frac{\partial x^l}{\partial \bar{x}^j}\frac{\partial A_l}{\partial \bar{x}^i} + \underbrace{\frac{\partial^2 x^l}{\partial \bar{x}^i \partial \bar{x}^j} A_l} - \frac{\partial x^l}{\partial \bar{x}^i}\frac{\partial A_l}{\partial \bar{x}^j} - \frac{\partial^2 x^l}{\partial \bar{x}^j \partial \bar{x}^i} A_l$$

$$= \frac{\partial x^l}{\partial \bar{x}^j}\frac{\partial A_l}{\partial x^r}\frac{\partial x^r}{\partial \bar{x}^i} - \frac{\partial x^l}{\partial \bar{x}^i}\frac{\partial A_l}{\partial x^r}\frac{\partial x^r}{\partial \bar{x}^j}$$

$$= \frac{\partial x^l}{\partial \bar{x}^i}\frac{\partial x^r}{\partial \bar{x}^j}(\partial_l A_r - \partial_r A_l)$$

$$= \frac{\partial x^l}{\partial \bar{x}^i}\frac{\partial x^r}{\partial \bar{x}^j} F_{lr} \quad . \quad \square$$

{ remember l,r are dummy summation indices.

Note that the underlined term is what prevents $(\partial_i A_j)$ from being a tensor.

- **Problem 15.** Suppose $v = v^i \dfrac{\partial}{\partial x^i}$ is a contravariant vector. Show by the transformation properties that $v_j = g_{ji} v^i$ yields a covariant vector. Does $\dfrac{\partial v^i}{\partial x^j}$ yield a tensor? Does $\dfrac{\partial v^i}{\partial x^j} - \dfrac{\partial v^j}{\partial x^i}$ yield a tensor?

(5) **Pull-backs, and a digression into elasticity.** In Lecture 6, §4, we defined the pull back of a covariant vector under a mapping. We may define, by the same

philosophy, the pull back of <u>any covariant tensor</u> ! Let $F : M^n \to V^r$ and let α be a covariant tensor of rank s at $y \in V^r$. Then $F^*\alpha$ is the rank s covariant tensor at each point of $F^{-1}(y)$ defined as follows. If $v_1, \ldots, v_s$ are tangent vectors at $x \in F^{-1}(y)$, put

$$(F^*\alpha)(v_1, \ldots, v_s) \equiv \alpha(F_* v_1, \ldots, F_* v_s) \quad . \tag{7.5}$$

In terms of coordinates $x^1, \ldots, x^n$ on M and $y^1, \ldots, y^r$ on V^r

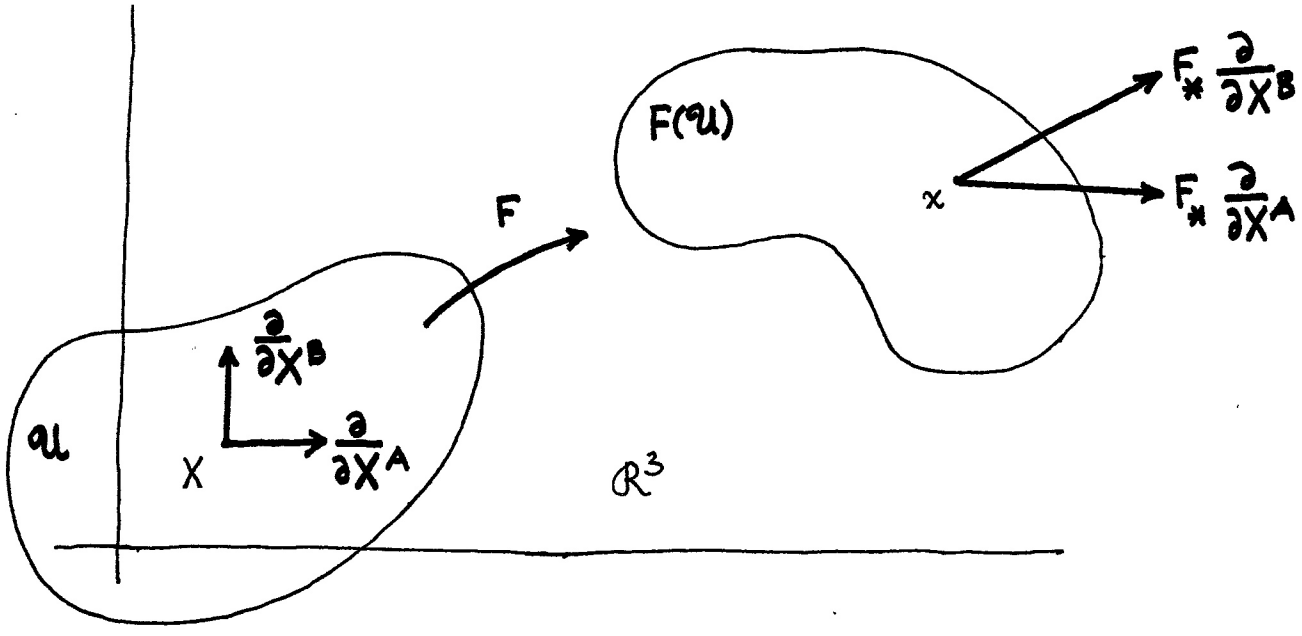
$$\begin{aligned}(F^*\alpha)_{i_1 \cdots i_s} &= (F^*\alpha)\left(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_s}}\right) = \alpha\left(F_* \frac{\partial}{\partial x^{i_1}}, \ldots, F_* \frac{\partial}{\partial x^{i_s}}\right) \\ &= \alpha\left(\frac{\partial y^{j_1}}{\partial x^{i_1}} \frac{\partial}{\partial y^{j_1}}, \ldots, \frac{\partial y^{j_s}}{\partial x^{i_s}} \frac{\partial}{\partial y^{j_s}}\right) \\ &= \frac{\partial y^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{j_s}}{\partial x^{i_s}} \alpha\left(\frac{\partial}{\partial y^{j_1}}, \ldots, \frac{\partial}{\partial y^{j_s}}\right) \quad ,\end{aligned}$$

i.e.,

$$(F^*\alpha)_{i_1 \cdots i_s}(x) = \frac{\partial y^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{j_s}}{\partial x^{i_s}} \alpha_{j_1 \cdots j_s} \quad . \tag{7.6}$$

<u>Note that when</u> $M = V$ <u>and</u> F <u>is the identity map</u> , <u>but expressed in different coor</u> - <u>dinates</u> x <u>and</u> $y = \bar{x}$, <u>then this pull back formula just expresses the transformation property of covariant tensors</u> !

Consider an elastic body $\mathcal{U} \subset \mathbb{R}^3$ and a deformation $F : \mathcal{U} \to \mathbb{R}^3$ of this body. We write F in the form $x^a = x^a(X^1, X^2, X^3)$, or $x^a = x^a(X^A)$; x^a are the new Cartesian coordinates of the point that originally had Cartesian coordinates X^a.



Under the deformation an orthonormal pair $\frac{\partial}{\partial X^A}$, $\frac{\partial}{\partial X^B}$ gets sent into a (usually) non-orthonormal pair $F_* \frac{\partial}{\partial X_A}$, $F* \frac{\partial}{\partial X^B}$. The metric tensor of $\mathbb{R}^3$ (like all sym - metric tensors) is written usually without the $\otimes$ sign .

$$dS^2 = \sum_{A,B} G_{AB}(X) dX^A \otimes dX^B = \sum_{A,B} \delta_{AB}\, dX^A dX^B = \sum_A (dX^A)^2$$

in coordinates X for $\mathcal{U}$. Likewise the same tensor at points x of $F(\mathcal{U})$ is

$$ds^2 = \sum_{ab} g_{ab}(x) dx^a \otimes dx^b = \sum_{a,b} \delta_{ab} dx^a dx^b = \sum_a (dx^a)^2 \quad .$$

For pull back under F we have

$$F^*(ds^2)\left(\frac{\partial}{\partial X^A}, \frac{\partial}{\partial X^B}\right)$$

$$= ds^2\left(F_* \frac{\partial}{\partial X^A}, \frac{\partial}{\partial X^B}\right) = ds^2\left(\frac{\partial x^a}{\partial X^A}\frac{\partial}{\partial x^a}, \frac{\partial x^b}{\partial X^B}\frac{\partial}{\partial x^b}\right)$$

$$= \frac{\partial x^a}{\partial X^A}\frac{\partial x^b}{\partial X^B} ds^2\left(\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right) = \frac{\partial x^a}{\partial X^A}\frac{\partial x^b}{\partial X^B} g_{ab} = \sum_a \frac{\partial x^a}{\partial X^A}\frac{\partial x^a}{\partial X^B}$$

This tensor

$$F^*(ds^2) = \sum_{\substack{a \\ A,B}} \frac{\partial x^a}{\partial X^A}\frac{\partial x^a}{\partial X^B} dX^A\, dX^B$$

when applied to $\frac{\partial}{\partial X^A}$, $\frac{\partial}{\partial X^B}$ reads off the scalar product of the deformed pair $F_* \frac{\partial}{\partial X^A}$, $F_* \frac{\partial}{\partial X^B}$ and is called, in elasticity the "right Cauchy - Green tensor"

$$C \equiv F^*(ds^2) \quad .$$

One measure of the deformation taking place, is given by the Lagrange deformation tensor

$$\frac{1}{2}\left(F^*(ds^2) - dS^2\right) = \frac{1}{2}\sum_{\substack{a \\ A,B}} \left(\frac{\partial x^a}{\partial X^A}\frac{\partial x^a}{\partial X^B} - \delta_{AB}\right) dX^A dX^B \quad .$$

When this deformation tensor vanishes at a point, the mapping is an infinitesimal isometry , i.e., scalar products are preserved.

Lecture 8. Exterior or Grassmann Algebra

Before the middle of the last century Grassmann introduced a new algebra whose product operation gave a vast generalization of scalar and vector products applicable in spaces of any dimension.

We shall primarily be concerned with taking products of covariant tensors.

(1) **The Tensor Product.** In the last lecture we defined the space $\otimes^p E^*$ of covariant p-tensors over the vector space E ($\dim E = n$); these are simply p-linear functions (real valued for the present) defined on p-tuples of vectors taken from E. If $\alpha^p \in \otimes^p E^*$ and $\beta^q \in \otimes^q E^*$ we then define their tensor product $\alpha^p \otimes \beta^q \in \otimes^{p+q} E^*$ by

$$\alpha^p \otimes \beta^q(v_1, \ldots, v_{p+q}) \equiv \alpha^p(v_1, \ldots, v_p) \cdot \beta^q(v_{p+1}, \ldots, v_{p+q}) \tag{8.1}$$

(we already talked about this product in the last lecture for the special case $p = 1 = q$). Grassmann discovered that there is much more geometrical significance attached to tensors that are skew-symmetric. We shall go directly to these.

(2) **The Grassmann or Exterior Algebra.** Let

$$\Lambda^p E^* \equiv E^* \wedge E^* \wedge \cdots \wedge E^* \subset \otimes^p E^*$$

be the vector space of <u>alternating</u>

(= skew symmetric = antisymmetric)

covariant p–tensors , i.e., tensors such that

$$\alpha(v_1, \ldots, v_r, \ldots, v_s, \ldots, v_p) = -\alpha(v_1, \ldots, v_s, \ldots, v_r, \ldots, v_p) .$$

The elements of $\Lambda^p E^*$ will be called <u>exterior</u> p- <u>forms</u> or simply p-forms. Note that if $\alpha^p \in \Lambda^p E^*$ and $\beta^q \in \Lambda^q E^*$, then $\alpha^p \otimes \beta^q \in \otimes^{p+q} E^*$ but $\alpha^p \otimes \beta^q$ will not be completely skew symmetric. Our goal will be to define a new product $\alpha^p \wedge \beta^q$ such that the result <u>will</u> be a $p+q$ form. This is easy to do in the special case $p = 1 = q$

$$\alpha^1 \wedge \beta^1 \equiv \alpha^1 \otimes \beta^1 - \beta^1 \otimes \alpha^1$$

i.e.,

$$\alpha^1 \wedge \beta^1(v_1,v_2) = \alpha^1(v_1)\beta^1(v_2) - \beta^1(v_1)\alpha^1(v_2) \quad .$$

Before doing this in general let us look more closely at p-forms.

By definition $\Lambda^1 E^* = E^*$ is just the space of 1-forms. It is convenient to introduce $\Lambda^0 E^* \equiv \mathbb{R}$, i.e., 0- forms are just elements of the coefficient field $\mathbb{R}$. A 0-form field α^0 on a manifold is simply a real valued function.

Let $\alpha^p \in \Lambda^p E^*$. Suppose $e_1, \ldots, e_n$ is a basis for E. Then, by multilinearity

$$\begin{aligned}\alpha^p(v_1, \ldots, v_p) &= \alpha^p\left(\sum_{i_1} e_{i_1} v_1^{\,i_1}, \ldots, \sum_{i_p} e_{i_p} v_p^{\,i_p}\right) \\ &= \sum_{i_1} v_1^{\,i_1} \alpha^p\left(e_{i_1}, \ldots, \sum_{i_p} e_{i_p} v_p^{\,i_p}\right) = \cdots \\ &= \sum_{i_1, \ldots, i_p} v_1^{\,i_1} v_2^{\,i_2} \cdots v_p^{\,i_p}\, \alpha(e_{i_1}, \ldots, e_{i_p}) \quad . \end{aligned} \tag{8.2}$$

The numbers $\alpha_{i_1, \ldots, i_p} \equiv \alpha(e_{i_1}, \ldots, e_{i_p})$ are the components of α. By skew symmetry $\alpha_{i_1 \cdots i_r \cdots i_s \cdots i_p} = -\alpha_{i_1 \cdots i_s \cdots i_r \cdots i_p}$. Since α is alternating we see that α is completely determined by the values $\alpha(e_{i_1}, \ldots, e_{i_p})$ where $i_1 < i_2 < \cdots < i_p$. The number of distinct $e_{i_1},\ldots,e_{i_p}$ with $i_1 < \cdots < i_p$ is $\binom{n}{p}$; we see that

$$\dim \Lambda^p E^* = \binom{n}{p}$$

and in particular $\dim \Lambda^n E^* = 1$; any n form on E is determined by its value on $(e_1, \ldots, e_n)$. If $p > n$, $\dim \Lambda^p E^* = 0$, since a repeated e_i will give $\alpha(e_{i_1}, \ldots, e_i, \ldots, e_i, \ldots, e_{i_p}) = 0$ by skew symmetry.

We now wish to define a product of exterior forms. First define the generalized Kronecker symbol

$$\delta_{i_1 \cdots i_r}^{j_1 \cdots j_r} = \begin{cases} 1 \text{ if } (j_1, \ldots, j_r) \text{ is even permutation of } (i_1 \cdots i_r) \\ -1 \text{ if } (j_1, \ldots, j_r) \text{ is odd permutation of } (i_1 \cdots i_r) \\ 0 \text{ if } (j_1, \ldots, j_r) \text{ is not a permutation of } (i_1 \cdots i_r) \end{cases} ,$$

(e.g., $\delta^{126}_{621} = -1$, $\delta^{126}_{623} = 0$, $\delta^{126}_{612} = 1$, etc.). This δ is the familiar <u>permutation symbol</u> that occurs in the definition of determinant

$$\det(A^j_i) = \sum_{j_1, \ldots, j_n} \delta^{j_1 \cdots j_n}_{1, \ldots, n} A^1{}_{j_1} A^2{}_{j_2} \cdots A^n{}_{j_n}$$

$$= \sum_{i_1 \cdots i_n} \delta^{1, \ldots, n}_{i_1 \cdots i_n} A^{i_1}{}_1 A^{i_2}{}_2 \cdots A^{i_n}{}_n \quad .$$

We further define

$$\epsilon_{i_1 \cdots i_n} = \epsilon^{i_1 \cdots i_n} = \delta^{i_1 \cdots i_n}_{1 \cdots n} \quad .$$

We now define the <u>exterior</u> or <u>Grassmann product</u>

$$\wedge : \Lambda^p E^* \times \Lambda^q E^* \longrightarrow \Lambda^{p+q} E^* \quad .$$

Let $\alpha^p \in \Lambda^p$, $\beta^q \in \Lambda^q$. Then we define $\alpha^p \wedge \beta^q \in \Lambda^{p+q}$ by

$$\alpha^p \wedge \beta^q(v_{i_1}, \ldots, v_{i_{p+q}}) \tag{8.3}$$

$$= \sum_{\substack{k_1 < \cdots < k_p \\ l_1 < \cdots < l_q}} \delta^{k_1 \cdots k_p l_1 \cdots l_q}_{i_1 \cdots \cdots \cdots i_{p+q}} \alpha(v_{k_1}, \ldots, v_{k_p}) \beta(v_{l_1}, \ldots, v_{l_q}) \quad .$$

For example if $\dim E = 5$ and $e_1, \ldots, e_5$ is basis for E, then

$$\alpha^2 \wedge \beta^1(e_5, e_2, e_3) = \sum_{\substack{k_1 < k_2 \\ l}} \delta^{k_1 k_2 l}_{523} \alpha(e_{k_1}, e_{k_2}) \beta(e_l)$$

$$= \delta^{235}_{523} \alpha(e_2, e_3)\beta(e_5) + \delta^{253}_{523} \alpha(e_2, e_5)\beta(e_3) + \delta^{352}_{523} \alpha(e_3, e_5)\beta(e_2)$$

$$= \alpha(e_2, e_3)\beta(e_5) - \alpha(e_2, e_5)\beta(e_3) + \alpha(e_3, e_5)\beta(e_2) \quad .$$

One should easily check that $\alpha \wedge \beta$ is indeed (p+q)-linear. Also since

$$\delta^{k_1 \cdots \cdots \cdots \cdots l_q}_{i_1 \cdots i_r \cdots i_s \cdots i_{p+q}} = -\delta^{k_1 \cdots \cdots \cdots \cdots l_q}_{i_1 \cdots i_s \cdots i_r \cdots i_{p+q}}$$

we see that $\alpha \wedge \beta$ is again alternating. However $\alpha^p \wedge \beta^q \neq \beta^q \wedge \alpha^p$. In fact

$$\beta^q \wedge \alpha^p(v_{i_1}, \ldots, v_{i_{p+q}}) = \sum_{\substack{l_1 < \cdots < l_q \\ k_1 < \cdots < k_p}} \delta_{i_1 \cdots i_s \cdots i_r \cdots i_{p+q}}^{l_1 \cdots l_q k_1 \cdots k_p} \beta(v_{l_1} \cdots v_{l_q}) \alpha(v_{k_1} \cdots v_{k_p})$$

$$= (-1)^{pq} \alpha^p \wedge \beta^q(v_{i_1} \cdots v_{i_{p+q}})$$

since $(l_1, \ldots, l_q, k_1 \cdots k_p) \to (k_1 \cdots k_p, l_1 \cdots l_q)$ requires pq transpositions

$$\boxed{\alpha^p \wedge \beta^q = (-1)^{pq} \beta^q \wedge \alpha^p} \tag{8.4}$$

i.e., the Grassmann product is said to be "anti-commutative." In particular for a form of odd degree

$$\alpha^{2p+1} \wedge \alpha^{2p+1} = 0$$

Let us consider the vector space of all forms, all degrees, over E^*

$$\Lambda^* E \equiv (\Lambda^0 E^* = \mathbb{R}) \oplus (\Lambda^1 E^* = E^*) \oplus (\Lambda^2 E^*) \oplus \cdots \oplus (\Lambda^n E^*) \quad .$$

This is the "Grassmann" or "Exterior algebra" over E^*

$$\dim \Lambda^* E = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n} = 2^n \quad .$$

We now introduce a simplifying notation for multi-indices. We shall write $I = (i_1, \ldots, i_p)$, $J = (j_1, \ldots, j_q)$, etc. the number of indices will be clear in each context. We also will frequently want to order the indices

$$\underset{\rightarrow}{I} = (i_1, \ldots, i_p) \quad \text{means} \quad i_1 < i_2 < , \ldots , < i_p \quad .$$

Instead of $(v_{i_1}, \ldots, v_{i_p})$ we shall also write v_I. Then our formula (8.3) becomes

$$\alpha^p \wedge \beta^q(v_I) = \sum_{\underset{\rightarrow}{K,L}} \delta_I^{KL}\, \alpha(v_K)\beta(v_L)$$

or , in components

$$(\alpha \wedge \beta)_I = \sum_{\underset{\rightarrow}{K,L}} \delta_I^{KL}\, \alpha_K\, \beta_L \tag{8.5}$$

where, as usual $\alpha_K = \alpha_{k_1 \cdots k_p} = \alpha(e_{k_1}, \ldots, e_{k_p})$.

The exterior product is easily seen to be distributive. We now show associativity. First a lemma (if you think this is bad, write it out without multi-indices)

Lemma. If $I = (i_1 \cdots i_p)$, $J = (j_1 \cdots j_{q+r})$, $K = (k_1 \cdots k_q)$, $L = (l_1 \cdots l_r)$

$$\sum_{\underset{\rightarrow}{J}} \delta_{1 \cdots (p+q+r)}^{IJ}\, \delta_J^{KL} = \delta_{1 \cdots (p+q+r)}^{IKL} \tag{8.6}$$

Proof. I, K, L are all fixed. Since $j_1 < \cdots < j_{q+r}$ there is at most one non-vanishing term on the left hand side, namely when J is some permutation of KL. Proof is then immediate.

Now we prove associativity. If $M = (m_1, m_2, \ldots, m_{p+q+r})$, then the above lemma says really

$$\sum_{\underset{\rightarrow}{J}} \delta_M^{IJ}\, \delta_J^{KL} = \delta_M^{IKL} \quad .$$

Now look at the component $[\alpha \wedge (\beta \wedge \gamma)]_{m_1 \cdots m_{p+q+r}} = [\alpha \wedge (\beta \wedge \gamma)]_M$. From (8.5)

$$\begin{aligned}
[\alpha \wedge (\beta \wedge \gamma)]_M &= \sum_{\underset{\rightarrow}{I},\underset{\rightarrow}{J}} \delta_M^{IJ}\, \alpha_I (\beta \wedge \gamma)_J \\
&= \sum_{\underset{\rightarrow}{I},\underset{\rightarrow}{J}} \delta_M^{IJ}\, \alpha_I \sum_{\underset{\rightarrow}{K,L}} \delta_J^{KL}\, \beta_K\, \gamma_L \\
&= \sum_{\underset{\rightarrow}{I},\underset{\rightarrow}{K},\underset{\rightarrow}{L}} \delta_M^{IKL}\, \alpha_I \beta_K\, \gamma_L \quad .
\end{aligned} \tag{8.7}$$

One can do a similar computation for products of any number of forms. It is clear that we shall get the same result for $[(\alpha \wedge \beta) \wedge \gamma]_M$, which proves associativity. Note that the same computation gives

$$(\alpha \wedge \beta \wedge \gamma)(v_M) = \sum_{\underset{\rightarrow}{I,K,L}} \delta_M^{IKL}\, \alpha(v_I)\beta(v_K)\gamma(v_L)$$

If we had taken the exterior product of r 1-forms $\alpha_{(1)}, \alpha_{(2)}, \ldots, \alpha_{(r)}$

$$\left.\begin{aligned} \alpha_{(1)} \wedge \alpha_{(2)} \wedge \ldots \wedge \alpha_{(r)}(v_1,\ldots,v_r) &= \sum_{i_1,i_2,\ldots,i_r} \delta_{1,\ldots,r}^{i_1\ldots i_r}\, \alpha_{(1)}(v_{i_1})\ldots\alpha_{(r)}(v_{i_r}) \\ &= \det\,[\alpha_{(i)}(v_j)] \end{aligned}\right\} \tag{8.8}$$

In particular, for the dual basis $\sigma^1, \ldots, \sigma^n$ to $e_1, \ldots, e_n$, we have

$$\sigma^{i_1} \wedge \sigma^{i_2} \wedge \cdots \wedge \sigma^{i_r}(e_{j_1}, \ldots, e_{j_r}) = \delta_{j_1 \cdots j_r}^{i_1 \cdots i_r} \tag{8.9}$$

since this is certainly true if $i_1 < \cdots < i_r$ and $j_1 < \cdots < j_r$, from (8.8).

- **Problem 16.** Show that if α^p is any p-form, we have the expansion

$$\alpha^p = \sum_{\underset{\rightarrow}{I}} \alpha(e_I)\sigma^I = \sum_{i_1 < \cdots < i_p} \alpha(e_{i_1}, \ldots, e_{i_p})\sigma^{i_1} \wedge \cdots \wedge \sigma^{i_p} \quad .$$

(Hint: Check values of both sides on $(e_{j_1}, \ldots, e_{j_p})$.)

We shall then usually write

$$\boxed{\begin{aligned} &\alpha^p = \sum_{\underset{\rightarrow}{I}} \alpha_I \sigma^I = \sum_{i_1 < \cdots < i_p} \alpha_{i_1 \cdots i_p}\, \sigma^{i_1} \wedge \cdots \wedge \sigma^{i_p} \\ &\alpha_{i_1 \cdots i_p} \equiv \alpha(e_{i_1}, \ldots, e_{i_p}) \text{ is skew in I} \end{aligned}} \tag{8.10}$$

and of course $\alpha_{i_1 \cdots i_p}$ are the "components" of α with respect to basis $\sigma^1, \ldots, \sigma^n$.

Thus the most general 2-form in $\mathbb{R}^3$ is of the form

$$\beta^2 = \sum_{i<j} \beta_{ij} dx^i \wedge dx^j = \beta_{12}\, dx^1 \wedge dx^2 + \beta_{13}\, dx^1 \wedge dx^3 + \beta_{23}\, dx^2 \wedge dx^3$$
$$= \beta_{23}\, dx^2 \wedge dx^3 + \beta_{31}\, dx^3 \wedge dx^1 + \beta_{12}\, dx^1 \wedge dx^2$$

(in this last expression we are using $dx^1 \wedge dx^3 = -dx^3 \wedge dx^1$ and $\beta_{13} = -\beta_{31}$). We shall see in a moment why we prefer this last form.

- **Problem 17.** Show that in $\mathbb{R}^3$ $(\alpha^1 \wedge \beta^2)_{ijk} = \alpha_i\beta_{jk} + \alpha_k\beta_{ij} + \alpha_j\beta_{ki}$, i.e., one writes down $\alpha_i\beta_{jk}$ and then adds all other terms after cyclically permuting i, j, k. Investigate $\alpha^1 \wedge \beta^{n-1}$ in $\mathbb{R}^n$.

(3) **Geometric meaning of forms.** Let us look at the geometric meaning of exterior forms in $\mathbb{R}^n$ in the special case when coordinates $x^1, \ldots, x^n$ are Cartesian (i.e., we shall be using the Euclidean metric of $\mathbb{R}^n$).

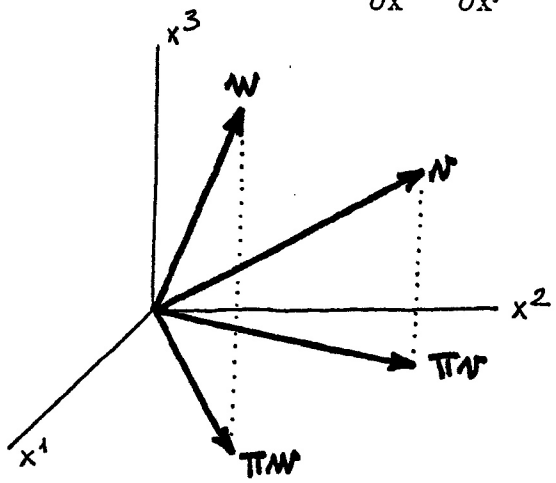
We have basis vectors for $E = \mathbb{R}^n$ given by $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$, and dual basis $dx^1, \ldots, dx^n$ for E^*. We already know for the 1-forms that

$$dx^i(v) = v^i \quad ,$$

i.e., dx^i reads off the i^{th} component of a vector v. Next

$$dx^i \wedge dx^j(v,w) = dx^i(v)dx^j(w) - dx^i(w)dx^j(v)$$
$$= \begin{vmatrix} v^i & w^i \\ v^j & w^j \end{vmatrix}$$

$= \pm$ area of parallelogram spanned by the projections $\pi(v)$, $\pi(w)$ of v and w into the $x^i x^j$ plane, the + sign being used if and only if πv and πw determine the same orientation as does $\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}$.



In this figure

$dx^1 \wedge dx^2(v,w)$

reads off the negative of the area of the parallelogram spanned by πv and πw.

Likewise

$$dx^{i_1} \wedge \cdots dx^{i_p}(v_1, \ldots, v_p) = \det(v_r^j)$$

$= \pm$ the p dimensional volume of the parallelopiped spanned by the projections of $v_1, \ldots, v_p$ into the coordinate plane spanned by $\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_p}}$, the + sign being used only if these projected vectors define the same orientation of $\mathbb{R}^p$ as does $\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_p}}$

(4) **Special cases of the exterior product.** Now let $\tau^1, \ldots, \tau^n$ be any collection of 1 - forms and expand each in terms of the basis $\sigma^1, \ldots, \sigma^n$

$$\tau^i = \sum_{j=1}^{n} A^i_j \sigma^j \quad .$$

Then

$$\tau^1 \wedge \cdots \wedge \tau^n = \sum_J A^1_{j_1} A^2_{j_2} \cdots A^n_{j_n} \sigma^{j_1} \wedge \sigma^{j_2} \wedge \cdots \wedge \sigma^{j_n}$$

$$= \sum_J (A^1_{j_1} \cdots A^n_{j_n} \delta_1^{j_1 \cdots j_n}) \sigma^1 \wedge \cdots \wedge \sigma^n$$

$$\therefore \qquad \boxed{\tau^1 \wedge \tau^2 \wedge \cdots \wedge \tau^n = (\det A)\, \sigma^1 \wedge \sigma^2 \wedge \cdots \wedge \sigma^n \quad .} \tag{8.11}$$

Proposition. The 1-forms $\tau^1, \ldots, \tau^p$ are <u>linearly dependent</u> if and only if

$$\tau^1 \wedge \tau^2 \wedge \cdots \wedge \tau^p = 0 \quad .$$

Proof. If

$$\tau^r = \sum_{i \neq r} a_i \tau^i$$

then $\tau^1 \wedge \cdots \wedge \tau^r \wedge \cdots \wedge \tau^p$ will be a sum of terms each having a repeated τ. From (8.4) each of these terms is 0. On the other hand if $\tau^1, \ldots, \tau^p$ are independent then we can complete them to a basis $\tau^1, \ldots, \tau^p, \tau^{p+1}, \ldots, \tau^n$. Let $f_1, \ldots, f_n$ be the dual basis to these for E. From (17-8)

$$\tau^1 \wedge \cdots \wedge \tau^n(f_1, \ldots, f_n) = +1$$

and so $\tau^1 \wedge \cdots \wedge \tau^p \neq 0$. □

(5) **Computations and relation with vector analysis.** For computations using forms we may use the usual rules of algebra except the commutative law is replaced by (8.4). In particular $dx^i \wedge dx^j = -dx^j \wedge dx^i$ and $dx^i \wedge dx^i = 0$. Consider $\mathbb{R}^3$ as a manifold with <u>any</u> coordinate system x^1, x^2, x^3. Let $f(x)$ be a 0-form, i.e., function, let

$$\alpha^1 = a_1 dx^1 + a_2 dx^2 + a_3 dx^3 \quad \text{and} \quad \beta^1 = b_1 dx^1 + b_2 dx^2 + a_3 dx^3 \quad ,$$

$$\gamma^2 = c_{23} dx^2 \wedge dx^3 + c_{31} dx^3 \wedge dx^1 + c_{12} dx^1 \wedge dx^2$$

$$= c_1 dx^2 \wedge dx^3 + c_2 dx^3 \wedge dx^1 + c_3 dx^1 \wedge dx^2$$

$$\omega^3 = dx^1 \wedge dx^2 \wedge dx^3 \quad .$$

We then have, for example $\alpha^1 \wedge \beta^1$

$$= (a_1dx^1 + a_2dx^2 + a_3dx^3) \wedge (b_1dx^1 + b_2dx^2 + b_3dx^3) =$$

$$= a_1b_1dx^1 \wedge dx^1 + \cdots + a_2b_3dx^2 \wedge dx^3 + \cdots + a_3b_2dx^3 \wedge dx^2 + \cdots$$

$$= 0 + \cdots + (a_2b_3 - a_3b_2)dx^2 \wedge dx^3 + \cdots$$

$$= (a_2b_3 - a_3b_2)\ dx^2 \wedge dx^3 + (a_3b_1 - a_2b_3)dx^3 \wedge dx^1 + (a_1b_2 - a_2b_1)dx^1 \wedge dx^2 .$$

Note that if x^1, x^2, x^3 are Cartesian coordinates this is similar to a result in vector analysis, namely put

$$\vec{A} = \begin{bmatrix} a^1 \\ a^2 \\ a^3 \end{bmatrix}, \vec{B} = \begin{bmatrix} b^1 \\ b^2 \\ b^3 \end{bmatrix}, \text{ then } \vec{A} \times \vec{B} = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

but these last three components are the components of $\vec{A} \times \vec{B}$ only in Cartesian coordinates! We shall discuss the precise rlation of forms and vector analysis later, but we emphasize now that our expression for $\alpha^1 \wedge \beta^1$ holds in any coordinates.

- **Problem 18.** Compute $\alpha^1 \wedge \gamma^2$ and $\alpha^1 \wedge \beta^1 \wedge \eta^1$, where η^1 is a 1-form. Interpret in terms of vectors.

Lecture 9. Exterior Differentiation

(1) **The Exterior Differential** d. Let M^n be a differentiable manifold and let $\Lambda^p M^n$ be the vector space of all differentiable p-forms on M^n. In local coordinates $\alpha^p \in \Lambda^p M^n$ is of the form

$$\alpha^p = \sum_{\substack{I \\ \rightarrow}} A_I dx^I = \sum_{i_1 < \cdots < i_p} a_{i_1 \cdots i_p}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

where $a_I = \alpha\left(\frac{\partial}{\partial x^{i_1}}, \ldots, \frac{\partial}{\partial x^{i_p}}\right)$ are differentiable functions. Since these coefficients are functions $\Lambda^p M$ is an infinite dimensional vector space over $\mathbb{R}$.

By definition a 0-form α^0 is merely a <u>function</u> on M and anticommutivity (8.4) says in this case $\alpha^0 \wedge \beta^q = \beta^q \wedge \alpha^0$. When one of the forms in a product is a 0-form we usually omit the $\wedge$ sign.

Theorem. There is a unique operator (<u>exterior differentiation</u>)

$$d : \Lambda^p M \to \Lambda^{p+1} M$$

satisfying

(1) d is additive, $d(\alpha^p + \beta^p) = d\alpha^p + d\beta^p$

(2) $d\alpha^0 =$ the differential of the function α^0 (Lecture 5,§2).

(3) $d(\alpha^p \wedge \beta^q) = d\alpha^p \wedge \beta^q + (-1)^p \alpha^p \wedge d\beta^q$

(4) $d^2\alpha^p = d(d\alpha^p) \equiv 0$, all α

Proof. We shall define d first by using local coordinates, then we will show that the result is independent of coordinates used. Let $(x^1, \ldots, x^n)$ be local coordinates.

<u>Step I</u>. If f is a 0-form, define

$$\underset{x}{df} = d\,f = \sum_i \frac{\partial f}{\partial x^i} dx^i \quad .$$

This definition, we know, is independent of coordinates; its intrinsic definition is $df(v) = v(f)$. Condition (2) has been satisfied.

Step II . If $a(x)$ is a function, define

$$\underset{x}{d}\,[a(x)dx^{i_1} \wedge \cdots \wedge dx^{i_p}]$$

$$= da \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} = \sum_j \frac{\partial a}{\partial x^j}\, dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \quad .$$

Then we define $\underset{x}{d}$ for any p-form by linearity

$$\underset{x}{d} \sum_{i_1 < \cdots < i_p} a_{i_1 \cdots i_p}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

$$\equiv \sum_{i_1 < \cdots < i_p} da_{i_1 \cdots i_p} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \quad .$$

Condition (1) is automatically satisfied. Consider condition (3). Let $I = (i_1, \ldots, i_p)$, $J = (j_1, \ldots, j_q)$

$$\underset{x}{d} \left(\sum_{\substack{I \\ \rightarrow}} a_I dx^I \wedge \sum_{\substack{J \\ \rightarrow}} b_J dx^J \right)$$

$$= \underset{x}{d} \sum_{\substack{I \\ \rightarrow}} \sum_{\substack{J \\ \rightarrow}} a_I b_J\, dx^I \wedge dx^J = \sum_{\substack{I \\ \rightarrow}} \sum_{\substack{J \\ \rightarrow}} d(a_I b_J) \wedge dx^I \wedge dx^J$$

$$= \sum_{\substack{I \\ \rightarrow}} \sum_{\substack{J \\ \rightarrow}} (da_I b_J + a_I db_J) \wedge dx^I \wedge dx^J$$

$$= \sum_{\substack{I \\ \rightarrow}} \sum_{\substack{J \\ \rightarrow}} da_I \wedge d\,x^I \wedge b_J dx^J + \sum_{\substack{I \\ \rightarrow}} \sum_{\substack{J \\ \rightarrow}} a_I dx^I \wedge (-1)^p db_J \wedge dx^J$$

(since $db_J \wedge dx^I = (-1)^p dx^I \wedge db_J$ involves p interchanges)

$$= d\alpha^p \wedge \beta^q + (-1)^p \alpha^p \wedge d\beta^q \text{ , as desired} \quad .$$

To verify condition (4), note that if f is a 0-form, i.e., function

$$\underset{x}{d}\,(\underset{x}{d}f) = \underset{x}{d}\left(\sum_i \frac{\partial f}{\partial x^i}\,dx^i\right) = \sum_i \underset{x}{d}\left(\frac{\partial f}{\partial x^i}\right) \wedge dx^i = \sum_i \sum_j \frac{\partial^2 f}{\partial x^j \partial x^i}\,dx^j \wedge dx^i$$

$$= \cdots \frac{\partial^2 f}{\partial x^r \partial x^s}\,dx^r \wedge dx^s + \cdots + \frac{\partial^2 f}{\partial x^s \partial x^r}\,dx^s \wedge dx^r + \cdots = 0 \ .$$

Then from (3) for any functions f and g

$$\underset{x}{d}\,(df \wedge dg) = 0$$

and by induction

$$\underset{x}{d}\,(df_{(1)} \wedge df_{(2)} \wedge \cdots \wedge df_{(p)}) = 0 \qquad f_{(1)}, \ldots, f_{(p)} \quad \text{functions} \ . \tag{9.1}$$

Then for any p-form α

$$\underset{x}{d^2} \sum_{i_1<\ldots<i_p} a_{i_1\ldots i_p} dx^{i_1} \wedge \ldots \wedge dx^{i_p} = \underset{x}{d} \sum_{i_1<\ldots<i_p} da_{i_1\ldots i_p} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p}$$

$$= 0 \ .$$

We have now defined an operator $\underset{x}{d}$ in each coordinate system x, satisfying (1), (2), (3), (4). Let $y = y^1, \ldots, y^n$ be a coordinate system that overlaps x and let $\underset{y}{d}$ be its associated operator. Then

$$\underset{y}{d} \sum_{i_1<\cdots<i_p} a_{i_1\cdots i_p}(x) dx^{i_1} \wedge \cdots \wedge dx^{i_p} = \text{(using (9.1))}$$

$$= \sum_{i_1<\cdots<i_p} \underset{y}{d}\, a_{i_1\cdots i_p}(x(y)) \wedge dx^{i_1} \wedge \cdots dx^{i_p}$$

$$= \sum_{i_1<\cdots<i_p} d\, a_{i_1\cdots i_p} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

$$= \underset{x}{d} \sum_{i_1<\cdots<i_p} a_{i_1\cdots i_p}\, dx^{i_1} \wedge \cdots \wedge dx^{i_p}$$

showing that d is well defined, independent of coordinates used. As to uniqueness, if any operator d satisfies (1), (2), (3), (4), then

$$d \sum_{i_1<\ldots<i_p} a_{i_1\ldots i_p}\, dx^{i_1} \wedge \ldots \wedge dx^{i_p} = \sum_{i_1<\ldots<i_p} da_{i_1\ldots i_p} \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p}$$

$$= d_x \sum_{i_1<\ldots<i_p} a_{i_1<\ldots<i_p}\, dx^{i_1} \wedge \ldots \wedge dx^{i_p} \quad . \qquad \square$$

(2) **Examples** in $\mathbb{R}^3$. Let x, y, z be any coordinates for $\mathbb{R}^3$. Then

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz \quad .$$

If coordinates are Cartesian, coefficients are the same as the gradient vector of f. If, in general coordinates

$$\alpha^1 = a_1\, dx + a_2\, dy + a_3\, dz$$

then

$$d\alpha^1 = da_1 \wedge dx + da_2 \wedge dy + da_3 \wedge dz$$

$$= \left(\frac{\partial a_1}{\partial x}\, dx + \frac{\partial a_1}{\partial y}\, dy + \frac{\partial a_1}{\partial z}\, dz \right) \wedge dx$$

$$+ \left(\frac{\partial a_2}{\partial x}\, dx + \frac{\partial a_2}{\partial y}\, dy + \frac{\partial a_2}{\partial z}\, dz \right) \wedge dy + \ldots$$

$$= \left(\frac{\partial a_3}{\partial y} - \frac{\partial a_2}{\partial z} \right) dy \wedge dz + \left(\frac{\partial a_1}{\partial z} - \frac{\partial a_3}{\partial x} \right) dz \wedge dx$$

$$+ \left(\frac{\partial a_2}{\partial x} - \frac{\partial a_1}{\partial y} \right) dx \wedge dy$$

whose coefficients in Cartesian coordindates correspond to curl $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$. Finally, for a 2-form (putting $b_1 = \beta_{23}$, etc.) $\beta^2 = b_1 dy \wedge dz + b_2 dz \wedge dx + b_3 dx \wedge dy$

$$d\beta^2 = db_1 \wedge dy \wedge dz + db_2 \wedge dz \wedge dx + db_3 \wedge dx \wedge dy$$

$$= \left(\frac{\partial b_1}{\partial x} + \frac{\partial b_2}{\partial y} + \frac{\partial b_3}{\partial z} \right) dx \wedge dy \wedge dz$$

whose coefficient in Cartesian coordinates corresponds to div $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

In Cartesian coordinates $d^2f = 0$ just corresponds to curl grad $f = 0$, while $d^2(\alpha^1) = 0$ corresponds to div curl $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = 0$.

(3) **A concise coordinate expression for** d. Let $\alpha^p = \sum_{\underset{\rightarrow}{L}} \alpha_L dx^L$. Then $d\alpha^p = \sum_{\underset{\rightarrow}{L}} (d\alpha_L) \wedge dx^L$. Now $d\alpha_L$ is the 1-form $(d\alpha_L)_j = \partial_j \alpha_L$. Also dx^L is the p-form with $(dx^L)_K = \delta^L_K$. Hence from (8.5)

$$(d\alpha)_I = \sum_{\underset{\rightarrow}{L}} (d\alpha_L \wedge dx^L)_I = \sum_{\underset{\rightarrow}{L}} \left(\sum_{j,\underset{\rightarrow}{K}} \delta_I^{jK} (\partial_j \alpha_L) \delta^L_K \right) ,$$

i.e.,

$$\boxed{(d\alpha)_I = \sum_{j,\underset{\rightarrow}{K}} \delta_I^{jK} \partial_j \alpha_K} \qquad (9.2)$$

For example, if $\alpha = \sum_{i<j} \alpha_{ij} dx^i \wedge dx^j$ is a 2-form in $\mathbb{R}^n$

$$d\alpha = \sum_{i<j<k} (\partial_i \alpha_{jk} + \partial_k \alpha_{ij} + \partial_j \alpha_{ki}) dx^i \wedge dx^j \wedge dx^k \quad . \qquad (9.3)$$

Lecture 10. Pull-Backs

In Lecture 7 we introduced the notion of pull-back of covariant tensors under a map $F : M^n \to V^r$. Since exterior differential forms are skew symmetric covariant tensors we may apply pull-back F^* to forms. As we shall see the pull-back of a form is again a form. We shall see that this pull-back operation is remarkable in that it commutes both with exterior products and exterior differentiation

$$F^*(\alpha \wedge \beta) = (F^*\alpha) \wedge (F^*\beta) \qquad F^*(d\alpha) = d(F^*\alpha) \quad .$$

There are no analogues of these results in vector analysis and that is one reason why exterior calculus is so much more powerful than vector analysis !.

If $F : M^n \to V^r$ and if $f : V^r \to \mathbb{R}$ is a function on V then the composition $f \circ F : M^n \to \mathbb{R}$, $(f \circ F)(x) = f(F(x))$ defines the "pull-back" of the 0-form f

$$F^*(f) \equiv f \circ F$$

$$M^n \xrightarrow{F} V^r \xrightarrow{f} \mathbb{R}$$

$$f \circ F$$

Lemma. If v is a vector at $x \in M$, then

$$\boxed{v(f \circ F) = df(F_* v) \quad .} \tag{10.1}$$

Proof. Let x^i be coordinates near x on M and let y^α be coordinates near $f(x)$ on V. Then by the chain rule, since F is given by $y = F(x)$,

$$v(f \circ F) = \sum v^i \frac{\partial}{\partial x^i} (f \circ F) = \sum_i v^i \frac{\partial}{\partial x^i} (f(y(x))$$

$$= \sum_i v^i \sum_\alpha \frac{\partial f}{\partial y^\alpha} \frac{\partial y^\alpha}{\partial x^i} = \sum_\alpha \left(\sum_i v^i \frac{\partial y^\alpha}{\partial x^i} \right) \frac{\partial f}{\partial y^\alpha}$$

$$\text{(see Lec.1,§3)} \quad = \sum_\alpha (F_* v)^\alpha \frac{\partial f}{\partial y^\alpha} = \left(\sum_\alpha \frac{\partial f}{\partial y^\alpha} \, dy^\alpha \right)(F_* v)$$

$$= df(F_* v) \quad . \qquad \square$$

We have defined the pull-back of a 0-form by $F^*(f) = f \circ F$. We define the pull-back of a $p > 0$ form as in Lecture 7:

$$(F^* \alpha^p)(v_1, \ldots, v_p) \equiv \alpha^p(F_* v_1, \ldots, F_* v_p) \quad .$$

again we have $F^* : \Lambda^p V \to \Lambda^p M$ for every $F : M^n \to V^r$. (Note the reversal of direction of arrows).

Let us now prove our basic properties.

• **Problem 19.** Show that F is an "algebra homomorphism," i.e.,

$$F^*(\alpha^p \wedge \beta^q) = (F^* \alpha) \wedge (F^* \beta) \quad .$$

(Hint: Just use (8.5).)

Now look at $F^* \circ d\alpha^p$ for α^p a p-form on V^r . First look at the case when α is a function (i.e., $p = 0$) on V^r .

$$(F^* df) = F^* \left(\sum_\alpha \frac{\partial f}{\partial y^\alpha} \, dy^\alpha \right) = \text{(from Problem 19)}$$

$$\sum_\alpha \left(F^* \frac{\partial f}{\partial y^\alpha} \right) F^*(dy^\alpha) = \sum_\alpha \frac{\partial f}{\partial y^\alpha}(y(x)) \cdot \sum_i \frac{\partial y^\alpha}{\partial x^i} \, dx^i$$

$$= \sum_i \left(\sum_\alpha \frac{\partial f}{\partial y^\alpha} \frac{\partial y^\alpha}{\partial x^i} \right) dx^i = \sum_i \left(\frac{\partial}{\partial x^i}(f \circ F) \right) dx^i$$

$$= d(f \circ F) = d(F^* f) \quad .$$

Hence for 0- forms we do have

$$F^* \circ d = d \circ F^* \quad . \tag{10.2}$$

Remark. It is somewhat amusing to prove this without using coordinates but using our Lemma 10.1 directly.

Proposition. For forms of any degree

$$F^* \circ d = d \circ F^* \quad . \tag{10.3}$$

Proof.

$$d \circ F^* \sum_{\substack{J \\ \rightarrow}} a_{j_1 \dots j_p}(y) dy^{j_1} \wedge \cdots \wedge dy^{j_p} \qquad \text{(from Problem 19)}$$

$$= d \sum_{\substack{J \\ \rightarrow}} (F^* a_{j_1 \dots j_p}) F^*(dy^{j_1}) \wedge \cdots \wedge F^*(dy^{j_p}) \qquad \text{(from (10.2))} \quad .$$

$$= d \sum_{\substack{J \\ \rightarrow}} (F^* a_{j_1 \dots j_p}) d(F^* y^{j_1}) \wedge \cdots \wedge d(F^* y^{j_p})$$

$$= \sum_{\substack{J \\ \rightarrow}} (dF^* a_{j_1 \dots j_p}) \wedge d(F^* y^{j_1}) \wedge \cdots \wedge d(F^* y^{j_p})$$

$$= \sum_{\substack{J \\ \rightarrow}} (F^* da_{j_1 \dots j_p}) \wedge F^*(dy^{j_1} \wedge \cdots \wedge F^*(dy^{j_p})$$

$$= F^* \sum_{\substack{J \\ \rightarrow}} (da_{j_1 \dots j_p}) \wedge dy^{j_1} \wedge \cdots \wedge dy^{j_p}$$

$$= F^* \circ d \sum_{\substack{J \\ \rightarrow}} (a_{j_1 \dots j_p}) dy^{j_1} \wedge \cdots \wedge dy^{j_p}$$

□

Note that in coordinates

$$F^* \sum_{\substack{J \\ \rightarrow}} a_J dy^{j_1} \wedge \cdots \wedge dy^{j_p} = \sum_{\substack{J \\ \rightarrow}} a_J(y(x)) \sum_I \frac{\partial y^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{j_p}}{\partial x^{i_p}} \, dx^{i_1} \wedge \cdots \wedge dx^{i_p} \quad .$$

But

$$dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} = \sum_{l_1 < \cdots < l_p} \delta_{l_1 \cdots l_p}^{i_1 \cdots i_p} \, dx^{l_1} \wedge \cdots \wedge dx^{l_p} \quad .$$

(Note that we are merely putting the dx's in increasing order; for each $(i_1, \ldots, i_p)$ there is only one nonzero term in the sum on the right.) Thus

$$\begin{aligned} F^*\alpha^p &= \sum_{\underset{\rightarrow}{J}} a_J(y(x)) \sum_{\underset{\rightarrow}{L}} \left\{ \sum_{I} \frac{\partial y^{j_1}}{\partial x^{i_1}} \cdots \frac{\partial y^{j_p}}{\partial x^{i_p}} \, \delta_L^I \right\} dx^{l_1} \wedge \cdots \wedge dx^{l_p} \\ &= \sum_{\underset{\rightarrow}{J}} a_J(y(x)) \sum_{\underset{\rightarrow}{L}} \det \frac{\partial(y^{j_1}, \ldots, y^{j_p})}{\partial(x^{l_1}, \ldots, x^{l_p})} \, dx^{l_1} \wedge \cdots \wedge dx^{l_p} \\ &= \sum_{\underset{\rightarrow}{L}} \tilde{a}_L(x) dx^{l_1} \wedge \cdots \wedge dx^{l_p} \end{aligned}$$

where

$$\boxed{\tilde{a}_{l_1 \cdots l_p}(x) = \sum_{\rightarrow} a_{j_1 \cdots j_p}(y(x)) \det \frac{\partial(y^{j_1}, \ldots, y^{j_p}}{\partial(x^{l_1}, \ldots, x^{l_p})} \quad .} \tag{10.4}$$

For example, if $F : \mathbb{R}^2 \to \mathbb{R}^3$ is a parameterized surface (with u,v coordinates in $\mathbb{R}^2$, x,y,z coordinates in $\mathbb{R}^3$) then for a 2-form β^2 on $\mathbb{R}^3$ its pull-back to the $\mathbb{R}^2$ plane is

$$\begin{aligned} F^*\beta^2 &= F^*(b_1 \, dy \wedge dz + b_2 \, dz \wedge dx + b_3 \, dx \wedge dy) \qquad (10.5) \\ &= \left[(b_1 \circ F) \frac{\partial(y,z)}{\partial(u,v)} + (b_2 \circ F) \frac{\partial(z,x)}{\partial(u,v)} + (b_3 \circ F) \frac{\partial(x,y)}{\partial(u,v)} \right] du \wedge dv \end{aligned}$$

For another example, if $F : \mathbb{R}^3 \to \mathbb{R}^3$, with u,v,w coordinates in domain space

$$F^*(dx \wedge dy \wedge dz) = \frac{\partial(x,y,z)}{\partial(u,v,w)} \, du \wedge dv \wedge dw \quad . \tag{10.6}$$

It usually is not worthwhile to use (10.4) in a computation. We did use it in example (10.5) above, but it is easy to do the computation directly. For example

$$F^*(b_1 dy \wedge dz) = (F^* b_1)\, F^*(dy \wedge dz) = (b_1 \circ F)\, F^*(dy) \wedge F^*(dz)$$

$$= (b_1 \circ F) \left(\frac{\partial y}{\partial u}\, du + \frac{\partial y}{\partial v}\, dv \right) \wedge \left(\frac{\partial z}{\partial u}\, du + \frac{\partial z}{\partial v}\, dv \right)$$

$$= (b_1 \circ F) \left(\frac{\partial y}{\partial u} \frac{\partial z}{\partial v} - \frac{\partial y}{\partial v} \frac{\partial z}{\partial u} \right) du \wedge dw \quad \text{,etc.}$$

Lecture 11. Orientation and Pseudo-forms

(1) **Orientation of a manifold.** Let $e = (e_1, \ldots, e_n)$ and $f = (f_1, \ldots, f_n)$ be two bases of a vector space E; we can then write uniquely $f = eA$, i.e., $f_i = e_j A^j_i$ for some nonsingular matrix A. If $\det A > 0$ (resp. < 0) we say that e and f have the same (resp. opposite) orientation. The collection of all bases of E then naturally falls into two equivalence classes of bases. For example in $\mathbb{R}^3$ we have the "right handed" bases and the "left handed" bases. To pick out one of the two equivalence classes in E is to pick out an "orientation of E" or to "orient E;" we do this merely by picking out one basis of E and declaring it "positively oriented." For example, in $\mathbb{R}^3$ we may choose some left handed basis and declare it positively oriented and any right handed basis would be "negatively" oriented. Except for our prejudices about right and left we can see that neither of the two possible choices for orienting E is more natural than the other. Of course if a vector space E comes equipped with a special basis than you may wish to use this basis for orienting E, but then again you may not.

Now consider a manifold M^n. If we wish we may orient each tangent space M^n_x rather arbitrarily, but for many purposes it is important to do this in a "coherent" or "continuous" fashion. What do we mean by this? First note that if $\mathcal{U}_\alpha, x^1_\alpha, \ldots, x^n_\alpha$ is a coordinate patch on M^n then we can orient each tangent space M^n_x for $x \in \mathcal{U}_\alpha$ by declaring $\dfrac{\partial}{\partial x^1_\alpha}, \ldots, \dfrac{\partial}{\partial x^n_\alpha}$ to be positively oriented. We shall say that the entire manifold M^n is orientable provided we can cover M^n by coordinates patches $\{\mathcal{U}_\alpha\}$ such that in each overlap $\mathcal{U}_\alpha \cap \mathcal{U}_\beta$ the Jacobian determinant

$$\frac{\partial(x^1_\beta, \ldots, x^n_\beta)}{\partial(x^1_\alpha, \ldots, x^n_\alpha)} > 0 \quad . \tag{11.1}$$

Note then that

$$\frac{\partial}{\partial x^i_\alpha} = \sum_j \frac{\partial x^j_\beta}{\partial x^i_\alpha} \frac{\partial}{\partial x^j_\beta}$$

says that the two orientations of M^n_x, coming from $\mathcal{U}_\alpha$ and $\mathcal{U}_\beta$, agree!

Proposition. The tangent bundle TM of any manifold is an orientable manifold.

Proof. Let (u,x), $(\bar{u},\bar{x})$ be two coordinate patches in M. Then (x,v) and $(\bar{x},\bar{v})$ are local coordinates in TM and (see (6.1)),

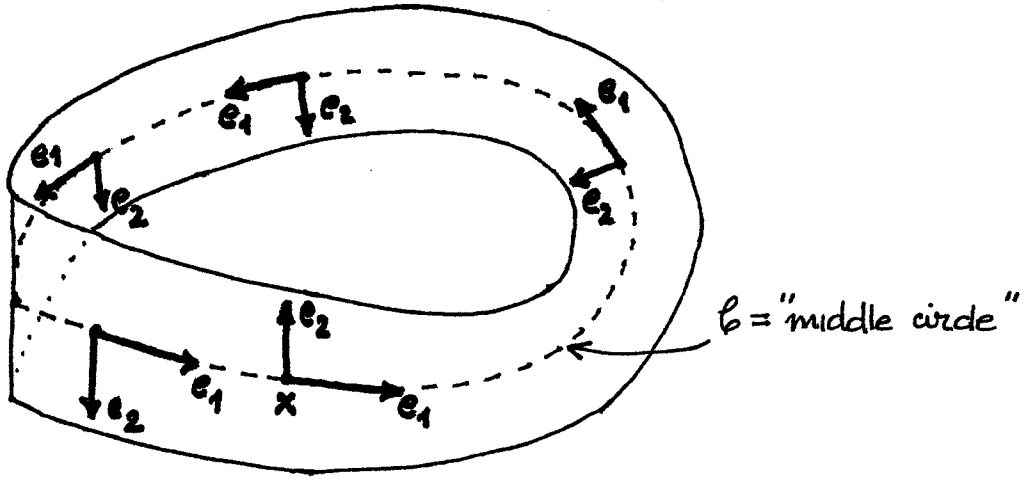
$$\frac{\partial(\bar{x},\bar{v})}{\partial(x,v)} = \begin{vmatrix} \frac{\partial\bar{x}}{\partial x} & \frac{\partial\bar{x}}{\partial v} \\ \frac{\partial\bar{v}}{\partial x} & \frac{\partial\bar{v}}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial\bar{x}}{\partial x} & 0 \\ * & \frac{\partial\bar{x}}{\partial x} \end{vmatrix}$$

$$= \left| \frac{\partial(\bar{x}^1, \ldots, \bar{x}^n)}{\partial(x^1, \ldots, x^n)} \right|^2 . \qquad \square$$

- **Problem 20.** Discuss orientability of T*M.

If M^n is orientable and if C is a closed curve starting and ending at x, and if we continuously transport a basis from x around C to x again then the final basis at x must define the same orientation of M_x^n as the original; this is evident from comparing the moving basis with the coordinate patch basis u_α as C moves through these patches (it is assumed that the patches satisfy (11.1). On the other hand, if on traversing a closed path in a manifold M^n the final orientation of M_x^n disagrees with the initial one, then M^n cannot be orientable !

Certainly then the Möbius band $M\ddot{o}^2$ is not orientable, since



transport around the middle circle reverses orientation.

From Problem 5 (Lecture 2) we know that Mö $\subset P^2(\mathbb{R})$; consequently transport around the mid circle of Mö reverses orientation in $P^2(\mathbb{R})$, $P^2(\mathbb{R})$ <u>is not orientable</u> !

$\mathbb{R}^n$ is clearly <u>orientable</u> since it can be covered by a single coordinate patch.

Mö is a 2-dimensional submanifold of $\mathbb{R}^3$ (provided we omit the edges) and it is famous that Mö: $\subset \mathbb{R}^3$ is <u>1 - sided</u>, i.e., there is no continuous, nonvanishing <u>normal</u> vector field to Mö in $\mathbb{R}^3$. Most of the familiar surfaces in $\mathbb{R}^3$ are 2-sided, i.e., possess continuous nonvanishing normal vector fields. Is there a relation between the notion of 1-sidedness and non-orientability?

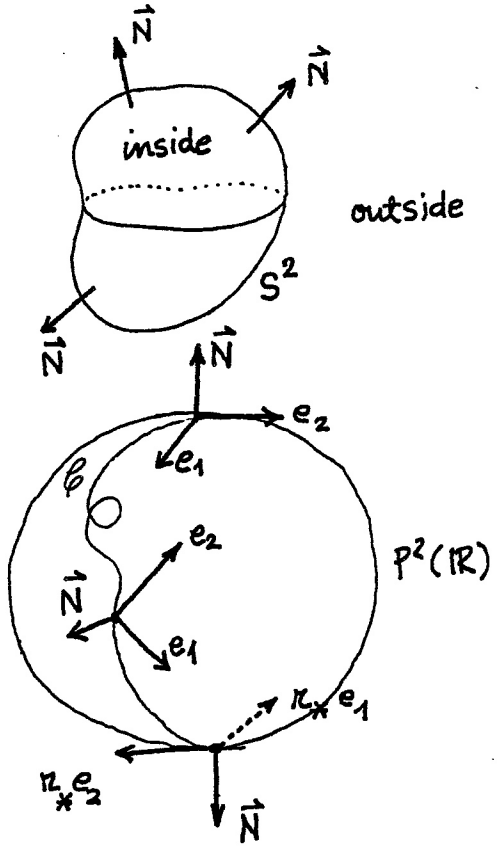
If $M^{n-1} \subset \mathbb{R}^n$ is a 2-sided submanifold of $\mathbb{R}^n$ then M^{n-1} is orientable! <u>Proof.</u> Choose an orientation of $\mathbb{R}^n$. Let N be a continuous nonvanishing normal to M^{n-1}. We may orient each tangent space M_x^{n-1} as follows. The basis $e_1,\ldots,e_{n-1}$ of M_x^{n-1} will be declared positively oriented provided $e_1,\ldots,e_{n-1},N$ is a positively oriented basis of $\mathbb{R}_x^n$. Then cover M^{n-1} with coordinate patches u_α, $x_\alpha^1,\ldots,x_\alpha^{n-1}$, but permute a pair of coordinates if necessary to ensure that $\frac{\partial}{\partial x_\alpha^1},\ldots,\frac{\partial}{\partial x_\alpha^{n-1}}$ are positively oriented. Such patches will orient M^{n-1}.

More generally, by exactly the same type of reasoning we have that if M^{n-1} <u>is a 2 - sided hypersurface of an orientable manifold</u> W^n <u>then</u> M^{n-1} <u>is itself orientable</u> !

For example the embedded 2-sphere $S^2 \subset \mathbb{R}^3$, since it has no self-intersections, is the boundary of both a bounded region of $\mathbb{R}^3$ (the "inside") and an unbounded "outside" region.

The <u>outward pointing normal</u> N shows us that this $S^2 \subset \mathbb{R}^3$ is 2-sided, hence S^2 is orientable.

It should be clear from our discussions that a manifold is orientable provided we can continuously orient each tangent space and this usually results from having some unambiguous method for assigning these orientations. S^2 inherits an orientation from $\mathbb{R}^3$ via the "outward normal." From Lecture 2, section d we know that $P^2(\mathbb{R})$ results



from S^2 by identifying points via the antipodal map $r : \mathbb{R}^3 \to \mathbb{R}^3$, $r(x) = -x$. But it is clear that a positively oriented frame e_1, e_2 at the north pole of S^2 is sent into a negatively oriented frame $r_* e_1$, $r_* e_2$ at the south pole of S^2. Both the north and south poles of S^2 correspond to the same point x of $P^2(\mathbb{R})$. It is clear then that if we have a curve C on S^2 from the north to the south pole and if we transport a basis along C, then considered as a closed curve on $P^2(\mathbb{R})$ (from x to x) the orientation of $P^2_x(\mathbb{R})$ gets reversed. This is another proof that $P^2(\mathbb{R})$ is non-orientable.

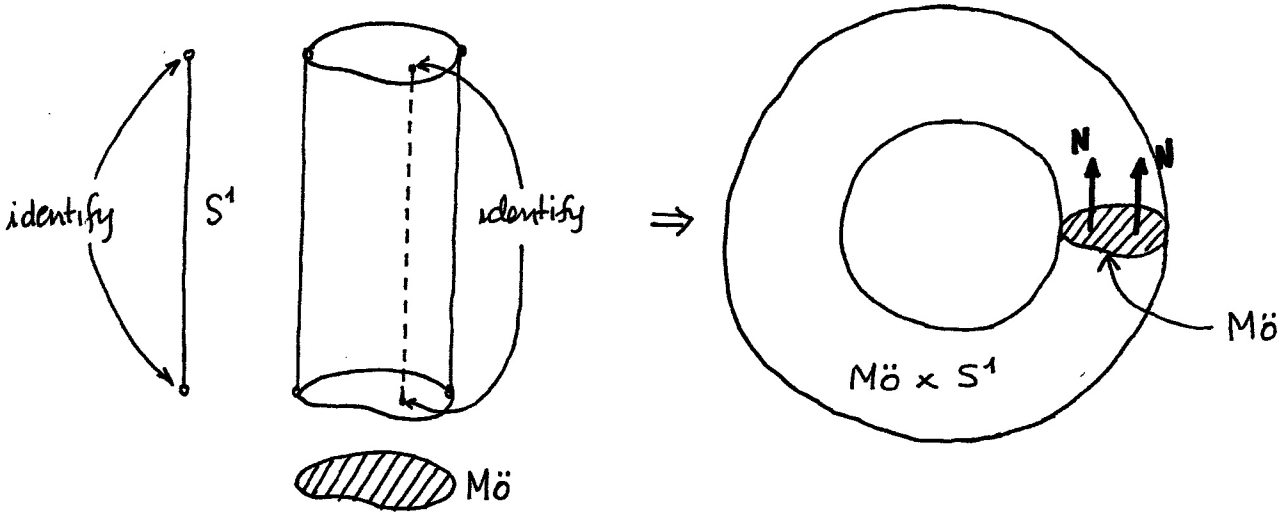
- **Problem 21.** Discuss orientability of $P^n(\mathbb{R})$. (Hint: Look at $r : S^n \to S^n$ and worry about n even, odd.)

(2) **1-sided, 2-sided.** Note carefully that the orientability of a manifold M^n is an intrinsic property of a manifold.

- **Problem 22.** Show that if $F : M^n \to V^n$ is a diffeomorphism then M is orientable if and only if V is orientable.

On the other hand one can discuss 1- or 2-sidedness of a manifold M^n only with regard to the way M^n is sitting in another W^{n+1} (since it is a question of existence of a continuous normal vector field).

Mö is nonorientable, and in its usual embedding in $\mathbb{R}^3$ it is (necessarily) 1-sided. We can also consider the 3-manifold $W^3 = \text{Mö} \times S^1$, the Cartesian product of the Möbius band and the unit circle (1-torus) (schematically, a weird solid doughnut).



I have drawn a continuous normal to Mö in Mö $\times$ S^1. In this embedding Mö is a 2 - sided surface in $W^3 = M\ddot{o}^2 \times S^1$. Of course Mö is still non-orientable.

On the other hand note that S^1 is an orientable curve (1-manifold) (in particular $S^1 \subset \mathbb{R}^2$ is 2-sided). However S^1, considered as the middle circle on the Möbius band $S^1 \subset$ Mö is clearly 1 - sided (see first drawing, §1).

(3) **Pseudo-forms $\equiv$ twisted forms $\equiv$ forms of odd kind.** The differential forms considered in our previous lectures were independent of the notion of orientability; $\alpha^p(v_1, \ldots, v_p)$ is a number which has nothing to do with the orientation of the particular tangent space in which $v_1, \ldots, v_n$ live.

There is a different notion of exterior form which is involved with orientation.

Consider $\mathbb{R}^3$ with its usual Euclidean structure. The volume pseudo form is an orientation dependent 3-form vol^3 such that

(i) given an orientation and a positively oriented o.n. frame e_1, e_2, e_3

$$\text{vol}^3(e_1,e_2,e_3) = 1$$

(and of course $\text{vol}^3(e_2,e_1,e_3) = -1$, etc.).

(ii) If we reverse the orientation of $\mathbb{R}^3$, then for the same frame e_1,e_2,e_3 we have (it is of course now negatively oriented)

$$\text{vol}^3(e_1,e_2,e_3) = -1 \quad .$$

Let us denote an orientation of $\mathbb{R}^3$ by o; its opposite orientation will be $-o$. Then if x^1, x^2, x^3 is a Cartesian coordinate system, the volume pseudo-form is

$$\text{vol}^3_o = o(x)\, dx^1 \wedge dx^2 \wedge dx^3$$

where

$$o(x) = \begin{cases} +1 \text{ if } \dfrac{\partial}{\partial x^1}, \dfrac{\partial}{\partial x^2}, \dfrac{\partial}{\partial x^3} \text{ is positively oriented with respect to } o \\ +1 \text{ if } \dfrac{\partial}{\partial x^1}, \dfrac{\partial}{\partial x^2}, \dfrac{\partial}{\partial x^3} \text{ is negatively oriented with respect to } o \end{cases} .$$

A pseudo form of degree p on a manifold M^n assigns to each orientation o of the tangent space M^n_x an exterior p- form α^p_o in such a way that

$$\alpha^p_{-o} = -\alpha^p_o$$

i.e., reversing the orientation changes the form α to its negative !

(4) **The Volume Pseudo n-form on a Riemannian Manifold** Let M^n be a Riemannian manifold, let $x_0 \in M^n$ and let $x^1, \ldots, x^n$ be a coordinate system about x that is orthonormal at the point x_0, i.e., $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^n}$ are o.n. at x_0, $\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right\rangle_{x_0} = \delta_{ij}$. (This can be achieve by taking any coordinate system with x_0 as origin and then taking a suitable linear change of coordinates, via Gram-Schmidt). In these coordinates, the volume pseudo n-form, at x_0, is simply (see Lecture 8, §3).

$$\text{vol}^n = o(x)dx^1 \wedge \cdots \wedge dx^n \quad .$$

Now let $y^1, y^2, \ldots, y^n$ be any coordinate system near x_0. From (8.11)

$$\begin{aligned} \text{vol}^n &= o(x) \frac{\partial(x^1, \ldots, x^n)}{\partial(y^1, \ldots, y^n)} dy^1 \wedge \cdots \wedge dy^n \quad \text{at} \quad x_0 \\ &= o(y) \left| \frac{\partial(x^1, \ldots, x^n)}{\partial(y^1, \ldots, y^n)} \right| dy^1 \wedge \cdots \wedge dy^n \quad \text{at} \quad x_0 \quad \text{(why?)} \end{aligned}$$

Now at x_0, $ds^2 = \delta_{rs}\, dx^r dx^s = g_{ij}(y)dy^i dy^j$, where (see (5.11))

$$g_{ij}(y) = \frac{\partial x^r}{\partial y^i} \frac{\partial x^s}{\partial y^j} \delta_{rs} = \frac{\partial x^r}{\partial y^i} \frac{\partial x^r}{\partial y^j} \tag{11.2}$$

and so if we define

$$g(y) \equiv \det g_{ij}(y) \tag{11.3}$$

we have

$$g(y) = \det g_{ij}(y) = \det \frac{\partial x^r}{\partial y^i} \frac{\partial x^r}{\partial y^j}$$

$$= \det \left[\left(\frac{\partial x}{\partial y} \right)^t \left(\frac{\partial x}{\partial y} \right) \right] = \det \left(\frac{\partial x}{\partial y} \right)^t \det \left(\frac{\partial x}{\partial y} \right) = \left[\det \left(\frac{\partial x}{\partial y} \right) \right]^2 .$$

Consequently

$$\left| \frac{\partial(x^1, \ldots, x^n)}{\partial(y^1, \ldots, y^n)} \right| = \sqrt{g(y)} \tag{11.4}$$

and so

$$\boxed{\text{vol}^n = o(y)\, \sqrt{g(y)}\, dy^1 \wedge \cdots \wedge dy^n \quad .} \tag{11.5}$$

Since x does not appear here, we see that this expression gives vol^n at every point of the coordinate system y.

• **Problem 23.** Use problem 10 to write down vol^3 for $\mathbb{R}^3$ in spherical coordinates r, θ, ϕ.

Lecture 12. Interior Products and Vector Analysis in $\mathbb{R}^3$

(1) **Interior product (contraction).** If X is a vector and α^p is a p-form we define their interior product , or contraction to be the $(p-1)$ form $i_X\alpha^p = X \lrcorner \alpha^p$ as follows:

$$\begin{cases} i_X\alpha^0 = 0 & \text{if } \alpha \text{ is a 0–form} \\ i_X\alpha^1 = \alpha^1(x) & \text{if } \alpha \text{ is a 1–form} \\ i_X\alpha^p(X_2,X_3,\ldots,X_p) = \alpha^p(X,X_2,\ldots,X_p) & \text{if } p > 1 \end{cases} .$$

Proposition. $i_X : \Lambda^p \to \Lambda^{p-1}$ is an anti - derivation , i.e.,

$$i_X(\alpha^p \wedge \beta^q) = (i_X\alpha^p) \wedge \beta^q + (-1)^p \alpha^p \wedge (i_X\beta^q) \quad .$$

Proof. Let us call $X = X_1$. Then

$$i_{X_1}(\alpha \wedge \beta)\,(X_2,\ldots,X_{p+q}) = \alpha \wedge \beta\,(X_1,X_2,\ldots,X_{p+q})$$

$$= \sum_{\substack{I\ J \\ \rightarrow,\rightarrow}} \delta^{IJ}_{1\ldots p+q}\, \alpha(X_{i_1}\cdots X_{i_p})\beta(X_{j_1}\cdots X_{j_q}) = \sum_{\substack{I\ J \\ \rightarrow,\rightarrow \\ 1\in I}} + \sum_{\substack{I\ J \\ \rightarrow,\rightarrow \\ 1\in J}}$$

$$= \sum_{\substack{i_2<\cdots<i_p \\ \rightarrow \\ J}} \delta^{1\,i_2\cdots i_p\,j_1\cdots j_q}_{1\,2\cdots\cdots p+q}\, \alpha(X_1,X_{i_2}\cdots X_{i_p})\beta(X_{j_1}\cdots X_{j_q})$$

$$+ \sum_{\substack{\rightarrow \\ I \\ j_2<\cdots<j_q}} \delta^{i_1\cdots i_p 1 j_2\cdots j_q}_{1\cdots\cdots p+q}\, \alpha(X_{i_1},\ldots,X_{i_p})\beta(X_1,X_{j_2},\ldots,X_{j_q})$$

$$= \sum_{\substack{1 \neq i_2 < \cdots < i_p \\ 1 \notin J}} \delta_{2 \cdots \cdots (p+q)}^{i_2 \cdots i_p j_1 \cdots j_q} (i_X\alpha)(X_{i_1} \cdots X_{i_p})\beta(X_{j_1}, \ldots, X_{j_q})$$

$$+ \sum_{\substack{1 \neq j_2 < \cdots < j_q \\ 1 \notin I}} (-1)^p \delta_{2 \cdots \cdots (p+q)}^{i_1 \cdots i_p j_2 \cdots j_q} \alpha(X_{i_1}, \ldots, X_{i_p})(i_X\beta)(X_{j_2}, \ldots, X_{j_q})$$

$$= [(i_X\alpha) \wedge \beta + (-1)^p \alpha \wedge (i_X\beta)](X_2, \ldots, X_{p+q}) \quad . \qquad \square$$

- **Problem 24.** In components $\alpha = \sum_{\substack{I \\ \rightarrow}} a_{i_1 \cdots i_p} \sigma^{i_1} \wedge \cdots \wedge \sigma^{i_p}$ and $X = \sum e_i X^i$, show

$$i_X\alpha = \sum_{i_2 < \cdots < i_p} b_{i_2 \cdots i_p} \sigma^{i_2} \wedge \cdots \wedge \sigma^{i_p} \quad \text{where}$$

$$\boxed{b_{i_2 \cdots i_p} = \sum_j X^j a_{j i_2 \cdots i_p}} \qquad (12.1)$$

It should be clear that if X is a vector and β is a pseudo form then $i_X\beta$ is again a pseudo form, and $d\beta$ is a pseudo form.

One should also keep in mind that if M^n is orientable then one can choose a coordinate cover for which $\dfrac{\partial(x_\alpha^1, \ldots, x_\alpha^n)}{\partial(x_\beta^1, \ldots, x_\beta^n)} > 0$. Then if β^p is a pseudo form, by restricting ourselves to this "positive coordinate cover" β^p yields a true p-form. Thus, in an orientable space one can always do without pseudo forms, but this is an unnatural advantage and we shall always try to keep the distinction. This is especially important in physics, as we shall see.

(2) **Vector analysis.** Now we can begin to identify vectors and forms in $\mathbb{R}^3$! In vector analysis, vectors appear in line and surface integrals. In Cartesian coordinates we have two appearances:

(i) $\underline{A} \leftrightarrow a_1\, dx^1 + a_2\, dx^2 + a_3\, dx^3$, and

(ii) $\underline{A} \leftrightarrow a_1\, dx^2 \wedge dx^3 + a_2\, dx^3 \wedge dx^1 + a_3\, dx^1 \wedge dx^2$.

These are the following in general coordinates:

(i) $\underline{A} \leftrightarrow$ the 1-form $\alpha = \left\langle\ , \underline{A}\right\rangle$, i.e., the covariant version of the contravariant vector $\underline{A}$. Recall that

$$\text{if } \underline{A} = a^1 \frac{\partial}{\partial u^1} + a^2 \frac{\partial}{\partial u^2} + a^3 \frac{\partial}{\partial u^3} \text{ , then } a_i = \sum_j g_{ij}(u) a^j \quad .$$

(ii) $\underline{A} \leftrightarrow$ the pseudo-2-form $i_{\underline{A}}\, vol_o$.

To see what this says in curvilinear coordinates

$$i_{\sum_j a^j \frac{\partial}{\partial u^j}}\, o(u)\sqrt{g(u)}\, du^1 \wedge du^2 \wedge du^3 = \sum_j a^j\, i_{\frac{\partial}{\partial u^j}}\, o(u)\sqrt{g(u)}\, du^1 \wedge du^2 \wedge du^3$$

$$= o(u)\sqrt{g(u)} \sum_j i_{\frac{\partial}{\partial u^j}} (du^1 \wedge du^2 \wedge du^3) \quad .$$

Now use the fact (Prop., §1) that $i_{\frac{\partial}{\partial u^j}}$ is an anti-derivation. Repeated use of this shows

$$i_{\frac{\partial}{\partial u^j}} (du^1 \wedge du^2 \wedge du^3) = i_{\frac{\partial}{\partial u^j}} (du^1) \wedge du^2 \wedge du^3$$

$$- du^1 \wedge i_{\frac{\partial}{\partial u^j}} (du^2) \wedge du^3 + du^1 \wedge du^2 \wedge i_{\frac{\partial}{\partial u^j}} (du^3)$$

$$= du^1 \left(\frac{\partial}{\partial u^j}\right) du^2 \wedge du^3 - du^2 \left(\frac{\partial}{\partial u^j}\right) du^1 \wedge du^2 + du^3 \left(\frac{\partial}{\partial u^j}\right) du^1 \wedge du^2$$

$$= \delta_j^{\,1}\, du^2 \wedge du^3 - \delta_j^{\,2}\, du^1 \wedge du^3 + \delta_j^{\,3}\, du^1 \wedge du^2 \text{ , and so}$$

$$\boxed{i_{\underline{A}}\, vol_o = o(u)\sqrt{g(u)}\,(a^1\, du^2 \wedge du^3 + a^2\, du^3 \wedge du^1 + a^3\, du^1 \wedge du^2)} \qquad (12.2)$$

Note the appearance of the factor $\sqrt{g(u)}$ in general coordinates!

If we write $i_A vol_o = \sum_{i<j} a_{ij}\, du^i \wedge du^j$ then we see

$$a_{23} = o(u)\sqrt{g}\, a^1\ ,\ a_{31} = o(u)\sqrt{g}\, a^2\ ,\ a_{12} = o(u)\sqrt{g}\, a^3\quad . \tag{12.3}$$

Conversely, given a 2-form (not a pseudo-form)

$$\beta^2 = \sum_{i<j} b_{ij}\, du^i \wedge du^j$$

we can associate to it a pseudo-vector $\underline{B}$ (meaning that the vector is orientation dependent and becomes its negative when the orientation is reversed) again by

$$\beta^2 = i_{\underline{B}}\, vol_o \tag{12.4}$$

where $b^1 = o(u)\dfrac{b_{23}}{\sqrt{g}}$, $b^2 = o(u)\dfrac{b_{31}}{\sqrt{g}}$, $b^3 = o(u)\dfrac{b_{12}}{\sqrt{g}}$.

Next, given a pair of vectors $\underline{A},\underline{B}$, using identification (i) above we have

$$\underline{A}\cdot\underline{B} = \alpha^1(\underline{B}) = i_{\underline{B}}\alpha^1\ ,\ \text{etc.} \tag{12.5}$$

Note in Cartesian coordinates

$$\begin{aligned}\alpha^1 \wedge \beta^2 &= (a_1dx^1+a_2dx^2+a_3dx^3) \wedge (b_1dx^2 \wedge dx^3+b_2dx^3 \wedge dx^1+b_3dx^1 \wedge dx^2)\\ &= (a_1b_1 + a_2b_2 + a_3b_3)dx^1 \wedge dx^2 \wedge dx^3\end{aligned}$$

so it should not be surprising that in general coordinates

$$\boxed{\alpha^1 \wedge \beta^2 = (\underline{A}\cdot\underline{B})\, vol_o} \tag{12.6}$$

For proof, note

$$\begin{aligned}\alpha^1 \wedge \beta^2 &= \alpha^1 \wedge i_{\underline{B}} vol_o = (i_{\underline{B}} vol_o) \wedge \alpha^1 \quad \text{(why?)} \\ &= i_B(vol_o \wedge \alpha^1) + vol_o \wedge i_{\underline{B}}(\alpha^1) \quad \text{(why?)} \\ &= i_B(\alpha^1) vol_o \quad \text{(why?)} \\ &= (\underline{A} \cdot \underline{B}) vol_o \quad .\end{aligned}$$

□

What about the $\times$ product of vectors? In Cartesian coordinates one gets immediately (see Lecture 8, §4)

$$\alpha^1 \wedge \beta^1 \leftrightarrow \underline{A} \times \underline{B} \quad .$$

$$\boxed{\alpha^1 \wedge \beta^1 = i_{\underline{A} \times \underline{B}} \, vol_o} \tag{12.7}$$

This is true in any coordinate system since both sides are true 2-forms, (why?) and the result is true in Cartesian coordinates.

In (12.7) the pseudo vector $\underline{A} \times \underline{B}$ is identified with a 2-form, however we can ask for a pseudo-1-form (the covariant expression for $\underline{A} \times \underline{B}$) instead. Note that the official definition of $\underline{A} \times \underline{B}$ is as follows:

$\underline{A} \times \underline{B}$ is that unique pseudovector such that

$(\underline{A} \times \underline{B}) \cdot \underline{C} = vol_o(\underline{A},\underline{B},\underline{C})$, for all vectors $\underline{C}$.

Note then that the form $-i_{\underline{A}}\beta^2$ has the following property

$$\begin{aligned}(-i_{\underline{A}}\beta^2)(\underline{C}) = -\beta^2(\underline{A},\underline{C}) &= (-i_{\underline{B}} vol_o)(\underline{A},\underline{C}) = -vol_o(\underline{B},\underline{A},\underline{C}) \\ &= vol_o(\underline{A},\underline{B},\underline{C}) = (\underline{A} \times \underline{B}) \cdot \underline{C}\end{aligned}$$

and so

$$\boxed{-i_{\underline{A}}\beta^2 \text{ is the covariant version of } \underline{A} \times \underline{B}} \tag{12.8}$$

Thus we have

$$\underline{A} \times \underline{B} \leftrightarrow \begin{cases} -\, i_A \beta^2 & \text{a pseudo–1–form} \\ \alpha^1 \wedge \beta^1 & \text{a true 2–form} \end{cases} .$$

If γ^1 is the 1-form corresponding to vector $\underline{C}$, note

$$\alpha^1 \wedge \beta^1 \wedge \gamma^1 = (i_{\underline{A} \times \underline{B}}\, \mathrm{vol}_o) \wedge \gamma^1 = (\underline{A} \times \underline{B}) \cdot \underline{C}\, \mathrm{vol}_o \quad \text{(why?)} .$$

Note that the identities of vector algebra are usually easily done via forms, the only decision being whether to use 1-forms or 2-forms. For example

$$\underline{A} \times (\underline{B} \times \underline{C}) \leftrightarrow -\, i_{\underline{A}}(\beta^1 \wedge \gamma^1) = -\,(i_A \beta^1)\gamma^1 + \beta^1\, i_{\underline{A}}(\gamma^1)$$
$$\leftrightarrow -\,(\underline{A} \cdot \underline{B})\underline{C} + \underline{B}(\underline{A} \cdot \underline{C})$$

i.e.,

$$A \times (\underline{B} \times \underline{C}) = (\underline{A} \cdot \underline{C})\underline{B} - (\underline{A} \cdot \underline{B})\underline{C} .$$

So much for vector <u>algebra</u>, now for <u>calculus</u>.

We already know that if f is a function, then df is the covariant version of grad $f = \nabla f$

$$df = \langle\, , \nabla f \rangle , \qquad (\nabla f)^i = \sum g^{ij}(u)\, \frac{\partial f}{\partial u^j} .$$

From the simple calculations (Lecture 9, §2) we see

$$\boxed{d\alpha^1 = i_{\mathrm{curl}\, \underline{A}}\, \mathrm{vol}_o} \tag{12.9}$$

(note that both vol_o and $\mathrm{curl}\, \underline{A} = \nabla \times \underline{A}$ are pseudo quantities), and

$$\boxed{d\beta^2 = (\mathrm{div}\, \underline{B})\, \mathrm{vol}_o} \tag{12.10}$$

Note that (12.10) actually <u>defines</u> the notion $\mathrm{div}\, \underline{B}$, the divergence of a vector, i.e., div B is uniquely defined by

$$d(i_{\underline{B}}\, \mathrm{vol}_o) = (\mathrm{div}\, B) \cdot \mathrm{vol}_o \tag{12.11}$$

(and so div B is a function, not a pseudo function!)

To compute $\operatorname{div} \underline{B} = \operatorname{div}\left[b^1 \frac{\partial}{\partial u^1} + b^2 \frac{\partial}{\partial u^2} + b^3 \frac{\partial}{\partial u^3}\right]$, we know from (12.2) that

$$d(i_{\underline{B}} \operatorname{vol}_o) = d\, o(u) \cdot (\sqrt{g(u)}\, b^1\, du^2 \wedge du^3 + \sqrt{g(u)}\, b^2\, du^3 \wedge du^1 + \sqrt{g(u)}\, b^3\, du^1 \wedge du^2)$$

$$= o(u)\left\{\frac{\partial}{\partial u^1}(\sqrt{g}\, b^1) + \frac{\partial}{\partial u^2}(\sqrt{g}\, b^2) + \frac{\partial}{\partial u^3}(\sqrt{g}\, b^3)\right\} du^1 \wedge du^2 \wedge du^3$$

$$= \frac{1}{\sqrt{g}}\left\{\frac{\partial}{\partial u^i}(\sqrt{g}\, b^i)\right\} \cdot o(u)\sqrt{g}\, du^1 \wedge du^2 \wedge du^3$$

$$= \frac{1}{\sqrt{g}}\left\{\sum_i \frac{\partial}{\partial u^i}(\sqrt{g}\, b^i)\right\} \cdot \operatorname{vol}_o, \quad \text{and so}$$

$$\boxed{\operatorname{div} \underline{B} = \frac{1}{\sqrt{g}} \sum_i \frac{\partial}{\partial u^i}(\sqrt{g}\, b^i)} \tag{12.12}$$

It is traditional to define the Laplacian $\nabla^2 f = \Delta f$ by

$$\Delta f = \operatorname{div}(\operatorname{grad} f) = \operatorname{div}(\nabla f) \tag{12.13}$$

From (12.12) we get in general coordinates

$$\boxed{\Delta f = \frac{1}{\sqrt{g}} \sum_{i,j} \frac{\partial}{\partial u^i}(\sqrt{g}\, g^{ij} \frac{\partial f}{\partial u^j}).} \tag{12.14}$$

Again we can do vector identities: For example

$$\operatorname{div}(\underline{A} \times \underline{B}) \leftrightarrow d(\alpha^1 \wedge \beta^1) = d\,\alpha^1 \wedge \beta^1 - \alpha^1 \wedge d\beta^1 \leftrightarrow (\operatorname{curl} \underline{A}) \cdot \underline{B} - \underline{A} \cdot (\operatorname{curl} \underline{B}) \quad .$$

• **Problem 25.** Derive the following identities using exterior methods:

(i) $\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f$, (ii) $\operatorname{div}(f\underline{A}) = f \operatorname{div} \underline{A} + \operatorname{grad} f \cdot \underline{A}$

(iii) $\operatorname{curl}(f\underline{A}) = f \operatorname{curl} \vec{B} + (\operatorname{grad} f) \times \underline{A}$, (iv) $(\underline{A} \times \underline{B}) \cdot (\underline{C} \times \underline{D}) = ???$

Lecture 13. Integration of Forms Over Parameterized Subsets

(1) **Integration of a (true) p-form in $\mathbb{R}^p$.** A p-dimensional region $u \subset \mathbb{R}^p$ is compact provided it is closed (see Lecture 2, Appendix) and bounded.

Let u_o be a compact oriented region in $\mathbb{R}^p$ (the same region without orientation will be denoted by u or $|u|$). Let $u^1, \ldots, u^p$ be a coordinate system for $\mathbb{R}^p$ (or at least u). Recall that

$$o(u) \equiv \begin{cases} +1 \text{ if } \dfrac{\partial}{\partial u^1}, \ldots, \dfrac{\partial}{\partial u^p} \text{ is positively oriented w.r.t. } o \\ -1 \text{ if } \dfrac{\partial}{\partial u^1}, \ldots, \dfrac{\partial}{\partial u^p} \text{ is negatively oriented w.r.t. } o \end{cases}$$

If α^p is an exterior form of maximal degree p on we can write

$$\alpha^p = a_{1 \ldots p}(u)\, du^1 \wedge \cdots \wedge du^p = a(u)\, du^1 \wedge \cdots \wedge du^p \quad .$$

We then define

$$\boxed{\int_{u_o} \alpha^p = \int_{u_o} a(u)\, du^1 \wedge \cdots \wedge du^p \equiv o(u) \int_{|u|} \ldots \int a(u)\, du^1 \cdots du^p} \qquad (13.1)$$

where the last integral is the usual multiple Riemann integral (which is defined independent of any notion of exterior form (there are no wedges "$\wedge$") and of any orientation).

Clearly if we reverse the orientation o of u the integral will change sign

$$\int_{u_{-o}} \alpha^p = -\int_{u_o} \alpha^p \quad . \qquad (13.2)$$

(2) **Integration over parameterized subsets.** We shall define a p- dimensional (oriented) parameterized subset of a manifold M^n to be a differentiable map

$$\phi : u \longrightarrow M^n$$

of a compact (oriented) region $u \subset \mathbb{R}^p$ into M^n.

Familiar examples when $M^n = \mathbb{R}^3$ are parameterized curves $(p = 1)$ $\vec{x} = \vec{x}(t)$ and parameterized surfaces $(p = 2)$, $\vec{x} = \vec{x}(u,v)$. In these cases, for example u may be a closed interval $[a,b] \subset \mathbb{R}^1$ or perhaps a closed rectangle $[a,b] \times [c,d] \subset \mathbb{R}^2$.

> We emphasize that ϕ need not be 1:1. Consequently, the image subset $\phi(u)$ may have, e.g., smaller dimension than p. In the most important cases ϕ is 1:1.

We also emphasize that our parameterized subset of M^n consists of the pair (ϕ,u), and not just the image $\phi(u) \subset M^n$.

Now given a p-dimensional oriented parameterized subset (ϕ,u_o) of M^n, and given a (true) p-form α^p on M, we define

$$\int_{(\phi,u_o)} \alpha^p \equiv \int_{u_o} \phi^*\alpha^p \tag{13.3}$$

the right hand integral having been defined in (13.1). In detail

$$\int_{(\phi,u_o)} \alpha^p = \int_{u_o} \phi^*\alpha^p = \int_{u_o} \phi^*\alpha^p \left(\frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^p} \right) du^1 \wedge \cdots \wedge du^p$$

$$= o(u) \int\!\!\cdots\!\!\int_{|u|} \phi^*\alpha^p \left(\frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^p} \right) du^1 \cdots du^p \quad . \tag{13.4}$$

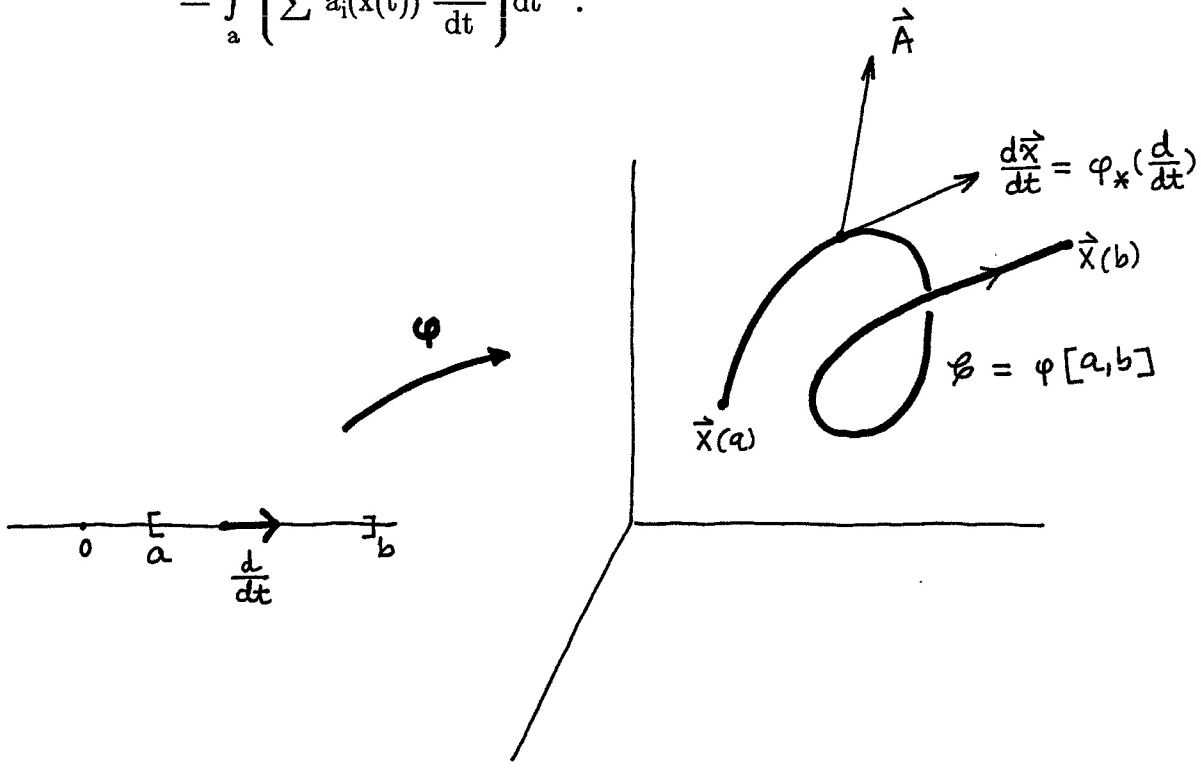
Note that this last Riemann integral can also be written

$$o(u) \int\!\!\cdots\!\!\int_{|u|} \alpha^p \left(\phi_* \frac{\partial}{\partial u^1}, \ldots, \phi_* \frac{\partial}{\partial u^p} \right) du^1 \cdots du^p \quad . \tag{13.4'}$$

For example, for a parameterized curve $\vec{x} = \vec{x}(t)$, $a \leq t \leq b$ in $\mathbb{R}^3$, oriented such that $\frac{d}{dt}$ on $\mathbb{R}^1$ defines the positive orientation, if $\alpha^1 = a_1(x)dx^1 + a_2(x)dx^2 + a_3(x)dx^3$ is a 1-form on $\mathbb{R}^3$ then

$$\int_{(\phi,[a,b])} \alpha^1 = \int_{(\phi,[a,b])} \sum_i a_i(x)dx = \int_{[a,b]} \phi^* \left(\sum a_i(x)dx^i \right)$$

$$= \int_a^b \left(\sum a_i(x(t)) \frac{dx^i}{dt} \right) dt \quad .$$



Usually the parameterized curve $(\phi,[a,b])$ is abbreviated C (with an arrow in the figure to indicate the orientation). This is the familiar rule for evaluating a line integral over an oriented curve. Note that if we use the usual Riemannian (i.e., Euclidean) structure of $\mathbb{R}^3$ then we may identify the 1-form α^1 with a contravariant vector $\vec{A}$ (see Lecture 5) and then we can write

$$\int_a^b \left(\sum_i a_i(x(t)) \right) dt = \int_a^b \left\langle \vec{A}, \frac{d\vec{x}}{dt} \right\rangle dt$$

i.e., one simply integrates the scalar product of the vector $\vec{A}$ with the velocity vector $\frac{d\vec{x}}{dt}$. Of course, if the coordinates are not Cartesian

$$\left\langle \vec{A}, \frac{d\vec{x}}{dt} \right\rangle = \sum_{i,j} A^i g_{ij} \frac{dx^j}{dt} = \sum_j a_j \frac{dx^j}{dt} \quad .$$

The lesson here is that one integrates covectors over curves, not vectors (remember $\phi_* \frac{d}{dt} = \frac{d\vec{x}}{dt}$)

$$\boxed{\int_C \alpha^1 = \int_a^b \alpha^1 \left(\frac{d\vec{x}}{dt} \right) dt} \tag{13.5}$$

If one insists on integrating (contravariant) vectors $\vec{A}$ over curves then one is going to need a Riemannian metric to change $\vec{A}$ to a covector; this is the meaning of the familiar

$$\int_C \vec{A} \cdot d\vec{x} = \int_a^b \left\langle \vec{A}, \frac{d\vec{x}}{dt} \right\rangle dt = \int_a^b \alpha^1 \left(\frac{dx}{dt} \right) dt = \int_C \alpha^1 \quad .$$

Integrals of forms require no Riemannian metric! Finally, note

$$\left\langle \vec{A}, \frac{d\vec{x}}{dt} \right\rangle dt = \| A \| \left\| \frac{d\vec{x}}{dt} \right\| \cos \measuredangle \left(\vec{A}, \frac{d\vec{x}}{dt} \right) dt$$

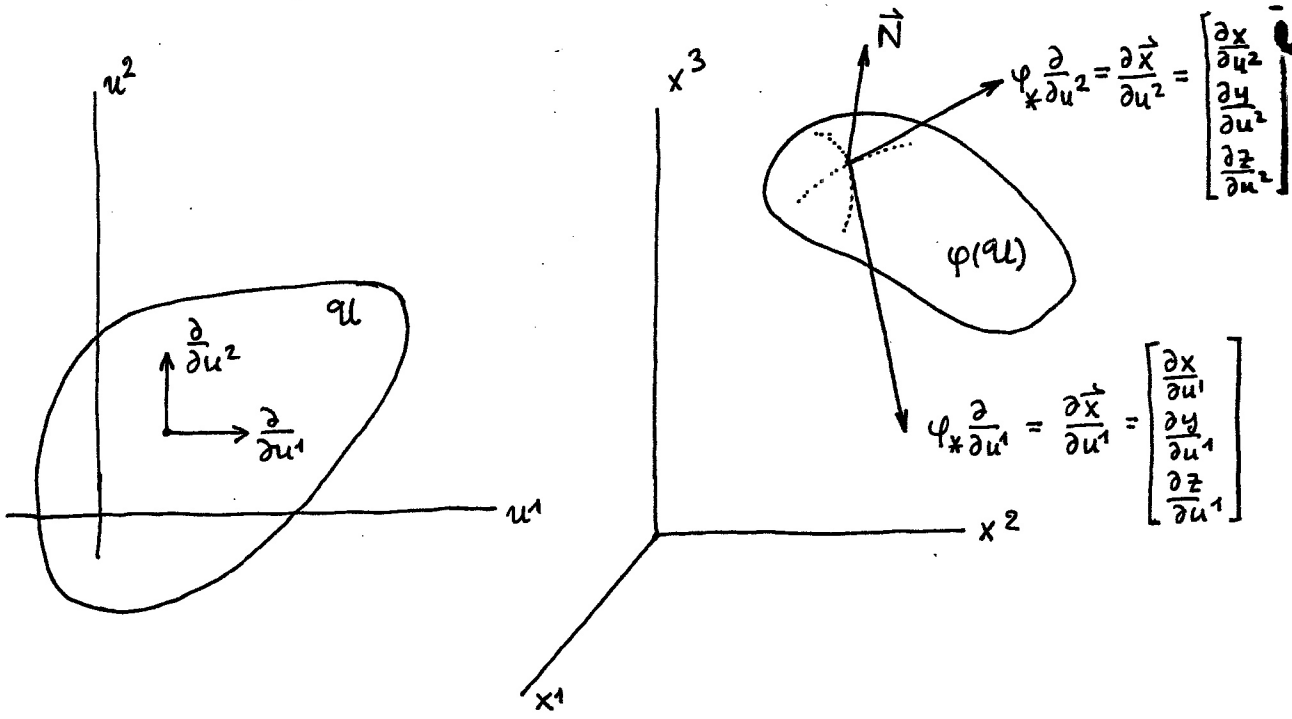
$$= A_t \, ds$$

where $A_t = \| \vec{A} \| \cos \measuredangle \left(\vec{A}, \frac{d\vec{x}}{dt} \right)$ is the tangential component of $\vec{A}$ along the curve, and $ds = \left\| \frac{d\vec{x}}{dt} \right\| dt$ is the element of arc length. If the curve has length L

$$\boxed{\int_C \alpha^1 = \int_0^L A_t \, ds} \tag{13.6}$$

as usual. However $\int \alpha^1 = \int_a^b \sum_i a_i(x(t)) \frac{dx^i}{dt} \, dt$ is much simpler !

Next consider a parameterized oriented surface.



We suppose that coordinates u^1, u^2 are chosen so that $o(u) = +1$. Then if $\beta^2 = b_1(x)dx^2 \wedge dx^3 + b_2(x)dx^3 \wedge dx^1 + b_3(x)dx^1 \wedge dx^2$ is a 2-form on $\mathbb{R}^3$ (i.e., $b_1 = B_{23}$, $b_2 = B_{31}$, $b_3 = B_{12}$)

$$\iint_{(\phi,\mathcal{U})} \beta^2 = \iint_{(\phi,\mathcal{U})} b_1\, dx^2 \wedge dx^3 + b_2\, dx^3 \wedge dx^1 + b_3\, dx^1 \wedge dx^2$$

$$= \iint_{\mathcal{U}} \left[\frac{\partial(x^2,x^3)}{\partial(u^1,u^2)} + b_2 \frac{\partial(x^3,x^1)}{\partial(u^1,u^2)} + b_3 \frac{\partial(x^1,x^2)}{\partial(u^1,u^2)} \right] du^1 du^2 \qquad (13.7)$$

which is a rather complicated (but computationally simple) way of writing

$$\iint_{(\phi,\mathcal{U})} \beta^2 = \iint_{\mathcal{U}} \beta^2 \left(\frac{\partial \vec{x}}{\partial u^1}, \frac{\partial \vec{x}}{\partial u^2} \right) du^1 du^2 \quad . \qquad (13.8)$$

If one insists on writing this in terms of vectors, as one does in textbooks we could proceed as follows. Use the familiar Riemannian metric for $\mathbb{R}^3$. From (12.4) (here o denotes orientation of $\mathbb{R}^3$)

$$\beta^2 = i_{\vec{B}_o} \mathrm{vol}^3_o$$

for a unique pseudo-vector $\vec{B}_o$. Then

$$\iint_{(\phi,\mathcal{U})} \beta^2 = \iint_{(\phi,\mathcal{U})} i_{\vec{B}_o} \operatorname{vol}_o^3 \left(\frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) du^1 \wedge du^2$$

$$= \iint_{(\phi,\mathcal{U})} \operatorname{vol}_o^3 \left(\vec{B}_o , \frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) du^1 \wedge du^2 \quad . \tag{13.9}$$

(It is clear that this answer is indèpendent of the orientation o of $\mathbb{R}^3$ since both vol_o^3 and $\vec{B}_o$ change sign under $o \to -o$.)

We shall only be concerned with <u>regular</u> parameterized surfaces, i.e., $\phi_* : \mathcal{U}_p \to \mathbb{R}^3_{\phi(p)}$ is of rank 2. Thus $\frac{\partial \vec{x}}{\partial u^1}$ and $\frac{\partial \vec{x}}{\partial u^2}$ are linearly independent.

<u>Pick now arbitrarily one of the two unit normals</u> $\vec{N}$ to the surface. This choice of $\vec{N}$ will give us a choice of o for $\mathbb{R}^3$ by demanding

$$o\left(\vec{N} , \frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) = +1 \quad .$$

We then have a unique decomposition $\vec{B} = \langle \vec{B} , \vec{N} \rangle \vec{N} + \vec{T}$ where $\vec{T}$ is tangent to the surface. But then $\operatorname{vol}^3 \left(T , \frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) = 0$ and so, from (13.9)

$$\iint_{(\phi,\mathcal{U})} \beta^2 = \iint_{(\phi,\mathcal{U})} \operatorname{vol}^3 \left(\langle B,N \rangle N , \frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) du^1 \wedge du^2$$

$$= \iint_{\mathcal{U}} \langle \vec{B},\vec{N} \rangle \operatorname{vol}^3 \left(\vec{N} , \frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) du^1 du^2 \quad .$$

But

$$\operatorname{vol}^3 \left(\vec{N} , \frac{\partial \vec{x}}{\partial u^1} , \frac{\partial \vec{x}}{\partial u^2} \right) du^1 du^2 = dS \tag{13.10}$$

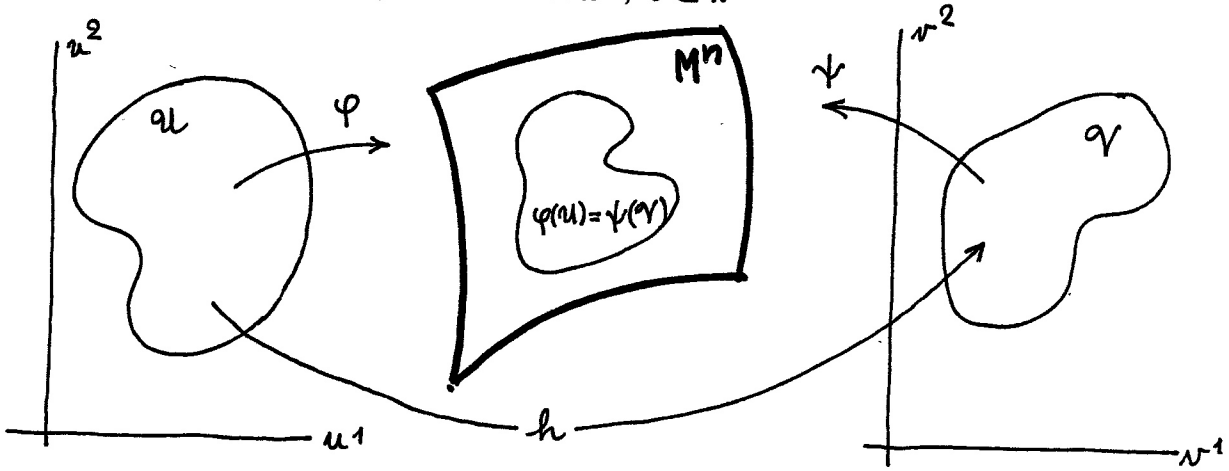
is simply the area element on the surface $V = \phi(u)$, since $\mathrm{vol}^3\left(\vec{N}, \frac{\partial \vec{x}}{\partial u^1}, \frac{\partial \vec{x}}{\partial u^2}\right) = +$ (area of the parallelogram spanned by $\frac{\partial \vec{x}}{\partial u^1}$ and $\frac{\partial \vec{x}}{\partial u^2}$). Hence we have the usual expression

$$\boxed{\iint_{(\phi,u)} \beta^2 = \iint_V B_N \, dS} \tag{13.11}$$

where B_N is the normal component of $\vec{B}$.

(3) **The integrals are independent of the parameterization.** It is evident from (13.6) and (13.11) that the line and surface integrals are in fact independent of the parameterization used in their evaluation; this is because, using our Riemannian metrics, these integrals are expressible in terms of geometrical quantities (integration with respect to arc length, surface area). Let us look at the general situation.

Let (ϕ,u) and (ψ,v) be two parameterized oriented subsets of M^n; $\phi : u \to M^n$, $\psi : v \to M^n$, where $u \subset \mathbb{R}^p$, $v \subset \mathbb{R}^p$



We assume that coordinates u in $\mathbb{R}^p$ and v in $\mathbb{R}^p$ are chosen so that $\frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^p}$ is positively oriented for u and $\frac{\partial}{\partial v^1}, \ldots, \frac{\partial}{\partial v^p}$ is positively oriented for v. We shall say that (ψ,v) is a reparameterization of (ϕ,u) provided there is an orientation preserving diffeomorphism $h : u \to v$ such that

$$\phi = \psi \circ h \quad .$$

Since h is orientation preserving, h is described by $v = v(u)$ such that

$$\frac{\partial(v^1, \ldots, v^p)}{\partial(u^1, \ldots, u^p)} > 0 \quad .$$

Then $\int_{(\psi,U)} \alpha^p = \int_U \psi^*\alpha^p = \int_U f(v)dv^1 \wedge \cdots \wedge dv^p = \int_{V=h(U)} f(v)dv^1 \cdots dv^p$ But Jacobi's rule for change of variables in a multiple integral says

$$\int_{h(U)} f(v)dv^1 \cdots dv^p$$

$$= \int_U f(v(u)) \left| \frac{\partial(v^1, \ldots, v^p)}{\partial(u^1, \ldots, u^p)} \right| du^1 \cdots du^p$$

$$= \int_U f(v(u)) \frac{\partial(v^1, \ldots, v^p)}{\partial(u^1, \ldots, u^p)} du^1 \wedge \cdots \wedge du^p$$

$$= \int_U h^*(\psi^*\alpha) = \text{(see Problem 12)}$$

$$= \int_U (\psi \circ h)^*\alpha = \int_{(\phi,U)} \phi^*\alpha \quad .$$

Hence Jacobi's rule insures us that

$$\int_{(\psi,U)} \alpha^p = \int_{(\phi,U)} \alpha^p$$

and so our integral is invariant under reparameterization.

Two final remarks. First, we have yet to define integrals of pseudo-forms. This will be done in Lecture 16. Second, parameterized subsets of M^n are rather special; for example they are by definition covered by a single "coordinate system." There are many interesting objects (for example submanifolds) that cannot be parameterized by a single parameterization. We shall discuss this in Lecture 14.

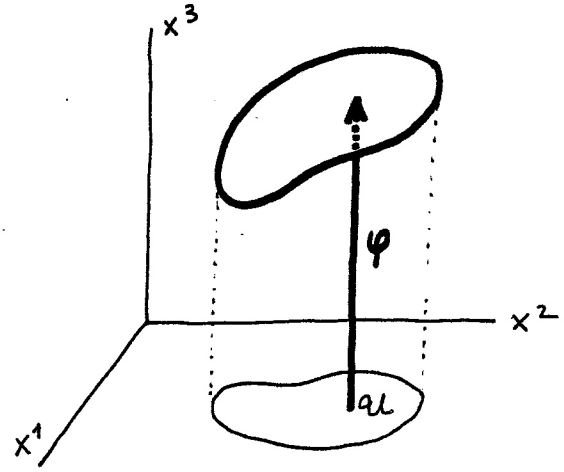
- **Problem 26.** We have seen that the "area form" for a parameterized surface in $\mathbb{R}^3$ is the pull-back ϕ^* of the 2-form

$$i_{\vec{N}} \, vol^3 = \frac{1}{\|\vec{n}\|} \, i_{\vec{n}} \, vol^3$$

where $\vec{n} = \frac{\partial \vec{x}}{\partial u^1} \times \frac{\partial \vec{x}}{\partial u^2}$ is the (non-unit) normal. Using Cartestian coordinates

$$x^1,x^2,x^3 \text{ , } \vec{n} = \frac{\partial(x^2,x^3)}{\partial(u^1,u^2)} \frac{\partial}{\partial x^1} + \frac{\partial(x^3,x^1)}{\partial(u^1,u^2)} \frac{\partial}{\partial x^2} + \frac{\partial(x^1,x^2)}{\partial(u^1,u^2)} + \frac{\partial}{\partial x^3} \text{ .}$$

Show that in the case where $(\phi,\mathcal{U})$ is simply the graph of a function over the x^1,x^2 plane, i.e., $x^1 = u^1, x^2 = u^2$, $x^3 = f(x^1,x^2)$, then one recovers the classical expression for area element for a surface.

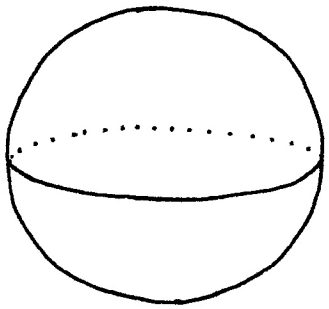


- **Problem 27.** Let us say that a p-dimensional parameterized subset $(\phi,\mathcal{U})$ of M^n is "irregular" provided ϕ_* has rank $< p$ at some point $u_0 \in \mathcal{U}$. Show that

$$\phi^* \alpha^p = 0 \quad \text{at} \quad u_0 \text{ .}$$

Lecture 14. Integration and Partitions of Unity

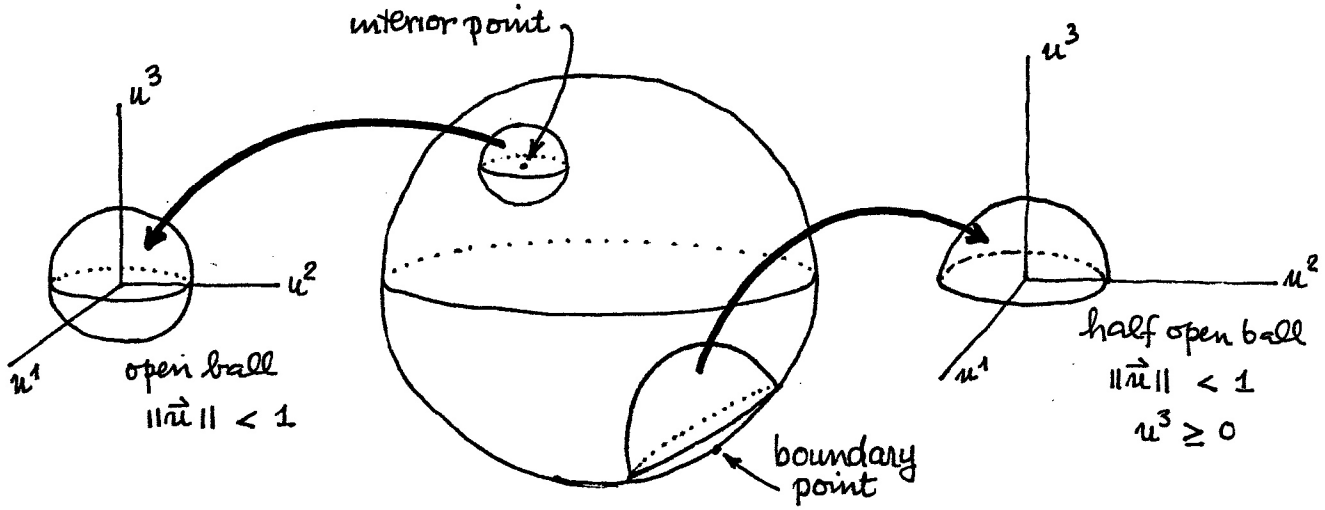
(1) **Manifolds with boundary.** In Lecture 13 we integrated (true) p-forms over parameterized oriented subsets of a manifold M^n. In a sense these objects are very special; by definition they are described by a single coordinate system $u^1, \ldots, u^p$. We wish to integrate over more complicated objects, mainly p-dimensional submanifolds of M^n (for instance we would like to take surface integrals over a 2-sphere in $\mathbb{R}^3$). In elementary vector analysis (and also in "algebraic topology"; this will be discussed later) this is accomplished by breaking up the submanifold V^p into a union of essentially non-overlapping parameterized subsets, doing each integral as before, and then adding the results.



Writing the 2-sphere as a union of two parameterized subsets (a northern and a southern hemisphere).

We prefer to use at this stage, a different, very useful theoretical device, a "partition of unity," but first we shall discuss an extension of the notion of manifold to "manifold with boundary." These will be our principal "domains of integration."

The closed 3-ball in $\mathbb{R}^3$ is <u>not</u> a 3-manifold, for while points in the interior, i.e., $\vec{x}$ for which $\|\vec{x}\| < 1$, do have neighborhoods diffeomorphic to open balls in $\mathbb{R}^3$, points on the boundary 2-sphere have neighborhoods that resemble half open balls.



Let us check that a point on the boundary of this closed ball does have a neighborhood described by a half open set where $u^3 \geq 0$ in 3-space with coordinates u^1, u^2, u^3. For simplicity suppose the point $\vec{x}_0$ in question is on the lower hemisphere $x_0^2 + y_0^2 + z_0^2 = 1$, $z_0 < 0$ of the boundary 2-sphere. This hemisphere can be described by

$$0 = F(x,y,z) = z + \sqrt{1-(x^2+y^2)} = z + f(x,y) \quad .$$

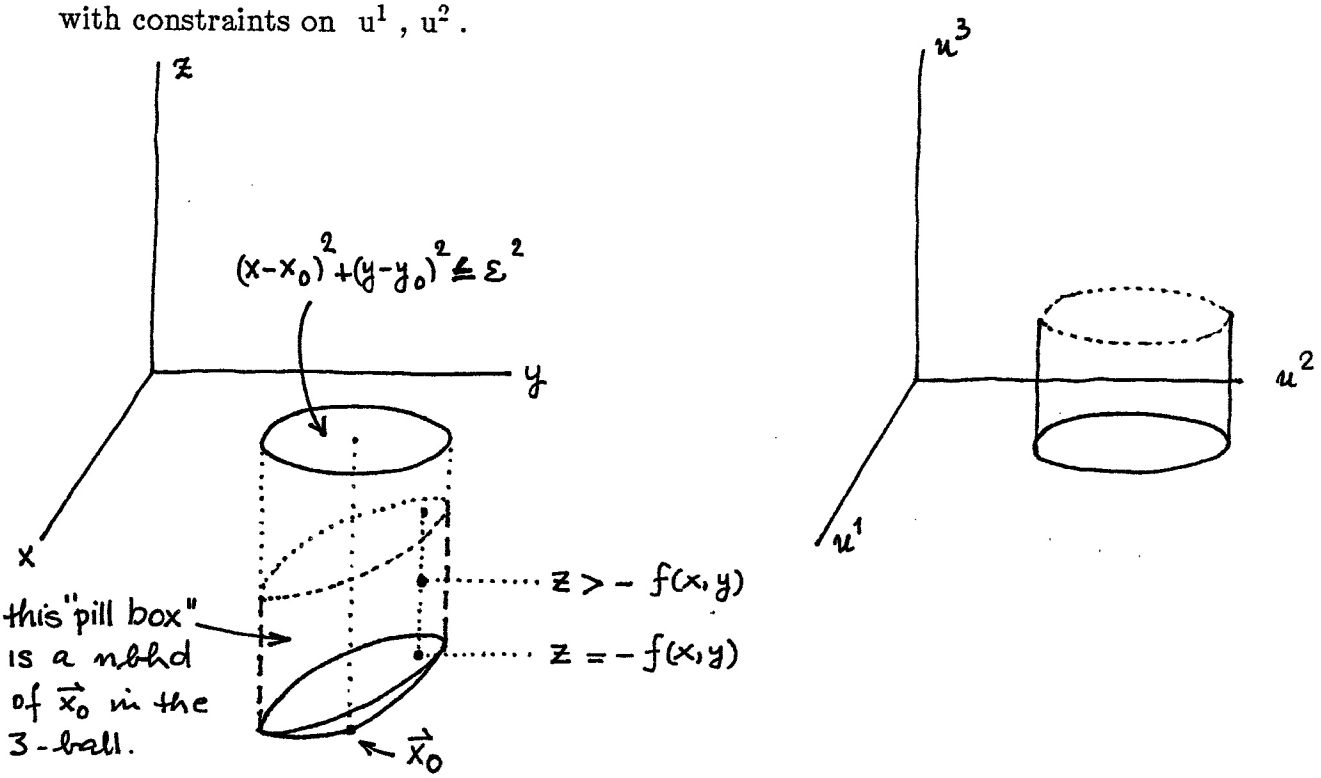
Introduce new coordinates in $\mathbb{R}^3$ in a neighborhood of this point by

$$\begin{aligned} u^1 &= x \\ u^2 &= y \\ u^3 &= z + f(x,y) \end{aligned}$$

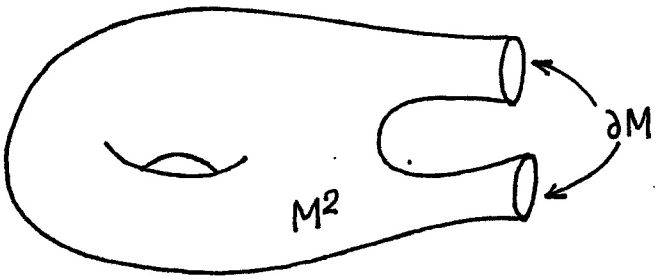
These do form a legitimate local coordinate system because (see Lecture 3, §(3))

$$\frac{\partial(u^1,u^2,u^3)}{\partial(x,y,z)} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ f_x & f_y & 1 \end{vmatrix} = 1 \neq 0 \quad .$$

We then note that a "pill box" neighborhood of this point $\vec{x}_0$ in the closed ball is described in these new coordinates by $z \geq -f(x,y)$, i.e., $u^3 \geq 0$ together with constraints on u^1, u^2.



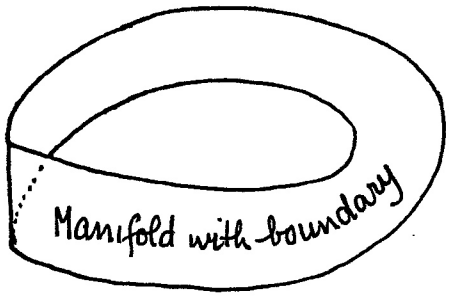
Very briefly speaking, an n-manifold with boundary M^n has an "interior" which is a genuine n-manifold, together with a "boundary" or "edge". Points on the boundary do not have neighborhoods diffeomorphic to open sets in $\mathbb{R}^n$ but rather "half open" sets in $\mathbb{R}^n$, described say by $(u^1)^2 + \cdots + (u^{n-1})^2 < \epsilon$ and $0 \leq u^n < \delta$. For more details, see e.g. G-P, p. 57ff or A-M-R, p.406ff. One should note that the boundary, written ∂M turns out always to be a true (n–1) manifold (without boundary) but it need not be connected (i.e., it may consist of more than one piece).

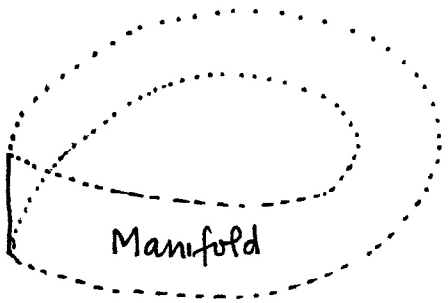


A 2-manifold, i.e., surface, with boundary consisting of two "circles".

Note also that if the boundary is empty, $\partial M = \phi$, then M is simply an n-manifold.

Orientability, one-sided-ness etc., are exactly as in the case of a manifold. An actual Möbius band constructed from a strip of paper is a surface with boundary; if the edge is omitted it is a surface.





(2) **Partition of Unity.** To integrate a p-form α^p over a manifold V^p (with or without boundary), instead of breaking up V^p into "non-overlapping" parameterized subsets we will break up the form α into a sum of forms.

We shall almost entirely deal with manifolds V^p that are covered by a finite number of coordinate patches (we have to use "half" coordinate patches for boundary points; we still call them coordinate patches). What we say can be generalized to the general case (for details, see G-P, pp. 52ff or A-M-R, pp. 309ff).

Given a finite coordinate covering $\{u_\alpha, x_\alpha\}$. $\alpha = 1,2,\ldots,N$, a partition of unity will exhibit N real valued functions $f_\alpha : V \to \mathbb{R}$ that are differentiable and have the following properties. First define the support of a function, form, tensor, etc. to be

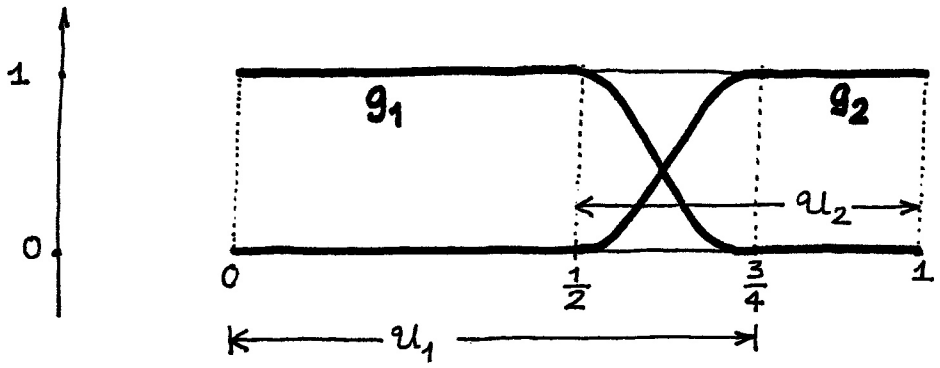
the closure of the set of points at which the function , . . . , is not zero. Thus support of f consists of all points of V^p where $f \neq 0$ together with all limit points of this subset. (The problem here is that the subset where $f \neq 0$ is automatically an open subset). We demand

(i) $f_\alpha(x) \geq 0 \qquad \alpha = 1, \ldots, N$.

(ii) The support of f_α is a closed subset $\subset u_\alpha$, in particular $f_\alpha(x) = 0$ for $x \notin u_\alpha$.

(iii) $\sum_{\alpha=1}^{N} f_\alpha(x) = 1$, all $x \in V^p$.

Such partitions of unity always exist; they are constructed in the general case in the above references. We shall illustrate the construction in the simplest possible case. Suppose that $V^p = V^1 = [0,1]$ is the unit interval on $\mathbb{R}^1$. This is a 1-manifold with boundary (the boundary consists of two points, i.e., two 0-manifolds). We suppose that [0,1] comes equipped with two coordinate patches, $u_1 : \{x \mid 0 \leq x < \frac{3}{4}\}$ and $u_2 : \{x \mid \frac{1}{2} < x \leq 1\}$ which of course must overlap. We first construct two differentiable functions g_1, g_2 whose supports are in u_1, u_2, respectively. We simply draw their graphs (the analytical constructions of g_1, g_2 would be in the references).



We have drawn g_1 and g_2 so that their supports are in u_1 and u_2 respectively, and yet these functions never vanish simultaneously. We now define (since $g_1(x) + g_2(x) > 0$ everywhere)

$$f_1(x) = \frac{g_1(x)}{g_1(x)+g_2(x)} \qquad f_2(x) = \frac{g_2(x)}{g_1(x)+g_2(x)}$$

yielding the required partition of unity, $1 = \sum_{\alpha=1}^{2} f_\alpha(x)$.

Point set topology allows one to go through an analogous construction in higher dimensions with far more complicated coordinate patches.

(3) **Integration of forms over compact oriented submanifolds of M^n.** Let M^n be a manifold, perhaps with boundary, and let V^p be an oriented submanifold (perhaps with boundary) of M^n. We of course may think of V^p as a manifold in its own right; we then have the inclusion map

$$i : V^p \longrightarrow M^n$$

where $i(x)$ says think of the point $x \in V^p$ as a point in M^n. If $u^1, \ldots, u^p$ are local coordinates for V^p and $x^1, \ldots, x^n$ those for M^n then $i : V^p \to M^n$ is simply given by $x^i = x^i(u^1, \ldots, u^p)$, $i = 1, \ldots, n$.

Now let ω^p be a p-form on M^n. We may restrict this form to V^p by taking the pull-back $i^*\omega^p$. Note that this restriction is a very simple minded operation; the value of $i^*\omega^p$ on a p- tuple of vectors tangent to V^p is exactly the same as the value of ω^p on these vectors.

We now wish to define $\int_{V^p} \omega^p$,. For this purpose we shall assume that V^p is compact. This notion can be explained in two equivalent ways in our situation.

(a) V^p is compact provided from any covering of V^p by open sets one can extract a finite subcovering.

(b) V^p is compact provided any infinite sequence of points of V^p has a "limit point" in V^p.

If V^p is a submanifold of $\mathbb{R}^n$ then V^p is compact provided it is a bounded and closed subset of $\mathbb{R}^n$.

Now let $\{\mathcal{U}_\alpha\}$ be a finite covering of V^p by coordinate charts that are positively oriented for the orientation of V^p given. Let $\{f_\alpha\}$ be a partition of unity for this coordinate cover. We then define

where

$$\boxed{\int_{V^p} \omega^p \equiv \int_{V^p} i^*\omega^p \equiv \sum_\alpha \int_{\mathcal{U}_\alpha} \omega_\alpha \qquad \omega_\alpha = f_\alpha \cdot i^*\omega} \tag{14.1}$$

i.e., we construct the forms $\omega_\alpha = f_\alpha\, i^*\omega$ by simply multiplying by the appropriate partition function f_α. Since $\mathcal{U}_\alpha$ can be considered an oriented parameterized subset of V^p, the individual integrals

$$\int_{U_\alpha} \omega_\alpha$$

are already defined (see (13.3)). The crucial fact that makes all this work of course is

$$\sum_\alpha \omega_\alpha = \sum_\alpha f_\alpha \, i^*\omega = i^*\omega \sum_\alpha f_\alpha = i^*\omega \quad .$$

It is not difficult to check that the resulting value for $\int_{V^p} \omega^p$ is in fact independent of the cover $\{U_\alpha\}$ and the partition $\{f_\alpha\}$ subordinate to this cover. See the references.

(4) **Manifolds admit Riemannian metrics.** We give now an immediate application of partitions of unity.

If M^n is a submanifold of $\mathbb{R}^n$, some N, then $i : M^n \to \mathbb{R}^N$ is the inclusion and if we let ds^2 be the usual Riemannian metric for $\mathbb{R}^N$, $ds^2 = \sum_{k=1}^{N} (dy^k)^2$, then the pull-back or <u>restriction</u> i^*ds^2 will yield a Riemannian metric for M^n. This is the usual method for getting a Riemannian metric on a surface $M^n \subset \mathbb{R}^3$; if the surface is given by $y^3 = f(y^1,y^2)$ then

$$\begin{aligned} i^*((dy^1)^2 + (dy^2)^2 + (dy^3)^2) &= (dy^1)^2 + (dy^2)^2 + \{f_{y^1}dy^1 + f_{y^2}dy^2\}^2 \\ &= [1 + (f_{y_1})^2](dy^1)^2 + 2\, f_{y_1}f_{y_2}dy^1dy^2 + [1 + (f_{y_2})^2](dy^2)^2 \quad . \end{aligned}$$

How do we get a metric if M^n is not given as a submanifold of $\mathbb{R}^N$?

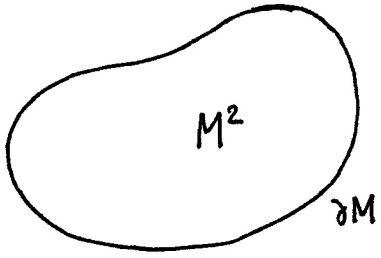
Let $\{U_\alpha, x_\alpha\}$ be a coordinate cover for M^n. In U_α we have a metric $ds_\alpha^2 = \sum_{i=1}^{n} (dx_\alpha^i)^2$ but of course in $U_\alpha \cap U_\beta$, $ds_\alpha^2 \neq ds_\beta^2$, so these don't fit together to give a metric for M^n. How about putting $ds^2 = \sum ds_\alpha^2$? This won't work because ds_α^2 is only defined in U_α. It would give some highly discontinuous metric on M^n. Let $\{f_\alpha\}$ be a partition of unity. Then $f_\alpha ds_\alpha$ can be defined to be differentiable everywhere because $f_\alpha = 0$ outside of U_α; put $f_\alpha ds_\alpha^2 = 0$ outside of U_α. Now define a Riemannian metric on all of M^n by

$$ds^2 = \sum_\alpha f_\alpha ds_\alpha^2 \quad . \tag{14.2}$$

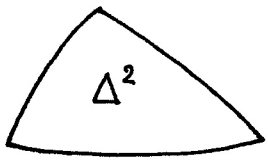
This works, although it <u>is</u> arbitrary!

Lecture 15. Stokes theorem for True Forms

(1) **Orienting the boundary of an oriented manifold.** Let M^n be an oriented manifold with non-empty boundary ∂M^n ; ∂M is an (n–1)-dimensional manifold and we state again, without proof, that ∂M does not have a boundary (we shall discuss a "homological" version of this at a later time). Since the changes of coordinates in overlapping patches are differentiable the boundary really <u>is</u> a differentiable manifold. For example the triangle is <u>not</u> a 2-manifold; its boundary is only piecewise differentiable

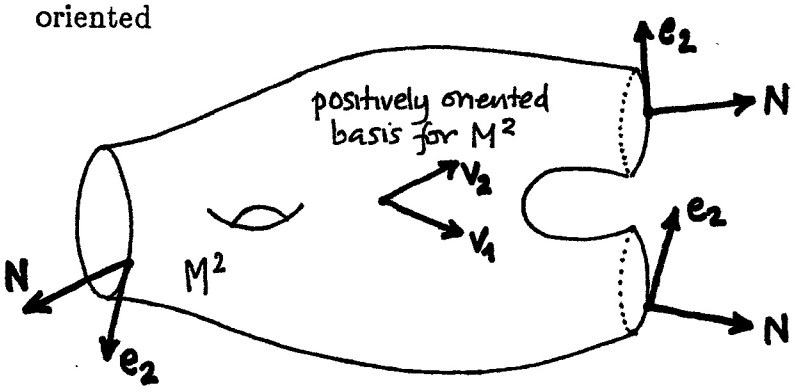


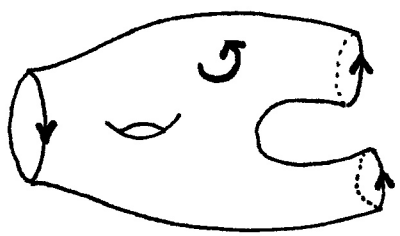
2-manifold with boundary



<u>not</u> a manifold with boundary

Given the orientation on M^n we can orient ∂M as follows. Let $e_2, \ldots, e_n$ span the <u>tangent space</u> to ∂M at x. Let N be a tangent vector to M^n at x that is transverse to ∂M to x and points <u>out</u> of M. Then we say $e_2 \cdots e_n$ is positively oriented





for ∂M at x provided $N, e_2, \ldots, e_n$ is positively oriented with respect to the given orientation for M^n . In the above figure we have indicated the positive orientation for M^2 by v_1, v_2 ; then the indicated e_2 is positively oriented on ∂M^2 . In the right-hand figure we indicate the orientations of the boundary curves by simply giving an arrow (this works only in this dimension).

(2) **Stokes' Theorem.** Let $V^p \subset M^n$ be a compact oriented submanifold with boundary ∂V in a manifold M^n. Let ω^{p-1} be a continuously differentiable $p-1$ form on M^n. Then

$$\boxed{\int_{V^p} d\omega^{p-1} = \int_{\partial V} \omega^{p-1}} \tag{15.1}$$

Proof. Let $i : V^p \to M^n$ be the inclusion map. Then since, (see 10.3),

$$\int_{V^p} d\omega^{p-1} = \int_{V^p} i^* d\omega^{p-1} = \int_{V^p} d\, i^* \omega^{p-1} \quad ,$$

and

$$\int_{\partial V} \omega^{p-1} = \int_{\partial V} i^* \omega^{p-1} \quad .$$

We see that to prove (15.1) we may forget M^n altogether and merely prove (15.1) for the form $i^*\omega^{p-1}$ instead. In other words we need only prove

$$\int_{V^p} d\beta^{p-1} = \int_{\partial V} \beta^{p-1}$$

for any continuously differentiable form β^{p-1} on V^p.

Introduce a partition of unity $1 = \sum f_\alpha$ for a finite coordinate cover $\{U_\alpha\}$ of V^p. Then writing $\beta = \sum \beta_\alpha$, $\beta_\alpha = f_\alpha \beta$

$$\int_{V^p} d\beta^{p-1} = \int_{V^p} d \sum_\alpha \beta_\alpha = \sum_\alpha \int_{U_\alpha} d\beta_\alpha^{p-1}$$

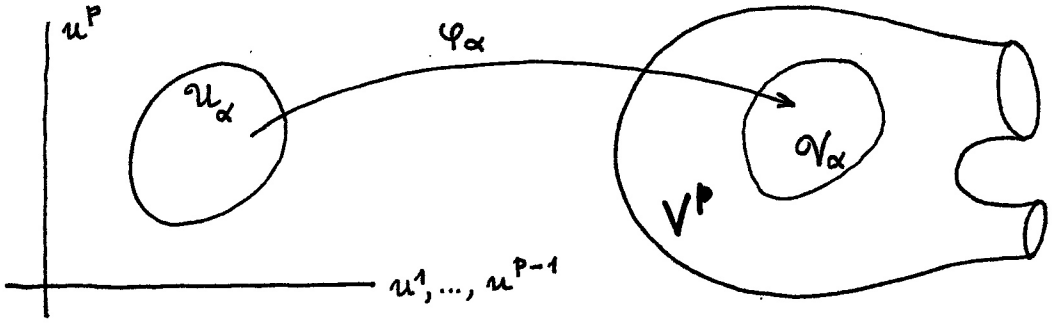
$$\int_{\partial V} \beta^{p-1} = \sum_\alpha \int_{\partial V} \beta_\alpha^{p-1} \quad .$$

We see that we need only prove

$$\int_{U_\alpha} d\beta_\alpha^{p-1} = \int_{\partial V} \beta_\alpha^{p-1} \tag{15.2}$$

for the form β_α whose support lies in U_α. There are two cases to consider.

Case (i). U_α is a full coordinate patch lying in the interior of V^p (see Lecture 14). Then, when everything is expressed in terms of this coordinate patch



$$\int_{U_\alpha = \phi_\alpha \mathcal{U}_\alpha} d\beta_\alpha = \int_{\mathcal{U}_\alpha} \phi_\alpha^* \, d\beta_\alpha = \int_{\mathcal{U}_\alpha} d(\phi_\alpha^* \beta_\alpha) \quad .$$

Now the support of $\phi_\alpha^* \beta_\alpha$ lies entirely in $\mathcal{U}_\alpha$; thus since $\phi_\alpha^* \beta_\alpha = 0$ outside of $\mathcal{U}_\alpha$ we may replace the domain of integration $\mathcal{U}_\alpha$ in this ordinary multiple integral by all of $\mathbb{R}^p$. If we call $\phi_\alpha^* \beta_\alpha = \gamma^{p-1}$

$$\phi_\alpha^* \beta_\alpha = \gamma^{p-1} = \sum (-1)^{i-1} \gamma_i \, du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^p \tag{15.3}$$

where $\widehat{a}$ signifies that the term a is to be omitted, then

$$\begin{aligned} \int_{\mathcal{U}_\alpha} d(\phi_\alpha^* \, \beta_\alpha) = \int_{\mathcal{U}_\alpha} d\gamma^{p-1} &= \sum_i \int_{\mathcal{U}_\alpha} (-1)^{i-1} \, d(\gamma_i \, du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^p) \\ &= \sum_i (-1)^{i-1} \int_{\mathcal{U}_\alpha} \frac{\partial \gamma_i}{\partial u^i} \, du^i \wedge du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^p \\ &= \sum_i \int_{\mathcal{U}_\alpha} \frac{\partial \gamma_i}{\partial u^i} \, du^1 \wedge \cdots \wedge du^p \quad . \end{aligned} \tag{15.4}$$

We may assume that the coordinates patch $\mathcal{U}_\alpha$ is positively oriented with respect to the given orientation of V^p. Then

$$\int_{\mathcal{U}_\alpha} d(\phi_\alpha^* \beta_\alpha) = \sum_i \int_{\mathcal{U}_\alpha} \frac{\partial \gamma_i}{\partial u^i} \, du^1 \cdots du^p = \sum_i \int_{\mathbb{R}^p} \frac{\partial \gamma_i}{\partial u^i} \, du^1 \cdots du^p = 0 \quad .$$

since

$$\int_{\mathbb{R}^p} \frac{\partial \gamma_i}{\partial u^i}\, du^1 \cdots du^p = \int_{\mathbb{R}^{p-1}} du^1 \cdots \widehat{du^i} \cdots du^p \int_{-\infty}^{\infty} \frac{\partial \gamma_i}{\partial u^i}\, du^i$$

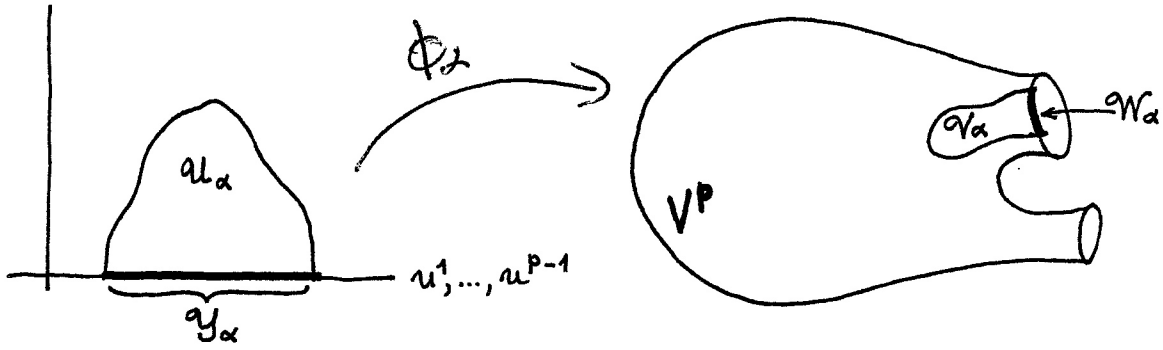
$$= \int_{p-1} du^1 \cdots \widehat{du^i} \cdots du^p (\gamma_i \Big|_{-\infty}^{\infty})$$

and γ_i vanishes outside of U_α. Of course we also have

$$\int_{\partial V} \beta_\alpha = 0 \quad .$$

Since in this case β_α vanishes outside U_α, and so $\beta_\alpha = 0$ on ∂V. Thus case (i) is finished, both sides of (15.2) vanishing.

Case (ii). U_α is a coordinate patch that meets the boundary.



We proceed exactly as in the first case, reaching (15.4).

$$\int_{U_\alpha} d\beta_\alpha = \int_{U_\alpha} d\gamma^{p-1} = \sum_i \int_{U_\alpha} \frac{\partial \gamma_i}{\partial u^i}\, du^1 \wedge \cdots \wedge du^p \quad .$$

Now the only non-vanishing term here is the one with $i = p$ since the other terms will involve $\int_{-\infty}^{\infty} \frac{\partial \gamma_i}{\partial u^i}\, du^i$ which vanishes if $i < p$. On the other hand, for $i = p$ we get

$$\int_{U_\alpha} d\beta_\alpha = \int_{U_\alpha} \frac{\partial \gamma_p}{\partial u^p}\, du^1 \wedge \cdots \wedge du^p = \int_{\mathbb{R}^p} du^1 \cdots du^{p-1} \int_0^{\infty} \frac{\partial \gamma_p}{\partial u^p}\, du^p$$

$$= \int_{\mathbb{R}^{p-1}} du^1 \cdots du^{p-1} (\gamma_p(\infty) - \gamma_p(0)) \qquad (15.5)$$

$$= -\int_{\mathbb{R}^{p-1}} \gamma_p(u^1, \ldots, u^{p-1}, 0) du^1 \cdots du^{p-1} \quad .$$

Now if we restrict $\phi_\alpha : \mathcal{U}_\alpha \to V$ to the subset of $\mathcal{U}_\alpha$ where $u^p = 0$ we get a (p–1)-dimensional coordinate patch for ∂V, call it $\mathcal{W}_\alpha = \phi_\alpha(\mathcal{U}_\alpha \cap \{u^p = 0\}) \equiv \phi_\alpha(\mathcal{Y}_\alpha)$; see the figure above. Then we have, since support of β_α lies in $\mathcal{W}_\alpha$,

$$\int_{\partial V} \beta_\alpha^{p-1} = \int_{\mathcal{W}_\alpha} \beta_\alpha^{p-1} = \int_{\phi_\alpha(\mathcal{Y}_\alpha)} \beta_\alpha^{p-1} = \int_{\mathcal{Y}_\alpha} \phi_\alpha^* \beta_\alpha^{p-1}$$

$$= \int_{\mathcal{Y}_\alpha} \gamma^{p-1} = \int_{\mathcal{Y}_\alpha} \sum (-1)^{i-1} \gamma_i(u^1 \cdots u^p) du^1 \wedge \cdots \wedge \widehat{du^i} \wedge \cdots \wedge du^p \quad .$$

But on $\mathcal{Y}_\alpha$ $u^p = 0$ and so $du^p = 0$ and the only surviving term is

$$\int_{\partial V} \beta_\alpha^{p-1} = (-1)^{p-1} \int_{\mathcal{Y}_\alpha} \gamma_p(u^1, \ldots, u^{p-1}, 0) du^1 \wedge \cdots \wedge du^{p-1} \quad .$$

Now $\dfrac{\partial}{\partial u^1}, \ldots, \dfrac{\partial}{\partial u^p}$ is a positively oriented coordinate system for V^p and $-\dfrac{\partial}{\partial u^p}$ is the outward pointing transversal to ∂V, consequently (see §1) $\dfrac{\partial}{\partial u^1}, \ldots, \dfrac{\partial}{\partial u^{p-1}}$ carries orientation $(-1)^p$ on $\mathcal{Y}_\alpha$ (there is one minus sign for $-\dfrac{\partial}{\partial u^p}$, and (p–1) minus signs to get $\dfrac{\partial}{\partial u^p}$ in the last position). Consequently

$$\int_{\partial V} \beta_\alpha^{p-1} = (-1)^{p-1} (-1)^p \int_{\mathcal{Y}_\alpha} \gamma_p(u^1, \ldots, u^{p-1}, 0) du^1 \cdots du^{p-1}$$

$$= -\int_{\mathbb{R}^{p-1}} \gamma_p(u^1, \ldots, u^{p-1}, 0) du^1 \cdots du^{p-1} \quad .$$

Since this coincides with (15.5) case (ii) is proven. □

Lecture 16. Integration of Pseudo-forms. Maxwell's Equations

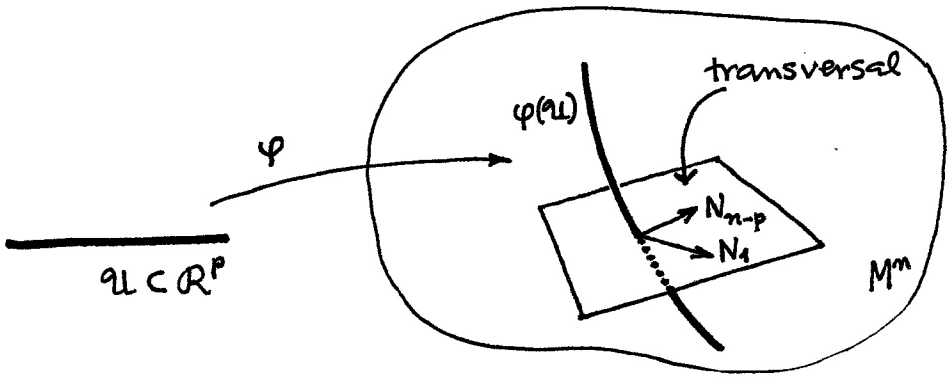
(1) **Integration of pseudo-forms.** Integration of p-forms α^p over parameterized oriented subsets of any manifold M^n was defined by pulling the form back to $\mathbb{R}^p$

$$\int_{(\phi,\mathcal{U}_o)} \alpha^p \equiv \int_{\mathcal{U}_o} \phi^*\alpha^p \equiv o(u) \int\!\!\cdots\!\!\int_{|\mathcal{U}|} (\phi^*\alpha^p)\Big(\frac{\partial}{\partial u^1}, \ldots, \frac{\partial}{\partial u^p}\Big)\, du^1 \cdots du^p$$

(see (13.3), (13.4)). This has nothing to do with orientation of M^n. Suppose we try to do the same for a <u>pseudo</u> - <u>form</u> ω^p. Suppose either that M^n is non-orientable or else that M^n is orientable but we haven't picked one of the two orientations for M^n. Then the <u>pull</u> - <u>back</u> $\phi^*\omega^p$ <u>makes no sense</u>; at each point $\phi(u) \in M^n$ there are two choices for ω^p and we don't know which to choose to pull-back! An orientation o of $\mathcal{U} \subset \mathbb{R}^p$ does <u>not</u> (ordinarily) yield an orientation of M^n near $\phi(\mathcal{U})$.

To pull-back a pseudo-form ω via a map $\phi : \mathcal{U} \to M$ we need to equip the map ϕ with some extra structure so that an orientation o of $\mathcal{U}$ will give some sort of M^n orientation in some neighborhood of $\phi(\mathcal{U})$ (de Rham calls such maps "oriented maps").

We shall <u>not</u> consider the most general situation, but rather concentrate on the most practical occurrence. Let $\phi : \mathcal{U} \to M^n$ be a parameterized subset, $\mathcal{U} \subset \mathbb{R}^p$ and assume that ϕ_* has rank p everywhere on $\mathcal{U}$ (thus $\phi_* : \mathcal{U}_x \to M^n_{\phi(x)}$ is 1:1)



Suppose that one can find (n-p) continuous linearly independent vector fields defined along $\phi(\mathcal{U})$ and transverse (e.g. normal) to $\phi(\mathcal{U})$. If we can do this we shall say that the parameterized $\phi : \mathcal{U} \to M^n$ is <u>transversally oriented</u>. Given such a map we <u>can</u> pull-back a pseudo-form ω^p on M^n to $\mathcal{U}$, the answer being a pseudo-form $\phi^*\omega$ on $\mathcal{U}$, as follows.

Let $e_1, \ldots, e_p$ define an orientation of $\mathcal{U}_x$ at $x \in \mathcal{U}$. Then we define an orientation o_M of $M^n_{\phi(x)}$ by declaring

$$N_1, \ldots, N_{n-p}, \phi_* e_1, \ldots, \phi_* e_p \tag{16.1}$$

to be positively oriented. With this orientation we have a specific choice of $\omega^p_{o_M}$. Then $\phi^*\omega \equiv \phi^*\omega^p_{o_M}$. Note that if we reverse the orientation of $\mathcal{U}$, say by interchanging e_1 and e_2, then automatically $\phi^*\omega$ would be sent into its negative. Thus if ω is a pseudo-form on M^n then $\phi^*\omega$ is indeed a pseudo-form on $\mathcal{U}$. Thus, in contrast with true forms we deal with maps $\phi : |\mathcal{U}| \to M^n$ of <u>un</u> -oriented $\mathcal{U}$ but require that $\phi(\mathcal{U})$ be <u>transversally</u> oriented.

For such a transversally oriented map and pseudo-form ω^p on M^n we define the integral as follows:

$$\int_{(\phi,|\mathcal{U}|)} \omega^p \equiv \int_{\mathcal{U}_o} \phi^*\omega_{o_M} \tag{16.2}$$

and this last integral means the following. Choose <u>any</u> orientation o of $|\mathcal{U}|$ that you wish. Then a specific o_M is determined and

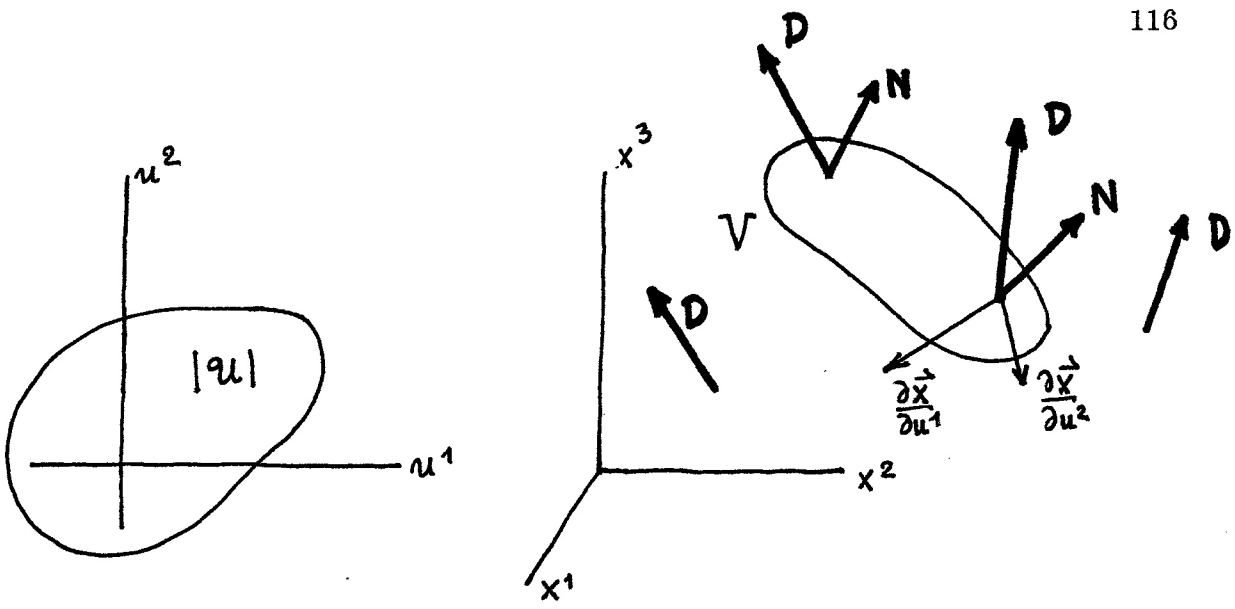
$$\int_{\mathcal{U}_o} \phi^*\omega_{o_M}$$

is an integral of a specific form $\phi^*\omega_{o_M}$ over an oriented $\mathcal{U}_o$ and we know what this integral is. Note that if we had chosen the opposite orientation $-o$, then we would get $-o_M$ and

$$\int_{\mathcal{U}_{-o}} \phi^*\omega_{-o_M} = \int_{\mathcal{U}_o} \phi^*\omega_{o_M} \quad ,$$

since we would be integrating the opposite form over a region with the opposite orientation.

For example, let $\phi : |\mathcal{U}| \subset \mathbb{R}^2 \to \mathbb{R}^3$ be an unoriented parameterized surface in $\mathbb{R}^3$ with a <u>given</u> <u>transverse</u> <u>orientation</u> ; e.g. we may be given a unit normal vector field $\vec{N}$ to $\phi(|\mathcal{U}|) = V$.



Suppose $\vec{D}$ is a (true) continuous vector field in $\mathbb{R}^3$. We wish to compute

$$\iint_{(V,\vec{N})} \mathfrak{D}^2 = \iint_{(V,\vec{N})} i_{\vec{D}}\, \mathrm{vol}^3_o$$

where again o denotes orientation for $\mathbb{R}^3$.

Choose <u>arbitrarily</u> an orientation for $|u|$, say, the one for which $\frac{\partial}{\partial u^1}, \frac{\partial}{\partial u^2}$ is positively oriented. Assume again that ϕ_* is 1:1 (i.e., V is a <u>regularly</u> parameterized surface). Then $\vec{N}, \frac{\partial \vec{x}}{\partial u^1}, \frac{\partial \vec{x}}{\partial u^2}$ picks out a specific orientation for $\mathbb{R}^3$. As in the computation for (13.11)

$$\begin{aligned} \iint_{(V,\vec{N})} \mathfrak{D}^2 &= \iint_{u} \langle \vec{D},\vec{N} \rangle \, \mathrm{vol}^3\Big(\vec{N}, \frac{\partial \vec{x}}{\partial u^1}, \frac{\partial \vec{x}}{\partial u^2}\Big) du^1 du^2 \\ &= \iint_{V} \langle \vec{D},\vec{N} \rangle \, dS \quad . \end{aligned} \tag{16.3}$$

The <u>answer</u> is the same as in (13.11) but note that the <u>given</u> <u>data</u> were <u>different</u> ! In (13.11) one was given an <u>oriented</u> surface and <u>pseudo</u> -vector $\vec{B}$. One then chose <u>arbi</u> -<u>trarily</u> a unit normal $\vec{N}$ and proceeded to (13.11). In (16.3) one was given an <u>un</u> -oriented surface and a <u>given</u> <u>normal</u> $\vec{N}$, together with a <u>true</u> vector $\vec{D}$. One then chose <u>arbitrarily</u> an orientation for V and proceeded to (16.3).

One important remark. Suppose one is given an n-dimensional manifold M^n. If V^p is a p-dimensional, compact submanifold with boundary ∂V^p in M^n and if V^p is transversally oriented then one can again define, for each pseudo p-form ω^p

$$\int_{V^p \text{ with transverse orientation}} \omega^p$$

again by using a partition of unity; the construction is essentially the same as in Lecture 14 and you should convince yourself of this. In particular if M^n is itself a compact manifold with boundary then it is automatically transverse oriented (there isn't any transverse direction!). Consequently one can always integrate a pseudo -n form over any compact n- manifold M^n with boundary, whether or not it is orientable ! In particular, if M^n is a compact Riemannian manifold

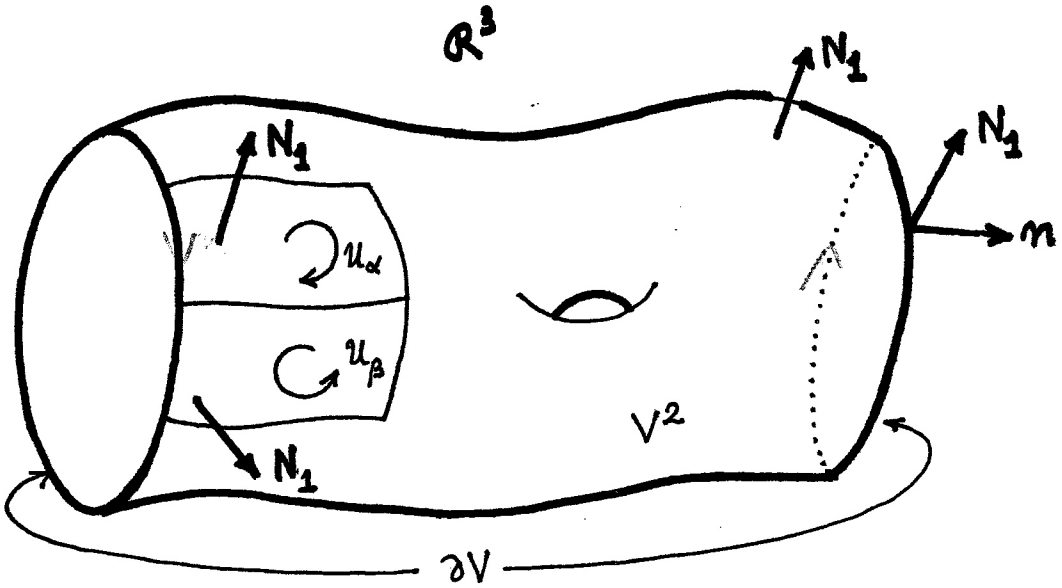
$$\text{the volume of } M^n \equiv \int_{M^n} \text{vol}^n$$

always makes sense !

(2) Stokes' theorem for pseudo-forms. Let $V^p \subset M^n$ be a transversally oriented compact submanifold-with-boundary of M^n. Let ω^{p-1} be a continuously differentiable pseudo (p-1) form on M^n. Then

$$\int_{V^p} d\omega^{p-1} = \int_{\partial V} \omega^{p-1} \quad . \tag{16.4}$$

The proof is very much the same as for (15.1); you should convince yourself of this. One should note two things. First, if at a point $x \in \partial V$, $\vec{N}_1, \ldots, \vec{N}_{n-p}$ give the transverse orientation for V then $\vec{N}_1, \ldots, \vec{N}_{n-p}, \vec{n}$ will give the transverse orientation for ∂V (here $\vec{n}$ is a tangent vector to V that points out of V). This describes the transverse orientation for ∂V in (16.4)!



Second, if you are used to proving Stokes' theorem by breaking V up into "non-overlapping" patches u_α, u_β,... you are familiar with cancellations in $\int \omega^{p-1}$ over boundaries common to adjacent patches. This still happens with pseudo forms in spite of the arbitrariness in picking orientations for u_α and u_β. In our figure we have given opposite orientations to u_α and u_β for the evaluations of $\int_{u_\alpha} d\omega$ and $\int_{u_\beta} d\omega$. It <u>looks</u> like the common boundary integrals $\int_{\partial u_\alpha} \omega$ and $\int_{\partial u_\beta} \omega$ would <u>not</u> cancel, but this is <u>not</u> true since the ω used in u_α would be the negative of the ω used in u_β! Saved!

(3) **Electromagnetism and Maxwells equations.** We shall work for the present in $\mathbb{R}^3$ but we shall <u>not</u> use an orientation for $\mathbb{R}^3$ (how could we communicate our notion of right handed or left handed to some strange creature in a distant part of the universe?)

We accept as basic the notion of "charge" of a test particle. If a small particle with charge e is moving in an electromagnetic field, the <u>electric</u> field vector $\vec{E}$ and the <u>magnetic</u> field vector $\vec{B}$ are defined as follows. The force $\vec{f}$ on the particle due to the fields, called the <u>Lorentz</u> force, is given by

$$\vec{f} = e(\vec{E} + \vec{v} \times \vec{B}) \tag{16.5}$$

where $\vec{v}$ is the velocity of the particle. Thus by knowing the charge e and by measuring $\vec{f}$ when the particle is at rest, $\vec{v} = 0$, one gets $\vec{E}$. Then by quickly measuring the force on particles moving in three independent directions one can determine $\vec{B}$ (since $\vec{E}$ is already known). Thus (16.5) really just serves to define the vector fields $\vec{E}$ and $\vec{B}$. It is remarkable that the correct expression for the magnetic force, $e\vec{v} \times \vec{B}$ was not known 'till Heaviside ($\sim$ 1890).

We shall consider (16.5) as a contravariant equation; $\vec{f}$ is a contravariant vector and so $e(\vec{E} + \vec{v} \times \vec{B})$ is a contravariant vector. The charge e is a scalar, hence $\vec{E} + \vec{v} \times \vec{B}$ is a vector. Then $\vec{E}$ must be a vector and $\vec{v} \times \vec{B}$ is a vector. But velocity $\vec{v}$ is certainly a vector, hence $\vec{B}$ must be a <u>pseudo</u> vector. $\vec{B}$ must change into its negative if the orientation of space is reversed. We would like to get rid of this orientation dependence of $\vec{B}$. We know that we can associate to this pseudo-vector $\vec{B}$ a true 2-form (see (12.4))

$$\mathcal{B}^2 = i_{\vec{B}} \, \mathrm{vol}^3 \quad . \tag{16.6}$$

Let us use Cartesian coordinates. We started out with

$$\vec{E} = E^1 \frac{\partial}{\partial x} + E^2 \frac{\partial}{\partial y} + E^3 \frac{\partial}{\partial z}$$

$$\vec{B} = B^1 \frac{\partial}{\partial x} + B^2 \frac{\partial}{\partial y} + B^3 \frac{\partial}{\partial z} \quad .$$

Then

$$\mathcal{B}^2 = B^1 \, dy \wedge dz + B^2 \, dz \wedge dx + B^3 \, dx \wedge dy \quad . \tag{16.7}$$

Note then that we can write the <u>covariant</u> version of the Lorentz force law (16.5) as (see (12.8))

$$f^1 = e(\mathcal{E} - i_{\vec{v}} \mathcal{B}^2) \tag{16.8}$$

where $\mathcal{E}$ is the electric field covector

$$\mathcal{E} = E_1 \, dx + E_2 \, dy + E_3 \, dz \quad . \tag{16.9}$$

Note that (16.8) is completely <u>orientation independent</u> !

We now state the laws satisfied by $\vec{E}$ and $\vec{B}$, the Maxwell equations. The fields $\vec{E}$ and $\vec{B}$ are in general time dependent $E_i = E_i(x,y,z,t)$. We define the time derivative of a form by just differentiating the coefficients. For example

$$\frac{\partial \mathfrak{B}^2}{\partial t} = \frac{\partial B^1}{\partial t}\, dy \wedge dz + \frac{\partial B^2}{\partial t}\, dz \wedge dx + \frac{\partial B^3}{\partial t}\, dx \wedge dy \quad .$$

We shall call a compact manifold M^3 with boundary ∂M in $\mathbb{R}^3$ a compact region with boundary. Likewise we have a compact surface M^2 with boundary (∂M^2 is a collection of closed curves).

Law of no magnetic monopoles. Let $M^3 \subset \mathbb{R}^3$ be a compact oriented domain with boundary ∂M^3. Then

$$\iint_{\partial M^3} \mathfrak{B}^2 = 0 \quad . \tag{16.10}$$

Faraday's Law. Let M^2 be a compact oriented surface with boundary ∂M^2 (with its induced orientation, see §2). Then

$$\int_{\partial M^2} \mathcal{E}^1 = -\iint_{M^2} \frac{\partial \mathfrak{B}^2}{\partial t} \quad . \tag{16.11}$$

We shall now need to deal essentially with materials that are not affected by the electromagnetic fields ("non-inductive" material, no polarization, etc.). For these materials we shall consider the 2-form associated to $\vec{E}$ and the 1-form associated with $\vec{B}$, and traditionally we denote them by different letters. First we have the electric pseudo 2-form, commonly called "displacement"

$$\mathfrak{D}^2 = i_{\vec{E}}\, vol^3 = E^1\, dy \wedge dz + E^2\, dz \wedge dx + E^3\, dx \wedge dy \quad . \tag{16.12}$$

Next we have the magnetic pseudo 1-form; this is simply the covector associated with the pseudo vector $\vec{B}$

$$\mathfrak{H} = B_1\, dx + B_2\, dy + B_3\, dz \quad . \tag{16.13}$$

We also have two forms associated with charge. If $\rho(x,y,z,t)$ is the charge density (it is a scalar), then

$$\rho\, vol_o^3 = \rho\, dx \wedge dy \wedge dz \tag{16.14}$$

is the charge pseudo 3-form (we omit the orientation symbol).

If we have charged particles (a "fluid") with charge density ρ and velocity $\vec{v}$

$$\mathcal{J}^2 = i_{\rho\vec{v}}\, vol^3 = \rho(v^1\, dy \wedge dz + v^2\, dz \wedge dx + v^3\, dx \wedge dy) \tag{16.15}$$

is the current pseudo 2-form.

Gauss' Law. If M^3 is a compact (un -oriented) region and if ∂M^3 is transversally oriented by using the outward pointing normal, then

$$\iint_{\partial M^3} \mathcal{D}^2 = 4\pi \iiint_{M^3} \rho\, vol^3 \tag{16.16}$$

measures the total charge in M^3 .

Ampere-Maxwell Law. If M^2 is a compact (un-oriented) surface with given transverse orientation, then

$$\int_{\partial M^2} \mathcal{H}^1 = 4\pi \iint_{M^2} \mathcal{J}^2 + \iint_{M^2} \frac{\partial \mathcal{D}}{\partial t} \quad . \tag{16.17}$$

(Ampere's original law omitted $\iint_{M^2} \frac{\partial \mathcal{D}}{\partial t}$; Maxwell added this "displacement current" term.).

(4) **Maxwell's equations in "differential form".** Maxwell's equations, as stated above hold, as far as I know, in very general circumstances. They make mathematical sense when the coefficients E^i , B^i of the forms are simply integrable functions! Let us state briefly the classical meaning of these laws, and also express them in the special case when these coefficients are continuously differentiable and hence amenable to Stokes' theorem.

(i) **No magnetic monopoles.** For every compact oriented region M^3

$$0 = \iint_{\partial M^3} \mathcal{B}^2 = \iiint_{M^3} d\mathcal{B}^2 = \iiint_{M^3} d(i_{\vec{B}}\, vol^3) \quad .$$

Since this is true for every such M^3 it must be that $di_B\, vol^3 = 0$. But (from (12.10)), $d(i_V\, vol^3) = \operatorname{div} \vec{B}\, vol^3$, and so

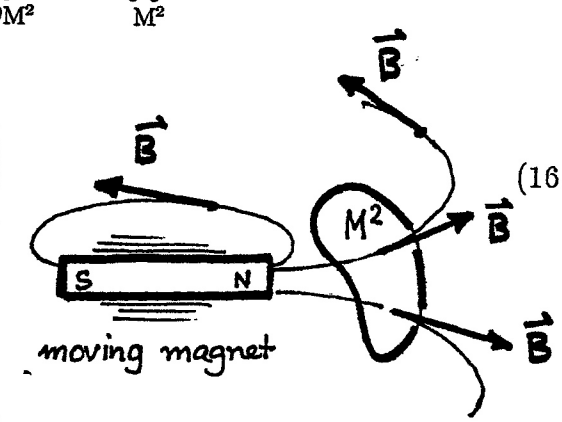
$$\boxed{d\mathcal{B}^2 = 0\ ,\ \operatorname{div} \vec{B} = 0} \qquad (16.10')$$

(ii) **Faradays Law.** For every compact oriented surface M^2

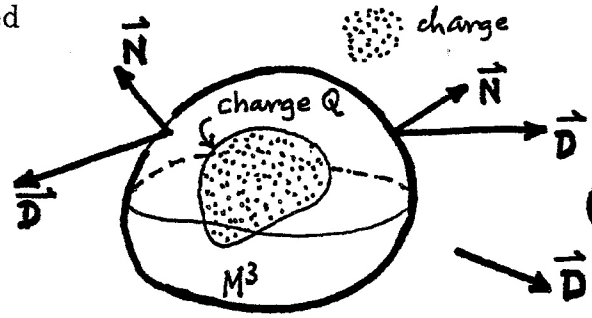
$$-\iint_{M^2} \frac{\partial \mathcal{B}^2}{\partial t} = \int_{\partial M^2} \mathcal{E}^1 = \iint_{M^2} d\mathcal{E}^1$$

and so

$$\boxed{d\mathcal{E}^1 = -\frac{\partial \mathcal{B}^2}{\partial t}\ ,\ \operatorname{curl} \vec{E} = -\frac{\partial \vec{B}}{\partial t}} \qquad (16.11')$$



says a changing magnetic field produces an electric field. The line integral $\int \mathcal{E}$ is called the "emf" or "potential drop," or "voltage."



(iii) **Gauss law**; for every compact M^3

$$Q = \text{charge inside } M^3 = \iiint_{M^3} \rho\ \text{vol}^3 = \frac{1}{4\pi} \iint_{\partial M^3} \mathcal{D}^2$$

$$= \frac{1}{4\pi} \iiint_{M^3} d\mathcal{D}^2 = \frac{1}{4\pi} \iiint \operatorname{div} \vec{D}\ \text{vol}^3$$

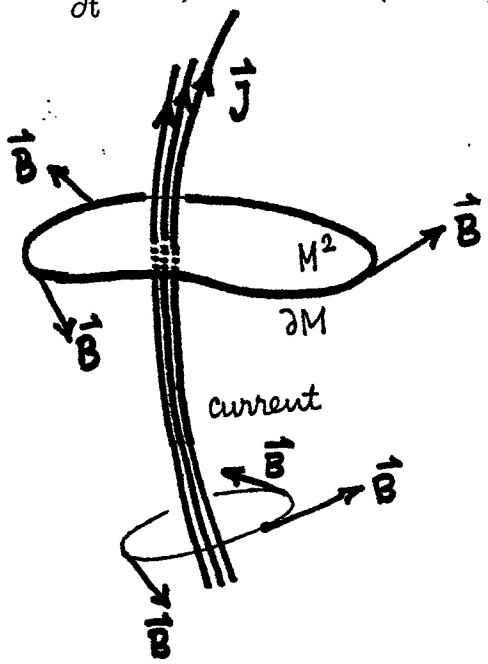
$$\boxed{d\mathcal{D}^2 = 4\pi\, \rho\ \text{vol}^3\ ,\ \operatorname{div} \vec{D} = 4\pi\rho} \qquad (16.16')$$

(iv) **Ampere's Law** (assume $\frac{\partial \mathcal{D}}{\partial t} = 0$). For compact M^2 with normal $\vec{N}$

$$4\pi \iint_{M^2} \mathcal{J}^2 = \int_{\partial M^2} \mathcal{H}^1 = \iint_{M^2} d\mathcal{H}^1$$

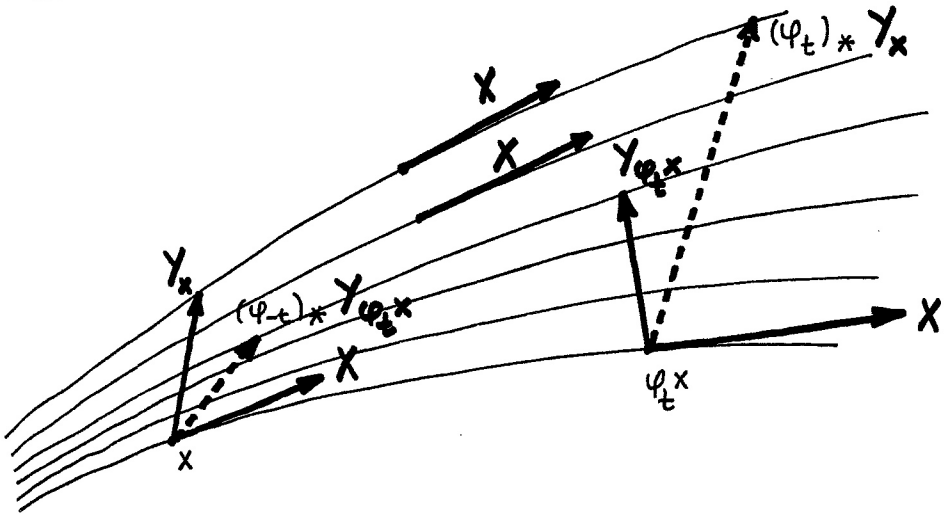
$$\boxed{d\mathcal{H}^1 = 4\pi j^2 \text{ , curl } \vec{H} = 4\pi \vec{j}} \qquad \text{(assumes } \frac{\partial \mathcal{D}}{\partial t} = 0\text{)} \qquad (16.17')$$

A current flowing through M^2 causes a magnetic field to "circulate about ∂M^2". The direction of $\vec{B}$ (i.e., $\pm$) depends on the orientation of $\mathcal{R}^3$ chosen.



Lecture 17. The Lie Derivative of a Vector Field

(1) The Lie bracket. Let X and Y be vector fields on M^n and let ϕ_t be the local flow generated by X (see Lecture 4). We shall compare the vector field $Y_{\phi(t)}$ (at points along an orbit $\phi_t(x)$ of X) with the result of "pushing" Y be means of the differential $(\phi_t)_*$.



The "Lie derivative" or "Lie bracket" of Y with respect to X is defined to be the vector field

$$(\mathfrak{L}_X Y)_x \equiv \lim_{t\to 0} \frac{Y_{\phi_t x} - (\phi_t)_* Y_x}{t} \tag{17.1}$$

$$= \lim_{t\to 0} (\phi_t)_* \frac{(\phi_{-t})_* Y_{\phi_t x} - Y_x}{t}$$

$$= \lim_{t\to 0} \frac{(\phi_{-t})_* Y_{\phi_t x} - Y_x}{t} \quad . \tag{17.2}$$

We must show that the limit actually exists. For this we shall invoke a very useful version of "Taylor expansion" essentially due to Hadamard.

Let f be a continuously differentiable function defined near x.

Lemma. There is a continuously differentiable function $g(t,x) = g_t(x)$ such that

$$g_0(x) = X_x(f) \qquad \text{and}$$

$$f \circ \phi_t = f + t\, g_t$$

i.e., $f(\phi_t x) = f(x) + t\, g(t,x)$.

Proof. Fix t and x. Define $F(t,x) = (f \circ \phi_t)(x)$ and $\hat{F}(s) = F(st,x)$. Then

$$(f \circ \phi_t)x - f(x) = F(t,x) - F(0,x)$$

$$= \hat{F}(1) - \hat{F}(0) = \int_0^1 \frac{d}{ds}\, \hat{F}(s)ds$$

$$= \int_0^1 \frac{d}{ds}\, F(st,x)ds = \int_0^1 F_{(1)}(st,x)t\, ds$$

where $F_{(1)}$ denotes partial derivative with respect to the first variable. Hence if we define

$$g_t(x) = g(t,x) = \int_0^1 F_{(1)}(st,x)ds$$

we see

$$(f \circ \phi_t)(x) - f(x) = t\, g_t(x) \quad .$$

Furthermore

$$g_0(x) = \int_0^1 F_{(1)}(0,x)ds = F_{(1)}(0,x)$$

$$= \lim_{t\to 0} \frac{F(t,x) - F(0,x)}{t}$$

$$= \lim_{t\to 0} \frac{(f \circ \phi_t)(x) - f(x)}{t} = X_x(f) \quad . \qquad \square$$

We now can show that $\mathfrak{L}_X Y$ exists. Let f be of class C^2 , then

$$(\mathfrak{L}_X Y)_x f = \lim_{t\to 0} \frac{1}{t}\, [Y_{\phi_t x} - (\phi_t)_* Y_x](f)$$

$$= \lim_{t\to 0} \frac{1}{t}\, [Y_{\phi_t x}(f) - d[(\phi_t)_* Y_x]$$

$$= \text{(see Lemma, Lecture 10)}$$

$$= \lim_{t\to 0} \frac{1}{t} [Y_{\phi_t x}(f) - Y_x(f \circ \phi_t)]$$

$$= \lim_{t\to 0} \frac{1}{t} [Y_{\phi_t x}(f) - Y_x\{f + t\, g_t\}]$$

$$= \lim_{t\to 0} \frac{1}{t} [Y_{\phi_t x}(f) - Y_x(f)] - \lim_{t\to 0} Y_x(g_t)$$

$$= X_x(Y(f)) - Y_x(\lim_{t\to 0} g_t)$$

$$= X_x(Y(f)) - Y_x(X(f)) \quad . \tag{17.3}$$

$\bar{f}(\phi_t x) = Y_{\phi_t x}(f)$ where $X_x(\bar{f}) = \left[\frac{d}{dt} \bar{f}(\phi_t x)\right]_{t=0}$ from lecture 4

$\lim_{t\to 0} \frac{1}{t}\left(Y_{\phi_t x}(f) - Y_x(f)\right) = \frac{d}{dt}\left[Y_{\phi_t}(f)\right]_{t=0} = X_x(Y_x(f))$

Thus we have shown in fact that if one defines the commutator of the two differential operators X and Y by $[X,Y]_p f = X_p(Yf) - Y_p(Xf)$ then

$$\boxed{\mathfrak{L}_X Y = [X,Y] \quad .} \tag{17.4}$$

• **Problem 28.** By expressing $X_x(Y(f)) - Y_x(X(f))$ in local coordinates show that

$$(\mathfrak{L}_X Y)^i = [X,Y]^i = \sum_j X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \quad . \tag{17.5}$$

Finally, note that definition (17.2) can be written in the form

$$\boxed{\mathfrak{L}_X Y_x = \left(\frac{d}{dt} (\phi_{-t})_* Y_{\phi_t(x)}\right]_{t=0} \quad .} \tag{17.6}$$

Note that the vector $(\phi_{-t})_* Y_{\phi_t(x)}$ is always located at the point x!

(2) Invariant vector fields; Jacobi's variational equation. In §1, the orbits $\phi_t x$ are the solutions of the differential equations $\frac{dx^i}{dt} = X^i$ and so we may write, using Problem 28,

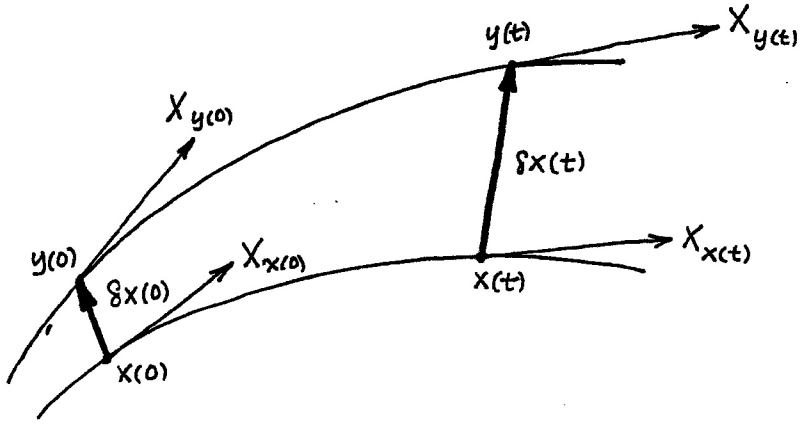
$$(\mathfrak{L}_X Y)_x = \sum_j \frac{dx^j}{dt} \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} = \frac{dY^i}{dt} - \sum_j \frac{\partial X^i}{\partial x^j} Y^j \quad . \tag{17.7}$$

Note that the right hand side makes sense when the vector field Y is not defined in a neighborhood of x; Y need only be defined along the orbit of X through x. It is clear from the definition (17.1) that $\mathfrak{L}_X Y$ should make sense also in this situation; we shall not hesitate to use (17.6) in this more general setting. It can be interpreted as follows. If Y is defined along the orbit $\phi_t x$, extend Y to a smooth vector field in a neighborhood of x; call the resulting field $\tilde{Y}$. Then apply (17.1) to $\tilde{Y}$ and note that the answers, given by (17.6) depends only on Y along the orbit, not on the extension $\tilde{Y}$ used!

We shall say that a vector field $Y_{\phi_t x}$ defined along an orbit of the vector field X is " <u>invariant under the flow</u> ϕ_t" or simply "<u>invariant under</u> X" provided $Y_{\phi_t x} = (\phi_t)_* Y_p$. Thus

$$\boxed{\mathfrak{L}_X Y = 0\,, \text{ i.e., } \frac{dY^i}{dt} = \sum_j \frac{\partial X^i}{\partial x^j} Y^j \qquad \text{Jacobi's equation of variation.}} \tag{17.8}$$

The reason for the name is as follows. Classically one considers two infinitesimally close orbits $x(t)$ and $y(t) = x(t) + \delta x(t)$.



Then $\frac{dx}{dt} = X(x)$ and $\frac{d}{dt}(x+\delta x) = X(x+\delta x)$, since both are solution curves. Subtracting, $\delta \frac{dx}{dt} \equiv \frac{d(x+\delta x)}{dt} - \frac{dx}{dt} = \frac{d}{dt}\delta x$, becomes

$$\frac{d}{dt}(\delta x)^i = X^i(x+\delta x) - X^i(x) = \sum_j \frac{\partial X^i}{\partial x^j}(x)\delta x^k + \text{ higher order in } \delta x\,.$$

If we throw away the higher powers of δx we get the linearized equations of variation

$$\frac{d}{dt}\,\delta x^i = \sum_j \frac{\partial X^i}{\partial x^j}\,\delta x^j$$

a linear system of differential equations for the vector δx (with matrix $\frac{\partial X^i}{\partial x^j}(x(t))$) along the given solution $x = x(t)$. Mathematicians prefer to consider $Y = \delta x$ as a tangent vector rather than an infinitesimal displacement δx. This then is (17.8).

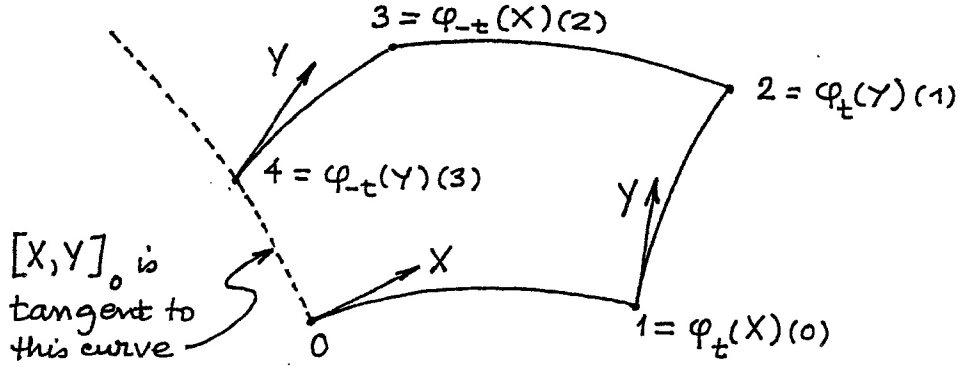
To actually exhibit a vector field Y that is invariant under the field X along a solution curve $\phi_t x$ for X one would have to actually solve the system (17.8). Frequently a knowledge of the matrix $\frac{\partial X^i}{\partial x^j}(x(t))$ is sufficient for drawing important conclusions about the behavior of nearby solutions (for example "linearized stability").

~~ **WARNING.** The equation (17.8) has invariant meaning, independent of coordinates, since $\mathcal{L}_X Y$ is indeed a vector. However, the individual terms $\frac{dY^i}{dt}$ and $\frac{\partial X^i}{\partial x^j} Y^i$ do not have such meaning since, e.g. $\frac{\partial X^i}{\partial x^j}$ do not form the components of a tensor (recall Problem 15).

(3) The Flow Associated to $\mathcal{L}_X Y = [X,Y]$. What is the flow generated [X,Y]? Let $\phi_t(X)$, $\phi_t(Y)$ be the flows generated by X and Y. We claim that the flow generated by $[X,Y]$ is in a sense the commutator of the flows $\phi_t(X)$ and $\phi_t(Y)$.

Proposition. Let $\sigma(t)$ be the curve

$$\sigma(t) = \phi_{-t}(Y) \cdot \phi_{-t}(X) \circ \phi_t(Y) \circ \phi_t(X)(0)$$



Then for any smooth function defined near 0 we have

$$\lim_{t\to 0} \frac{f(\sigma(\sqrt{t}) - f(\sigma(0))}{t} = [X,Y]_0(f) \quad .$$

Proof (Richard Faber). Look at a smooth function f restricted to the broken curve portions (we let 0,1,2,3,4 be the vertices). Form $f(4) - f(0) = f(\sigma(t)) - f(\sigma(0)) = [f(4) - f(3)] + [f(3) - f(2)] + [f(2) - f(1)] + [f(1) - f(0)]$. By Taylor's theorem,

$$f(1) - f(0) = t\, X_0(f) + \frac{t^2}{2} X_0(X(f)) + 0(3) \tag{i}$$

where $\frac{0(3)(t)}{t^2} \to 0$ as $t \to 0$.

Also

$$f(2) - f(1) = t\, Y_1(f) + \frac{t^2}{2} Y_1(Y(f)) + 0(3) \quad .$$

Note

$$Y_1(Y(f)) = Y_0(Y(f)) + t\, X_0(Y_{\phi_t(x)(0)}(Y(f))) + 0(2)$$

therefore

$$f(2) - f(1) = t\, Y_1(f) + \frac{t^2}{2} Y_0(Y(f)) + 0(3) \quad . \tag{ii}$$

Similarly

$$f(3) - f(2) = -t\, X_2(f) + \frac{t^2}{2} X_0(X(f)) + 0(3) \tag{iii}$$

$$f(4) - f(3) = -t\, Y_3(f) + \frac{t^2}{2} Y_0(Y(f)) + 0(3) \quad . \tag{iv}$$

Adding (i) + (ii) + (iii) + (iv) we get

$$f(4) - f(0) = t[X_0(f) + Y_1(f) - X_2(f) - Y_3(f)] + t^2[X_0(X(f)) + Y_0(Y(f))] + 0(3) \quad .$$

But

$$\begin{aligned} X_2(f) - X_0(f) &= X_2(f) - X_1(f) + X_1(f) - X_0(f) \\ &= t\, Y_1(X(f)) + 0(2) + t\, X_0(X(f)) + 0(2) \\ &= t\, Y_0(X(f)) + t\, X_0(X(f)) + 0(2) \quad . \end{aligned} \tag{v}$$

Also

$$\begin{aligned} Y_3(f) - Y_1(f) &= Y_3(f) - Y_2(f) + Y_2(f) - Y_1(f) \\ &= -t\, X_2(Y(f)) + 0(2) + t\, Y_1(Y(f)) + 0(2) \\ &= -t\, X_0(Y(f)) + t\, Y_0(Y(f)) + 0(2) \quad \text{(by (v))} \quad . \end{aligned}$$

Thus

$$f(4) - f(0) = t^2[X_0(Y(f)) - Y_0(X(f))] + 0(3)$$

and so

$$\frac{f(\sigma(t)) - f(\sigma(0))}{t^2} \xrightarrow[t\to 0]{} X_0(Y(f)) - Y_0(X(f)) \quad . \qquad \square$$

If we denote the derivative for $t \to 0^+$ by $\dfrac{d}{dt_+}$ then we can write

$$\boxed{[X,Y] = \frac{d}{dt_+}\, \sigma(\sqrt{t})\Big|_{t=0}} \tag{17.9}$$

~~ **WARNING.** Many books (including Arnold's book on mechanics) define the Lie bracket [X,Y] to be the negative of ours. Also many books call [X,Y] the "Poisson bracket"; we however, shall reserve the name Poisson bracket for a different (but related) concept.

Lecture 18. Lie Derivatives of Exterior Forms

(1) **Lie derivatives of forms.** If X is a vector field (with local flow ϕ_t) and f is a function we shall define the Lie derivative of f with respect to X at x by

$$\mathcal{L}_X(f) \equiv X_p(f) = \frac{d}{dt}\, f(\phi_t x)\Big|_{t=0} = \frac{d}{dt}\, (\phi_t{}^* f)\Big|_{t=0} \tag{18.1}$$

(we have used Lecture 4, §1 and Lecture 10). This is just telling us how f changes along the flow lines of X.

If α^p is a p-form we define

$$\boxed{\mathcal{L}_X\alpha^p \equiv \frac{d}{dt}\, (\phi_t{}^*\alpha^p)\Big|_{t=0} = \lim_{t\to 0} \frac{\phi_t{}^*\alpha - \alpha}{t}} \tag{18.2}$$

We mean by this

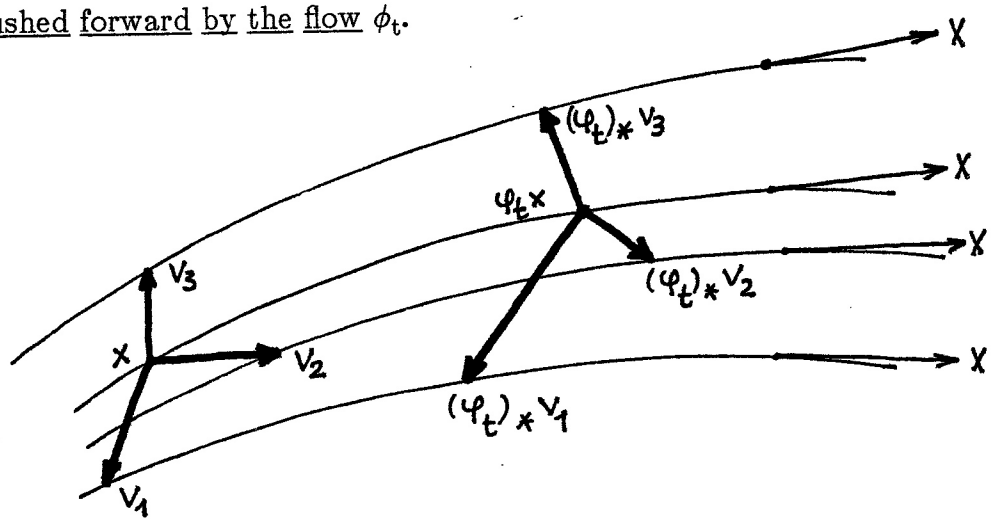
$$\begin{aligned}\Big(\frac{d}{dt}\, \phi_t{}^*\alpha\Big)_{t=0}(v_1, \ldots, v_p) &= \frac{d}{dt}\, \alpha\big((\phi_t)_* v_1, \ldots, (\phi_t)_* v_p\big)\Big|_{t=0} \\ &= \frac{d}{dt}\, [(\phi_t{}^*\alpha)(v_1, \ldots, v_p)]_{t=0} \quad .\end{aligned} \tag{18.3}$$

In particular, if $v_1, \ldots, v_p$ are <u>invariant</u> under the flow ϕ_t generated by X, i.e., $(\phi_t)_* v_x = v_{\phi_t x}$, then

$$(\mathcal{L}_X\alpha^p)(v_1, \ldots, v_p) = \frac{d}{dt}\, (\alpha^p_{\phi_t x}(v_1, \ldots, v_p))\Big|_{t=0} \quad . \tag{18.3$'$}$$

$\mathcal{L}_X\alpha$ measures the derivative of the value of α evaluated on a p-tuple of vector fields that are invariant under the flow generated by X. In particular, if $\alpha^n = \mathrm{vol}^n = \sqrt{g}\, dx^1 \wedge \cdots \wedge dx^n$ is the volume form for a Riemannian metric on M^n, and if X is a vector field on M^n then $\mathcal{L}_X \mathrm{vol}^n$ <u>is the</u> n- <u>form that reads off</u>

the rate of change of volume of a parallelopiped spanned by n vectors that are pushed forward by the flow ϕ_t.



In other words $\mathcal{L}_X \text{vol}^n$ measures how volumes are changing under the flow ϕ_t generated by X! One usually thinks of vol^n as a given form; then $\mathcal{L}_X \text{vol}^n$ is describing a property of the flow generated by X, namely how this flow is distorting volumes.

We need convenient methods of computing $\mathcal{L}_X \alpha$. First note

$$\mathcal{L}_X(\alpha^p \wedge \beta^q)(v_1, \ldots, v_{p+q}) = \frac{d}{dt}\,[\alpha \wedge \beta((\phi_t)_* v_1, \ldots, (\phi_t)_* v_p)]_{t=0}$$

$$= \frac{d}{dt} \sum_{\substack{I \\ \to}, \substack{J \\ \to}} \delta^{I\,J}_{1\ldots(p+q)}\ \alpha\Big((\phi_t)_* v_I\Big)\beta\Big((\phi_t)_* v_J\Big)\ \Big|_{t=0}$$

$$= \sum_{\substack{I \\ \to}, \substack{J \\ \to}} \delta^{I\,J}_{1\ldots(p+q)}\ \frac{d}{dt}\ \alpha\Big((\phi_t)_* v_I\Big)\ \Big|_{t=0}\ \beta(v_J)$$

$$+ \sum_{\substack{I \\ \to}, \substack{J \\ \to}} \delta^{I\,J}_{1\ldots(p+q)}\ \alpha(v_I)\ \frac{d}{dt}\ \beta\Big((\phi_t)_* v_J\Big)\ \Big|_{t=0}$$

and so $\mathcal{L}_X$ is a derivation

$$\boxed{\mathcal{L}_X(\alpha^p \wedge \beta^q) = (\mathcal{L}_X \alpha^p) \wedge \beta^q + \alpha^p \wedge (\mathcal{L}_X \beta^q)} \tag{18.4}$$

Furthermore we have the important

$$\boxed{\mathfrak{L}_X \circ d = d \circ \mathfrak{L}_X} \tag{18.5}$$

Proof. First verify for 0-forms, i.e., functions f. If $v = v_x$ is a vector at $x \in M^n$

$$\mathfrak{L}_X(df)(v) = \frac{d}{dt}(\phi_t^* df)v\Big|_{t=0} = \frac{d}{dt}(d\phi_t^* f)(v)\Big|_{t=0}$$

$$= \frac{d}{dt} v(\phi_t^* f)\Big|_{t=0} = \frac{d}{dt} v(f \circ \phi_t)\Big|_{t=0} =$$

(since v is a constant operator)

$$= v_p\Big(\frac{d}{dt}\phi_t^* f\Big)\Big|_{t=0} = v_p(X(f)) = v_p(\mathfrak{L}_X(f)) = d(\mathfrak{L}_X f)(v_p)$$

$$\therefore \mathfrak{L}_X \circ d = d \circ \mathfrak{L}_X$$

when applied to 0-forms. When applied to p-forms

$$\mathfrak{L}_X d \sum a_I dx^{i_1} \wedge \ldots \wedge dx^{i_p} = \mathfrak{L}_X \sum da_I \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p} = \text{(by (18.4))}$$

$$= \sum \mathfrak{L}_X da_I \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p} + \sum da_I \wedge \mathfrak{L}_X(dx^{i_1}) \wedge \ldots \wedge dx^{i_p} + \ldots$$

$$= \sum d(\mathfrak{L}_X a_I) \wedge dx^{i_1} \wedge \ldots \wedge dx^{i_p} + \sum da_I \wedge d(\mathfrak{L}_X x^i) \wedge \ldots \wedge dx^{i_p} + \ldots$$

$$= d\Big(\sum (\mathfrak{L}_X a_I) dx^{i_1} \wedge \ldots \wedge dx^{i_p} + \sum a_I d\mathfrak{L}_X(x^i) \wedge \ldots \wedge dx^{i_p} + \ldots\Big)$$

$$= d\Big(\sum (\mathfrak{L}_X a_I) dx^{i_1} \wedge \ldots \wedge dx^{i_p} + \sum a_I \mathfrak{L}_X(dx^i) \wedge \ldots \wedge dx^{i_p} + \ldots\Big)$$

$$= d\,\mathfrak{L}_X \sum a_I dx^{i_1} \wedge \ldots \wedge dx^{i_p} \quad .$$

□

A homomorphism $A : \Lambda^p \to \Lambda^{p+r}$ is

a derivation if r is even and $A(\alpha^p \wedge \beta^q) = (A\alpha^p) \wedge \beta^q + \alpha^p \wedge A(\beta^q)$

(e.g. $\mathfrak{L}_X$)

an anti-derivation if r is odd and

$$A(\alpha^p \wedge \beta^q) = (A\alpha^p) \wedge \beta^q + (-1)^p \alpha^p \wedge A(\beta^q) ,$$

(e.g. d and i_X)

Suppose we know the values of a derivation or antiderivation on functions and on d(functions). Then we see that we know its values on any p-forms α^p, for $\alpha^p = \sum_I a_I(x) dx^{i_1} \wedge \ldots \wedge dx^{i_p}$. If A and B are both either deviations or antiderivations, to prove $A\alpha^p = B\alpha^p$ for all forms α it is sufficient to verify this formula for functions α^0 and for d of functions !

- **Problem 29.**

θ derivation and A an antiderivation $\Rightarrow \theta \circ A - A \circ \theta$ is an antiderivation
A an antiderivation and B an antiderivation $\Rightarrow A \circ B + B \circ A$ is a derivation .

(2) **Cartan's formula.**

Proposition (H. Cartan). $\boxed{\mathfrak{L}_X = i_X \circ d + d \circ i_X}$ (18.6)

when operating on p-forms.

Proof. Both sides are derivations. On functions $i_X(f) = 0$ and $i_X df = X(f) = \mathfrak{L}_X(f)$ and both sides of (18.6) agree on functions. On differentials of functions

$$(i_X \circ d + d \circ i_X)(df) = d i_X df = d\mathfrak{L}_X(f) = \mathfrak{L}_X(df) .$$

Thus both sides agree on functions and d(functions) and so agree on p-forms. □

A direct calculational proof of this formula would be very messy.

• **Problem 30.** If X and Y are vector fields, then when applied to forms

$$\mathfrak{L}_X \circ i_Y - i_Y \circ \mathfrak{L}_X = i_{[X,Y]} \quad . \tag{18.7}$$

As an application of Problem 30 we have the following very useful formula for the exterior derivative of a 1-form α.

Proposition. Let α be a 1-form and let u_x, v_x, be vectors at a point $x \in M^n$. Extend these vectors in any way to be smooth vector fields u and v near x. Then

$$\boxed{d\alpha(u_x,v_x) = u_x(\alpha(v)) - v_x(\alpha(u)) - \alpha([u,v]) \quad .} \tag{18.8}$$

Proof.

$$\begin{aligned} d\alpha(u,v) &= (i_u d\alpha)(v) = (\mathfrak{L}_u\alpha - di_u\alpha)(v) \\ &= i_v\mathfrak{L}_u\alpha - v(i_u\alpha) \\ &= (\mathfrak{L}_u i_v - i_{[u,v]})(\alpha) - v(\alpha(u)) \\ &= \mathfrak{L}_u(\alpha(v)) - \alpha([u,v]) - v(\alpha(u)) \\ &= u(\alpha(v)) - \alpha([u,v]) - v(\alpha(u)) \quad . \end{aligned}$$

□

• **Problem 31.** Prove this proposition by writing both sides out in terms of coordinates.

The following proposition says that one may differentiate the function $a_{i_1 \cdots i_p} u_1^{i_1} \cdots u_p^{i_p}$ by using a product type rule for Lie differentiation of everything in sight. This is a "Leibnitz rule".

Proposition. For vector fields X, $u_1, \ldots, u_p$ we have

$$\boxed{X(\alpha^p(u_1, \ldots, u_p)) = (\mathfrak{L}_X\alpha^p)(u_1, \ldots, u_p) + \sum_r \alpha^p(u_1, \ldots, [X,u_r],\ldots u_p)} \quad (18.9$$

(this is easy to remember since $[x,u_r] = \mathfrak{L}_X(u_r)$).

Proof. First for 1-forms

$$(\mathfrak{L}_X\alpha)(u) = i_u(\mathfrak{L}_X\alpha) = \mathfrak{L}_X i_u\alpha - i_{[X,u]}\alpha$$
$$= X(\alpha(u)) - \alpha([X,u]) \quad ,$$

as desired. By induction, assume true for (p-1)-forms. Then for α^p

$$\mathfrak{L}_X\alpha(u_1 \cdots u_p) = i_{u_1}(\mathfrak{L}_X\alpha)(u_2, \ldots, u_p) = (\mathfrak{L}_X i_{u_1}\alpha - i_{[X_1,u_1]}\alpha)(u_2 \cdots u_p) \quad .$$

But $i_{u_1}\alpha$ is a (p-1)-form, by induction hypothesis then

$$\mathfrak{L}_X(i_{u_1}\alpha)(u_2 \cdots u_p) = X((i_{u_1}\alpha)(u_2 \cdots u_p)) - \sum_{r=2}^{p} (i_{u_1}\alpha)(u_2 \cdots [X,u_r],\ldots u_p)$$

and so

$$(\mathfrak{L}_X\alpha)(u_1, \ldots, u_p) = X(\alpha(u_1, \ldots, u_p)) - \alpha([X,u_1],u_2 \cdots u_p)$$
$$- \sum_{r=2}^{p} \alpha(u_1, \ldots, [X,u_r], \ldots, u_p) \quad . \qquad \square$$

Finally, we have a formula that exhibits the fact that $d\alpha^p$ is independent of coordinates, as follows. For vector fields $u_0, \ldots, u_p$

$$\boxed{\begin{aligned} d\alpha^p(u_0, \ldots, u_p) = \quad & \sum_{r=0}^{p} (-1)^r u_r\alpha(u_0, \ldots, \hat{u}_r, \ldots, u_p) \\ & + \sum_{r<s} (-1)^{r+s} \alpha^p([u_r,u_s],u_0, \ldots, \hat{u}_r, \ldots, \hat{u}_s, \ldots, u_p) \end{aligned}} \quad (18.10)$$

This formula generalizes the proposition (18.8) and can be proved by induction again. Note that the left hand side depends only on the vectors at the given point, while the right hand side seems to depend on the vector fields !

• **Problem 32.** Show that if $\alpha = \sum a_i dx^i$ and $u = \sum u^j \frac{\partial}{\partial x^j}$, then

$$\mathfrak{L}_u \alpha^1 = \sum_j \left[\sum_i (u^i \frac{\partial a_j}{\partial x^i} + a_i \frac{\partial u^i}{\partial x^j}) \right] dx^j \quad . \tag{18.11}$$

This should be compared with (17.5).

(3) **Vector analysis again.** Let vol^n be a volume form for M^n (Lecture 11, §4). If X is a vector field on M^n define the divergence of X (with respect to vol^n) to be the function $\text{div}\, X$ such that

$$\boxed{\mathfrak{L}_X \text{vol}^n = \text{div}\, X \cdot \text{vol}^n}$$

If we let $u_1, \ldots, u_n$ be vector fields along an orbit of X that are <u>invariant</u> under X, $u_\alpha(t) \equiv u_\alpha(\phi_t x) = (\phi_t)_* u_\alpha(x)$, then (see (18.3′))

$$(\mathfrak{L}_X \text{vol}^n)_x(u_1, \ldots, u_n) = \frac{d}{dt} \text{vol}^n (u_1(t), \ldots, u_n(t)) \Big|_{t=0}$$

and so $\text{div}\, X$ <u>measures the logarithmic rate of change of volume along the flow generated by</u> X. In local coordinates

$$\text{vol}^n = \rho(x) dx^1 \wedge \cdots \wedge dx^n , \quad \rho(x) > 0 \quad .$$

$$\begin{aligned}
\mathfrak{L}_X \text{vol}^n &= \mathfrak{L}_X(\rho\, dx^1 \wedge \ldots \wedge dx^n) \\
&= \mathfrak{L}_X(\rho) dx^1 \wedge \ldots \wedge dx^n + \sum_{r=1}^{n} \rho\, dx^1 \wedge \ldots \wedge \mathfrak{L}_X(dx^r) \wedge \ldots \wedge dx^n \\
&= \sum_{i=1}^{n} \left(X^i \frac{\partial \rho}{\partial x^i} \right) dx^1 \wedge \ldots \wedge dx^n + \sum_{r=1}^{n} \rho\, dx^1 \wedge \ldots \wedge d(\mathfrak{L}_X x^r) \wedge \ldots \wedge dx^n \\
&= \sum_{i=1}^{n} \left(X^i \frac{\partial \rho}{\partial x^i} \right) dx^1 \wedge \ldots \wedge dx^n + \rho \sum_{r=1}^{n} dx^1 \wedge \ldots \wedge dX^r \wedge \ldots \wedge dx^n \\
&= \sum_{i=1}^{n} \left(X^i \frac{\partial \rho}{\partial x^i} \right) dx^1 \wedge \ldots \wedge dx^n + \rho \sum_{r=1}^{n} dx^1 \wedge \ldots \wedge \sum_{s=1}^{n} \frac{\partial X^r}{\partial x^s} dx^s \wedge \ldots \wedge dx^n
\end{aligned}$$

$$= \sum_{i=1}^{n} \left(X^i \frac{\partial \rho}{\partial x^i}\right) dx^1 \wedge \ldots \wedge dx^n + \rho \sum_{r=1}^{n} dx^1 \wedge \ldots \wedge \frac{\partial X^r}{\partial x^r} dx^r \wedge \ldots \wedge dx^n$$

$$= \sum_{r=1}^{n} \left(X^r \frac{\partial \rho}{\partial x^r} + \rho \frac{\partial X^r}{\partial x^r}\right) dx^1 \wedge \ldots \wedge dx^n = \sum_{r=1}^{n} \frac{1}{\rho} \frac{\partial}{\partial x^r} (\rho X^r) \rho \, dx^1 \wedge \ldots \wedge dx^n$$

$$\therefore \qquad \boxed{\operatorname{div} X = \frac{1}{\rho} \sum_{r=1}^{n} \frac{\partial}{\partial x^r} (\rho X^r)} \qquad (18.12)$$

as in (12.12).

Note that Cartan's formula (18.6) shows that our definition of div X coincides with that given in (12.11), since

$$\mathfrak{L}_X \operatorname{vol}^n = i_X d \operatorname{vol}^n + d i_X \operatorname{vol}^n = d \, i_X \operatorname{vol}^n \quad .$$

Let us use our formalism to complete our form-description of vector analysis in $\mathbb{R}^3$ started in Lecture 12.

Consider for example curl $(\vec{A} \times \vec{B})$

$$\operatorname{curl} (\vec{A} \times \vec{B}) \leftrightarrow d(i_{\vec{B}} \alpha^2)$$

$$= \mathfrak{L}_{\vec{B}} \alpha^2 - i_{\vec{B}} d\alpha^2$$

$$= \mathfrak{L}_{\vec{B}} \alpha^2 - i_{\vec{B}} d(i_{\vec{A}} \operatorname{vol}^3)$$

$$= \mathfrak{L}_{\vec{B}} \alpha^2 - i_{\vec{B}} \cdot \operatorname{div} A \operatorname{vol}^3$$

$$= \mathfrak{L}_{\vec{B}} \alpha^2 - \operatorname{div} \vec{A} \, i_{\vec{B}} \operatorname{vol}^3 \quad .$$

Now use the result of Problem 30.

$$\mathfrak{L}_{\vec{B}} \alpha^2 = \mathfrak{L}_{\vec{B}} i_{\vec{A}} \operatorname{vol}^3 = i_{\vec{A}} \mathfrak{L}_{\vec{B}} \operatorname{vol}^3 + i_{[\vec{B},\vec{A}]} \operatorname{vol}^3$$

$$= \operatorname{div} \vec{B} \, i_{\vec{A}} \operatorname{vol}^3 + i_{[\vec{B},\vec{A}]} \operatorname{vol}^3$$

$$\leftrightarrow (\operatorname{div} \vec{B}) \vec{A} + [\vec{B},\vec{A}] \quad .$$

Thus $$\boxed{\text{curl}(\vec{A} \times \vec{B}) = (\text{div}\, \vec{B})\vec{A} + [\vec{B},\vec{A}] - (\text{div}\, A)\vec{B}\,.}$$

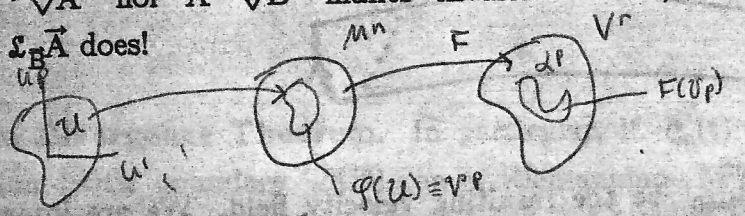
In vector analysis books the term $\mathcal{L}_{\vec{B}}\vec{A} = [\vec{B},\vec{A}]$ is written differently. We can write

$$[\vec{B},\vec{A}]^i = B^j \frac{\partial A^i}{\partial x^j} - A^j \frac{\partial B^i}{\partial x^j} = (D_B A)^i - (D_A B)^i$$

where $(D_B A)^i = \vec{B} \cdot (\text{grad}\, A^i))$, etc. In vector analysis books one then sees the mysterious

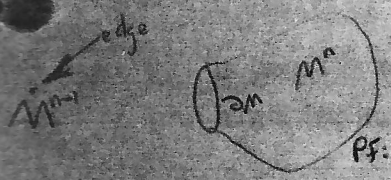
$$\text{curl}\,(\vec{A} \times \vec{B}) = (\text{div}\, \vec{B})\vec{A} - (\text{div}\, \vec{A})\vec{B} + \vec{B} \cdot \nabla\vec{A} - \vec{A} \cdot \nabla\vec{B}$$

as if it made sense to take the gradient $\nabla\vec{A}$ of a vector! In curvilinear coordinates neither $\vec{B} \cdot \nabla\vec{A}$ nor $\vec{A} \cdot \nabla\vec{B}$ makes invariant sense, but of course their difference $[\vec{B},\vec{A}] = \mathcal{L}_{\vec{B}}\vec{A}$ does!



$$\int_{F(U_p)} \alpha^p = \int_{(F\circ\varphi)(U_p)} \alpha^p \equiv \int_{U_p} (F\circ\varphi)^* \alpha^p$$

$$= \int_{U_p} \varphi^* \circ (F^* \alpha) \equiv \int_{\varphi(U_p) = V_p} F^* \alpha \qquad \text{so in general} \quad \int_{F(U_p)} \alpha^p = \int_{V_p} F^* \alpha^p$$

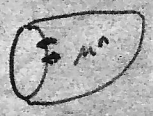


Thm: $\nexists$ continuous retraction of compact manifold with bndy onto its bndy. (i.e. keep bndy fixed)

$i(\partial M) = \partial M$ $\quad r : M \to \partial M$ $\quad r \circ i : \partial M \to \partial M$ (identity) $\Rightarrow \nexists\, r$.

Pf: Assume Consider special case when M^n is orientable & r diff'able (not a serious restriction).

Suppose r exists



An orientation for M^n orients edge

Let ω^{n-1} be a form on ∂M s.t. $\int_{\partial M} \omega^{n-1} > 0$

$$0 < \int_{\partial M} \omega^{n-1} = \int_{(r\circ i)\partial M} \omega^{n-1} = \int_{i\partial M = \partial M^n} r^* \omega^{n-1} = \int_{M^n} dr^* \omega^{n-1} = \int_{M^n} r^* d\omega^{n-1} = 0 \Rightarrow \nexists\, r.$$

$r^* d\omega^{n-1} = 0$

Brower Fixed Point Theorem: Every continuous map of the closed n-ball has a fixed pt.

$f : B^n \to B^n$

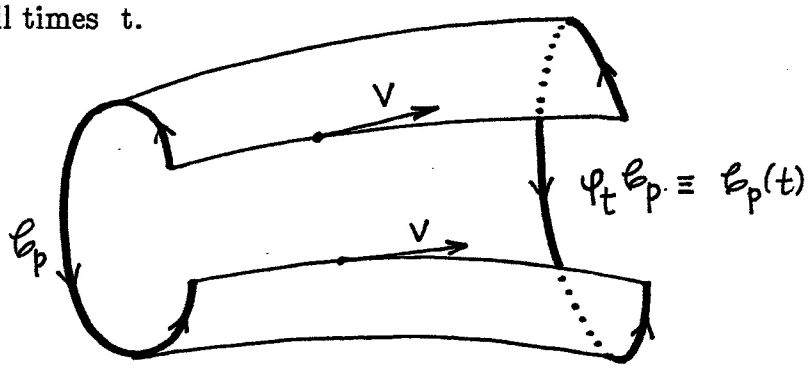
Pf: Assume $\nexists$ fixed pt with bndy pts fixed. $r(x) = x \;\; \forall x \in \partial B^n$ $r(x)$ is continuous & maps maps all points into bndy

$\therefore$ r is a retraction of this compact manifold onto bndy.

$\therefore$ it can't exist.

Lecture 19. Differentiation of Integrals. The Divergence Theorem

(1) **Time independent case.** Let α^p be a p-form, C_p a p dimensional submanifold (perhaps with boundary ∂C_p) of a manifold M^n. We shall consider a "variation" of C_p arising as follows. We suppose that there is a flow $\phi_t : M^n \to M^n$, i.e., a 1-parameter "group" of diffeomorphisms defined in a neighborhood of C_p for small times t.



Let $\vec{v}_x = \left(\frac{d}{dt}\phi_t x\right)_{t=0}$ be the resulting velocity field. We are interested in the time variation of the integral

$$I(t) \equiv \int_{C_p(t)} \alpha^p = \int_{\phi_t C_p} \alpha^p = \int_{C_p} \phi_t^* \alpha^p \quad .$$

Differentiating

$$\begin{aligned} I'(t) &= \lim_{h\to 0} \frac{1}{h} [I(t+h) - I(t)] \\ &= \lim_{h\to 0} \frac{1}{h} \left[\int_{C_p} \phi_{t+h}^* \alpha^p - \int_{C_p} \phi_t^* \alpha^p \right] \\ &= \lim_{h\to 0} \frac{1}{h} \left[\int_{C_p} \phi_t^* (\phi_h^* \alpha^p - \alpha^p) \right] \\ &= \lim_{h\to 0} \frac{1}{h} \int_{\phi_t C_p} (\phi_h^* \alpha^p - \alpha^p) = \int_{\phi_t C_p} \lim_{h\to 0} \frac{1}{h} (\phi_h^* \alpha^p - \alpha^p) \end{aligned}$$

i.e.,

$$\boxed{\frac{d}{dt} \int_{C_p(t)} \alpha^p = \int_{C_p(t)} \mathcal{L}_v \alpha^p} \qquad (19.1)$$

and in particular

$$\left(\frac{d}{dt}\int_{C_p(t)}\alpha^p\right)_{t=0} = \int_{C_p}\mathfrak{L}_{\bar{v}}\alpha^p \quad .$$

Using Cartan's formula (18.6) we may continue

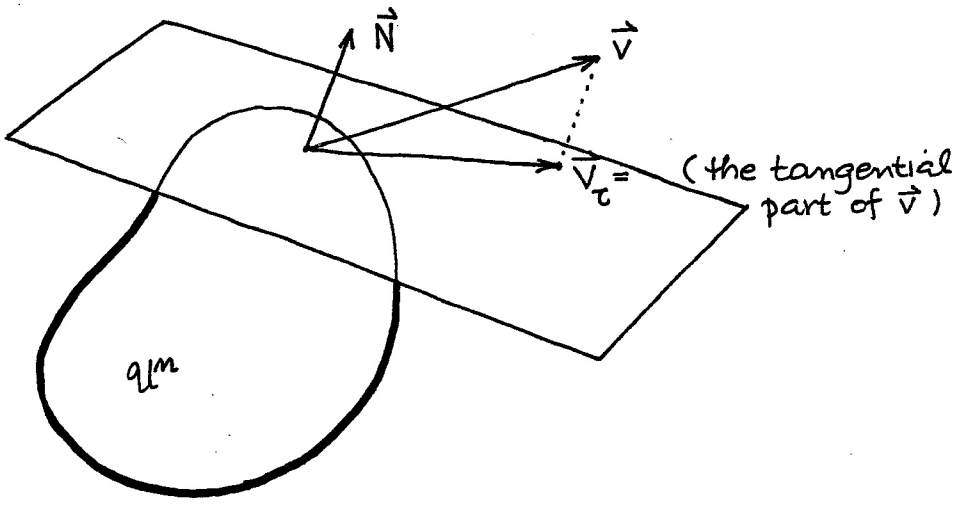
$$\frac{d}{dt}\int_{C_p(t)}\alpha^p = \int_{C_p(t)}\mathfrak{L}_{\bar{v}}\alpha^p = \int_{C_p(t)} i_{\bar{v}}d\alpha^p + di_{\bar{v}}\alpha^p = \int_{C_p(t)} i_{\bar{v}}d\alpha^p + \int_{\partial C_p(t)} i_{\bar{v}}\alpha^p \qquad (19.2)$$

$$\boxed{\frac{d}{dt}\int_{C_p(t)}\alpha^p = \int_{C_p(t)} i_{\bar{v}}d\alpha^p + \int_{\partial C_p(t)} i_{\bar{v}}\alpha^p} \qquad (19.3)$$

(2) **The Divergence Theorem.** In particular, if $C_p(t) = C_n(t) = U^n$ is a compact region (with boundary ∂U^n) on a Riemannian M^n, and if $\alpha^n = \text{vol}^n$ is the volume n-form, then

$$\frac{d}{dt}\int_{U^n}\text{vol}^n = \int_{U^n}\mathfrak{L}_{\bar{v}}\text{vol}^n = \int_{U^n}(\text{div }\bar{v})\text{vol}^n \qquad (19.4)$$

is just the usual geometrical interpretation of the divergence of $\bar{v}$. On the other hand, letting N be the unit <u>outward</u> pointing normal to ∂U^n



$$\int_{U^n}\mathfrak{L}_{\bar{v}}\,\text{vol}^n = \int_{\partial U^n} i_{\bar{v}}\,\text{vol}^n = \int_{\partial U^n} i_{\langle\bar{v},\vec{N}\rangle\vec{N}+\bar{v}_\tau}\,\text{vol}^n$$

$$= \int_{\partial U^n}\langle\bar{v},\vec{N}\rangle\, i_{\vec{N}}\,\text{vol}^n + \int_{\partial U^n} i_{\bar{v}_\tau}\,\text{vol}^n \quad .$$

But if $T_1, \ldots, T_{n-1}$ is an o.n. basis for tangent vectors to ∂U at x, we have

$$i_{\vec{v}_\tau} \mathrm{vol}^n(T_1, \ldots, T_{n-1}) = 0$$

and

$$i_{\vec{N}} \mathrm{vol}^n(T_1, \ldots, T_{n-1}) = \pm 1 \quad ,$$

i.e., $i_{\vec{N}} \mathrm{vol}^n$ is the $(n-1)$ volume form for ∂U

$$\boxed{i_{\vec{N}} \mathrm{vol}^n = \mathrm{vol}^{n-1}_{\partial U}} \tag{19.5}$$

Consequently

$$\int_{U^n} \mathfrak{L}_{\vec{v}} \mathrm{vol}^n = \int_{\partial U^n} \langle \vec{v}, \vec{N} \rangle \mathrm{vol}^{n-1}_{\partial U}$$

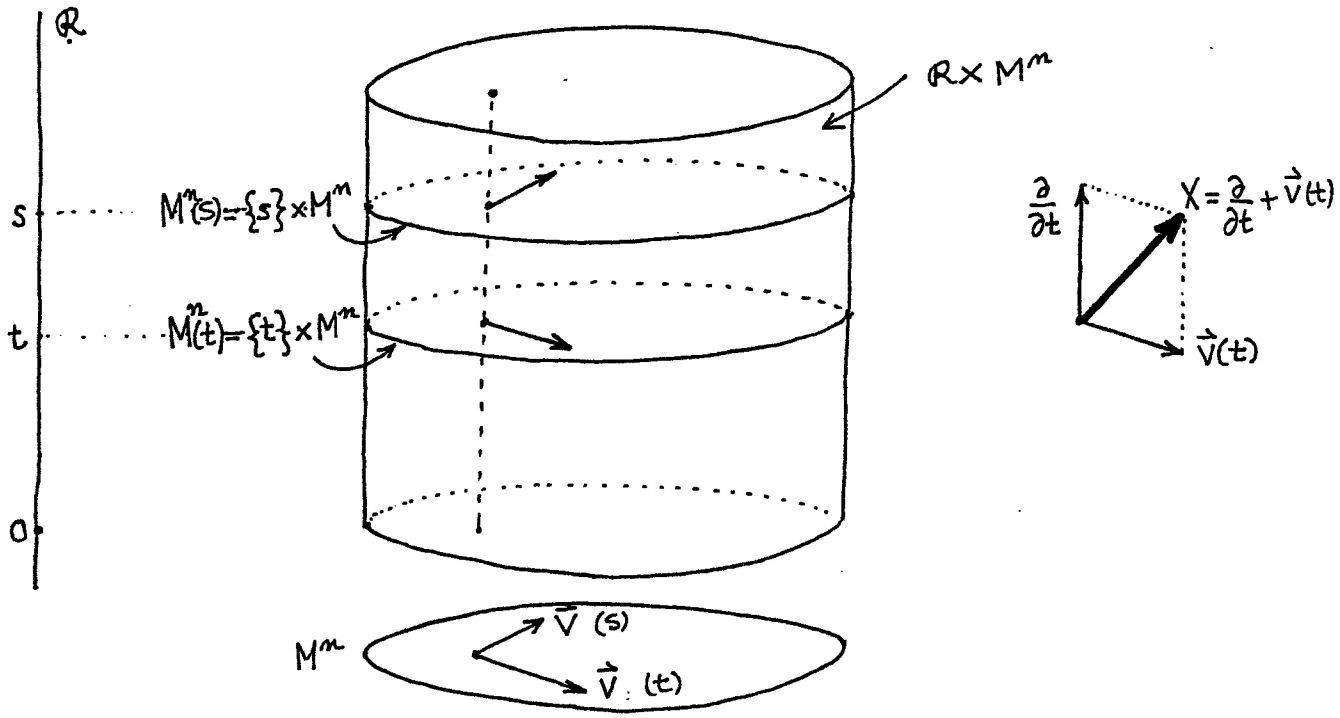
and we have the familiar version of the divergence theorem

$$\boxed{\int_{U^n} \mathrm{div}\, \vec{v}\, \mathrm{vol}^n = \int_{\partial U^n} \langle \vec{v}, \vec{N} \rangle \mathrm{vol}^{n-1}_{\partial U} \qquad \vec{N} \text{ outward unit normal to } \partial U^n} \tag{19.6}$$

allowing us to compute a "flux" (a boundary integral) in terms of a volume integral of $\mathrm{div}\, \vec{v}$.

We emphasize that (19.6) is true even if M^n is not orientable, because both vol^n and $\mathrm{vol}^{n-1}_{\partial U}$ are pseudo forms. $\vec{v}$ is assumed to be a true vector.

(3) **The time dependent case.** A time dependent function $f(t,x)$, or vector $\vec{v}(t,x)$,..., on a manifold M^n can (and should) be thought of as a function, or vector ,..., on the product manifold $\mathbb{R} \times M^n$, with t a coordinate for the real line $\mathbb{R}$



If on M^n we wish to solve the differential equations

$$\begin{cases} \dfrac{dx^i}{dt} = v^i(t,x) \\ x^i(0) = x_0^i \end{cases} \tag{19.7}$$

we may proceed as follows. Use x^0 for coordinate in $\mathbb{R}$, then $x^0,x^1,..,x^n$ are local coordinates in $\mathbb{R}\times M^n$ and we may consider the differential equations

$$\begin{cases} \dfrac{dx^0}{dt} = 1 \\ \dfrac{dx^i}{dt} = v^i(x^0,x^1,...,x^n) \quad i = 1,...,n \\ x^0(0) = 0\,, \quad x^i(0) = x_0^i \end{cases} \tag{19.8}$$

In other words we look at the new vector field on $\mathbb{R}\times M^n$

$$X = \frac{\partial}{\partial x^0} + \sum_{i=1}^{m} v^i(x^0,\vec{x})\,\frac{\partial}{\partial x^i} = \frac{\partial}{\partial t} + \vec{v}(t,\vec{x}) \quad .$$

Then the (local) flow (1 parameter group of diffeomorphisms of $\mathbb{R}\times M$), $\psi_t : \mathbb{R}\times M^n \to \mathbb{R}\times M^n$ generated by the vector field X yields the

solution curves to (19.8). Since $\frac{dx^0}{dt} = 1$ it is clear that if we let $M^n(s)$ be the section of $\mathbb{R} \times M^n$ where $x^0 = s$ then this flow operates so that

$$\psi_t : M^n(s) \to M^n(s+t) \quad .$$

By projecting down to M^n we get solutions to the original system (19.7).

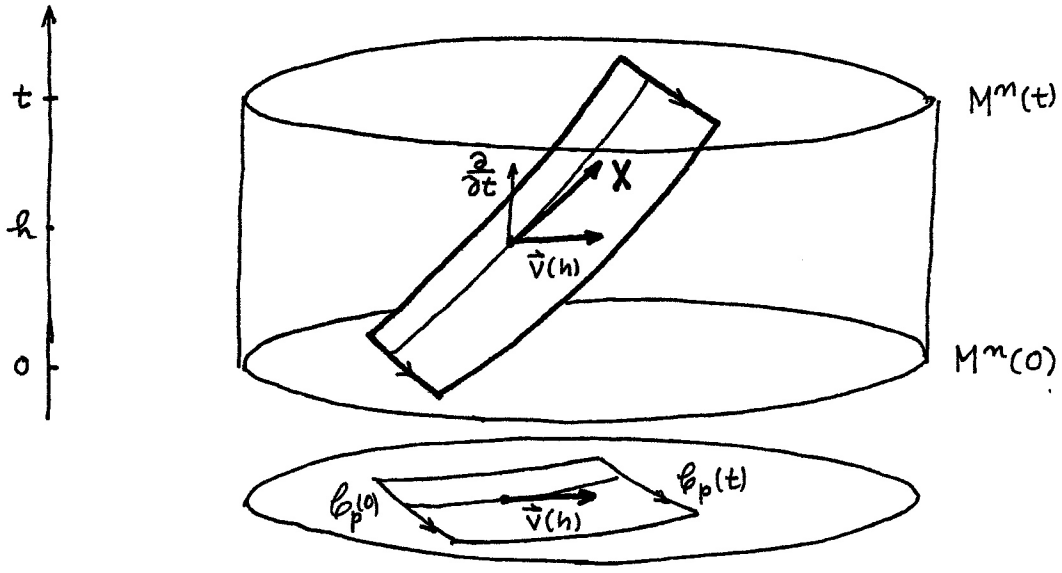
Suppose now that α is time dependent p-form on M^n

$$\alpha = \sum_{i_1 < \ldots < i_p} a_{i_1 \ldots i_p}(t,x) dx^{i_1} \wedge \ldots \wedge dx^{i_p}$$

(note that there are no dt terms). Again we consider it as a p-form on $\mathbb{R} \times M^n$ (having no dt terms). Suppose we have a time dependent "flow" of fluid in M^n, it has a time dependent velocity field $\vec{v}(t,\vec{x})$ on M^n. A p dimensional submanifold $C_p = C_p(0)$ at time 0 is pushed by the fluid to a submanifold $C_p(t)$ after t seconds and we again wish to compute

$$\frac{d}{dt} \int_{C_p(t)} \alpha^p \quad .$$

We pass to $\mathbb{R} \times M^n$, construct the vector field $X = \frac{\partial}{\partial t} + \vec{v}$ and get the resulting flow ψ_t, all as above



We then have

$$\frac{d}{dt} \int_{C_p(t)} \alpha^p = \int_{C_p(t)} \mathcal{L}_X \alpha^p \quad .$$

Let us write $\vec{d}$ for exterior differentiation in M^n, and d for exterior differentiation in $\mathbb{R} \times M^n$. Thus for forms on $\mathbb{R} \times M^n$

$$d\alpha = d \sum_{\substack{I \\ \rightarrow}} a_I dx^I = \sum_{\substack{I \\ \rightarrow}} \frac{\partial A_I}{dt}\, dt \wedge dx^I + \vec{d}\alpha \text{ , i.e.,}$$

$$d = dt \wedge \frac{\partial}{\partial t} + \vec{d} \quad . \tag{19.9}$$

Then $\mathfrak{L}_X = i_X d + d i_X = i_{\frac{\partial}{\partial t} + \vec{v}} d + d i_{\frac{\partial}{\partial t} + \vec{v}}$

$$\mathfrak{L}_X = i_{\vec{v}} d + d i_{\vec{v}} + i_{\frac{\partial}{\partial t}} d + d i_{\frac{\partial}{\partial t}} \quad . \tag{19.10}$$

We are going to apply $\mathfrak{L}_X$ to $\alpha = \sum a_I(t,x)dx^i$, no dt terms. Hence $i_{\frac{\partial}{\partial t}} \alpha = 0$ and last term in (19.10) vanishes. We are left with

$$\mathfrak{L}_X \alpha = [i_{\vec{v}}\left(\vec{d} + dt \wedge \frac{\partial}{\partial t}\right) + \left(\vec{d} + dt \wedge \frac{\partial}{\partial t}\right) i_{\vec{v}} + i_{\frac{\partial}{\partial t}} \left(\vec{d} + dt \wedge \frac{\partial}{\partial t}\right)](\alpha)$$

$$= (i_{\vec{v}} \vec{d}\alpha + \vec{d} i_{\vec{v}} \alpha) + i_{\frac{\partial}{\partial t}} \left(dt \wedge \frac{\partial \alpha}{\partial t}\right)$$

$$+ i_{\vec{v}} dt \wedge \frac{\partial}{\partial t}(\alpha) + dt \wedge \frac{\partial}{\partial t}(i_{\vec{v}} \alpha) + i_{\frac{\partial}{\partial t}} \vec{d}\alpha \quad .$$

Again $i_{\frac{\partial}{\partial t}} \vec{d}\alpha = 0$ since $\vec{d}\alpha$ has no dt term. Also we are going to be integrating over $C_p(t)$ which lies in the section of constant t, hence $dt = 0$ there and so $\int i_{\vec{v}} dt \wedge \frac{\partial \alpha}{\partial t} = 0 = \int dt \wedge \frac{\partial}{\partial t}(i_{\vec{v}} \alpha)$. Thus

$$\int_{C_p(t)} \mathfrak{L}_X \alpha = \int_{C_p(t)} i_{\vec{v}} \vec{d}\alpha + \vec{d} i_{\vec{v}} \alpha + i_{\frac{\partial}{\partial t}}\, dt \wedge \frac{\partial \alpha}{\partial t} \quad .$$

Now

$$i_{\frac{\partial}{\partial t}}\, dt \wedge \frac{\partial}{\partial t} \sum_{\substack{I \\ \rightarrow}} a_I dx^I = i_{\frac{\partial}{\partial t}}\, dt \wedge \sum_{\substack{I \\ \rightarrow}} \frac{\partial a_I}{\partial t}\, dx^I = \sum_{\substack{I \\ \rightarrow}} \frac{\partial a_I}{\partial t}\, dx^I$$

$$= \frac{\partial \alpha}{\partial t} \quad .$$

We conclude

$$\boxed{\frac{d}{dt}\int_{C_{p(t)}} \alpha^p = \int_{C_{p(t)}} \left(\frac{\partial \alpha}{\partial t} + i_{\vec{v}}\vec{d}\alpha + \vec{d}i_{\vec{v}}\alpha\right) .} \tag{19.11}$$

(Note that $i_{\vec{v}}\vec{d} + \vec{d}i_{\vec{v}}$ is simply the Lie derivative with respect to the vector field $\vec{v}$ "frozen" at time t.)

- **Problem 33.** Let $\vec{A}$ and $\vec{B}$ be time dependent vector fields on $\mathbb{R}^3$ and let $\rho(t,\vec{x})$ be smooth. Show that we have

$$\begin{cases} \dfrac{d}{dt}\displaystyle\int_{C_1} \vec{A}\cdot d\vec{r} = \int_{C_1}\left[\frac{\partial \vec{A}}{\partial t} - \vec{v}\times \operatorname{curl}\vec{A} + \operatorname{grad}(\vec{v}\cdot\vec{A})\right]\cdot d\vec{r} \\ \dfrac{d}{dt}\displaystyle\iint_{C_2} \vec{B}\cdot d\vec{S} = \iint_{C_2}\left[\frac{\partial \vec{B}}{\partial t} + (\operatorname{div}\vec{B})\vec{v} - \operatorname{curl}(\vec{v}\times\vec{B})\right]\cdot d\vec{S} \\ \dfrac{d}{dt}\displaystyle\iiint_{C_3} \rho dV = \iiint_{C_3}\left[\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho\vec{v})\right]dV \end{cases}$$

(4) **Euler's equations for a perfect fluid** If you are not familiar with Euler's equations for a perfect fluid (or the elements of continuum mechanics in general) may I suggest the treatment given on pages 64-68 of [E]. Euler's equations are (in the case where the external body forces are derivable from a potential, $\vec{b} = \rho \operatorname{grad}\phi$)

$$\frac{\partial v_\alpha}{\partial t} + v^\beta \frac{\partial v_\alpha}{\partial x^\beta} = \frac{\partial \phi}{\partial x^\alpha} - \frac{1}{\rho}\frac{\partial p}{\partial x^\alpha} . \tag{19.12}$$

Here v_α and v^α are the covariant and contravariant components of the (time dependent) fluid velocity vector $\vec{v}$, ρ is the density and p is the pressure. Let

$$\nu = v_\alpha dx^\alpha$$

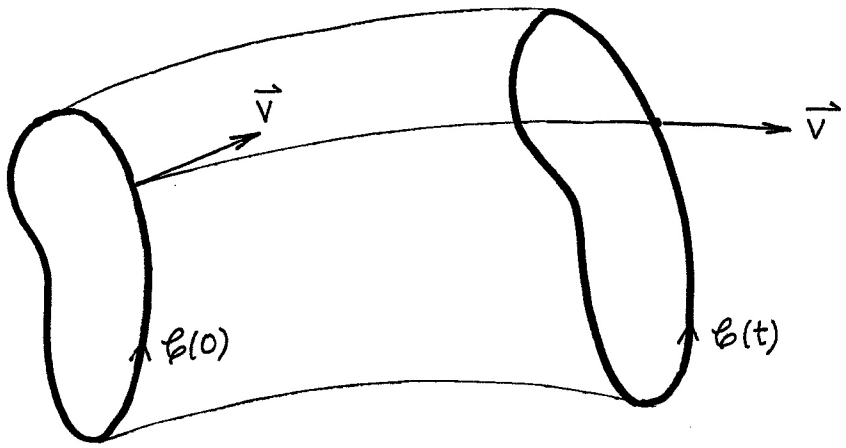
be the velocity covector.

- **Problem 34.** Assume that p is functionally related to the density ρ, i.e., there is an "equation of state" $p = p(\rho)$, and so $dp \wedge d\rho = 0$.

(i) Show that Euler's equations can be written

$$\frac{\partial \nu}{\partial t} + \mathfrak{L}_{\vec{v}}(\nu) = \vec{d}\left(\frac{\|v\|^2}{2} + \phi - \frac{p}{\rho}\right) .$$

(ii) Conclude Lord Kelvin's theorem: If $\mathbf{c}(t)$ is a closed curve that follows the motion of the fluid then the <u>circulation</u> $\oint_{\mathbf{c}} \nu$ is constant in time.



Lecture 20. A Problem Set on Hamiltonian Mechanics

(1) **Time-independent Hamiltonians.** Let M^n be the configuration space of a dynamical system; M has local coordinates $q^1,...,q^n$. The phase space is the cotangent bundle T*M, with local coordinates

$$x^1 = q^1, \ldots, x^n = q^n, \quad x^{n+1} = p_1, \ldots, x^{2n} = p_n \ .$$

On T*M there is the Poincaré 1- form λ with local coordinate expression

$$\lambda = \sum_{i=1}^{n} p_i \, dq^i$$

and the resulting Poincaré 2- form

$$\omega = d\lambda = \sum_i dp_i \wedge dq^i \ .$$

~~ **Warning.** Many books call the negative of this form ω.

Definition. A 2-form $\tilde{\omega}$ on an even dimensional manifold M^{2n} is called symplectic (and M^{2n} is then a symplectic manifold) provided $d\tilde{\omega} = 0$ and the linear transformation $(M_x^{2n})_* \to (M_x^{2n})^*$ taking tangent vectors to M^{2n} at x to cotangent vectors there given by

$$X \longmapsto i_X\tilde{\omega}$$

is nonsingular (i.e., 1-1). From Problem 24 (Lecture 12) $(i_X\tilde{\omega})_j = X^k\tilde{\omega}_{kj}$.

- **Problem M.1.** Show that ω is a symplectic form on T^*M^n. (To show 1:1 you need only check what happens to a basis $\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}, \frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_n}$ for tangent vectors.)

- **Problem M.2.** Show that

$$\omega^n \equiv \omega \wedge \cdots \wedge \omega = \pm \, n! \, dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n \ .$$

ω^n is called the (symplectic) volume form for phase space T^*M^n. Usually we shall forget the factor and use

$$dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n \quad .$$

as volume form.

Let $L = L(q,\dot{q},t)$ be a Lagrangian function on the tangent bundle TM. Recall (Lecture 9) that one has a map: $TM \to T^*M$ given by

$$p_i \equiv \frac{\partial L}{\partial \dot{q}^i} \quad .$$

Lagrange's equations $\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} = 0$ on TM translate to Hamilton's equations on T^*M

$$\begin{cases} \dfrac{dq^i}{dt} = \dfrac{\partial H}{\partial p_i} \\ \dfrac{dp_i}{dt} = -\dfrac{\partial H}{\partial q^i} \end{cases} \tag{20.1}$$

where the Hamiltonian function is defined by

$$H(q,p,t) \equiv \sum_i p_i \dot{q}_i - L(q,\dot{q},t) \quad .$$

(It is assumed in this definition that we can solve $p_i = \frac{\partial L}{\partial \dot{q}^i}$ for $\dot{q}^i$ in terms of p and q.)

We shall assume in this section (1) that H is actually independent of time t

$$H = H(q,p) \quad .$$

Then from (20.1) we can define the time independent vector field

$$\begin{aligned} X &= \sum_i X^i \frac{\partial}{\partial q^i} + X^{i+n} \frac{\partial}{\partial p_i} \\ &= \sum_i \frac{dq^i}{dt}\frac{\partial}{\partial q^i} + \frac{dp_i}{dt}\frac{\partial}{\partial p_i} \\ &= \sum_i \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} \end{aligned} \tag{20.2}$$

called the Hamiltonian vector field . Since X is time independent it defines a flow $\phi_t : T^*M^n \to T^*M^n$ called the Hamiltonian flow .

- **Problem M.3.** Show that Hamilton's equations are equivalent to

$$\boxed{i_X\omega = -dH \quad .}$$

- **Problem M.4.** Show that if X is Hamiltonian then

$$\boxed{\begin{array}{c} \mathfrak{L}_X\,\omega = 0 \\ \text{and} \\ \mathfrak{L}_X\,\omega^n = 0 \quad . \end{array}}$$

The equation $\mathfrak{L}_X\,\omega^n = 0$ shows that volumes in phase space are invariant under the Hamiltonian flow . This is Liouville's theorem!

Note that under an autonomous (i.e., time independent) Hamiltonian flow, H is a "constant of the motion"

$$\frac{dH}{dt} = X(H) = i_X dH = i_X(-i_X\omega) = 0 \quad .$$

(This is just a fancy way of saying $\dfrac{dH}{dt} = \dfrac{\partial H}{\partial q^i}\dfrac{dq^i}{dt} + \dfrac{\partial H}{\partial p_i}\dfrac{dp_i}{dt} = 0$ from (20.1).) Look now at the "level sets" of H in T^*M^n

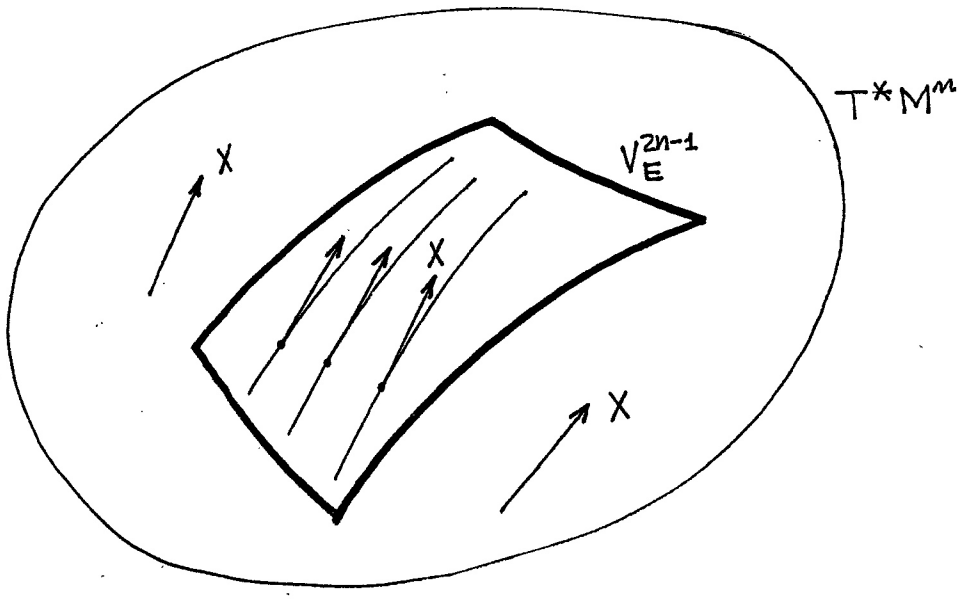
$$V_E^{2n-1} \equiv \{x \in T^*M^n \mid H(q,p) = \text{constant } E\} \quad .$$

If $dH \neq 0$ on V_E^{2n-1} (and we shall assume this is true for the constant E in question) then V_E^{2n-1} is a (2n-1)-dimensional submanifold of phase space, called the

hypersurface of constant energy $E : V_E^{2n-1}$.

• **Problem M.5.** Show that for almost all real E, $dH \neq 0$ on V_E^{2n-1}.

Since $\frac{dH}{dt} = 0$ along the integral curves of the Hamiltonian X, we conclude that X is tangent to V_E^{2n-1}.



We already know that $\mathfrak{L}_X\omega = \frac{d}{dt}\,\phi_t^*\omega \mid_{t=0} = 0$. Now note that for all small t

$$\frac{d}{dt}\,(\phi_t^*\omega) = \lim_{h\to 0} \frac{\phi_{t+h}^*\omega - \phi_t^*\omega}{h} = \phi_t^* \lim_{h\to 0} \frac{\phi_h^*\omega - \omega}{h}$$

$$\boxed{\frac{d}{dt}\,\phi_t^*\omega = \phi_t^*\,\mathfrak{L}_X\omega} \tag{20.3}$$

and consequently

$$\phi_t^*\omega = \omega \tag{20.4}$$

holds for all small t for a Hamiltonian flow ϕ_t.

Definition A differentiable map $F : T^*M \to T^*M$ that preserves the symplectic form

$$F^*\omega = \omega \,, \qquad \text{i.e.,} \quad F^*\omega_{F(x)} = \omega_x$$

is called a canonical map.

From (20.3) we see that ϕ_t is canonical. We say that the Hamiltonian vector field generates a 1- parameter group of canonical transformation .

We have seen that X is tangent to V_E^{2n-1} and consequently the integral curves that start on V_E^{2n-1} remain on V_E^{2n-1}. Consequently

$$\phi_t : V_E^{2n-1} \longrightarrow V_E^{2n-1} \quad .$$

We know that ϕ_t preserves Liouville volume ω^n on T^*M^n; we now show that there is a (2n-1) form τ_V on each V_E^{2n-1} that is also invariant under ϕ_t.

Since $dH \neq 0$ on V_E^{2n-1} by assumption on E, we claim that we can find locally, near $x \in V_E^{2n-1}$, a (2n-1) form σ^{2n-1} in T^*M^n such that

$$\boxed{\omega^n = dH \wedge \sigma^{2n-1} \quad .} \tag{20.5}$$

Proof. Since $dH \neq 0$ we know $\dfrac{\partial H}{\partial x^i} \neq 0$ for some i. For simplicity assume that $\dfrac{\partial H}{\partial q^1} \neq 0$. Introduce then a change of coordinates

$$y^1 = H \,, \; y^2 = q^2, \ldots, y^n = q^n \,, \; y^{n+1} = p_1, \ldots, y^{2n} = p_n \quad .$$

Then

$$dH \wedge dq^2 \wedge \ldots \wedge dq^n \wedge dp_1 \wedge \ldots \wedge dp_n = \frac{\partial H}{\partial q^1} dq^1 \wedge \ldots \wedge dq^n \wedge dp_1 \wedge \ldots \wedge dp_n$$

show that this is an admissible change of coordinates. Put then

$$\sigma^{2n-1} = \frac{dq^2 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n}{\dfrac{dH}{\partial q^1}} \quad . \tag{20.6}$$

Note that the choice for σ^{2n-1} is far from being unique. However, we claim

• **Problem M.6.** Let $i : V_E^{2n-1} \to T^*M^n$ be the inclusion map. Then the restriction (see Lecture 14, §(3))

$$\tau^{2n-1} \equiv i^*\sigma^{2n-1} \tag{20.7}$$

of σ to V_E^{2n-1} is a well defined form on V_E^{2n-1}, independent of the choice made in (20.5). (Hint: let σ , σ' be two choices satisfying (20.5); then show $i^*\sigma = i^*\sigma'$ by evaluating

$$dH \wedge (\sigma - \sigma')$$

on a 2n-tuple of vectors $(N,T_2,...,T_{2n})$ where N is transverse to V_E^{2n-1} and $T_2,...,T_{2n}$ are tangent to V_E^{2n-1}).

We now claim that $\tau^{2n-1} = i^*\sigma^{2n-1}$ is invariant under $\phi_t : V_E^{2n-1} \to V_E^{2n-1}$. First note that since H is invariant under ϕ_t i.e., $(\phi_t^*H))(x) = (H \circ \phi_t)(x) = H(x)$, then so is dH

$$\phi_t^*dH = d(\phi_t^*H) = dH \quad .$$

Next, $\phi_t^*\omega = \omega$ implies $\phi_t^*(\omega^n) = \omega^n$. Hence

$$\omega^n = \phi_t^*\omega^n = \phi_t^*(dH \wedge \sigma^{2n-1}) = dH \wedge \phi_t^*\sigma^{2n-1} \quad .$$

But $\omega^n = dH \wedge \sigma^{2n-1}$, and from Problem M.6, we conclude that

$$i^*\phi_t^*\sigma = i^*\sigma = \tau \quad .$$

We claim that $i^*\phi_t^*\sigma = \phi_t^*\tau$. To see this, evaluate on a (2n-1) tuple of vectors $(T_2,...,T_{2n})$ tangent to V_E^{2n-1}

$$\begin{aligned}(i^*\phi_t^*\sigma)(T_2,...,T_{2n}) &= \phi_t^*\sigma(i_*T_2,...,i_*T_{2n}) \\ &= \phi_t^*\sigma(T_2,...,T_{2n}) = \sigma((\phi_t)_*T_2,...,(\phi_t)_*T_{2n}) \\ &= (\text{since } (\phi_t)_*T_\alpha \text{ is again tangent to } V_E^{2n-1}) \\ &= \sigma(i_*(\phi_t)_*T_2,...,i_*(\phi_t)_*T_{2n}) \\ &= i^*\sigma((\phi_t)_*T_2,...,(\phi_t)_*T_{2n})\end{aligned}$$

$$= \tau((\phi_t)_* T_2, \ldots, (\phi_t)_* T_{2n})$$

$$= \phi_t^* \tau(T_2, \ldots, T_{2n})$$

and so $i^* \phi_t^* \sigma = \phi_t^* \tau$ and we do have

$$\boxed{\phi_t^* \tau = \tau \quad .} \tag{20.8}$$

We now write down a specific (but rather arbitrary) expression for τ^{2n-1} that is used in books on statistical mechanics.

Let ds^2 be any Riemannian metric on T^*M, in a coordinate patch, for which ω^n is the volume form vol^{2n}. (Usually T^*M does not come equipped with a Riemmanian metric; however locally we can always write in a coordinate patch (q,p)

$$ds^2 = (dq^1)^2 + \cdots + (dq^n)^2 + (dp_1)^2 + \cdots + (dp_n)^2$$

and then

$$vol^{2n} = \sqrt{g}\, dq^1 \wedge \cdots \wedge dq^n \wedge dp_1 \wedge \cdots \wedge dp_n$$

will do since $g = 1$.) In any metric $\operatorname{grad} H = \nabla H$ is normal to V_E^{2n-1}. Thus $\nabla H / \| \nabla H \|$ is a unit normal to V_E^{2n-1}. Then the (2n-1) volume form for V_E^{2n-1} (see the discussion leading to (19.5)) is

$$vol_V^{2n-1} = i_{\frac{\nabla H}{|\nabla H|}}\, vol^{2n} \quad .$$

However, although dH is invariant under ϕ_t, ∇H is not, since the rather arbitrary metric is not invariant under ϕ_t. We claim

$$\boxed{\tau = \frac{vol_V^{2n-1}}{|\nabla H|} = i_{\frac{\nabla H}{|\nabla H|^2}}\, vol^{2n}} \tag{20.9}$$

is an invariant (2n-1) form for V_E^{2n-1}.

• **Problem M.7.** Show that this τ is invariant by showing that if $\frac{\nabla H}{\|\nabla H\|}$, $T_2,\ldots,T_{2n}$ are orthonormal, then

$$dH \wedge i_{\frac{\nabla H}{\|\nabla H\|^2}} \operatorname{vol}^{2n}\left(\frac{\nabla H}{\|\nabla H\|}, T_2, \ldots, T_{2n}\right) = \pm 1$$

and consequently $dH \wedge i_{\frac{\nabla H}{\|\nabla H\|^2}} \operatorname{vol}^{2n} = \pm \operatorname{vol}^{2n}$ as needed.

(2) **Time dependent Hamiltonians.** When $H = H(q,p,t)$ we consider H as a function on the "extended phase space" $T^*M^n \times \mathbb{R}$. Hamilton's equations are still

$$\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$$

but note now that $\frac{dH}{dt} = \frac{\partial H}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial H}{\partial p_i}\frac{dp_i}{dt} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}$, and so H is no longer a constant of the motion.

Define the new Hamiltonian vector field on $T^*M \times \mathbb{R}$ by

$$X = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial}{\partial t} \tag{20.10}$$

and so $X^{2n+1} = 1$, where $x^{2n+1} = t$.

Define a new Poincaré 1 form on $T^*M \times \mathbb{R}$ by

$$\boxed{\Lambda = \sum_{i=1}^{n} p_i dq^i - H\,dt} \tag{20.11}$$

and

$$\boxed{\Omega = d\Lambda = \sum_{i=1}^{n} dp_i \wedge dq^i - dH \wedge dt} \tag{20.12}$$

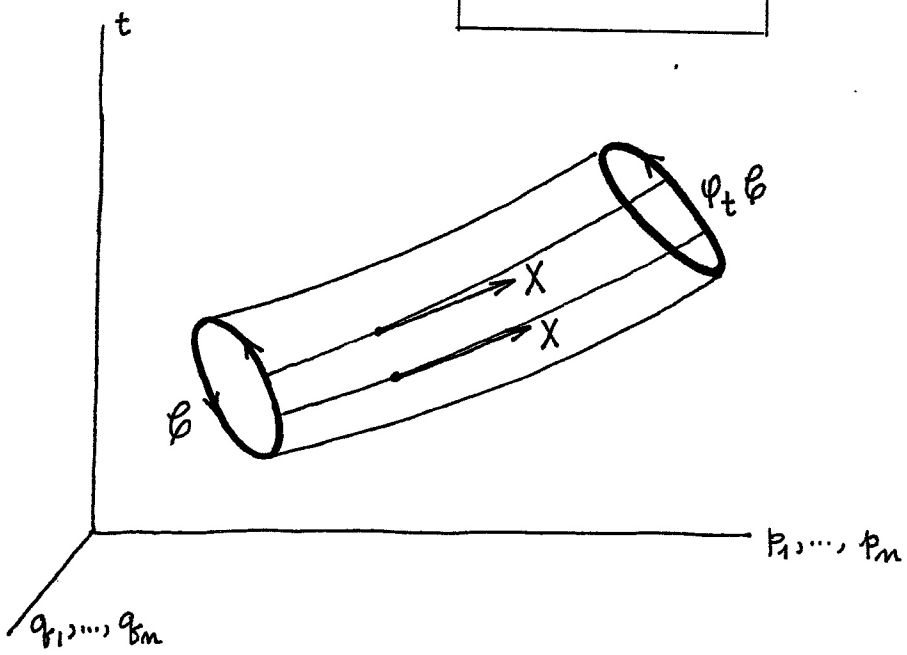
where now of course $df = \frac{\partial f}{\partial q^i}dq^i + \frac{\partial f}{\partial p_i}dp_i + \frac{\partial f}{\partial t}dt$, etc.

- **Problem M.8.** Show that Hamilton's equations and $\frac{dH}{dt} = \frac{\partial H}{\partial t}$ are equivalent to

$$\boxed{i_X \Omega = 0 \quad .}$$

Let $\phi_t : T^*M^n \times \mathbb{R} \to T^*M^n \times \mathbb{R}$ be the Hamiltonian flow generated by the new Hamiltonian vector field (20.10).

- **Problem M.9.** Show $\boxed{\mathcal{L}_X \Omega = 0}$ if X is Hamiltonian.



- **Problem M.10.** Let C be a closed curve in $T^*M \times \mathbb{R}$. Show that

$$\boxed{\oint_C p_i dq^i - H\, dt = \oint_{\phi_t C} p_i\, dq^i - H\, dt} \qquad (20.13)$$

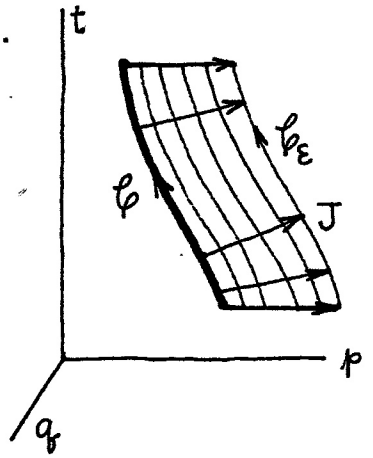
Hint: Look at the 2-manifold with boundary swept out by C under ϕ_t and consider Ω evaluated on a pair of vectors tangent to this surface.

Definition. Let C be a (not closed) curve in $T^*M \times \mathbb{R}$. Then the **Action** S associated with C is the integral

$$\boxed{S = \int_C \Lambda = \int_C p_i \, dq^i - H \, dt \quad .}$$

We shall compute the "first variation of action" under variations of C, demanding that the variations of q and t (but not p) vanish at the endpoints of C. Let J be a vector field on $T^*M \times \mathbb{R}$ whose flow ψ_ϵ generates such a variation. Then we know

$$\delta S \equiv \left(\frac{d}{d\epsilon} \int_{C_\epsilon} \Lambda \right)_{\epsilon=0} = \int_C \mathcal{L}_J \Lambda \quad .$$



- **Problem M.11.** Show

$$\boxed{\delta S = \int_C i_J \Omega \quad .}$$

Suppose now that the first variation vanishes, $\delta S = 0$ for all such variations, i.e., for all such J.

Problem M.12. Show that if $\delta S = 0$ for all such variation vector fields J then the tangent vector T to C must satisfy

$$i_T \Omega = 0$$

and so C^* must be a solution to Hamilton's equations. This is Hamilton's Principle of Stationary Action.

You may use the "fundamental lemma of the calculus of variations"; if f is continuous

*Use time t as parameter for C.

and if

$$\int_a^b f(t)\eta(t)dt = 0$$

for all continuous η for a which $\eta(a) = \eta(b) = 0$, then $f(t) \equiv 0$.

Lecture 21. Volterra's Theorem

(1) **Stokes' theorem again.** We shall accept (for a proof, see [A,M,R], p.418 ff.) the fact that Stokes' theorem (15.1)

$$\int_{V^p} d\beta^{p-1} = \int_{\partial V} \beta^{p-1}$$

holds, in fact, if V^p is merely a compact oriented p-manifold with <u>piecewise smooth</u> boundary ∂V. Thus we can apply Stokes' theorem to triangles, cubes, etc. Of course for something like a cube the theorem can be proved directly; this is the proof given in most calculus books when $p = 3$.

(2) **Closed forms and exact forms.** A form β^p is <u>closed</u> provided $d\beta^k = 0$. Thus

(i) $d\beta^0 = 0 \Rightarrow \beta^0$ is a constant function.

(ii) $d\beta^1 = 0 \Rightarrow (\partial_i b_j - \partial_j b_i) = 0.$

(iii) $d\beta^2 = 0 \Rightarrow (\partial_i b_{jK} + \partial_k b_{ij} + \partial_j b_{ki}) = 0.$

A form β^p is <u>exact</u> provided $\beta^p = d\alpha^{p-1}$ for some (p-1)-form α. <u>Certainly exact forms are closed</u>,

$$\beta^p = d\alpha^{p-1} \Rightarrow d\beta^p = d^2\alpha^{p-1} = 0 \quad .$$

• **Problem 35.** Show that the product of two closed forms is closed and that the product of a closed and an exact form is exact.

Are, conversely, closed forms exact? We shall need the useful and simple

Proposition. (i) If $\beta^p = d\alpha^{p-1}$ is <u>exact</u> and if V^p is a compact oriented manifold without boundary, then

$$\int_{V^p} \beta^p = \int_{V^p} d\alpha^{p-1} = \int_{\partial V^p = 0} \alpha^{p-1} = 0 \quad .$$

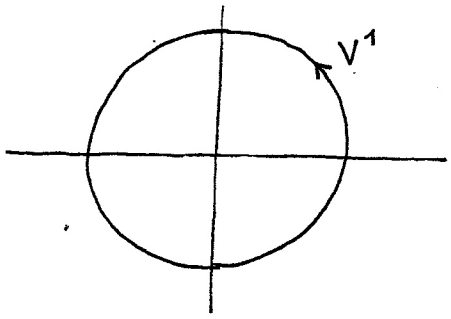
(ii) If β^p is closed and the piecewise smooth V^p is the boundary of a compact oriented W^{p+1}, $\partial W^{p+1} = V^p$, then

$$\int_{V^p} \beta^p = \int_{\partial W^{p+1}} \beta^p = \int_{W^{p+1}} d\beta^p = 0 \quad .$$

Example. Let $M^2 = \mathbb{R}^2 - 0$ be the plane with the origin deleted and let $\beta^1 = \frac{x\,dy - y\,dx}{x^2+y^2}$. β is a smooth form on $\mathbb{R}^2 - 0$ and one easily checks that $d\beta^1 = 0$. In fact we all know that

$$\beta^1 = \text{"}d\theta\text{"} = d \tan^{-1}\left(\frac{y}{x}\right) \quad .$$

It looks like β^1 is exact, but this is not clear since θ is not a differentiable single valued function on all of $\mathbb{R}^2 - 0$. In fact if we let V^1 be the unit circle



$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

then V^1 is a compact oriented 1-manifold without boundary and

$$\oint_{V^1} \beta^1 = \oint_{V^1} \frac{x\,dy - y\,dx}{x^2+y^2} = \oint \text{"}d\theta\text{"} = 2\pi \neq 0$$

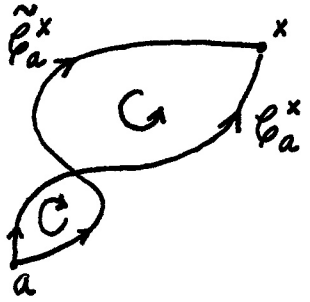
and from our proposition we conclude that β^1 is closed but not exact ! Note that V^1 is not a boundary, because we have had to delete the origin from the unit disc.

Theorem. Let M^n be a manifold having the property that every closed oriented piecewise smooth curve V^1 on M^n is the boundary of a compact oriented 2- dimensional submanifold W^2 (we then say that "M^n has 1st Betti number $= 0$"). Then every closed 1- form β^1 on M^n is exact .

Proof. The proof is the same as is given in calculus books. We wish to exhibit a function f on M^n such that $\beta^1 = df$, i.e., $b_i = \dfrac{\partial f}{\partial x^i}$ in terms of local coordinates. Let a be any given point of M^n. Define the function $f : M^n \to \mathbb{R}$ by

$$f(x) = \int_{C_a^x} \beta^1$$

where C_a^x is <u>any</u> piecewise smooth curve from a to x. First note that $f(x)$ is in fact <u>independent of the path</u> C_a^x chosen, for if $\tilde{C}_a^x$ is another such path then $C_a^x - \tilde{C}_a^x$ is a closed



oriented curve starting and ending at a. Since M^n has 1st Betti number $= 0$ we know there is some compact oriented surface W^2 such that $\partial W^2 = C_a^x - \tilde{C}_a^x$. Hence

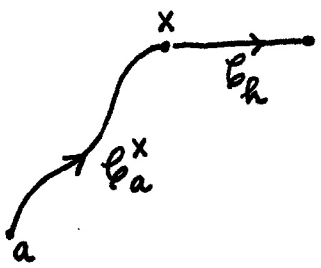
$$\int_{C_a^x} \beta^1 - \int_{\tilde{C}_a^x} \beta^1 = \int_{C_a^x - \tilde{C}_a^x} \beta^1 = \int_{\partial W^2} \beta^1 = \int_{W^1} d\beta^1 = 0 \qquad (21.1)$$

as desired.

We can now compute df, knowing that $f(x)$ is well defined independent of the path joining a to x. Let the point x lie in a coordinate patch and have coordinates $x^1,\ldots,x^n$.

$$\frac{\partial f}{\partial x^i} = \lim_{h \to 0} \frac{f(x^1,\ldots,x^i+h,\ldots,x^n) - f(x^1,\ldots,x^n)}{h} \quad .$$

To compute $f(x^1,\ldots,x^i+h,\ldots,x^n)$ we need a path joining a to the point with coordinates $(x^1,\ldots,x^i+h,\ldots,x^n)$. For this path we choose C_a^x followed by the straight line in the coordinate system joining $(x^1,\ldots,x^n)$ to $(x^1,\ldots,x^i+h,\ldots,x^n)$; call this last path C_h.



$$C_h: \quad x^j(t) = x^j \quad j \neq i$$
$$x^i(t) = x^i + th$$

Then

$$\frac{f(x^1,\ldots,x^i+h,\ldots,x^n)-f(x^1,\ldots,x^n)}{h} = \frac{1}{h}\left(\int_{C_a^x+C_h} \beta^1 - \int_{C_a^x} \beta^1\right)$$

$$= \frac{1}{h}\int_{C_a^x+C_h-C_a^x} \beta^1 = \frac{1}{h}\int_{C_h} \beta^1$$

$$= \frac{1}{h}\int_{C_h} \sum_j b_j\, dx^j = \frac{1}{h}\int_{C_h} b_i\, dx^i \qquad \text{(no sum)}$$

$$= \frac{1}{h}\int_0^1 b_i(x^1,\ldots,x^i+th,\ldots,x^n)h\,dt$$

$$= \int_0^1 b_i(x^1,\ldots,x^i+th,\ldots,x^n)dt = b_i(x^1,\ldots,x^i+\tau h,\ldots,x^n)$$

where $0 \leq \tau \leq 1$. Clearly then

$$\frac{\partial f}{\partial x^i} = \lim_{h\to 0} b_i(x^1,\ldots,x^i+\tau h,\ldots,x^n) = b_i(x^1,\ldots,x^n) \quad . \qquad \square$$

For future purposes (the next lecture) we shall restate this theorem. The essential ingredient was the statement (21.1). Using the fact that $C_a^x - \tilde{C}_a^x$ is a closed curve we may say

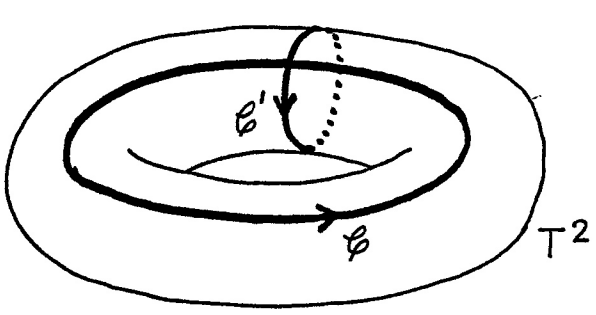
Corollary. If M^n is <u>any</u> manifold, if β^1 is a <u>closed</u> 1-form whose integrals over <u>every closed</u> curve C vanishes

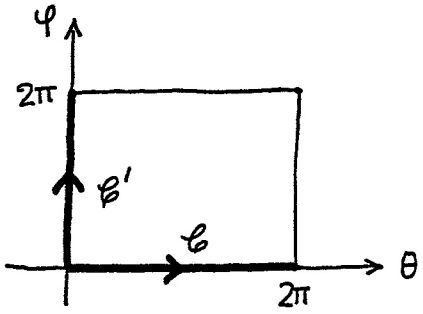
$$\int_C \beta^1 = 0 \qquad \text{all closed } C$$

then β^1 is exact.

If $\beta^1 = df$, we say that "β is derivable from a (scalar) <u>potential</u> f."

Example. There are closed curves C and C' on the 2-dimensional torus T^2 that do <u>not</u> bound oriented surfaces;





we should not expect that every closed β is exact. In fact $\int_{\mathcal{C}} \text{“d}\theta\text{”} = 2\pi = \int_{\mathcal{C}^1} \text{“d}\phi\text{”}$ shows that neither “$d\theta$” nor “$d\phi$” is exact.

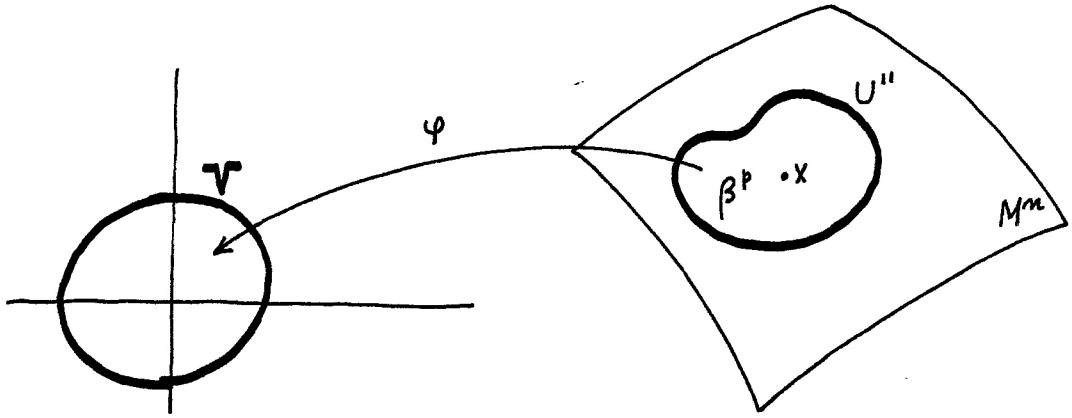
We need to consider closed forms β^p of higher degree. We shall state (but not prove) a very general theorem in the next lecture. Here we shall prove a weaker result, but it will be accompanied by an explicit formula.

(3) **The converse of the Poincaré lemma.** The statement that exact forms are closed

$$d(d\alpha^{p-1}) = d^2\,\alpha^{p-1} = 0$$

is usually called the Poincaré lemma. It should be appreciated that Poincaré utilized this result before the machinery of exterior calculus had been developed!

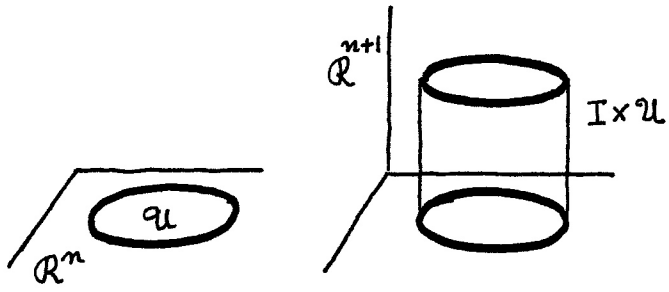
We claim that there is a partial converse of the Poincaré lemma, namely every closed p-form, $p > 0$, is <u>locally</u> exact. That is, if $d\beta^p = 0$ in a neighborhood U of a point x on a manifold M^n, then there is some perhaps smaller neighborhood U' of x and a (p-1)-form α^{p-1} defined on U' such that $\beta^p = d\,\alpha^{p-1}$ on U'. To prove this result it is sufficient to prove it for Euclidean space $\mathbb{R}^n$; this is because $x \in M^n$ always has a small neighborhood U'' that is diffeomorphic to an open ball $V \subset \mathbb{R}^n$ under a coordinate map ϕ.



Since $\phi : U'' \to V$ is a diffeomorphism, ϕ^{-1} exists and $\beta^p = (\phi \circ \phi^{-1})^* \beta^p = (\phi^{-1})^*(\phi^* \beta^p)$. Then if β^p is closed on M^n, $\phi^* \beta^p$ is closed on $V \subset \mathbb{R}^n$. If we have the converse of Poincaré on $V \subset \mathbb{R}^n$ then $\phi^* \beta^p = d\tilde{\alpha}^{p-1}$ shows $\beta^p = (\phi^{-1})^* d\tilde{\alpha}^{p-1} = d[(\phi^{-1})^* \tilde{\alpha}^{p-1}]$ as desired. We need only prove the theorem in $\mathbb{R}^n$. Before doing this we need some preliminaries.

We are working in Euclidean space. Let $u \subset \mathbb{R}^n$ be open, let $I = [0,1]$, and form the "cylinder" $I \times u \subset \mathbb{R}^{n+1}$

$$I \times u = \{(t,x) \mid t \in I,\ x \in u\}$$



We have the maps

$$j_1 : u \to I \times u;\quad x \mapsto (1,x)$$
$$j_0 : u \to I \times u;\quad x \mapsto (0,x)$$

Let β^{p+1} be a $(p+1)$ form on $I \times u$. β has the form

$$\beta^{p+1} = dt \wedge \sum_{\substack{I \\ \to}} a_I(t,x)dx^I + \sum_{\substack{J \\ \to}} b_J(t,x)dx^J$$

where $I = i_1, \ldots, i_p$ and $J = j_1, \ldots, j_{p+1}$.

Note that $j_1^* \beta^{p+1} = \sum_{\substack{J \\ \to}} b_J(1,x)dx^J$ just says put $t = 1$, while

$$j_0^*\beta^{p+1} = \sum_{\substack{J \\ \rightarrow}} b_J(0,x)dx^J \quad \text{just says put} \quad t = 0 \quad .$$

These are both forms on $u \subset \mathbb{R}^n$.

Define an operator (for all $p \geq 0$)

$$K : (p+1) \text{ forms on } I \times u \to p \text{ forms on } u$$

by

$$\begin{cases} K(b_J(t,x)dx^J) &= 0 \\ K(a_I(t,x)dt \wedge dx^I) &= \left(\int_0^1 a_I(t,x)dt \right) dx^I \end{cases} \tag{21.2}$$

and extend by linearity to all forms. We then have the

Homotopy Formula

$Kd\beta + dK\beta = j_1^*\beta - j_0^*\beta$ for each β^{p+1} on $I \times u$. (21.3)

Note that both sides are $(p+1)$ forms on u.

Proof. Case (1) β^{p+1} is a monomial of the form $b_J(t,x)dx^J$. Then $K\beta = 0$ and

$$d\beta = \frac{\partial b_J}{\partial t}\, dt \wedge dx^J + \text{terms not involving } dt$$

therefore

$$Kd\beta = \left(\int_0^1 \frac{\partial b_J(t,x)}{\partial t}\, dt \right) dx^J = b_J(1,x)dx^J - b_J(0,x)dx^J$$

$$= j_1^*\beta - j_0^*\beta$$

Case (2) $\beta^{p+1} = a_I(t,x)dt \wedge dx^I$, then $j_1^*\beta = 0 = j_0^*\beta$ since in both cases we are to put $dt = 0$. Further

$$K\beta = \left(\int_0^1 a_I(t,x)dt \right) dx^I \quad \text{is independent of} \quad t$$

and, therefore

$$dK\beta = \sum_k \left(\int_0^1 \frac{\partial a_I(t,x)}{\partial x^k}\, dt \right) dx^k \wedge dx^I$$

while

$$d\beta = -\sum_k \frac{\partial a_I(t,x)}{\partial x^k}\, dt \wedge dx^k \wedge dx^I$$

and

$$Kd\beta = -\sum_k \left(\int_0^1 \frac{\partial a_I(t,x)}{\partial x^k}\, dt \right) \wedge dx^k \wedge dx^I \quad . \qquad \square$$

Definition. A region $\mathcal{U} \subset M^n$ is said to be <u>deformable to a point</u> $x_0 \in \mathcal{U}$ (or contractible) if there is a differentiable map

$$\mathfrak{D} : I \times \mathcal{U} \longrightarrow \mathcal{U}$$

such that

$$\begin{cases} \mathfrak{D}(0,x) = x & x \in \mathcal{U} \\ \mathfrak{D}(1,x) = x_0 \end{cases}$$

(We can then define a 1-parameter family of maps $\mathfrak{D}_t : \mathcal{U} \to \mathcal{U}$ by $\mathfrak{D}_t(x) = \mathfrak{D}(t,x)$; then $\mathfrak{D}_0 =$ <u>identity</u> map while $\mathfrak{D}_1$ is the <u>constant</u> map $\mathcal{U} \to x_0$.)

Note that we then have

$$\mathcal{U} \xrightarrow{j_i} I \times \mathcal{U} \xrightarrow{\mathfrak{D}} \mathcal{U} \qquad i = 0,1$$

and $\mathfrak{D}_0 \equiv \mathfrak{D} \circ j_0 =$ identity and $\mathfrak{D}_1 \equiv \mathfrak{D} \circ j_1 =$ constant map.

Theorem (Volterra). Let $\mathcal{U} \subset \mathbb{R}^n$ be deformable to a point of $\mathcal{U}$. Let β^{p+1} be a $(p+1)$ form on $\mathcal{U}$ that is <u>closed,</u> $d\beta = 0$. Then $\exists$ a p-form β^p on $\mathcal{U}$ such that $d\alpha^p = \beta^{p+1}$, i.e., β is <u>exact</u> on $\mathcal{U}$.

Proof. $\mathfrak{D}_0 =$ identity $\Rightarrow$ $\mathfrak{D}_0^* =$ identity. $\mathfrak{D}_1$ constant map $\Rightarrow$ $\mathfrak{D}_1^* = 0$ ($\mathfrak{D}_1^*$ says put $x = x_0$, i.e., $dx^i = 0$). Then

$$\begin{aligned}\beta^{p+1} &= \mathfrak{D}_0^*\beta - \mathfrak{D}_1^*\beta = (\mathfrak{D}\circ j_0)^*\beta - (\mathfrak{D}\circ j_1)^*\beta \\ &= j_0^*\mathfrak{D}^*\beta - j_1^*\mathfrak{D}^*\beta = -dK(\mathfrak{D}^*\beta) - Kd(\mathfrak{D}^*\beta) \\ &= -dK(\mathfrak{D}^*\beta) - K\mathfrak{D}^*d\beta = -dK(\mathfrak{D}^*\beta)\end{aligned} \quad .$$

$$\therefore \qquad \boxed{\beta^{p+1} = d(-K\mathfrak{D}^*\beta)} \qquad (21.4)$$

an explicit formula for $\beta^{p+1} = -K\mathfrak{D}^*\beta$. □

Let now $\beta^p = \sum_{I}^{\rightarrow} b_I(y)dy^I$ be a closed form in $\mathbb{R}^n$, written in Cartesian coordinates. Consider the deformation $\mathfrak{D}: I \times \mathbb{R}^n \to \mathbb{R}^n$ given by

$$y = (1-t)(x), \quad \text{i.e.,} \quad y^i = (1-t)x^i \quad .$$

• **Problem 36.** Show that (21.4) gives then $\beta^p = d\alpha^{p-1}$ where

$$\alpha^{p-1} = \sum_{J}^{\rightarrow} a_J dx^J, \quad a_J(x) = \int_0^1 t^{p-1} \sum_{j=1}^{n} x^j b_{jJ}(tx)dt$$

and write this out in vector notation in the case of $\mathbb{R}^3$.

(4) **Finding potentials without Volterra's formula.** In some simple situations one may exhibit potentials with very little work.

Example. Consider the electric field $\vec{E}$ due to a charge q at the origin. Coulomb's law gives, in spherical coordinates

$$\vec{E} = \frac{q}{r^2}\frac{\partial}{\partial r} \quad .$$

Using, as usual, the metric $ds^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2)$ we see that the covector associated to $\vec{E}$ is

$$e^1 = \frac{q}{r^2}\,dr = d\left(-\frac{q}{r}\right) \qquad r \geq 0 \quad .$$

The function $-\frac{q}{r}$ is called the scalar potential. The 2-form associated to $\vec{E}$ is

$$\mathfrak{D}^2 = i_{\vec{E}}\, \mathrm{vol}^3$$

and since $d\mathfrak{D}^2 = 4\pi\rho\ \mathrm{vol}^3$ (Gauss' law) we see that $d\mathfrak{D}^2 = 0$, $\mathfrak{D}^2$ is closed, for $r > 0$ since charge density ρ vanishes outside the origin. Now the region ($\mathbb{R}^3$-origin) is not contractible, but if we consider any ball not containing $\vec{0}$ we know that we can find a potential $\mathfrak{A}^1$ there so that $d\mathfrak{A}^1 = \mathfrak{D}^2$. Rather than use Volterra's formula we shall just write it down. Obviously we should use spherical coordinates, center at origin. Then $\mathrm{vol}^3 = r^2 \sin\theta\ dr \wedge d\theta \wedge d\phi$ and so

$$\mathfrak{D}^2 = i_{\frac{q}{r^2}\frac{\partial}{\partial r}}\, r^2 \sin\theta\ dr \wedge d\theta \wedge d\phi = q \sin\theta\ d\theta \wedge d\phi \quad .$$

Thus, e.g. $\mathfrak{D}^2 = d(-q\cos\theta\ d\phi)$; $\mathfrak{A}^1 = -q\cos\theta\ d\phi$ is a choice for potential! Note that the spherical coordinate system behaves badly at $\theta = 0$ and $\theta = \pi$, i.e., along the entire z axis. Hence $\mathfrak{A}^1$ is a well defined potential everywhere except along the entire z axis. Note that we can also write

$$\mathfrak{D}^2 = d\{q(1 - \cos\theta)d\phi\} \quad .$$

Since $1 - \cos\theta = 0$ when $\theta = 0$, this expression

$$\mathfrak{A} = q(1 - \cos\theta)d\phi$$

is actually a well defined potential everywhere except on the negative z axis.

We get the same result when we consider the magnetic field $\mathfrak{B}^2$ due to a magnetic monopole at the origin, of strength μ. Then $\mathfrak{B}^2 = d\{\mu(1 - \cos\theta)d\phi\}$ and the vector potential has again a "Dirac string" of singularities along the negative z axis.

- **Problem 37.** Ampere's law (16.16) for current j in an infinite wire along z axis certainly is satisfied by magnetic field

$$\mathfrak{H}^1 = 2j\ d\phi \quad .$$

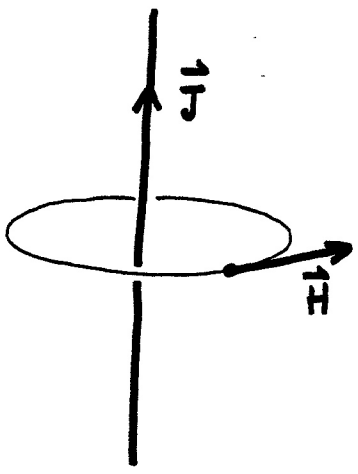
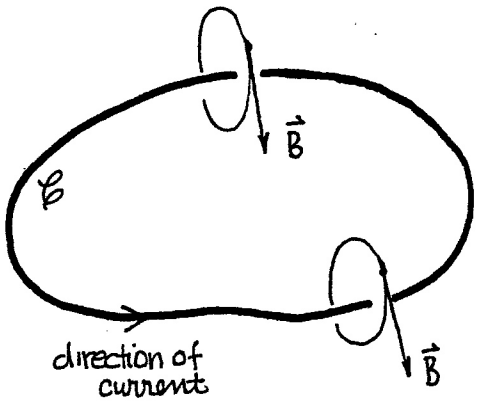
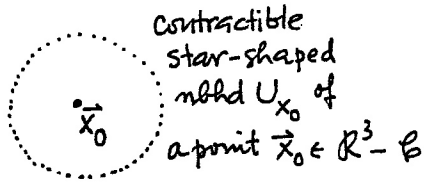


Exhibit a vector potential $\mathfrak{A}^1$, $\mathfrak{B}^2 = d\mathfrak{A}^1$ as we did above (in appropriate coordinates).

Lecture 22. Topology, deRham, and Electromagnetism

deRham's Theorem. Consider a current (of magnitude $j(t)$) flowing in wire loop C in $\mathbb{R}^3$. While the magnetic field may not be differentiable inside (and on the surface of) the wire, it is a differentiable vector field outside the wire, i.e., in $\mathbb{R}^3 - C$ (we use the right handed orientation of $\mathbb{R}^3$ to pick out $\vec{B}$)





This, of course, is highly idealized; the wire is "really" a 3-dimensional solid torus, but this has no effect on the problem we shall consider. Since $\operatorname{div} \vec{B} = 0$, we know that locally $\vec{B} = \operatorname{curl} \vec{A}$ for some local "vector potential" $\vec{A}$. For example, if $\vec{x}_0 \in \mathbb{R}^3 - C$, then there is a star shaped neighborhood U_{x_0} of x_0 lying in $\mathbb{R}^3 - C$, and the converse of Poincaré's lemma (i.e., Volterra's theorem of Lecture 21) assures us that such an $\vec{A}$ exists in U_{x_0}. Note, however, that the entire complement $\mathbb{R}^3 - C$ of the wire is not star shaped; any ttempt to contract $\mathbb{R}^3 - C$ to a single point is bound to bring some points of $\mathbb{R}^3 - C$ into contact with C. The converse to the Poincaré lemma will not allow us to find a single $\vec{A}$ for all of $\mathbb{R}^3 - C$.

· deRham's theorem (1931), will , however, allow us to, theoretically, find such an $\vec{A}$!

Before discussing this we should mention that there is a rather general integral formula exhibiting $\vec{A}$ in terms of the current j. This is given in most texts; a readable treatment is in, e.g. H. B. Phillips old book "*Vector Analysis.*" We shall discuss such things next quarter when we describe the "Hodge decomposition." We are concerned here with the general principle involved; for example, the above mentioned integral formula is specifically for Euclidean space and the singularities of the current must be handled carefully. Our goal is a principle that makes sense in any space.

Definition. A p- chain W^p on M^n is a finite, formal, linear combination (with real coefficients) of compact, oriented , p-dimensional submanifolds (with piecewise smooth boundaries)

$$W^p = \sum_i a_i V^p_{(i)} \qquad a_i \in \mathbb{R} \quad .$$

As usual we define

$$\int_{W^p} \alpha^p = \int_{\sum_i a_i V^p_{(i)}} \alpha^p = \sum_i a_i \int_{V^p_{(i)}} \alpha^p \quad .$$

Definition. The boundary ∂W^p of a p-chain is the (p-1) chain

$$\partial W^p = \sum a_i \, \partial V^p_{(i)} \quad .$$

Definition. A p-chain W^p with empty boundary

$$\partial W^p = 0$$

is called a p- cycle.

We have used many times the fact that $\partial(\partial V^p_{(i)}) = 0$; consequently every boundary ∂W^p is a cycle, $\partial(\partial W^p) = 0$. We already know that if β is exact and V^p is a p-cycle, then $\int_V \beta = 0$ (see Proposition, Lecture 21).

DeRham's First Theorem. Let β^p be a closed p-form, $d\beta^p = 0$, on a manifold M^n. Then β^p is exact , $\beta^p = d\alpha^{p-1}$, provided its integral over every p- cycle vanishes

$$\int_{V^p} \beta^p = 0 \quad .$$

The integral of β over V^p is called the period of β on V^p; thus β is exact if its periods on all p-cycles vanish.

For example, in our electrical case, $\beta^2 = i_{\vec{B}}\,\mathrm{vol}^3$ has $d\beta^2 = 0$ in $\mathbb{R}^3 - \mathbf{C}$. Let V^2 be any 2-cycle of $\mathbb{R}^3 - \mathbf{C}$, i.e., basically, any closed, oriented surface that does not touch the wire $\mathbf{C}$. We need to show that the period of β^2 on V^2 vanishes

$$\iint_{V^2} \vec{B} \cdot d\vec{S} = \iint_{V^2} \beta^2 = 0 \quad .$$

But Faraday's law (16.11) states that (for the fixed surface V^2)

$$\frac{d}{dt} \iint_{V^2} \beta^2 = \iint_{V^2} \frac{\partial \beta^2}{\partial t} = - \oint_{\partial V^2} \mathcal{E} \tag{22.1}$$

where $\mathcal{E}$ is the electric 1-form. But $\partial V^2 = 0$ since V^2 is a cycle, hence

$$\iint_{V^2} \beta^2 \qquad \text{is constant in time} \quad .$$

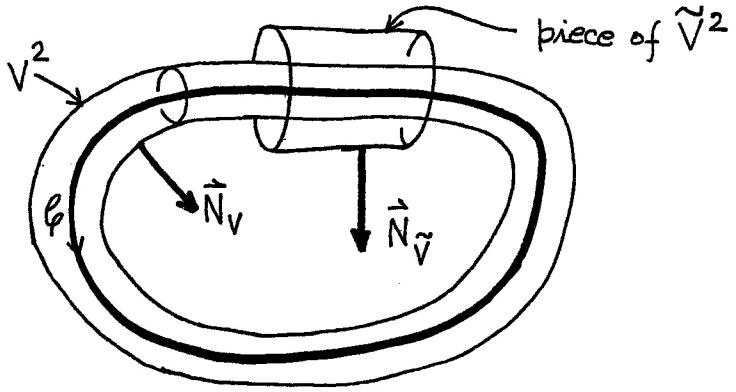
But if we shut off the current in $\mathbf{C}$ the field $\vec{B}$ will vanish, hence the constant must be zero. □

It should be mentioned that even when $\mathbf{C}$ is a planar circular loop in $\mathbb{R}^3$ it is not easy to exhibit the potential $\vec{A}$ (Morse and Feshback exhibit one, but it is a horror.)

Recall (Proposition, Lecture 21) that if the p-cycle V^p <u>bounds</u>, i.e., $V^p = \partial W^{p+1}$ then, of course, β has vanishing period on V^p if $d\beta = 0$

$$\int_{\partial W^{p+1}} \beta^p = \int_{W^{p+1}} d\beta^p = 0 \quad .$$

deRham's first theorem makes it clear that an important role is played by those p-cycles on M^n that do not bound. Consider again our loop $\mathbf{C}$ in $\mathbb{R}^3$, but consider a general closed 2-form γ^2, $d\gamma^2 = 0$ on $\mathbb{R}^3 - \mathbf{C}$, not necessarily the magnetic 2-form. Since we cannot invoke Faraday's law we need to look explicitly at the periods of γ^2. It seems clear that the only 2-cycles we need consider are those that in a sense surround $\mathbf{C}$; all the other will bound in $\mathbb{R}^3 - \mathbf{C}$ and, of course, will be of no interest.



For instance, a toroidal tube V^2 around $\mathbf{C}$ will not bound in $\mathbb{R}^3 - \mathbf{C}$ (of course it does bound in $\mathbb{R}^3$, but γ^2 is only defined in $\mathbb{R}^3 - \mathbf{C}$). I've oriented V^2 by using the right hand orientation of $\mathbb{R}^3$ and exhibiting a normal $\vec{N}_V$ to V. Then if $\tilde{V}^2$ is a second, larger toroid surrounding $\mathbf{C}$, it seems clear that $\tilde{V}^2 - V^2$ is the boundary of the 3-dimensional region W^3 between $\tilde{V}^2$ and V^2

$$\tilde{V}^2 - V^2 = \partial W^3 \quad .$$

Here I have oriented W^3 using the right hand orientation. I have put a minus sign before V^2 because the outward pointing normal to W^3 is $N_{\tilde{V}}$ on $\tilde{V}^2$ and $-N_V$ on V! Then

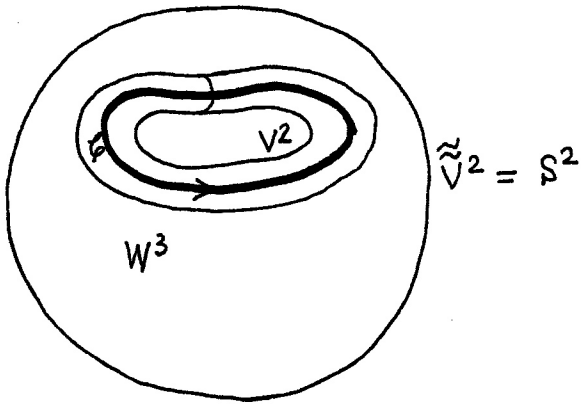
$$0 = \int_{\partial W^3} \gamma^2 = \int_{\tilde{V}^2 - V^2} \gamma^2 = \int_{\tilde{V}^2} \gamma^2 - \int_{V^2} \gamma^2 \quad .$$

Since $\tilde{V}^2 - V^2$ bounds in $\mathbb{R}^3 - \mathbf{C}$, a closed γ^2 will have the same periods on $\tilde{V}^2$ as it does on V^2!

Definition. Two p-cycles are homologous if they differ by a boundary

$$\tilde{V}^p \sim V^p \qquad \text{means} \qquad \tilde{V}^p - V^p = \partial W^{p+1} \quad .$$

Thus to check the periods of γ^p we only need look at the period on each representative of a " homology class". Thus once we know $\int_{V^2} \gamma^2$ we needn't look at $\int_{\tilde{V}^2} \gamma^2$ nor at $\int_{\tilde{\tilde{V}}^2} \gamma^2$, where $\tilde{\tilde{V}}^2$ is a sphere that encloses $\mathbf{C}$



because $S^2 - V^2$ is again the boundary of the $\tilde{W}^3$ between S^2 and V^2!

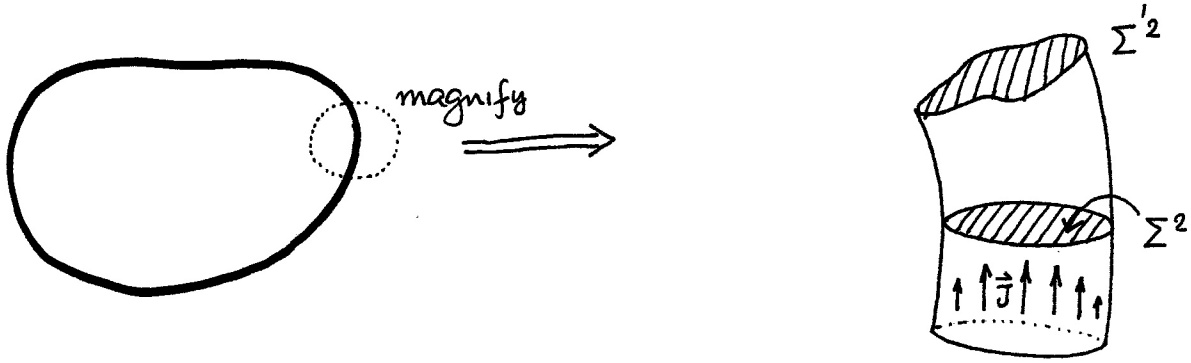
We see then that γ^2 is exact if and only if $\int_{V^2} \gamma^2 = 0$, and this of course depends on the specific form γ^2.

For a second example, consider again the same current loop $\mathbf{C}$ in $\mathbb{R}^3$, but this time consider the magnetic covector $\mathcal{H}^1$, a 1-form. For a constant current in $\mathbf{C}$ no fields will change in time and so the Ampere-Maxwell (16.16) reduces to Ampere's law

$$d\mathcal{H}^1 = 4\pi j^2$$

and so the 1-form $\mathcal{H}^1$ is closed, $d\mathcal{H}^1 = 0$, outside the wire, i.e., in $\mathbb{R}^3 - \mathbf{C}$.

If it makes you nervous to have a 2-form j^2 confined to the 1-dimensional wire it would be better to thicken the wire $\mathbf{C}$ to a 3-dimensional tube $\tilde{\mathbf{C}}^3$;



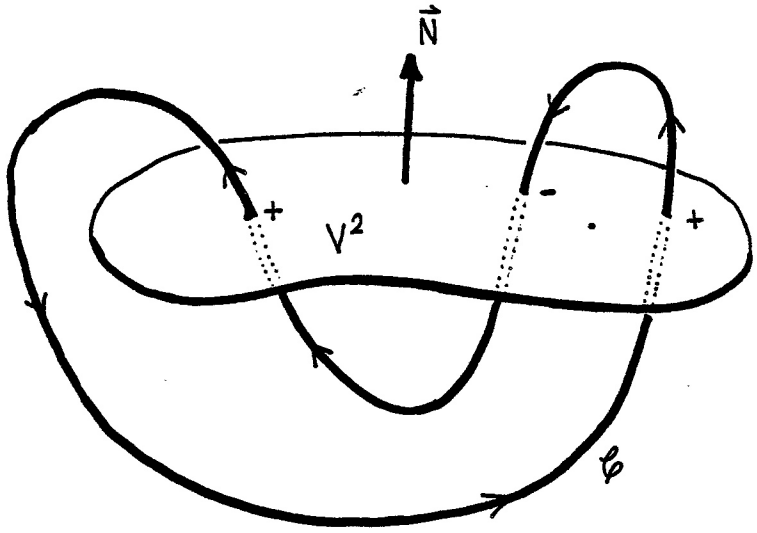
Then the 2-form $j^2 = i_{\vec{J}}\,\mathrm{vol}^3$, where $\vec{J}$ is some vector field, the "current" vector, that runs along the tube and vanishes outside this tube. To say that the magnitude of the current in the 1-dimensional wire $\mathbf{C}$ is j is to say that

$$\iint_{\Sigma^2} j^2 = \iint_{\Sigma^2} \vec{J} \cdot d\vec{S} = j$$

where Σ^2 is the 2-dimensional surface giving a cross-section of the 3-dimensional wire $\tilde{\mathbf{C}}^3$, whose normal is in the direction of the current flow (recall that $\mathcal{H}$ and j^2 are pseudo forms, Lecture 16).

Of course Ampere's law states that $dj^2 = 0$, and so $\operatorname{div} \vec{J} = 0$. Then $\iint_{\Sigma'^2} j^2 = j$ for <u>any</u> cross-section Σ' (why?).

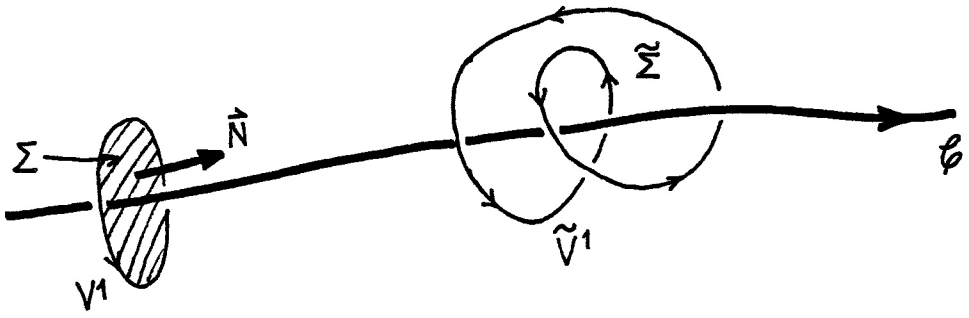
For <u>any</u> 2-dimensional surface V^2 <u>with</u> <u>transverse</u> <u>orientation</u> $\vec{N}$, we have, since $\vec{J}$ vanishes outside the wire,



$$\iint_{V^2} j^2 = \iint_{V^2} \vec{J} \cdot \vec{N}\, dS = j \cdot \quad \begin{array}{l}[(\text{number of times wire crosses } V^2 \text{ in} \\ \text{direction } \vec{N}) - (\text{number of times it} \\ \text{crosses in direction } -\vec{N})].\end{array} \tag{22.2}$$

We have oriented the wire in the direction of the given current. The algebraic number of times the oriented curve $\mathcal{C}$ pierces the transversally oriented surface V^2 is called the "intersection number" of $\mathcal{C}$ with V^2. In our picture $\iint_{V^2} j^2 = j$.

Now $d\mathcal{H}^1 = 0$ outside $\mathcal{C}$; does $\mathcal{H}$ have a scalar potential $\mathcal{H} = d\phi$? Locally, of course, it does. However, for a curve V^1 that loops once around $\mathcal{C}$,



we have $\oint_{V^1} \mathcal{H} = 4\pi \iint_{\Sigma} j^2 = 4\pi j \neq 0$, $\therefore \mathcal{H}$ has no global (single valued) scalar potential in $\mathbb{R}^3 - \mathcal{C}$. If $\tilde{V}^1$ is a closed curve that loops about $\mathcal{C}$ twice in the same direction then $\oint_{\tilde{V}^1} \mathcal{H} = 8\pi j$ since a surface $\hat{\Sigma}^2$ that spans V^1 will be pierced twice by the wire (algebraically). This is a reflection of the fact that $\tilde{V}^1 \sim 2V^1$, i.e., $\tilde{V}^1$ is homologous

to the curve V^1 traversed twice. It seems rather evident that any closed curve in $\mathbb{R}^3 - C$ is homologous to some multiple of our indicated V^1, and this is proven in courses in "algebraic topology."

Finally, we should note that deRham's first theorem is a global existence theorem for certain partial differential equations on manifolds. For example, given a skew symmetric covariant tensor on M^n

$$b_{ij}(x)$$

one can ask if there is a covariant vector

$$a_i(x)$$

that satisfies the system of partial differential equations

$$\partial_i a_j - \partial_j a_i = b_{ij} \quad .$$

A necessary condition is, since then $d\alpha^1 = \beta^2$, clearly $d\beta^2 = 0$, i.e.,

$$\partial_i b_{jk} + \partial_k b_{ij} + \partial_j b_{ki} = 0 \quad .$$

deRham says this condition is also sufficient provided

$$\iint_{V^2} \sum b_{ij}(x) dx^i \wedge dx^j$$

vanishes for every 2-cycle V^2.

Lecture 23. Topological Aspects of Maxwell's Equations

• **Problem 38.** In showing the existence of a global vector potential $\vec{A}$ for the circuit in Lecture 22 we assumed that the circuit was fixed in space. Show that we may draw the same conclusion even when the circuit is moving. (Hint: You will have to look at 2-cycles V^2 that move with the circuit.)

(1) **Compact spaces.** Einstein's general relativity shows that 3-dimensional space V^3 is usually not Euclidean $\mathbb{R}^3$ but must be curved (we shall discuss such matters next quarter). It might be curved so much, in fact, that it is no longer even topologically $\mathbb{R}^3$. One of Einstein's candidates for V^3 is in fact a cosmological model with $V^3 = S^3$, the 3-sphere. This in fact is a space with finite volume. Another model has V^3 a 3-dimensional torus $V^3 = T^3 = S^1 \times S^1 \times S^1$. This example is also popular in quantum mechanics when one deals with a large box with periodic boundary conditions. We shall not deal with relativity at present but we shall entertain ourselves with the possibility that space V^3 is perhaps some manifold differing from $\mathbb{R}^3$. In this paragraph we shall consider a general <u>compact</u> 3-dimensional Riemannian manifold V^3 without boundary, briefly a <u>closed</u> manifold. Maxwell's equations, as expressed in Lecture 16 make perfectly good sense in this situation, when written in terms of differential forms.

Our first result is a classic fact, as simple as it is disturbing.

Proposition. The total charge in a closed V^3 is zero!

Proof.

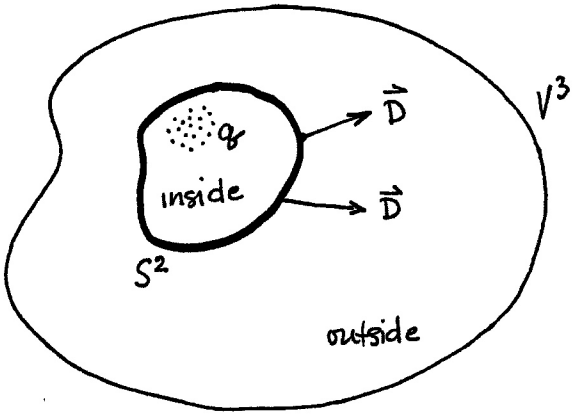
$$4\pi Q = \int_{V^3} \rho \, \mathrm{vol}^3 = \int_{V^3} d\mathfrak{D}^2 = \int_{\partial V^3 = 0} \mathfrak{D}^2 = 0 \quad .$$

Note that since $\mathfrak{D}^2$ is a pseudo-form, V^3 needn't be orientable.

A picture proof is just as simple. Let V^3_{inside} be a small region containing charge q and having S^2 as bounding surface. Then, if $\vec{N}$ is outward pointing normal

$$\int_{S^2} \vec{D} \cdot \vec{N} \, \mathrm{vol}_s^2 = \int_{S^2} \mathfrak{D}^2 = 4\pi \int_{V^3_{\text{inside}}} \rho \, \mathrm{vol}^3 = 4\pi q \quad .$$

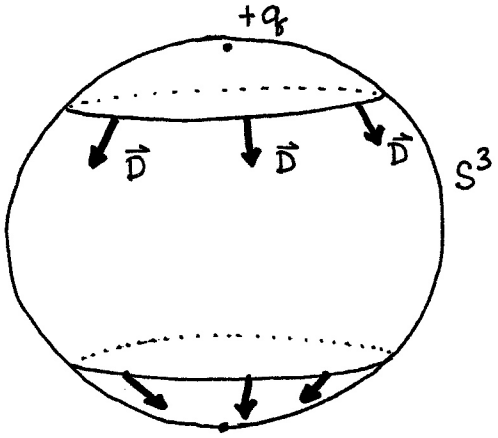
But since V^3 is compact , the region V^3_{outside} outside of S^2 is also compact, with the same boundary surface S^2, but with opposite normal.



$$\int_{S^2} \mathfrak{D}^2 = -4\pi \int_{V^3_{\text{outside}}} \rho \, \text{vol}^3 \quad ,$$

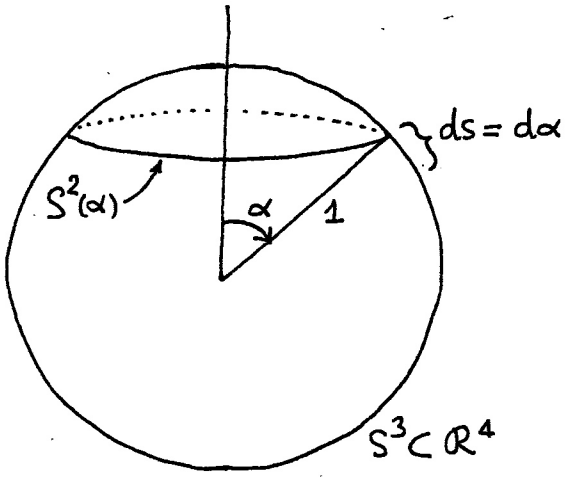
i.e., $\int_{\text{outside}} \rho \, \text{vol}^3 = - \int_{\text{inside}} \rho \, \text{vol}^3$.

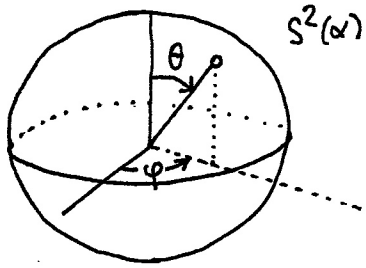
It is interesting to contemplate a thought experiment. Place a charge $+ q$ at the North Pole of an otherwise uncharged 3-sphere $V^3 = S^3$



Note then that by looking at the resulting $\vec{D}$ field (by symmetry) a negative charge $- q$ will appear (a " ghost" charge) at the South Pole !

Let's write down the field $\vec{D}$ on S^3 due to such a charge q.



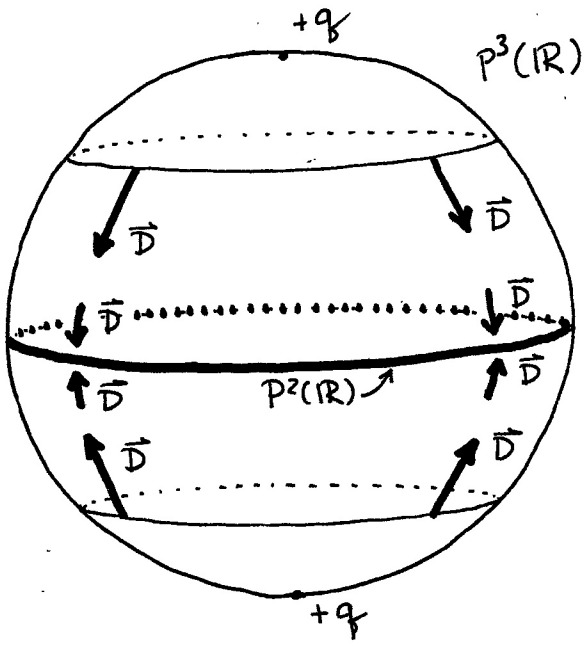


- **Problem 39.** Let $S^3 \subset \mathbb{R}^4$ be the unit 3-sphere in $\mathbb{R}^4$. Introduce coordinates α, θ, ϕ on S^3, where θ, ϕ are the usual spherical coordinates on the 2-sphere $S^2(\alpha)$. Since the "meridian" curves where $\theta = \text{const.}$, $\phi = \text{const.}$ are orthogonal to the 2-spheres $S^2(\alpha)$ where $\alpha = \text{const.}$, we have metric

$$ds^2 = d\alpha^2 + \sin^2\alpha(d\theta^2 + \sin^2\theta d\phi^2) \quad .$$

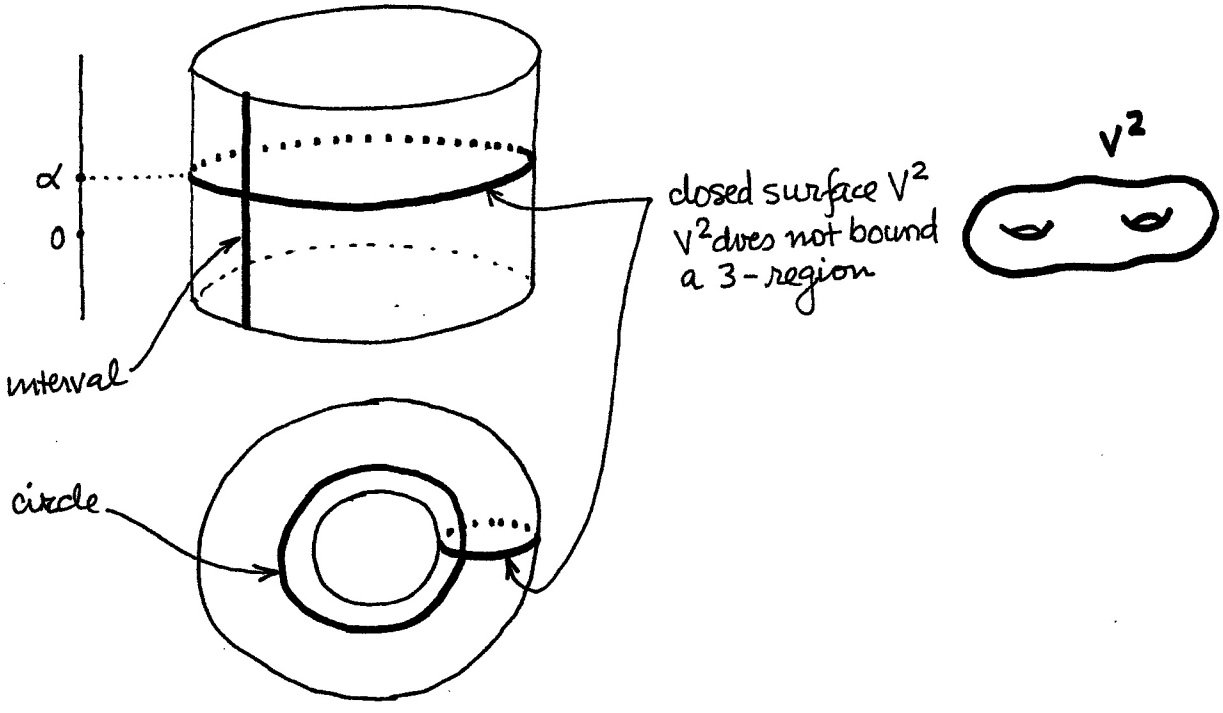
$\vec{D}$ will clearly have the form $\vec{D} = D(\alpha)\frac{\partial}{\partial\alpha}$. Write down $\mathcal{S}^2 = i_{\vec{D}}\,\text{vol}^3_{S^3}$. What form must the function $D(\alpha)$ have in order that $d\mathcal{S}^2 = 0$ for $\alpha \neq 0, \pi$? Finish the determination of $D(\alpha)$ by computing $\int_{S^2(\alpha)} \mathcal{S}^2$ (note; essentially no integration is needed if you know the area of the unit sphere). Write down (via the metric) the electric covector $\mathcal{E}^1$ and verify that $d\mathcal{E}^1 = 0$ and exhibit the scalar potential for $\mathcal{E}^1$. Put $\mathcal{B} \equiv 0$. You have just verified Maxwell's equations !

Another thought experiment. Put charge + q at the North Pole of the real projective 3-space $P^3(\mathbb{R})$. Recall this space is simply S^3 with antipodal points identified. Now + q appears at the South Pole because the South Pole represents the same point as the North Pole. One gets the following picture (I've drawn it on the 3-sphere but I am careful that everything is the same at antipodal points).



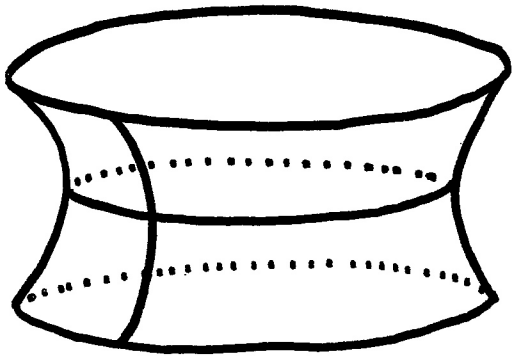
Now $\vec{D}$ cannot vanish as one approaches the "equatorial" $P^2(\mathbb{R})$ (why?). We see that $\vec{D}$ has a jump discontinuity on the equator. Thus there must be a surface charge distributed on the equator! If you place a charge at the North Pole of $P^3(\mathbb{R})$, a surface charge will appear on the entire equatorial $P^2(\mathbb{R})$!

(2) **Magnetic fields in a wormhole topology.** A large class of 3-manifolds can be obtained by taking the Cartesian product (Lecture 2, §2) of a 1-manifold (say an interval $I = \{-\pi < \alpha < +\pi\}$ or a 1-sphere (circle) and a 2-manifold (surface). For visualization we draw the surface as a curve. We are mainly interested in the case when the surface V^2 is closed, i.e., compact without boundary.

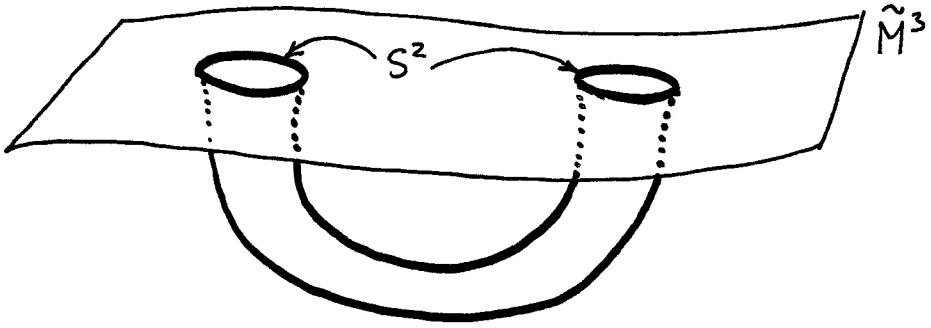


In the second case, when $V^2 = T^2$ is a torus, then $M^3 = S^1 \times T^2 = S^1 \times S^1 \times S^1$ is simply a 3-torus. Note that the "meridian" circle intersects the V^2 exactly once.

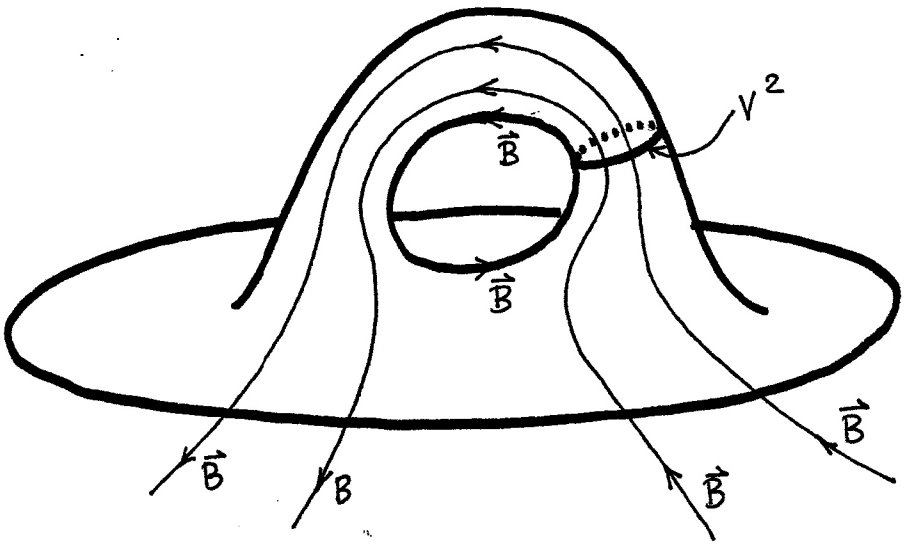
The first type $M^3 = I \times V^2$ is a very general type of "cylinder." If we flare out the ends we get some sort of "wormhole" (this is not a technical definition; any $I \times V^2$ can be considered a wormhole).



We can attach a wormhole to any other 3-manifold $\widetilde{M}^3$; an especially useful case is when the wormhole is $I \times S^2$, in which case it is called a "handle." (To attach a handle, remove two 3-balls from $\widetilde{M}^3$ leaving a pair of boundary 2-spheres; now sew these two spheres to the ends of the handle)

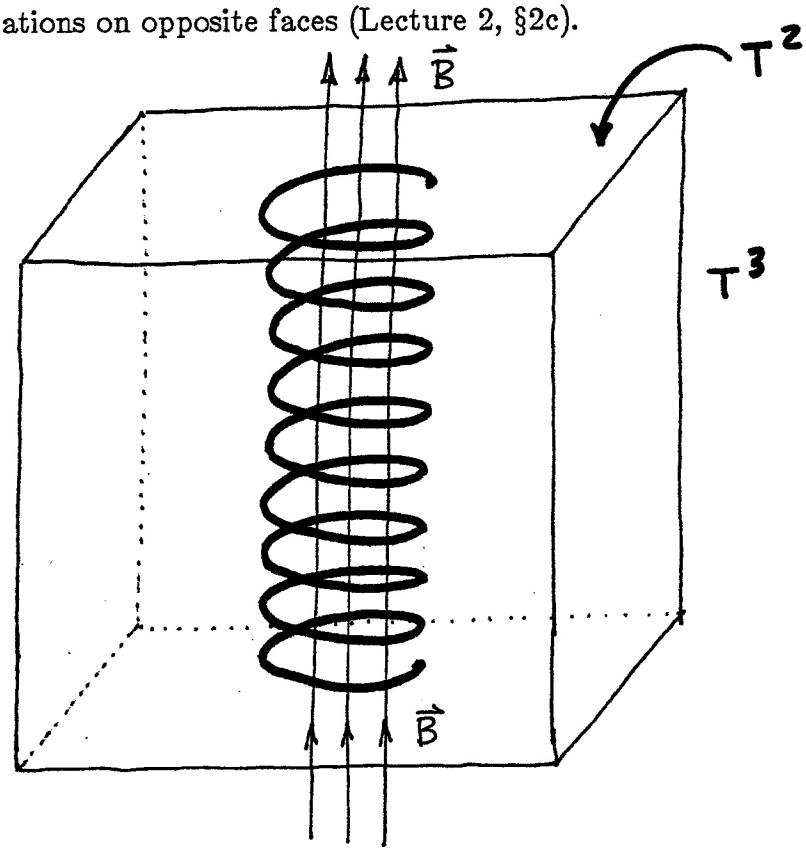


Note that perhaps the wormhole was born during the big bang and it might even have been born with a magnetic field fluxing through it.



In this case we can have $\int_{V^2} \mathcal{B}^2 \neq 0$. This doesn't violate our proposition on magnetic fields in Lecture 22 since the field $\vec{B}$ cannot be "turned off." In fact Faraday's law says that this field $\vec{B}$ is doomed to flux across the closed surface V^2 forever ! Also, since the period of $\mathcal{B}^2$ on V^2 does not vanish, there is no global vector potential $\vec{A}$ such that curl $\vec{A} = \vec{B}$.

Thought Problem (not to be handed in). The 3-torus is the usual box in $\mathcal{R}^3$ with identifications on opposite faces (Lecture 2, §2c).



We have drawn a solenoid. It has no ends due to the identifications. A current running through this coil will produce a magnetic field that seems to have a net flux through the closed torus surface T^2. But the current can be turned off. Explain!

(3) **Currents in a wormhole topology.** Run a constant current through a loop of wire in $\mathbb{R}^3$. There then results a time-independent electromagnetic field and this can be maintained "forever" (the wire might be a superconducting loop or else it could be an ordinary wire with replaceable batteries). I find it surprising that you can't do this in more complicated spaces!

Let us say that a current j in a wire can be " maintained" for all $t > 0$ if, for some $\epsilon > 0$, $|j(t)| \geq \epsilon$. Likewise, a current flux through a surface V^2 can be "maintained" if

$$\left| \int_{V^2} j^2 \right| \geq \epsilon > 0 , \qquad \text{all} \qquad t > 0 \quad .$$

Proposition. Let V^2 be a closed, transversally oriented surface in an M^3. Then no current flux through V^2 can be maintained; a time dependent electromagnetic field will result and the electric field $\vec{D}$ will blow up as $t \to \infty$.

Proof. By reversing the normal to V^2, if necessary, we may assume that

$$\int_{V^2} j^2 \leq -\epsilon < 0 \quad .$$

Ampere-Maxwell then gives

$$4\pi \int_{V^2} j^2 + \int_{V^2} \frac{\partial \mathcal{D}}{\partial t} = \int_{V^2} d\mathcal{H} = \oint_{\partial V^2 = 0} \mathcal{H} = 0$$

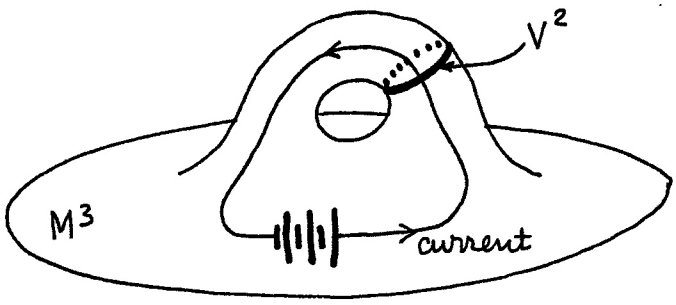
and so (since V^2 is fixed in space)

$$\frac{d}{dt} \int_{V^2} \mathcal{D} = \int_{V^2} \frac{\partial \mathcal{D}}{\partial t} = -4\pi \int_{V^2} j^2 \geq 4\pi\epsilon \quad \text{for} \quad t > 0$$

and so

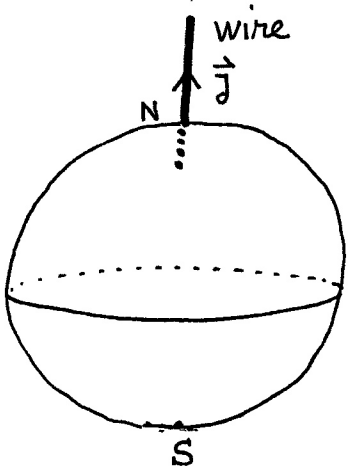
$$\int_{V^2} \mathcal{D} \geq \text{constant} + 4\pi\epsilon t \to \infty \quad \text{as} \quad t \to \infty \ . \qquad \square$$

Thus we cannot maintain a current loop as below!

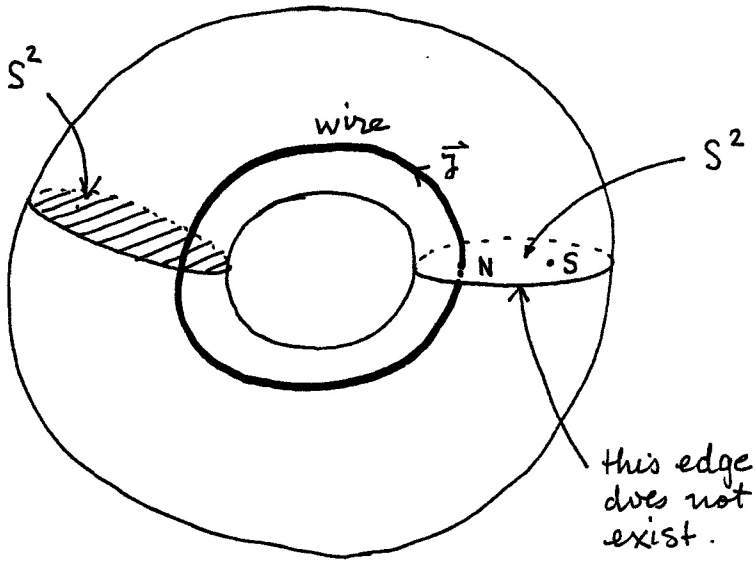


We have shown that if one attempts to construct such a current flux then a time varying electric field $\vec{D}$ is created. Bill Thompson pointed out to me that this gives the physical explanation for the impossibility of maintaining the current

Now let us look at a simple example when $M^3 = S^1 \times S^2$. Then V^2 is a 2-sphere and the wire loop pierces S^2, say, at the North Pole N. Careful ; this S^2 does not bound any region in M^3 and the wire pierces S^2 at just this one point! This is a little hard to draw, convincingly; I will give two imperfect pictures.

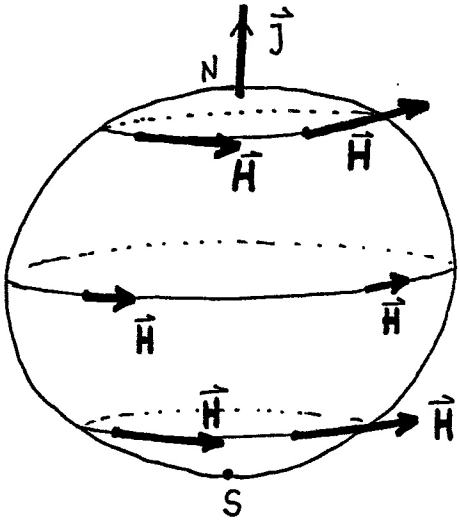


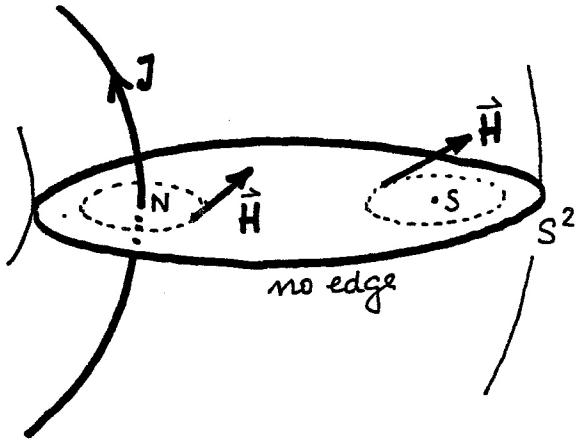
In this drawing it appears as if the wire must pierce S^2 in a second point. It doesn't.



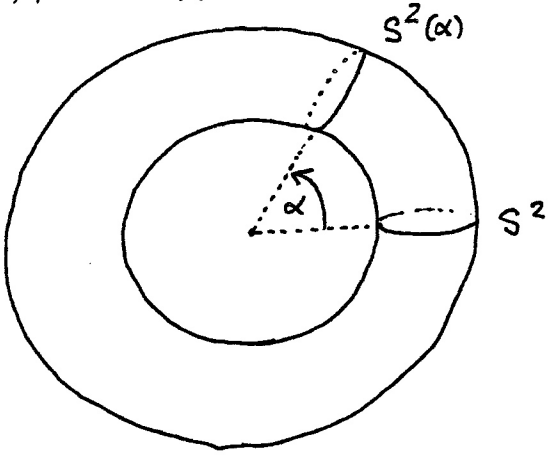
In the second diagram I have made S^2 2-dimensional but it <u>appears</u> to have an edge; it doesn't.

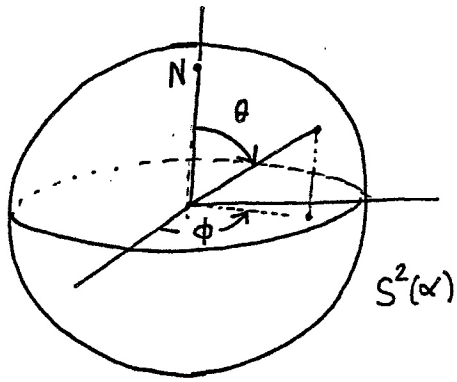
We will get a magnetic field $\vec{H}$ on M^3 that will be tangent to each S^2 and it will encircle the North Pole there.





Does this picture really work? Let's compute. It should be clear that in coordinates α , θ , ϕ for $S^1 \times S^2$





(here θ , ϕ, are the usual coordinates for S^2) that in fact (compare Problem 33)

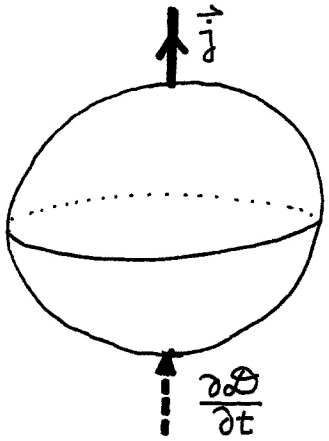
$$\mathcal{H}^1 = 2j\, d\phi \quad .$$

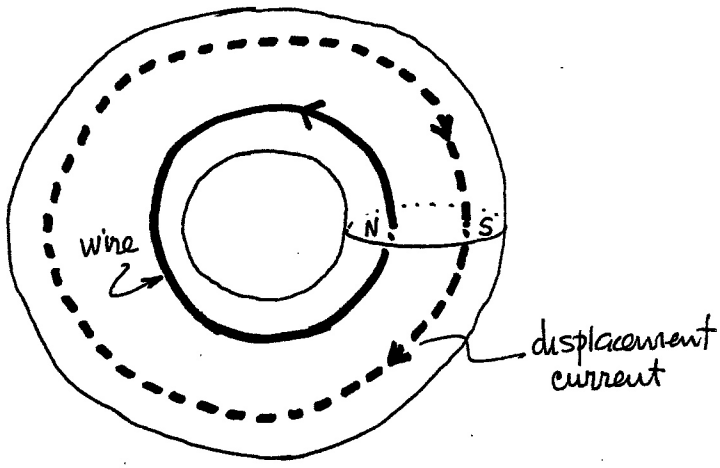
- **Problem 40.** It is immediate that for this $\mathcal{H}^1$ that $d\mathcal{H}^1 = 0$ and $\int_{V^1} \mathcal{H}^1 = 4\pi j$ when V^1 is any circle of latitude on $S^2(\alpha)$. The metric on $S^1 \times S^2$ can be chosen to be (this is called the <u>product metric</u> on $S^1 \times S^2$)

$$ds^2 = d\alpha^2 + d\theta^2 + \sin^2\theta \, d\phi^2 \quad .$$

Write down $\vec{B} \leftrightarrow \vec{H}$ and $\mathcal{B}^2$ and show $d\mathcal{B}^2 = 0$.

It is immediate that since $\vec{H}$ also automatically encircles the South Pole that there is a current of magnitude j piercing the S^2 at the South Pole with intersection number the negative of that of the wire at the North Pole. Since we are assuming that the only true current is in the wire, this "ghost current" can only be a Maxwell displacement current





Thus if one has a true current hit S^2 at the North Pole , automatically a displacement current appears, hitting S^2 at the South Pole !

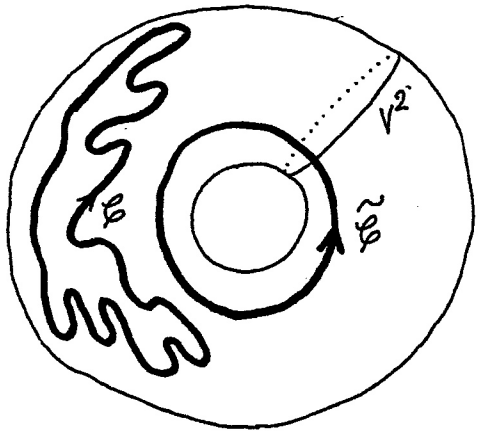
- **Thought Problem.** Consider the case $M^3 = S^1 \times P^2(\mathbb{R})$ (This is a bit more subtle since the projective plane $P^2(\mathbb{R})$ is not orientable. Remember that $\vec{H}$ must be a pseudo-vector.)

A final remark. When M^3 is a closed, manifold our proposition has a nicer statement.

Theorem. If M^3 is closed, then a current loop in M^3 can be maintained if and only if the loop $\mathcal{C}$ bounds a real 2-chain in M^3

$$\mathcal{C} = \partial W^2$$

Thus a current can be maintained in C but not in $\tilde{C}$.



Partial sketch of proof. Suppose loop $\tilde{C}$ does not bound. The first great theorem in algebraic topology, by its inventor, Poincare duality , asserts that if $\tilde{C}$ does not bound then one can find a closed, transversally oriented surface V^2 that has intersection number $= 1$ with $\tilde{C}$. Then our previous proposition shows that a current cannot be maintained.

Suppose on the other hand that the current loop C does bound. Using potential theory (to be discussed next quarter) it is not difficult to show that one can construct a j^2 confined to a "wire loop" and one can find time independent fields $\mathcal{H}$ and $\mathcal{D} = 0$ that will satisfy Maxwell's equations (with charge density $= 0$). This will correspond to a superconducting loop.

~~ **Warning.** We assume that the current loop is the only true current in M^3.

Lecture 24. Brouwer Degree and Kronecker Index

(1) **The Brouwer degree of a map.** Let $\phi : M^n \to V^n$ be a mapping of closed oriented n-manifolds. Let $\gamma^n = \gamma_V^n$ be any n-form on V such that

$$\int_{V^n} \gamma^n = 1$$

(of course if $\int_{V^n} \gamma^n \neq 0$ we may always normalize it so). The (Brouwer) degree of ϕ is defined to be

$$\boxed{\deg(\phi) = \int_{M^n} \phi^* \gamma_V^n \quad .} \tag{24.1}$$

Note that we may also write $\deg(\phi) = \int_{\phi(M^n)} \gamma^n$; this number tells us in some sense (see below) how many times (algebraically) M^n gets wrapped around V^n by ϕ.

Note that $\deg(\phi)$ is independent of the form γ_V^n chosen ; for is $\tilde{\gamma}$ is another such, then

$$d(\tilde{\gamma}^n - \gamma^n) = d\tilde{\gamma}^n - d\gamma^n = 0$$

and $\tilde{\gamma}^n - \gamma^n$ has zero period on V^n

$$\int_{V^n} \tilde{\gamma}^n - \gamma^n = 0 \quad .$$

But it seems clear (and is true) that the only nontrivial n-cycles on a closed oriented manifold V^n are multiples of V^n itself. Hence $\tilde{\gamma}^n - \gamma^n$ is a closed n-form on V^n and all its periods vanish; by de Rham's first theorem (Lecture 22)

$$\tilde{\gamma}^n - \gamma^n = d\alpha^{n-1}$$

for some (n–1) form α on V^n. Then

$$\int_M \phi^*\tilde{\gamma}^n - \int_M \phi^*\gamma^n = \int_M \phi^* d\alpha = \int_M d\phi^*\alpha = \int_{\partial M = 0} \phi^*\alpha = 0 \quad . \qquad \square$$

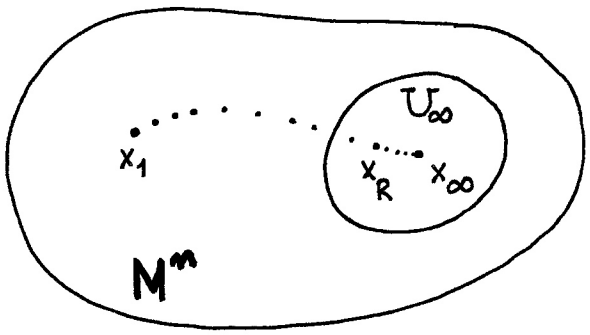
Theorem. Let $y \in V^n$ be a regular value of ϕ (see Lecture 3). (Recall that Sard's theorem says that the regular values are dense in V^n). For each $x \in \phi^{-1}(y)$, $\phi_* : M_x^n \to V_y^n$ is 1:1. Put

$$\text{sign } \phi_*(x) = \pm 1$$

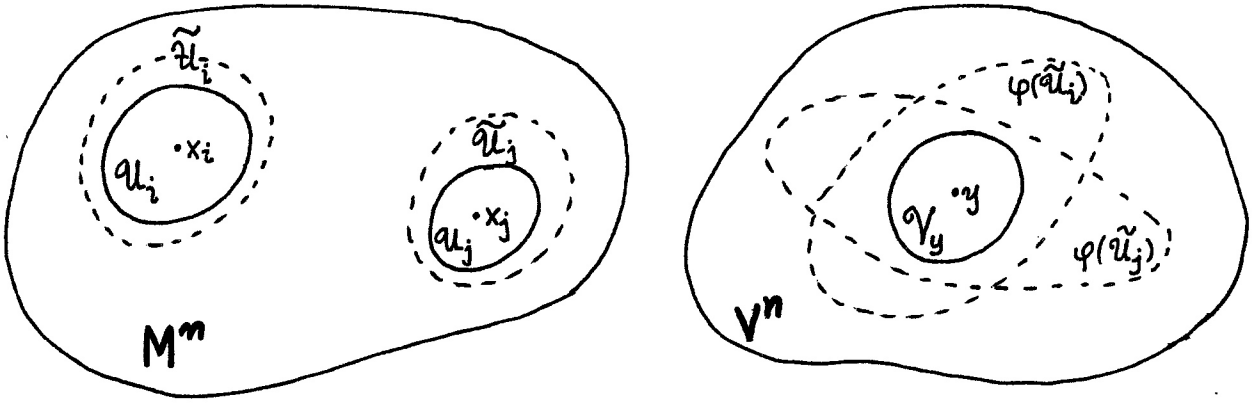
where the + sign is used only if $\phi_* : M_x^n \to V_y^n$ is orientation preserving. Then

$$\boxed{\deg(\phi) = \sum_{x \in \phi^{-1}(y)} \text{sign } \phi_*(x)} \tag{24.2}$$

Proof. Because y is regular there are only a finite number of $x \in \phi^{-1}(y)$. We can see this as follows. By compactness of M^n, if $\phi^{-1}(y)$ were infinite we could find a sequence $\{x_n\} \subset \phi^{-1}(y)$ that converged to some $x_\infty \in M^n$. But $\phi^{-1}(y)$ is a closed subset of M^n, (the inverse image of a closed set under a continuous map, is closed), and so x_∞ is also regular. Since $\phi_* : M_{x_\infty} \to V_y$ is 1:1 ϕ is a diffeomorphism of some neighborhood U_∞ of x_∞ by the inverse function theorem, Lecture 3, §(3). But $x_i \to x_\infty$ means $x_i \in U_\infty$ for all i sufficiently large, $i \geq R$. But then x_R and x_∞ are both in U_∞ and $\phi(x_R) = \phi(x_\infty) = y$. This contradicts ϕ being 1:1 on U_∞.



We now continue with our proof. Let $\{x_1,...,x_k\} = \phi^{-1}(y)$).



Let $\tilde{u}_1,\ldots,\tilde{u}_k$ be disjoint neighborhoods of $x_1,\ldots,x_k$ and such that ϕ is a diffeomorphism on each $\tilde{u}_i$. Put

$$v_y = \phi(\tilde{u}_1) \cap \cdots \cap \phi(\tilde{u}_k) \quad .$$

Then $\phi^{-1}(v_y) =$ a disjoint union $u_1 \cup u_2 \cup \ldots \cup u_k$ where $u_i \subset \tilde{u}_i$ and $\phi : u_i \to v_y$ is a diffeomorphism.

Let γ be any n-form on V^n whose support lies in v_y (i.e., $\gamma = 0$ outside v_y) and such that

$$\int_{V^n} \gamma^n = \int_{v_y} \gamma^n = 1 \quad .$$

Then

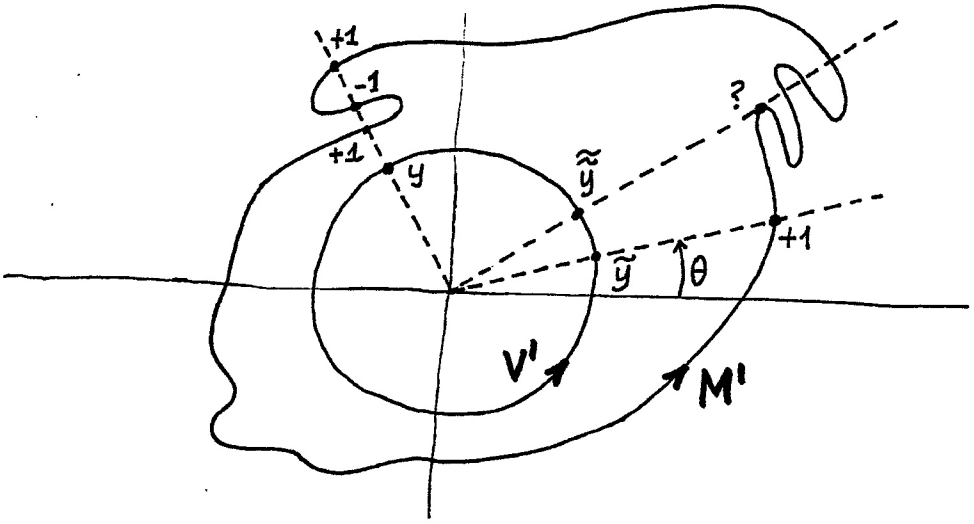
$$\deg(\phi) = \int_{M^n} \phi^*\gamma^n = \int_{\cup u_i} \phi^*\gamma^n = \sum_{i=1}^{k} \int_{u_i} \phi^*\gamma^n$$

$$= \sum_{i=1}^{k} \int_{\phi(u_i)} \gamma^n = \sum_{i=1}^{k} \operatorname{sign} \phi_*(x_i) \int_{v_y} \gamma^n$$

because orientation of $\phi(u_i)$ is (sign $\phi_*(x_i)$) $\cdot$ (orientation of v_y). But $\int_{v_y} \gamma^n = 1$. □

From (24.2) it is clear that deg (ϕ) is an integer !

Example. Let $V^1 =$ inner circle, $M^1 =$ outer closed curve, and $\phi : M^1 \to V^1$ radial projection. We have indicated orientations.



We have indicated three values $y, \tilde{y}, \tilde{\tilde{y}}$ ($\tilde{\tilde{y}}$ is not a regular value). We see deg (ϕ) = 1. Note that we (theoretically) could compute deg (ϕ) from original

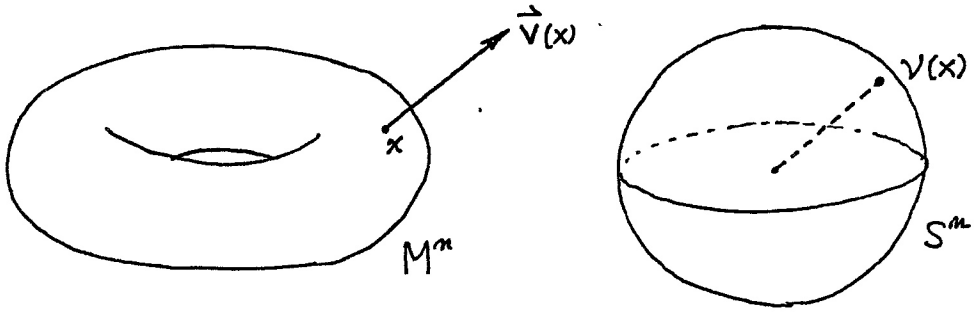
definition; $\gamma^1 = \frac{1}{2\pi}$ "$d\theta$" is an appropriate form on V^1. Then, letting, say $s =$ arc length parameter for M^1

$$\int_{M^1} \phi^*\Big(\frac{1}{2\pi}\text{ "}d\theta\text{"}\Big) = \frac{1}{2\pi}\int_{M^1}\phi^*\text{ "}d\theta\text{"} = \frac{1}{2\pi}\int_0^L \frac{d\theta}{ds}\,ds$$

$= \#$ of times radial +projection of M^1 wraps around unit circle

- **Problem 41.** Using any definition, show that the antipodal map $S^n \to S^n$; $\vec{x} \to -\vec{x}$ on the n-sphere $S^n \subset \mathbb{R}^{n+1}$ has degree $(-1)^{n+1}$. (Hint: try $n = 1$, $n = 2$ first.)

(2) **Kronecker index of a vector field.** Let M^n be a closed (i.e., compact and without boundary) submanifold of $\mathbb{R}^{n+1}$. M^n is the boundary of a compact region $u \subset \mathbb{R}^{n+1}$; $M^n = \partial u$, and an orientation of $\mathbb{R}^{n+1}$ (assumed given) together with the normal $\vec{N}$ to M^n that points out of u defines an orientation of M^n; u is then a compact oriented manifold with boundary M. Let $\vec{v}$ be a unit vector field defined along M^n (it need not be tangent to M). It then defines a map $\nu : M^n \to$ the unit sphere $S^n \subset \mathbb{R}^{n+1}$ by $x \in M^n \mapsto \vec{v}(x) \in S^n$



We define the (Kronecker) index of $\vec{v}$ on M^n to be the integer

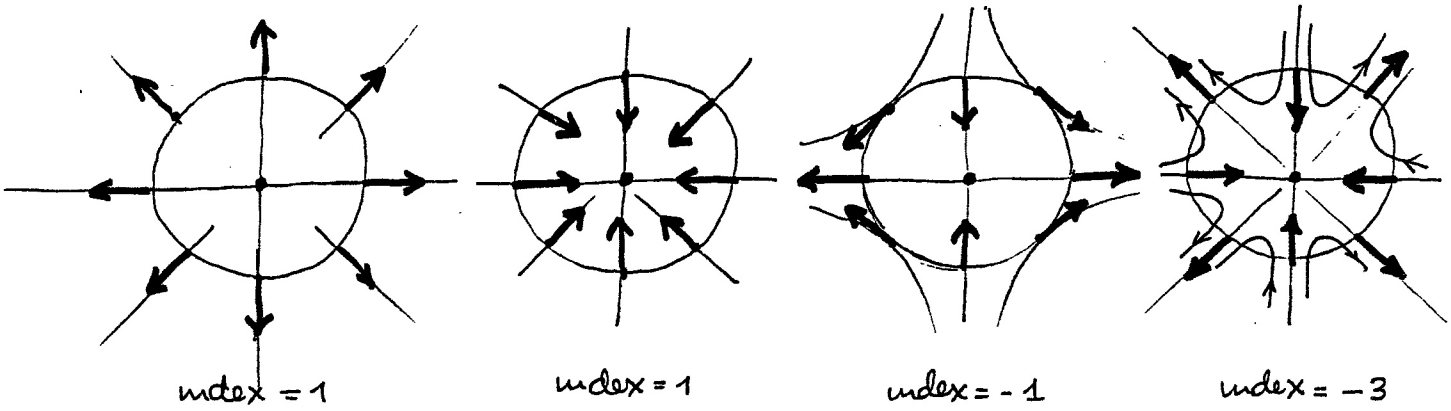
$$\boxed{\text{index of } \vec{v} \text{ on } M^n \equiv \deg \nu : M^n \to S^n = \frac{1}{A_n}\int_{M^n} \nu^* \,\text{vol}_S^n} \qquad (24.3)$$

where $A_n = \int_{S^n} \mathrm{vol}_S^n$ is the "area" of S^n. Of course we could use any n-form on S (with $\int_S = 1$) instead.

If $\vec{v}$ is any nonvanishing vector field on M^n we can define its Kronecker index trivially by

$$\text{index of } \vec{v} \equiv \text{index of } \frac{\vec{v}}{\|\vec{v}\|} \quad .$$

Examples in the plane, $M^1 =$ circle



- **Problem 42.** Show that if $\vec{v}$ can be extended to a smooth unit vector field on all of u, $v : u \to S^n$, then

$$\text{index of } \vec{v} \text{ on } M^n = 0 \quad .$$

Thus none of the four examples just given can be extended to be nonvanishing vector fields on the unit disc $u^2 \subset \mathbb{R}^2$.

- **Problem 43.** Suppose that the unit vector field $\vec{v}$ on M^n can be extended to a smooth unit field on all of u except for a finite number of points $P_1, P_2, \ldots, P_k$. Excise a small ball B_α centered about each P_α. Then $\vec{v}$ has an index on M^n and on each of the spheres ∂B_α (normal pointing out of B_α). Show that

$$\text{index of } \vec{v} \text{ on } M^n = \sum_{\alpha=1}^{k} \text{index of } \vec{v} \text{ on } (\partial B_\alpha) \quad .$$

We then say, "index of $\vec{v}$ on M^n = sum of indices of $\vec{v}$ inside M^n."

• **Problem 44.** Let $u^1,\ldots,u^n$ be coordinates for M^n in a patch.

(i) Why is $v_* \dfrac{\partial}{\partial u^i} = \dfrac{\partial \vec{v}}{\partial u^i}$?

(ii) Show that the contribution to index $\vec{v}$ in this patch is

$$\frac{1}{A_n} \int \mathrm{vol}^{n+1}_{\mathbb{R}^{n+1}} \left(\vec{v}, \frac{\partial \vec{v}}{\partial u^1}, \ldots, \frac{\partial \vec{v}}{\partial u^n}\right) du^1 \wedge \cdots \wedge du^n \quad . \tag{24.4}$$

This expression is for a unit $\vec{v}$. For nonvanishing $\vec{v}$, show contribution is

$$\boxed{\frac{1}{A_n} \int \frac{1}{\|\vec{v}\|^{n+1}} \mathrm{vol}^{n+1}_{\mathbb{R}^{n+1}} \left(\vec{v}, \frac{\partial \vec{v}}{\partial u^1}, \ldots, \frac{\partial \vec{v}}{\partial u^n}\right) du^1 \wedge \cdots \wedge du^n} \tag{24.5}$$

Note: (24.4) is a general formula for computing the degree of any map from a compact M^n to the unit sphere in $\mathbb{R}^{n+1}$

$$v : M^n \to S^n \subset \mathbb{R}^{n+1} \quad .$$

(24.5) has as corollary a famous formula of Kronecker (this was his motivation for introducing the notion of index). From Problem 41,

Corollary (Kronecker). Let (n+1) functions $f_1,\ldots,f_{n+1}$, be defined on M^n and its interior $U^{n+1} \subset \mathbb{R}^{n+1}$ with no common zeros on M^n. Then if

$$0 \neq \int_{M^n} (f_1^2 + \cdots + f_{n+1}^2)^{-\frac{n+1}{2}} \begin{vmatrix} f_1 & \dfrac{\partial f_1}{\partial u^1}, \ldots, \dfrac{\partial f_1}{\partial u^n} \\ \vdots & \\ f_{n+1} & \dfrac{\partial f_{n+1}}{\partial u^1}, \ldots, \dfrac{\partial f_{n+1}}{\partial u^n} \end{vmatrix} du^1 \wedge \cdots \wedge du^n$$

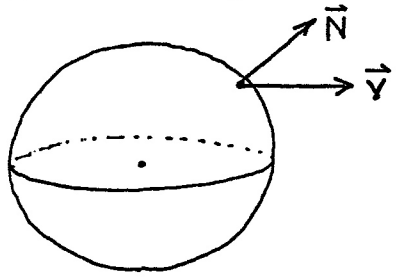
$f_1(x_1,\ldots,x_{n+1}) = 0,\ldots,f_{n+1}(x_1,\ldots,x_{n+1}) = 0$ must have a common solution in U^{n+1}! (When we write an integral $\int_{M^n} \cdots du^1 \wedge \cdots \wedge du^n$ we mean that this is the contribution from the given coordinate patch $u^1,\ldots,u^n$).

Note the important immediate fact

Proposition. If $\vec{v}_t$ is a smooth family of nonvanishing vector fields on M^n with $\vec{v}_0 = \vec{v}$ and $\vec{v}_1 = \vec{w}$, then since the index is an integer varying continuously with $\vec{v}_t$ we see

$$\text{index } \vec{v} = \text{index } \vec{w} \quad .$$

- **Problem 45.** Let $\vec{v}$ be a unit vector field on $M^n = S^n$ that never points to the center 0. Show that index of $\vec{v} =$ index of the unit outward normal $\vec{N} = +1$. Corollary. If $\vec{v}$ is tangent to S^n then index $= 1$.



(3) **Gauss looping integral.** Let $C_\alpha : S^1 \to \mathbb{R}^3$, $\alpha = 1,2$, be a pair of non-intersecting smooth closed curves in $\mathbb{R}^3$

$$C_1 \,, \vec{r} = \vec{r}_1(\theta) \qquad 0 \le \theta \le 2\pi$$

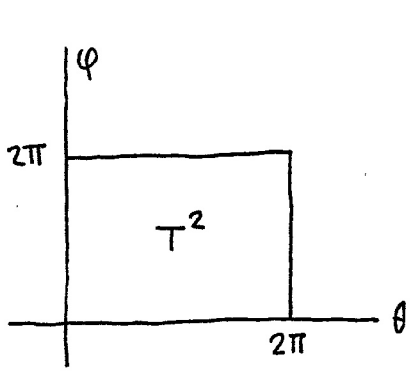
$$C_2 \,, \vec{r} = \vec{r}_2(\phi) \qquad 0 \le \phi \le 2\pi \quad .$$

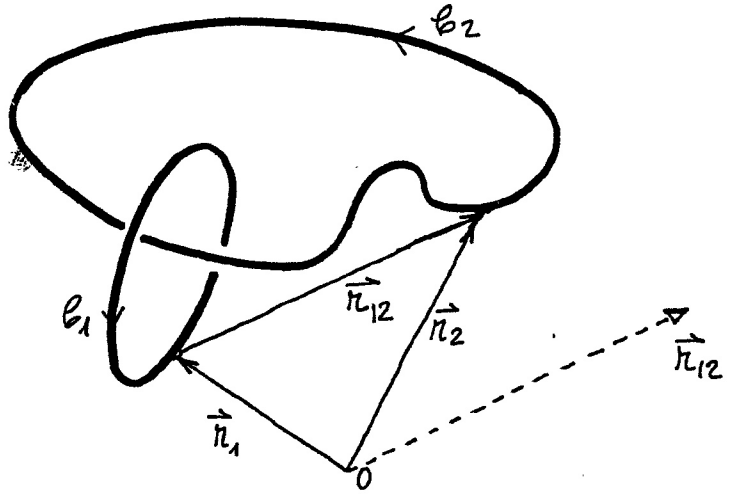
Consider the abstract torus $T^2 = S^1 \times S^1$ with coordinates θ, ϕ, and the map

$$\mathfrak{L} : T^2 \to S^2$$

defined by

$$\mathfrak{L}(\theta,\phi) = \frac{\vec{r}_2(\phi) - \vec{r}_1(\theta)}{\|\vec{r}_2(\phi) - \vec{r}_1(\theta)\|} \equiv \frac{\vec{r}_{12}}{\|\vec{r}_{12}\|} \equiv \frac{\vec{r}_{12}}{r_{12}} \quad .$$





The Gauss looping (or linking) number of C_1 and C_2 is the integer

$$\mathfrak{L}k(C_1, C_2) \equiv \text{degree } \mathfrak{L} : T^2 \to S^2 \quad .$$

• **Problem 46.** Show that $\mathfrak{Lk}(C_1, C_2)$ can be given by Gauss' integral

$$\mathfrak{Lk}(C_1, C_2) = \frac{1}{4\pi}\int_{C_1}\left[\int_{C_2}\frac{1}{r_{12}^3}(\vec{r}_{12}\times d\vec{r}_2)\right]\cdot d\vec{r}_1$$

$$= \frac{1}{4\pi}\int_{S^1}\left[\int_{S^1}\frac{1}{r_{12}^3}(\vec{r}_{12}\times\frac{d\vec{r}_2}{d\phi}\,d\phi)\right]\cdot\frac{d\vec{r}_1}{d\theta}\,d\theta \quad .$$

Now let W^2 be any orientable surface in $\mathbb{R}^3$ whose boundary is C_1. Choose the orientation of W^2 so that $\partial W^2 = C_1$. For a given orientation of $\mathbb{R}^3$ this picks out a preferred unit normal $\vec{N}$ to W^2. It is a fact that C_2 can always be moved slightly so that C_2 intersects W^2 transversally. We then can consider the intersection number $W^2 \circ C_2$ of the oriented curve C_2 and the transversally oriented surface W^2 (see discussion following (22.2)).

• **Problem 47.** Show that

$$\mathfrak{Lk}(C_1, C_2) = W^2 \circ C_2 \quad .$$

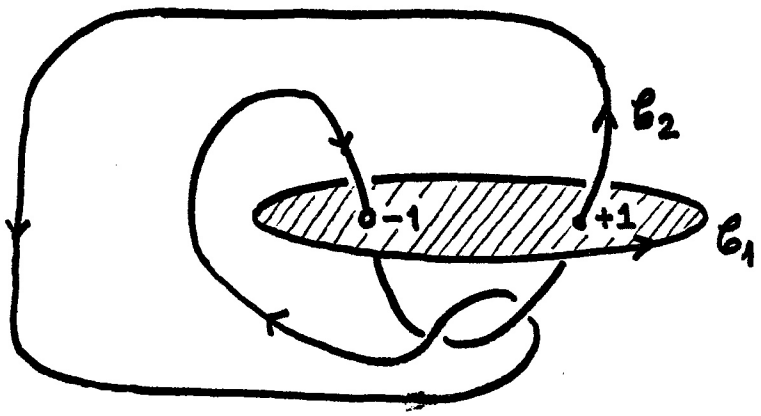
(Hint: A current of $I = 1$ in C_2 gives rise to a magnetic field at $\vec{r}_1$ given by

$$\vec{H}(\vec{r}_1) = \oint_{C_1}\frac{\vec{r}_{12}\times d\vec{r}_2}{r_{12}^3} \quad ;$$

see e.g. Feynman's lectures, vol. II, pp. 14-10).

This "shows" that $\mathfrak{Lk}(C_1, C_2)$ really does tell how many times (algebraically) C_2 loops around C_1.

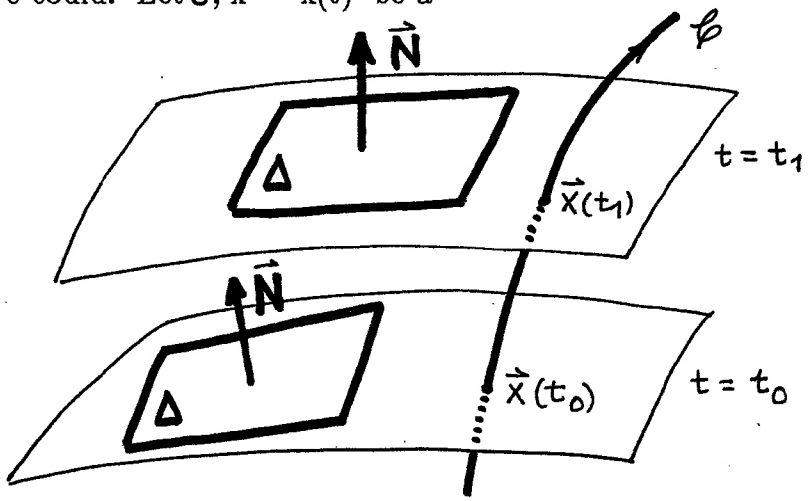
~~ **Warning.** Two wires can have linking number 0 and yet be physically "inseparable." For example



C_1 and C_2 are inseparable, yet have $\mathfrak{Lk}(C_1, C_2) = 0$.

Lecture 25. The Frobenius Integrability Theorem

(1) **Planes in $\mathbb{R}^3$.** Given a smooth vector field, one can always find a smooth family of integral curves, i.e., curves that are always tangent to the vector field (see Lecture 4). Given a smooth family of 2-planes in $\mathbb{R}^3$, Δ_2, can one always find a smooth family of "integral surfaces", i.e., surfaces everywhere tangent to the planes? Suppose that we could. Let $\mathcal{C}; \vec{x} = \vec{x}(t)$ be a



parametrized curve that is transverse to the family of supposed integral surfaces (we can certainly find such a curve locally). Then locally we can define a function $f(x,y,z)$ whose level surfaces are surfaces of the family; namely, the level surface where $f(x,y,z) = t_1$ consists of the integral surface that is pierced by the transversal curve $\mathcal{C}$ at parameter value $t = t_1$. But then grad f must be normal to the integral surface, and so we must have, for some function $\lambda(x,y,z)$

$$\lambda \text{ grad } f = \vec{N}$$

where $\vec{N}$ is a normal vector field to the surfaces. In Cartesian coordinates the covector corresponding to $\vec{N}$ is

$$\nu \equiv N_i dx^i = \lambda \, df \quad .$$

But then we would have

$$d\nu = d\lambda \wedge df = (d \log \lambda) \wedge \nu$$

and so we have recovered Euler's integrability condition

$$\nu \wedge d\nu = 0 , \quad \text{i.e.,} \quad \vec{N} \cdot \text{curl } \vec{N} = 0 \quad . \tag{25.1}$$

This condition, given entirely in terms of the normals $\vec{N}$ to the planes Δ_2 must be satisfied if Δ_2 is to admit a family of integral surfaces.

For example, consider the planes Δ_2 normal to the vectors

$$\vec{N} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} + \frac{\partial}{\partial z} \quad .$$

Then $\nu = y\,dx - x\,dy + dz$ and so

$$\nu \wedge d\nu = -2\,dx \wedge dy \wedge dz \neq 0 \quad ;$$

the vectors $\vec{N}$ are not normal to any family of surfaces !

Classically the plane orthogonal to the vector $\vec{N}$ above would be written

$$\nu = y\,dx - x\,dy + dz = 0$$

meaning that at each point (x_0,y_0,z_0) we look at all vectors $\vec{A} = A^1\frac{\partial}{\partial x} + A^2\frac{\partial}{\partial y} + A^3\frac{\partial}{\partial z}$ such that

$$0 = \nu(\vec{A}) = y_0A^1 - x_0A^2 + A^3 \quad ,$$

clearly a two dimensional plane at (x_0,y_0,z_0) (all vectors orthogonal to $\vec{N}$). The collection of all such planes, at all $(x,y,z) \in \mathbb{R}^3$ is called a distribution of planes (not to be confused with the generalized functions also called distributions).

Definition. A k dimensional distribution Δ_k on M^n assigns in a smooth fashion to each $x \in M^n$ a k-dimensional subspace $\Delta_k(x) \subset M_x^n$ of the tangent space to M^n at x. An integral manifold of Δ_k is a k-dimensional submanifold V^k of M^n whose tangent space $V_x^k = \Delta_k(x)$ at each $x \in V^k$. The distribution Δ_k is completely integrable if locally there are coordinates $x^1, \ldots, x^k, y^1, \ldots, y^{n-k}$ for M^n such that the "coordinates slices"

$$y^1 = \text{const.} = c^1, \ldots, y^{n-k} = c^{n-k}$$

are integral manifolds of Δ_k.

Fundamental question; when is Δ_k completely integrable?

(2) **Distributions and Vector Fields.** Suppose that Δ_k is completely integrable, with integral manifolds defined by the slices $y^1 = c^1, \ldots, y^{n-k} = c^{n-k}$. Let X and Y be a pair of vector fields that are tangent to the distribution, $X \in \Delta_k$, $Y \in \Delta_k$. Then since $\frac{\partial}{\partial x^1}, \ldots, \frac{\partial}{\partial x^k}$ span the distribution Δ_k we have

$$X = \sum_A X^A \frac{\partial}{\partial x^A}, \; Y = \sum_B Y^B \frac{\partial}{\partial x^B}, \; [X,Y] = \sum_{A,B} \left(X^A \frac{\partial Y^B}{\partial x^A} - Y^A \frac{\partial X^B}{\partial x^A} \right) \frac{\partial}{\partial x^B}$$

and so $[X,Y] \in \Delta_k$ also. We conclude

Proposition. If vector fields X and Y are in the completely integrable distribution, $X \in \Delta_k$, $Y \in \Delta_k$, then so is their Lie (commutator) bracket, $[X,Y] \in \Delta_k$.

We can then write, briefly that if Δ_k is completely integrable

$$[\Delta, \Delta] \subset \Delta \quad .$$

It will turn out that this condition is <u>also</u> <u>sufficient</u> for the distribution to be completely integrable (see §4).

(3) **Distributions and Pfaffians** (1-forms). Let θ be a nonzero Pfaffian (i.e., 1-form) at $x \in M^n$. The <u>annihilator</u> or <u>null space</u> of θ at x is the (n–1)-dimensional subspace of M_x^n defined by $\{X \in M_x^n \mid \theta(X) = 0\}$. Classically one writes $\theta = 0$ for this subspace. If $\theta_1, \ldots, \theta_r$ are $r = n-k$ independent 1-forms in an open subset of M^n, $\theta_1 \wedge \cdots \wedge \theta_r \neq 0$, then at each point the intersection of their null spaces forms an $n-r = k$ dimensional distribution Δ_k. Thus

$$\boxed{X \in \Delta_k \text{ if and only if } \theta_1(X) = \cdots = \theta_r(X) = 0}$$

Let $X_1, \ldots, X_k$ be smooth vector fields that locally span the distribution Δ_k.

Definition. A distribution Δ_k is in involution, or involutive if

$$[\Delta,\Delta] \subset \Delta \quad .$$

If Δ_k is involutive then we must have, for $\alpha = 1,\ldots,r$, from (18.8)

$$\begin{aligned} d\theta_\alpha(X_A,X_B) &= X_A(\theta_\alpha(X_B)) - X_B(\theta_\alpha(X_A) - \theta_\alpha([X_A,X_B]) \\ &= X_A(0) - X_B(0) - 0 = 0 \quad . \end{aligned}$$

Thus if Δ is in involution then $d\theta_\alpha = 0$ when restricted to Δ!

Conversely, suppose $d\theta_\alpha = 0$ when restricted to Δ, $\alpha = 1, \ldots, r$. Then $0 = d\theta_\alpha(X_A,X_B) = X_A(0) - X_B(0) - \theta_\alpha([X_A,X_B])$, $\alpha = 1, \ldots, r$ shows that $[X_A,X_B] \in \Delta$, and so $[\Delta,\Delta] \subset \Delta$. We now give several rewordings of this.

Theorem. The following conditions are equivalent:

(1) Δ is closed under brackets, i.e., $[\Delta,\Delta] \subset \Delta$.

(2) $d\theta_\alpha$, restricted to Δ, $= 0$.

(3) There are 1-forms $\lambda_{\alpha\beta}$ such that

$$d\theta_\alpha = \sum_\beta \lambda_{\alpha\beta} \wedge \theta_\beta.$$

(4) $d\theta_\alpha \wedge \Omega = 0$, $\alpha = 1, \ldots, r$, where

$$\Omega = \theta_1 \wedge \theta_2 \wedge \cdots \wedge \theta_r.$$

(25.2)

Proof. We have already proved (1) $\Leftrightarrow$ (2). Now (3) $\Rightarrow$ (2) because

$$\begin{aligned} d\theta_\alpha(X_A,X_B) &= \sum_\beta \lambda_{\alpha\beta} \wedge \theta_\beta\,(X_A,X_B) \\ &= \sum_\beta \lambda_{\alpha\beta}(X_A)\theta_\beta(X_B) - \lambda_{\alpha\beta}(X_B)\theta_\beta(X_A) = 0 \quad . \end{aligned}$$

Conversely, suppose that $d\theta_\alpha = 0$ when restrict to Δ, $\alpha = 1, \ldots, r$. Complete $\theta_1, \ldots, \theta_r$ locally to give a basis for 1-forms $\theta_1, \ldots, \theta_r, \theta_{r+1}, \ldots, \theta_n$. Let $e_1, \ldots, e_n$ be a dual basis for vector fields. Then

$$\theta_\alpha(e_i) = 0 \quad \text{for} \quad \alpha = 1, \ldots, r \quad \text{and} \quad i = r+1, \ldots, n \quad .$$

and hence $e_{r+1},\ldots,e_n$ span the distribution Δ_k. Now expand $d\theta_\alpha$ in terms of the basis $\theta_1, \ldots, \theta_n$

$$d\theta_\alpha = \sum_{\beta=1}^{r} \lambda_{\alpha\beta} \wedge \theta_\beta + \sum_{r<i<j} \mu_\alpha{}^{ij}\, \theta_i \wedge \theta_j \tag{25.3}$$

for some coefficients $\lambda_{\alpha\beta}$ and $\mu_\alpha{}^{ij}$, Then for $r < i < j$

$$0 = d\theta_\alpha(e_i,e_j) = u_\alpha{}^{ij} \quad ,$$

and so $d\theta_\alpha = \sum_\beta \lambda_{\alpha\beta} \wedge \theta_\beta$. This shows that (2) $\Rightarrow$ (3).

It is immediate that (3) $\Rightarrow$ (4). Assume (4). From (25.3)

$$\begin{aligned} 0 = d\theta_\alpha \wedge \Omega &= \sum_{r<i<j} \mu_\alpha{}^{ij}\, \theta_i \wedge \theta_j \wedge \Omega \\ &= \sum_{r<i<j} \mu_\alpha{}^{ij}\, \theta_i \wedge \theta_j \wedge \theta_1 \wedge \cdots \wedge \theta_r \quad . \end{aligned}$$

But the θ's are all independent, hence $\mu_\alpha{}^{ij} = 0$ for $r < i < j$ and so $d\theta_\alpha = \sum \lambda_{\alpha\beta} \wedge \theta_\beta$ and (4) $\Rightarrow$ (3). □

Remarks. We have seen that a k-dimensional distribution Δ_k can locally be described by either exhibiting k linearly independent vector fields

$$X_1, \ldots, X_k$$

that span Δ_k at each point, or by exhibiting $r = n-k$ linearly independent 1-forms

$$\theta_1, \ldots, \theta_{n-k}$$

whose common null space is Δ_k. The system is involutive if either

$$[\Delta_k,\Delta_k] \subset \Delta_k$$

or $d\theta_\alpha = \sum_\beta \lambda_{\alpha\beta} \wedge \theta_\beta$ for some 1-forms $\lambda_{\alpha\beta}$. In this case we write

$$d\theta_\alpha \equiv 0 \mod \theta \quad .$$

We have seen in the Proposition of §2 that if Δ_k is completely integrable then Δ_k is in involution. We now sketch the converse.

(4) **The Frobenius Theorem.** Let Δ_k be <u>any</u> smooth distribution of k-planes in M^n and let (locally) $\{X_A\}$, $A = 1,2, \ldots, k$, be k smooth vector fields that span Δ_k in some open set U of M^n. Let ϕ_A be the (local) flow generated by the vector field X_A. Given $x \in U$ we construct a k-dimensional submanifold of M^n passing through x as follows.

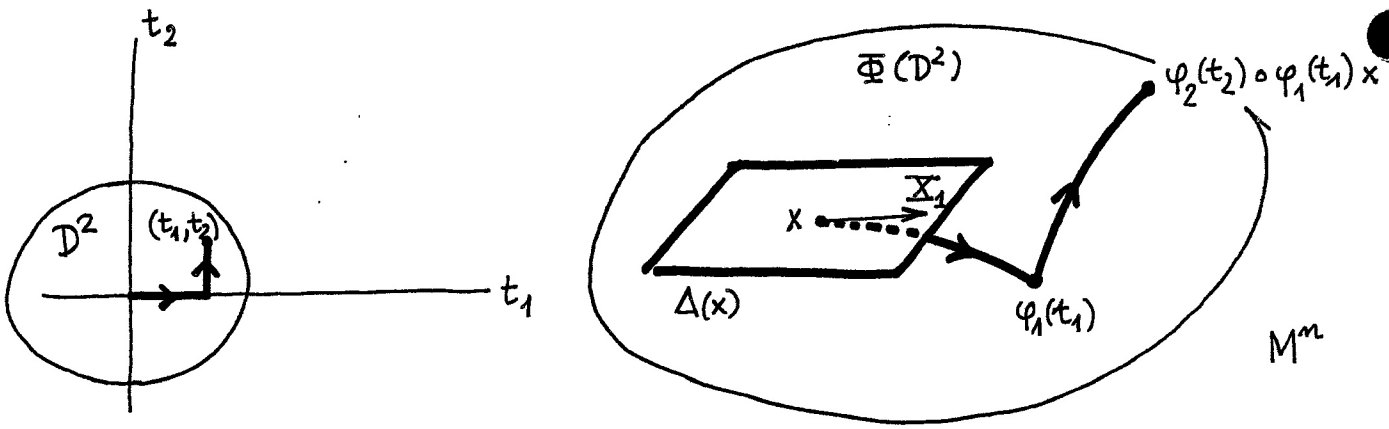
Let $D^k \subset \mathbb{R}^k$ be a small disc about the origin of $\mathbb{R}^k$ (with coordinates $t_1, \ldots, t_k$). Define

$$\Phi : D^k \longrightarrow M^n$$

by

$$\Phi(t_1, \ldots, t_k) = \phi_k(t_k) \circ \phi_{k-1}(t_{k-1}) \circ \cdots \circ \phi_1(t_1)(x) \quad .$$

This is well defined if $|t_1|^2 + \cdots + |t_k|^2$ is small enough.



It should be clear that for the differential of Φ at $t_1 = 0, \ldots, t_k = 0$

$$\Phi_* : \mathbb{R}_0^k \longrightarrow M_x^n$$

$$\Phi_*\left(\frac{\partial}{\partial t_A}\right) = X_A \quad \text{at} \quad x = \Phi(0) \quad .$$

48

- **Problem** . Verify this.

Definition. A smooth map $F : W^k \to M^n$ of manifolds is an <u>immersion</u> and $F(W^k)$ is an <u>immersed</u> submanifold of M^n provided

$$F_* : W^k_p \longrightarrow M_{F(p)}$$

is 1:1 (i.e., $\ker F_* = 0$) at each $p \in W^k$.

In our case Φ_* is 1:1 at $\vec{0} \in \mathbb{R}^k$ and consequently Φ_* is 1:1 in some neighborhood of $\mathbb{R}^k$. Thus the map $\Phi : D^k \to M^n$ defines an immersed submanifold $\Phi(D^k) \subset M^n$ provided D^k is small enough. This holds for <u>any</u> distribution, involutive or not.

It is also clear from $\Phi_* \dfrac{\partial}{\partial t_A} = X_A$ at $x = \Phi(0)$ that the tangent space to this submanifold <u>at</u> x is $\Delta(x)$. We do not know that Δ is everywhere tangent to $\Phi(D^k)$; if it were $\Phi(D^k)$ would be an integral manifold of the distribution!

> **Frobenius Theorem.** If the distribution Δ_k is in involution
>
> $$[\Delta,\Delta] \subset \Delta$$
>
> then each such immersed disc $\Phi(D^k)$ is an integral manifold of Δ and this distribution is completely integrable.

Proof Sketch. The essential point is to show that if Δ is involutive, $[\Delta,\Delta] \subset \Delta$, then Δ is tangent to $\Phi(D^k)$. We know, without any hypothesis, that this is true at $\Phi(0) = x$. From the definition of Φ

$$\Phi(\vec{t}) = \phi_k(t_k) \circ \cdots \circ \phi_1(t_1)(x)$$

we see that Φ_* takes the tangent vector $\dfrac{\partial}{\partial t_A}$ at $\vec{t}$ into the vector

$$\frac{\partial}{\partial h} \phi_k(t_k) \circ \cdots \circ \phi_A(t_A+h) \circ \cdots \circ \phi_1(t_1)(x) \Big]_{h=0}$$

$$= \phi_k(t_k)_* \circ \cdots \circ \phi_{A+1}(t_{A+1})_*[X_A(\phi_A(t_A) \circ \cdots \circ \phi_1(t_1)x)]$$

(you should convince yourself of this by using the 2-dimensional picture preceding Problem 47). But this simply says that the tangent space to $\Phi(D^k)$ at $\Phi(\vec{t})$ can be obtained from the tangent space to $\Phi(D^k)$ at $\Phi(0)$, i.e., $\Delta_k(x)$, by pushing it along by the flows. Thus this tangent space is simply

$$\phi_k(t_k)_* \circ \cdots \circ \phi_1(t_1)_* (\Delta_k(x)) \quad .$$

Thus we need only show that each flow $\phi_A(t_A)$ sends the distribution Δ_k into itself ! This follows from $[\Delta,\Delta] \subset \Delta$ in the following manner.

Let $Y \in \Delta(y)$. We must show $\phi_A(t)_* Y \in \Delta(\phi_A(t)y)$. Let the distribution be defined by the Pfaffians $\theta_1 = 0, \ldots, \theta_r = 0$. We know $\theta_\alpha(Y) = 0$, $\alpha = 1, \ldots, r$. Let $Y_t \equiv \phi_A(t)_* Y$. By construction then (see 17.8) $[X_A, Y] = \mathfrak{L}_{X_A}(Y) = 0$ along the orbit $\phi_A(t)y$. Consider the real valued functions

$$f_\alpha(t) = \theta_\alpha(Y_t) = i_{Y_t}\theta_\alpha \qquad \alpha = 1, \ldots, r \quad .$$

Then

$$\begin{aligned} f'_\alpha(t) &= X_A(i_{Y_t}\theta_\alpha) = \mathfrak{L}_{X_A} i_{Y_t}\theta_\alpha = \text{(see Problem 27)} \\ &= i_{Y_t}\mathfrak{L}_{X_A}(\theta_\alpha) = i_{Y_t}(i_{X_A} d\theta_\alpha + d i_{X_A}\theta_\alpha) \\ &= i_{Y_t} i_{X_A} d\theta_\alpha \quad \text{since} \quad i_{X_A}\theta_\alpha = \theta_\alpha(X_A) = 0 \quad . \end{aligned}$$

Now since Δ is in involution we have from (25.2) part (3)

$$\begin{aligned} f'_\alpha(t) &= i_{Y_t}\{i_{X_A} \Sigma\, \lambda_{\alpha\beta} \wedge \theta_\beta\} = i_{Y_t}\{\Sigma\, \lambda_{\alpha\beta}(X_A)\theta_\beta\} \\ &= \Sigma\, \lambda_{\alpha\beta}(X_A)\theta_\beta(Y_t) = \Sigma\, \lambda_{\alpha\beta}(X_A) f_\beta(t) \quad . \end{aligned}$$

Thus the functions $f_\alpha(t)$ satisfy the linear system

$$\begin{cases} f'_\alpha(t) = \sum\limits_\beta \lambda_{\alpha\beta}(X_A)\, f_\beta(t) \\ f_\alpha(0) = \theta_\alpha(Y) = 0 \end{cases} .$$

By uniqueness $f_\alpha(t) \equiv 0$ and so $\theta_\alpha(Y_t) \equiv 0$, i.e., $Y_t \in \Delta$ for all t. We have shown $\phi_A(t_A)_* \Delta_y = \Delta_{\phi_A(t_A)y}$ as desired.

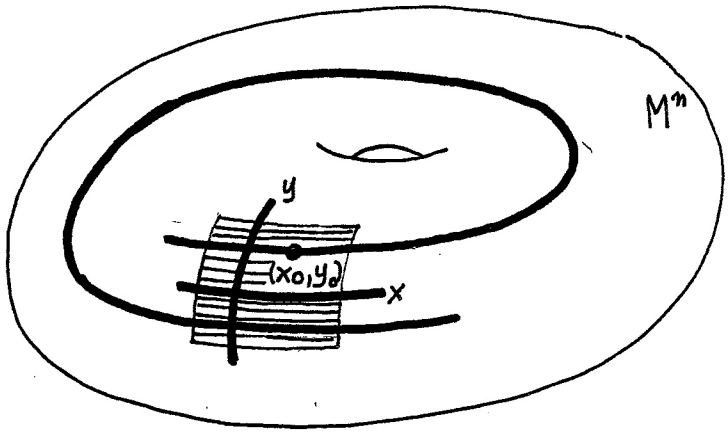
To show complete integrability we must introduce coordinates for which these discs are "slices" $y^1 = c^1, \ldots, y^{n-k} = c^{n-k}$. The procedure is very much like that followed in §1 where we introduced a "coordinate" f by considering a curve C transverse to the integral manifolds. Here we must introduce a transverse (n-k)-manifold W^{n-1}. We shall not go into the details.

Lecture 26. Integrability and Constraints

(1) **Foliations and the maximal leaves.** We know that if the distribution Δ_k on M^n is "involutive", $[\Delta,\Delta] \subset \Delta$, then the distribution is "completely integrable", i.e., in the neighborhood of any point of M^n one can introduce coordinates $x^1, \ldots, x^k, y^1, \ldots, y^{n-k}$, so that each coordinate "slice"

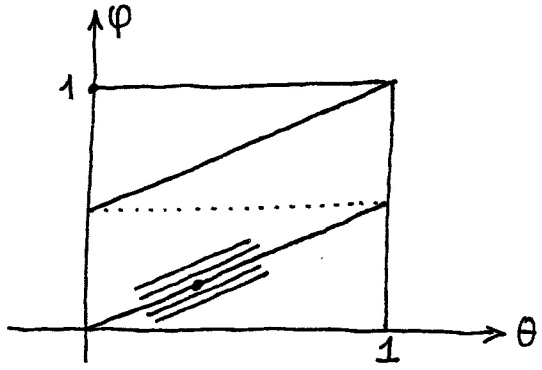
$$y^1 = c^1, \ldots, y^{n-k} = c^k$$

is a k-dimensional integral manifold of Δ. The integral manifold through a given point $(\underset{\sim}{x}_0, \underset{\sim}{y}_0)$ of course also exists outside of the given coordinate system



and it might even return to the coordinate patch. If so it will either appear as the same slice or a different one.

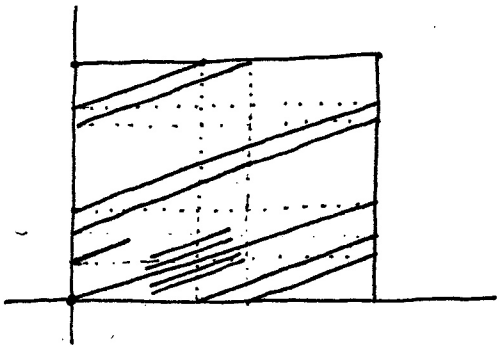
Example. In our usual model of the torus T^2 as a rectangle in the plane (this time with sides of length 1) with periodic identifications



$k = \frac{1}{2}$

consider the distribution Δ_1 defined by $\theta = d\phi - k d\theta$, k constant. The integral manifolds in this case are the curves (straight lines in the rectangle) $\frac{d\phi}{d\theta} = k$, i.e., straight

lines of slope k. If $k = \frac{p}{q}$ is a rational number (we have illustrated above the case $k = \frac{1}{2}$) then the slice through $\theta = 0$, $\phi = 0$ is a closed curve winding q times around the torus in the θ direction and p times in the ϕ direction. On the other hand if k is irrational it is clear that the integral curve leaving $\theta = 0$, $\phi = 0$ will never return to this point but (it turns out) will lie dense on the torus. The integral curve will leave and



k irrational

re-enter each slicing coordinate system (x,y) an infinite number of times.

If a distribution $\Delta_k \subset M^n$ is involutive the integral manifolds define a "foliation" of M^n and each connected integral manifold is called a "leaf" of the foliation. A leaf that is not properly contained in another leaf is called a "maximal leaf". It seems clear from the above example with k irrational that the maximal leaf through (0,0) is not an embedded submanifold; this is because the part of a maximal leaf that lies in any coordinate patch on T^2 will consist of an infinite number of "parallel" segments. There is no chance that we can describe all these segments by a single equation $y = f(x)$, see Lecture 3, §2. However, each "piece" of the leaf does look like a submanifold. Recall the notion of an immersed submanifold from Lecture 25, §(4). If $F : V^k \to M^n$ is a smooth map of a manifold V^k into M^n, then F is an immersion if $F_* : V^k_x \to M^n_{F(x)}$ is 1:1, and $F(V^k)$ is an immersed submanifold.

Here we illustrate two immersions of $\mathbb{R}^1$ into $\mathbb{R}^2$ that do not yield embedded submanifolds. The first fails because F is not 1:1. The second does have an F that is 1:1 but it still fails to be an embedding because the point F(0) is the limit of points F(t) for $t \to \infty$

As we have seen in the case of T^2 a maximal leaf need not be an embedded submanifold. Chevalley, however, has proven

Proposition. The maximal leaf through a point p of a foliated manifold M^n is an immersed submanifold; i.e., there is some manifold V^k and a 1:1 immersion $F : V^k \to M^n$ that realizes the maximal leaf through p.

(2) **Systems of Mayer-Lie.** Classically the Frobenius theory arose out of the study of partial differential equations. An important system of partial differential equations is the system of Mayer-Lie; we are to solve the system

$$\frac{\partial y^\beta}{\partial x^i} = b^\beta{}_i(x,y) \qquad \begin{matrix}\beta = 1, \ldots, r \\ i = 1,\ldots,k\end{matrix} \tag{26.1}$$

where $b^\beta{}_i(x,y)$ is a given matrix of functions and where we demand initial conditions

$$y^\beta = y^\beta{}_0 \quad \text{when} \quad x^i = x^i{}_0 \quad .$$

By equating mixed partial derivatives $\dfrac{\partial^2 y^\beta}{\partial x^j \partial x^i} = \dfrac{\partial^2 y^\beta}{\partial x^i \partial x^j}$ we immediately get the necessary integrability conditions

$$\left(\frac{\partial^2 b^\beta{}_i}{\partial x^j} - \frac{\partial^2 b^\beta{}_j}{\partial x^i}\right) + \sum_{d=1}^{r} \left(\frac{\partial b^\beta{}_i}{\partial y^\alpha}\, b^\alpha{}_j - \frac{\partial b^\beta{}_j}{\partial y^\alpha}\, b^\alpha{}_i\right) = 0 \quad . \tag{26.2}$$

We wish to show that (26.2) is also sufficient for a solution to exist. We restate our problem.

Let $x^1, \ldots, x^k$ be coordinates in $\mathbb{R}^k$, $y^1, \ldots, y^r$ coordinates in $\mathbb{R}^r$. Then in $M^n = \mathbb{R}^k \times \mathbb{R}^r$ we consider the distribution Δ_k defined by the Pfaffians

$$\theta_\beta \equiv dy^\beta - \sum_i b^\beta{}_i(x,y)dx^i = 0\,, \qquad \beta = 1, \ldots, r \quad . \tag{26.3}$$

- **Problem 49.** Verify that $\theta_1, \ldots, \theta_r$ are independent.

Now the Frobenius integrability condition (25.2) (3)

$$d\theta_\beta = \lambda_{\beta\gamma} \wedge \theta_\gamma$$

is simply the algebraic statement that $d\theta_\beta = 0$ after we put $\theta_1 = 0, \ldots, \theta_r = 0$ (that's why we say $d\theta_\beta \equiv 0 \bmod \theta_1, \ldots, \theta_r$). But in our case

$$d\theta_\beta = -d \sum_i b^\beta{}_i(x,y)dx^i = -\sum db^\beta{}_i \wedge dx^i$$

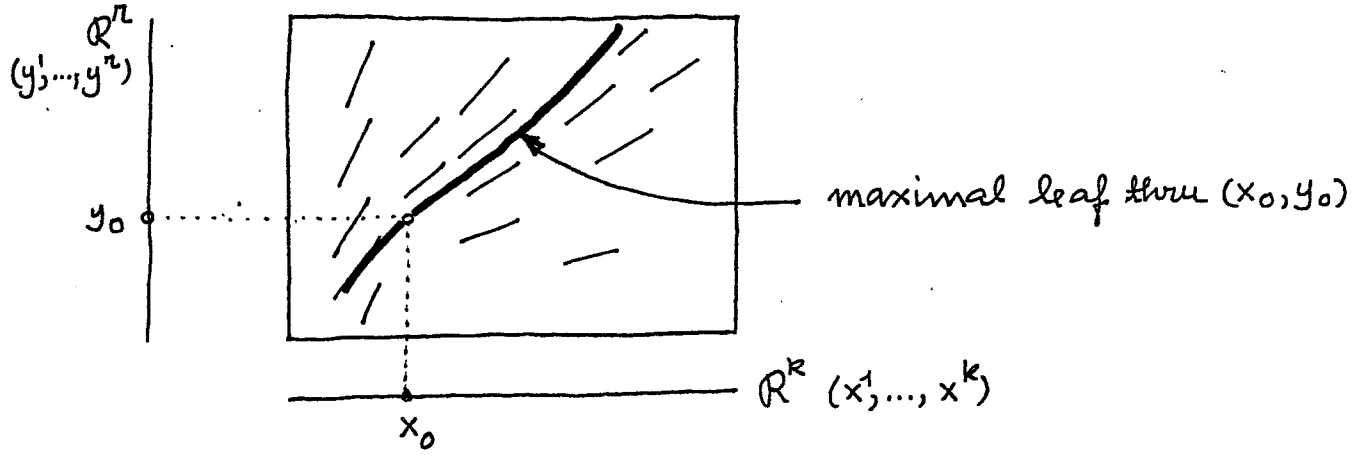
$$= -\sum_i \left(\sum_j \frac{\partial b^\beta{}_i}{\partial x^j}\, dx^j \wedge dx^i + \sum_\alpha \frac{\partial b^\beta{}_i}{\partial y^\alpha}\, dy^\alpha \wedge dx^i \right) \quad .$$

Now to put $\theta_\alpha = 0$ is to say $dy^\alpha = \sum b^\alpha{}_k\, dx^k$, and so

$$d\theta_\beta \equiv -\sum_{i,j} \frac{\partial b^\beta{}_i}{\partial x^j}\, dx^j \wedge dx^i - \sum_{i,\alpha,j} \frac{\partial b^\beta{}_i}{\partial y^\alpha}\, b^\alpha{}_j\, dx^j \wedge dx^i$$

$$= -\sum_{i,j,\alpha} \left(\frac{\partial b^\beta{}_i}{\partial x^j} + \frac{\partial b^\beta{}_i}{\partial y^\alpha}\, b^\alpha{}_j \right) dx^j \wedge dx^i$$

and so $d\theta_\beta \equiv 0 \bmod \theta_1 \cdots \theta_r$ is simply the statement that the coefficients of $dx^j \wedge dx^i$, made skew symmetric in (i,j), must vanish. This is exactly the naive condition (26.2). Hence the distribution defined by (26.3) is completely integrable on $\mathbb{R}^k \times \mathbb{R}^s$.



Let V^k be the maximal leaf through $(\underset{\sim}{x}_0 , \underset{\sim}{y}_0)$. One can see easily from (26.3) that the distribution is never "vertical", i.e., no nonzero vector of the form $\sum a^\beta \dfrac{\partial}{\partial y^\beta}$ ever lies in the distribution. It seems clear from our picture (and it is not difficult to prove) that this implies that the leaf through $(\underset{\sim}{x}_0 , \underset{\sim}{y}_0)$ can be written in the form $y^\beta = y^\beta(x)$ and for these functions we have $\theta_\beta = 0$ when restricted to the leaf, hence $dy^\beta = \sum b^\beta{}_i(x,y)dx^i$. Then

$$\frac{\partial y^\beta}{\partial x^i} = b^\beta{}_i(x,y)$$

as desired. □

(3) **Holonomic and nonholonomic constraints.** Consider a mechanical system with configuration space M^n and local coordinates $q^1, \ldots, q^n$. It may be that the system is constrained to lie on a submanifold of M^n. For example, a particle moving in $M^3 = \mathbb{R}^3$ may be constrained to move only on the unit sphere. In this case we have a single constraining equation $F(x,y,z) = x^2 + y^2 + z^2 = 1$. If we write this constraint in differential form

$$dF = 0 = x\,dx + y\,dy + z\,dz$$

we are assured that if the particle starts on $F = 1$ it will remain there if its coordinates obey the differential constraint. We have reduced the number of degrees of freedom from 3 to 2. In general we can consider on M^n imposing r constraints of the form

$$dF_1 = 0, \ldots, dF_r = 0$$

the system then being constrained to the submanifold V^{n-r} defined by $F_1 = c_1, \ldots, F_r = c_r$ (we are assuming that $dF_1 \wedge dF_2 \wedge \cdots \wedge dF_r \neq 0$).

More generally we can consider constraints of the form

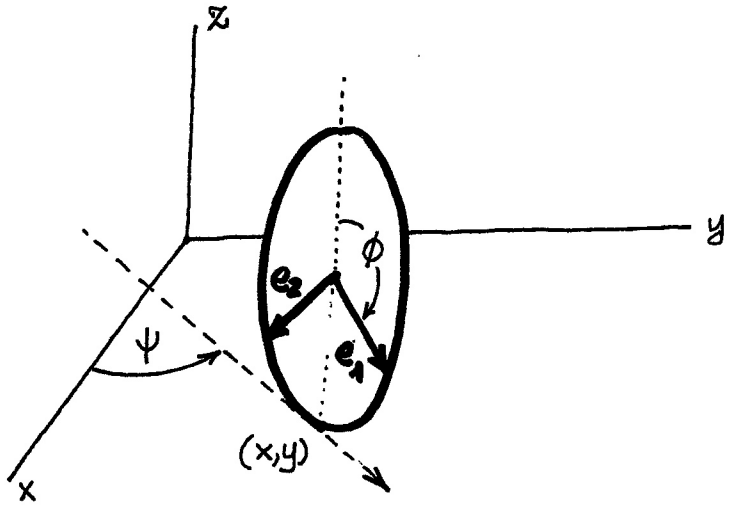
$$\theta_1 = 0, \ldots, \theta_r = 0 \tag{26.4}$$

where each $\theta_\beta = \sum a_{\beta i}(q)dq^i$ is a Pfaffian and again we assume $\theta_1 \wedge \theta_2 \wedge \cdots \wedge \theta_r \neq 0$.

Definition. If the distribution (26.4)) is completely integrable we say that the constraints are integrable or holonomic . If not the constraints are not integrable or non-holonomic .

Of course if the constraints are holonomic, we know from Frobenius' theorem that we can introduce new local coordinates $x^1, \ldots, x^k, y^1, \ldots, y^r$ such that the constraints are now expressed as $dy^1 = 0, \ldots, dy^r = 0$ and the system is (locally) constrained to the submanifolds $y^1 = c^1, \ldots, y^r = c^r$ and the system really only has r degrees of freedom. Non-holonomic constraints are more puzzling.

Consider the classic example of a vertical disc rolling on a horizontal plane "without slipping".



To completely describe the configuration of the disc we engrave a frame e_1 , e_2 of vectors in the disc and consider the point on the rim at the end of e_1 as a distinguished point. The disc is then completely described by

$$(x,y,\psi,\phi)$$

where (x,y) are the coordinates of the center of the disc, ϕ is the angle that e_1 makes with the vertical (positive rotations go from e_1 to e_2) and ψ is the angle that the plane of the disc makes with the x axis. (The line of intersection of this disc-plane and the xy plane is directed as follows; an increase of the angle ϕ will roll the disc in the positive direction along this line.)

It is then clear that the configuration space of this disc is simply

$$M^4 = \mathbb{R}^2 \times S^1 \times S^1 = \mathbb{R}^2 \times T^2 \quad .$$

The condition that the disc roll without slipping is expressed by looking at the motion of the center of the disc. It is

$$\frac{dx}{dt} = \cos\psi \frac{d\phi}{dt}, \qquad \frac{dy}{dt} = \sin\psi \frac{d\phi}{dt} \quad .$$

As it stands this is a condition in TM. We can write it as a pair of constraint equations on M by

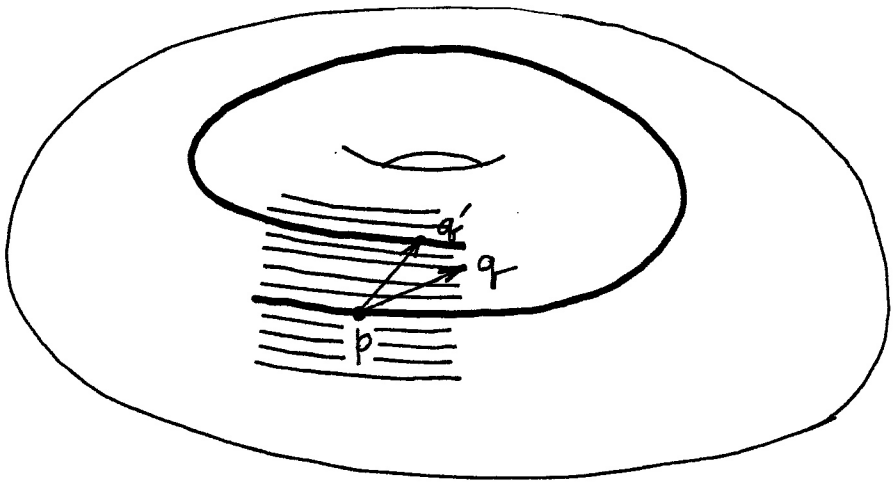
$$\begin{cases} \theta_1 \equiv dx - \cos\psi\, d\phi = 0 \\ \theta_2 \equiv dy - \sin\psi\, d\phi = 0 \end{cases} \quad . \tag{26.5}$$

It would seem that the system really has only 2 degrees of freedom, and this would be the case if the constraints were holonomic. However

$$d\theta_1 = \sin\psi\, d\psi \wedge d\phi$$
$$d\theta_2 = -\cos\psi\, d\psi \wedge d\phi$$

and so $d\theta_1 \wedge (\theta_1 \wedge \theta_2) = \sin\psi\, d\psi \wedge d\phi \wedge dx \wedge dy \neq 0$. From (25.2), part (4) we see that these constraints are non-holonomic. We claim that this mechanical system still has, in some sense, 4 degrees of freedom.

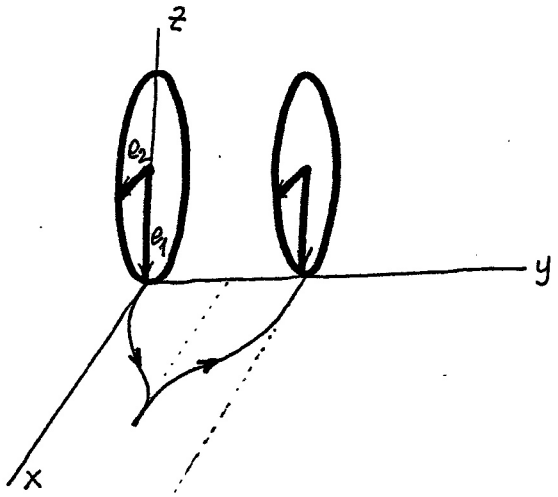
Recall that in the case of integrable constraints we have leaves V^k on which the system point must remain. If we move a small distance in a direction which violates the constraints, i.e., in a direction whose tangent vector is not annihilated by the Pfaffians $\theta_1, \ldots, \theta_r$,



then we automatically end at a point q on a different leaf. There is no way that one can move from p to q, obeying the constraints and remaining in the slicing coordinate

system (whose existence is assured by Frobenius). It is possible that the endpoint q' lies on the same maximal leaf as p, but to go from p to q' while obeying the constraints requires a "long" path, i.e., one leaving the coordinate patch. This is the meaning of the statement that in a holonomic system one has locally only $k = n - r$ degrees of freedom; we must stay on the (n–r) dimensional slice, locally.

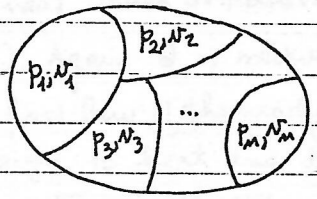
This is not the case in our non-holonomic disc! Consider the disc in an initial state, at the origin and lined up along the x-axis. Now violate the constraints by sliding the disc in the



y direction an arbitrarily small distance. If the systems were holonomic we could not obey the constraints and roll the disc along a small path from the initial to the final configuration. But we can! We have indicated a path in the figure. You should convince yourself that you can obey the constraints and end up at a final state that differs from the initial state by an increment of only one coordinate. In our example above Δy was the only change; a change in ψ is easy since $dx = dy = d\phi = 0$ satisfies the constraints (this is simply revolving the disc about its vertical axis). Thus, although the two constraints limit us "infinitesimally" to two degrees of freedom, we see that actually all neighboring states in a four dimensional region are "accessible" while obeying the constraints. In the general case of r nonholonomic constraints in an M^n there will be a set of states of dimension $> (n–r)$ neighboring a given state that will be accessible via short paths obeying the constraints. The dimension is given by "Chow's Theorem" (see e.g.). We shall discuss a very important special case for thermodynamics in the next lecture.

Lecture 27. Thermodynamics à la Carathéodory

(1) The first law of thermodynamics. Consider a system of regions of gases separated by "diathermous" membranes, i.e. membranes that allow the passage of heat. We assume system is connected.



We assume that each state of the system is a thermal equilibrium state. Let p_i, v_i be the (uniform) pressure and volume of the i^{th} region. The "equations of state" at thermal equilibrium will allow us to eliminate all but one pressure, say p_1; thus a state, instead of being described by $p_1, v_1, \dots, p_m, v_m$ can be described by $p_1, v_1, v_2, \dots, v_m$. It is more convenient for our purposes to assume that there is an "internal energy function" U of the system that can be used instead of p_1. Our state space then has coordinates

$$U = v_0, \ v_1, \ v_2, \ \dots, \ v_m.$$

State space is assumed to be an $n+1$ manifold with those coordinates (in more general situations one keeps U as a global coordinate and lets $v_1, \dots, v_n$ be only local coordinates). Call the manifold M^{n+1}.

A path M^{n+1} represents a sequance of states each in equilibrium. Practically speaking the states must change very slowly in time, so the sequence of states is said to represent a "quasi-static" transition. There are non-quasi-static transitions, for example "stirring". Such transitions start at some state x, and at some state y, but since the intermediate states are not equilibrium states there is no path in M^{n+1} joining x to y that represents the transition. These are "irreversible" processes.

On the manifold M^{n+1} of states we assume the existance of a "work 1-form" W describing work done by the system during a quasi-static adiabatic process

$$W = W^1 = \sum_{i=1}^{m} p_i \, dv_i = \sum_{i=1}^{m} p_i(U, v_1, \dots, v_m) \, dv_i .$$

W is usually neither exact nor closed ($\therefore \int W$ is path dependent)

We also assume the existence of a "heat 1-form"

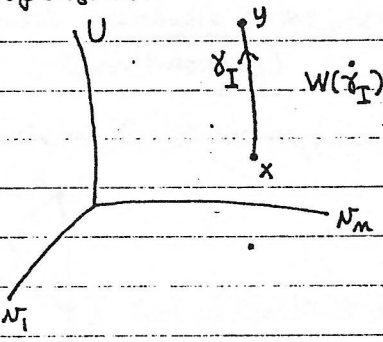
$$Q = Q^1 = \sum_{i=0}^{n} Q_i(U, N_1, \ldots, N_n)\, dN_i \qquad N_0 = U$$

representing heat added or removed from the system (quasi-statically). Again Q is neither exact nor closed. Assume Q never vanishes.

The first law of thermodynamics associates a "mechanical equivalent energy" to heat and demands conservation of energy

$$\boxed{dU = Q - W}$$

(2) Some elementary changes of state.



$W(\dot\gamma_I) = 0 \quad \therefore dU = Q$ along γ_I.

(I) Heating at constant volume. Since $dN_1 = \cdots = dN_n = 0$ along such a γ_I we have $W(\dot\gamma_I) = 0$, where $\dot\gamma_I$ is the tangent vector to γ_I. Then along γ_I we have

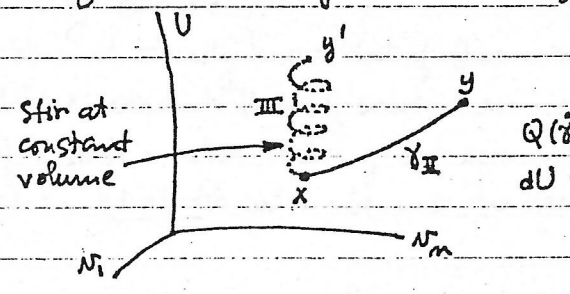
$$dU = Q$$

(II) Quasi-static "adiabatic" process. Here $Q = 0$ along γ_{II} and so

$$dU = -W$$

along γ_{II}.

(III) "Stirring" at constant volume. This is an adiabatic process but since it is not quasi-static we cannot represent it by a curve in M^{n+1}. We schematically indicate it by a dashed curve. Q and W make no sense for this process, but still work is being done on the system, the amount of work done being the difference of internal energy $U(y') - U(x)$.



$Q(\dot\gamma_{II}) = 0$ adiabatic

$dU = -W$ along γ_{II}

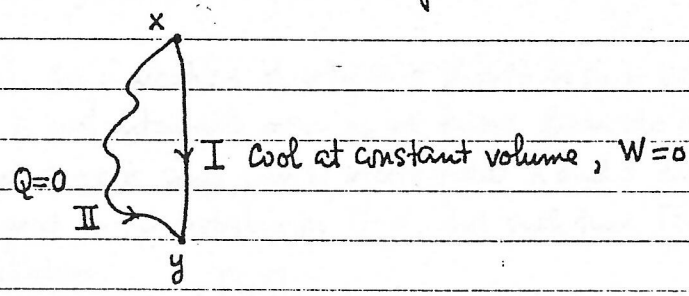
(3) The Second Law of Thermodynamics. According to Lord Kelvin this can be stated as follows

> In no quasi-static process can a quantity of heat be converted entirely into its mechanical equivalent of work. (Kelvin)

Caratheodory (1909) replaced this statement by a weaker one

> On every nbhd of every state x there are states y that are not accesible from x via quasi-static adiabatic paths, $Q=0$ (Caratheodory)

Proof that Kelvin $\Rightarrow$ Caratheodory.



Given state x take a process of type I by cooling at constant volume, $W=0$, ending at state y. We claim that there is no quasi-static adiabatic II going from x to y. For suppose otherwise. We would then have

$$\int_{II} W = \int_{II} Q - dU = -\int_{II} dU = \int_{-II} dU = \int_{-I} dU = \int_{-I} Q$$

But this would say that the heat energy pumped into the system by going from y to x along $-I$ (i.e. heating at constant volume) has been converted entirely into its mechanical equivalent of work $\int_{II} W$ by the hypothetical process II, contradicting Kelvin $\square$.

(Note that no state on I between x and y is quasi-statically adiabatically accesible from x!)

Now we already know that if $Q=0$ were a holonomic (i.e. integrable) constraint then of course there would exist, in any nbhd of a state x other states y not accesible from x along quasi-static adiabatic paths, because the accessible points would

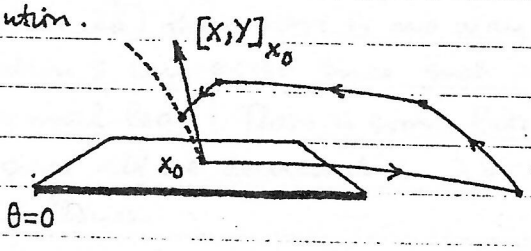
all lie on the maximal leaf (integral submanifold of codimension 1) through x. Does the existence of inaccessible points (ie the 2nd law) conversely imply that the distribution $Q=0$ (the "adiabatic" distribution) must be completely integrable? Caratheodory showed that this is indeed so by proving the following purely mathematical result.

Theorem (Caratheodory). *Let θ be a continuously differentiable non-vanishing 1-form on M^n and suppose that at $x_0 \in M^n$ we have that $\theta = 0$ is not integrable*

$$\theta \wedge d\theta \neq 0.$$

Then there is a nbhd U of x_0 such that any $y \in U$ can be joined to x_0 by a piecewise smooth curve that is always tangent to the distribution.

Proof (Sketch). An indication of why this should be so is easily given. Since $\theta = 0$ is not integrable near x_0 we know from the Theorem on p.199 that there is some pair of vector fields X and Y defined near x_0, always tangent to the distribution $\theta = 0$, but such that $[X,Y]$ is not in the distribution.



From the proposition on p.128 we see that the succession of integral curves,

$$\phi_{-t}(Y) \circ \phi_{-t}(X) \circ \phi_t(Y) \circ \phi_t(X)\,(x_0)$$

each lying in the distribution $\theta = 0$, has endpoint lying approximately in the direction $[X,Y]_{x_0}$ from x_0, ie lying in a direction transverse to the distribution $\theta = 0$. Thus, not only are points "along $\theta = 0$" accessible from x_0, but also some points transverse to $\theta = 0$. It is not difficult to show, using the machinery of Frobenius' theorem developed in lecture 25 that in fact *all* points in some nbhd of x_0 are accessible. For a rigorous proof along these lines see Herman [ ]. □

We thus conclude, from Carathéodory's mathematical theorem and also from his version of the 2nd law of thermodynamics that

> The adiabatic distribution $Q=0$ is completely integrable

(4) Entropy and Absolute Temperature. Since $Q=0$ is integrable, we know (see §(1), Lecture 25) that locally there are functions λ and s defined so that $\lambda \neq 0$ and

$$Q = \lambda\, ds.$$

Since

$$\frac{Q}{\lambda} = ds$$

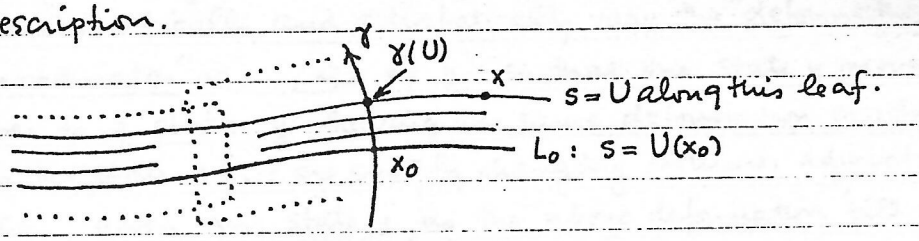
we say that the 1-form Q admits a (local) integrating factor λ (since ds is exact, $\int Q/\lambda$ is locally path independent, i.e. "integrable"). For thermodynamic purposes it is important that λ and s be globally defined. If, e.g., the foliation defined by $Q=0$ has leaves that wind densely in M^{n+1} (see §(1), Lecture 26) then there is no way that such a global function s can exist, since such an s must be constant on each maximal leaf! There is some latitude in the choices of λ and s since all we demand is $\lambda\, ds = \bar{\lambda}\, d\bar{s}$. We proceed as follows.

Pick a state x_0 once and for all and let L_0 be the maximal leaf thru x_0. The prescription in §(1), Lecture 25 asks that we pick out a curve in M^{n+1} passing thru x_0 and transversal to the leaves of $Q=0$. (We shall assume, from now on, that our systems are so simple that the assumptions made below are physically reasonable; for a careful axiomatic treatment see "An axiomatic approach to classical thermodynamics", by J.B. Boyling, Proc. R. Soc. London A, 329 35-70 (1972)).

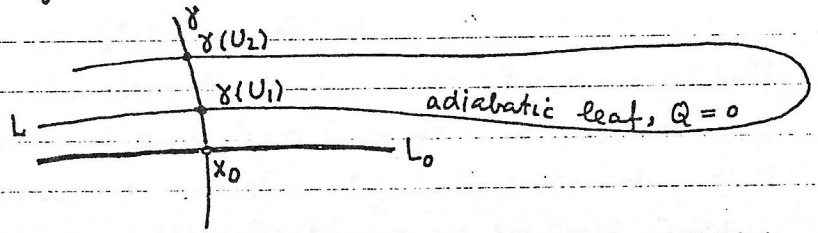
For transversal thru x_0 we choose the process γ "adding heat at constant volume"; $W=0$ along γ and

$dU(\dot\gamma) = Q(\dot\gamma) > 0$ for adding heat, $Q(\dot\gamma) < 0$ for cooling. Since γ is transverse to the leaves we may parameterize γ by the internal energy U. We shall assume that $\gamma(U)$ is defined for all $0 < U < +\infty$.

Now construct a function s near x_0 by our usual prescription.



In a Frobenius rectangular chart we define $s(x) = U$ iff the leaf thru x meet the transversal γ thru x_0 at parameter value U. We continue the function s into a neighboring rectangular chart meeting L_0 by keeping the function s constant on each maximal leaf, etc., etc. Can we run into trouble? This process will fail to define a function s in some nbhd of the leaf L_0 only if some leaf near L_0 strikes the transversal γ thru x_0 in more than one point.

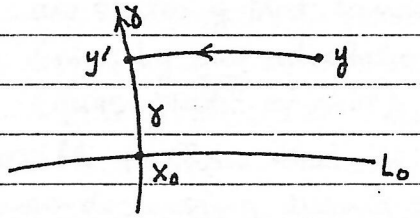


This is not pathology; it happens all the time with general foliations. Kelvin's formulation of the 2nd law, however, shows that this situation cannot arise in thermodynamics, for if γ struck an adiabatic leaf L at $\gamma(U_1)$ and $\gamma(U_2)$, $U_2 > U_1$, then the heat energy $U_2 - U_1 = \int_\gamma Q$ obtained from heating at constant volume, could be converted entirely into work by taking an adiabatic path from $\gamma(U_2)$ to $\gamma(U_1)$ lying on the leaf L

$$U_2 - U_1 = \int_{-\mathcal{C}} dU = \int_{-\mathcal{C}} Q - W = \int_{-\mathcal{C}} -W = \int_{\mathcal{C}} W.$$

Thus s is well defined in some nbhd of the maximal leaf L_0.

Is it possible to define s on the entire manifold M^{n+1} of states? We can do this if we know that every leaf of the foliation $Q=0$ strikes the transversal γ thru x_0 at some parameter value. We shall assume that the system is so simple that given any pair of states x and y one can quasi-statically and adiabatically vary the "deformation coordinates" $\nu_1, \dots, \nu_n$ of y so that the state y moves over to a state y' having the same deformation coordinates as the state x (we are simply changing volumes adiabatically) Then, given any state y, do the above deformation till we arrive at the same deformation coordinates $\nu_1, \dots, \nu_n$ as x_0.



We shall assume then that the state y' can be obtained from x_0 by heating or cooling at constant volume; then y' lies on γ and we have shown that each leaf on M^{n+1} does meet γ and so s is globally defined.

By construction then $Q = \lambda\, ds$ for some globally defined function λ and $\lambda \neq 0$ since we have assumed that the heat 1-form never vanishes. Again by construction $s=U$ along γ, and since $dU = Q$ along γ we see that $\lambda > 0$

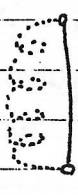
$$Q = \lambda\, ds \qquad \lambda > 0.$$

λ is called an empirical temperature and s is an empirical entropy.

(5) Concluding remarks. For details concerning our final remarks one can consult the article by Boyling mentioned above, and also the very readable treatment "The concepts of classical thermodynamics" by H.A. Buchdahl, Cambridge, 1966.

(i) Experience tells us that if we start at a state y and "stir" the system adiabatically (this process is not quasi-static)

we arrive at a state x having the property that no adiabatic process (quasi-static or not) can return us to y (we cannot "un-stir" the system)!



$U(x) > U(y)$

(Note that x can also be reached from y by heating at constant volume).

If we assume basically this, i.e. if we assume that Carathéodory's version of the 2nd law is strengthened to read

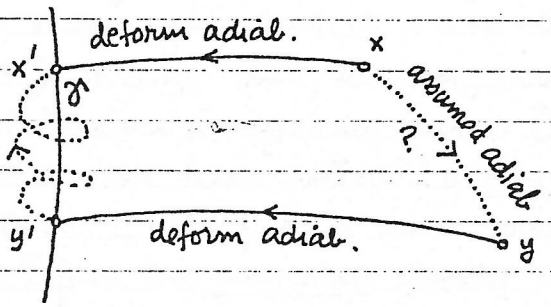
> In every nbhd of every x_0 there are states y that cannot be reached from x_0 by any adiabatic process (quasi-static or not).

then it is possible to show that in our simple systems entropy is non-decreasing during adiabatic transitions, i.e.,

> If y results from x via any adiabatic process, then $S(x) \leq S(y)$

(Of course $S(y) = S(x)$ if process is quasi-static, $dS = \frac{Q}{\lambda} = 0$).

To see this



Suppose that $S(x) > S(y)$ and that there is some adiabatic process $x \longrightarrow y$. By deforming adiabatically (see p. 219) we may move x and y quasi-statically and adiabatically to x' and y' respectively, where both x' and y' lie on our basic transversal γ thru x_0. Then

$$S(x') = S(x) > S(y) = S(y').$$

But along our basic transversal γ, we have by construction

of our function s, that $\partial s/\partial U > 0$, and so $U(x') > U(y')$. We can then stir adiabatically to go from y' to x'. But then the process

$$x' \longrightarrow x \xrightarrow{?} y \longrightarrow y'$$

would be adiabatic and would "un-stir" $y' \to x'$. Thus adiabatic $x \xrightarrow{?} y$ cannot exist. □.

(ii) By combining simple systems into a single compound system (while never introducing internal adiabatic membranes) one can show (Buchdahl, §§ 41, 42) that there is a specific universal choice for the temperature λ called the absolute temperature T. For any simple system, T will be a universal integrating factor

$$\frac{Q}{T} = dS$$

and the resulting empirical entropy will be a "metrical entropy" that is again never decreasing during adiabatic transitions.

(iii) Apparently it is possible to deal with "non-standard" systems (e.g. allowing adiabatic membranes), to define entropy and to show again that it is non-decreasing (Buchdahl § 46). It is not true, however, that for such systems Q will have an integrating factor (Buchdahl § 43)!

(iv) The procedure we have followed is roughly that of Caratheodory (1909), with further improvements by many workers. I find the geometric clarity of the program to be both impressive and beautiful; however, few people know less about thermodynamics than I. For a scathing criticism of Caratheodory's program see C. Truesdell's book "Rational Mechanics", 2nd Ed., Springer, 1984. It seems to me, however, that Truesdell's remarks at the end of section 2, p. 51, obviate his own criticism!

Lecture 28. Poincaré's Theorem: A Preview of Characteristic Classes

In this lecture we shall consider vector fields on surfaces; all that we shall discuss can be generalized to n-manifolds.

(1) **The sum of the indices of a vector field.** Let $\vec{v}$ be a tangent vector field on a surface M^2 and let $z \in M^2$ be an isolated zero of $\vec{v}$; thus $\vec{v}(z) = 0$ but $\vec{v}$ does not vanish elsewhere in some neighborhood of z. As in §(2) of Lecture 24 we can define the (Kronecker) index $j_{\vec{v}}(z)$ of $\vec{v}$ at the zero z as follows.

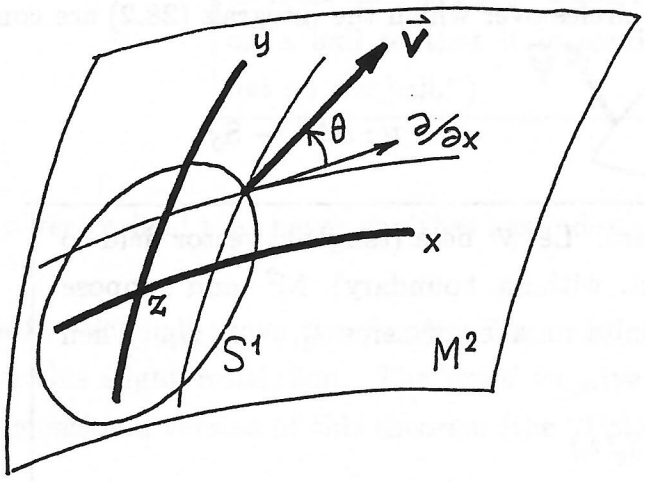
Introduce some Riemannian metric on M^2 (see e.g. (14.2)). Let $(x,y) = (x^1,x^2)$ be some coordinate system with origin at z. Let

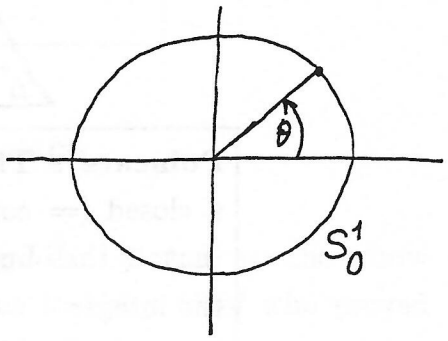
$$S^1 = \{(x,y) \mid x^2 + y^2 = \epsilon^2\} \quad .$$

Then if $\epsilon > 0$ is small enough, S^1 is diffeomorphic to a circle whose interior contains $z = (0,0)$ as only zero of $\vec{v}$. Let

$$\theta(x,y) = \measuredangle\left(\frac{\partial}{\partial x}, \vec{v}\right)$$

be the angle from the vector $\frac{\partial}{\partial x}$ to the vector $\vec{v}$, where $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ defines a positive orientation in the patch. Here $\cos\theta = \dfrac{\langle \partial/\partial x, \vec{v}\rangle}{\|\partial/\partial x\|\,\|\vec{v}\|}$.





We then let S_0^1 be a fixed unit circle in some $\mathbb{R}^2$, parameterized by an angle θ. Then the vector field $\vec{v}$ gives rise to a map

$$S^1 \xrightarrow{\nu} S_0^1$$

$$(x,y) \longrightarrow \theta(x,y) = \measuredangle\left(\frac{\partial}{\partial x}, \vec{v}\right)$$

and we define the (Kronecker) index $j_{\vec{v}}(z)$ by

$$j_{\vec{v}}(z) = \text{Brouwer degree of } \nu : S^1 \to S_0^1 \quad . \tag{28.1}$$

Of course,

$$j_v(z) = \frac{1}{2\pi} \oint_{S^1} d\theta(x,y) \tag{28.2}$$

where $d\theta(x,y) = \nu^* d\theta$ is the pullback of $d\theta$ on S_0^1.

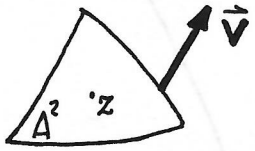
We have made a lot of rather arbitrary choices (Riemannian metric, $\epsilon > 0$) but $j_v(z)$ is independent of the choices since it is an integer and rather clearly is continuous as a function of metric, ϵ, etc.).

Note that if the circle S^1 contains no zero of $\vec{v}$ inside we can still consider the integral and it vanishes

$$\frac{1}{2\pi} \oint_{S^1} d\theta(x,y) = 0 \quad .$$

This is because we may let ϵ shrink down to 0 if $\vec{v}$ never vanishes inside, but as we have just said, the integral is really independent of ϵ. In other words, if the vector field does not vanish at z then the Kronecker index of $\vec{v}$ at z automatically vanishes!

Note also that we could have defined $j_{\vec{v}}(z)$ using a piecewise smooth circle, for example the boundary of a triangle, $\partial\Delta^2$, instead of the smooth S^1; such curves are limits of smooth circles over which the integrals (28.2) are constant.



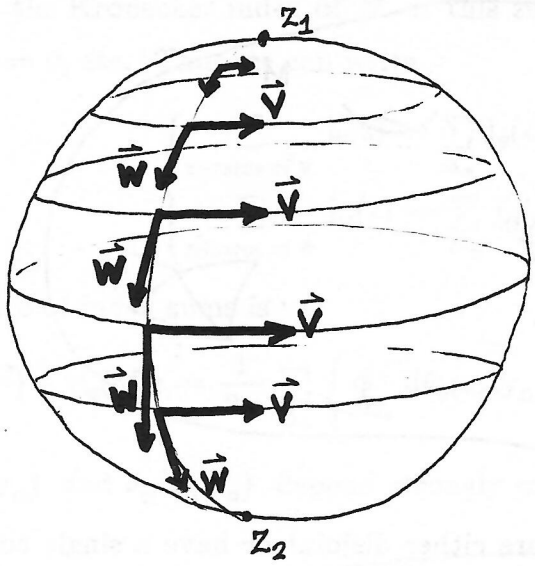
$$\nu : \partial\Delta^2 \to S_0^1$$

Poincaré's Theorem. Let $\vec{v}$ be a (tangent) vector field to a closed (= compact, without boundary) M^2 and suppose that $\vec{v}$ has but a finite number of zeros $z_1, \ldots, z_N$. Then the integer

$$\chi(M^2) = \sum_{i=1}^{N} j_{\vec{v}}(z_i) \tag{28.3}$$

is in fact independent of the vector field $\vec{v}$; it is a topological invariant of the surface M^2 called the Euler- Poincaré characteristic of M^2.

For example we indicate two vector fields $\vec{v}$ and $\vec{w}$ on the 2-sphere $M^1 = S^2$, each having index sum $= 2$. $\vec{v}$ is the



$j_{\vec{v}}(z_1) = 1 = j_{\vec{w}}(z_1)$

$j_{\vec{v}}(z_2) = 1 = j_{\vec{w}}(z_2)$

velocity field of a rotation, while $\vec{w}$ represents the velocity field of molasses oozing over S^2 under the influence of gravity.

We have an easy, famous corollary:

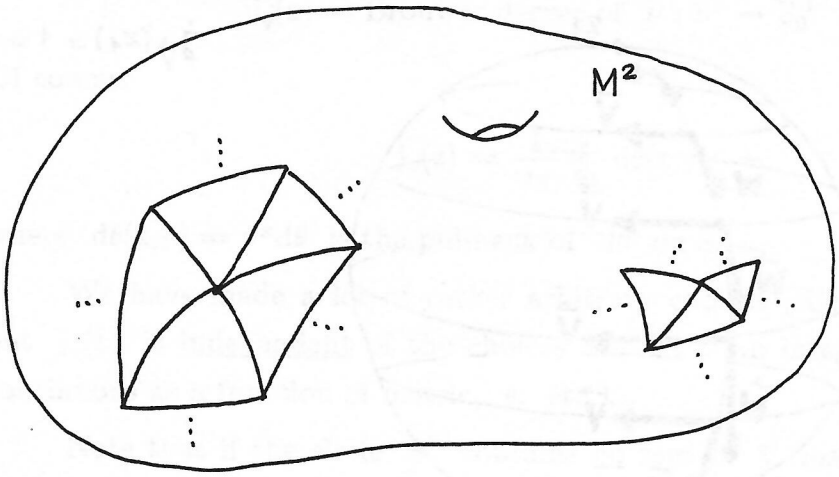
> **Corollary.** Every vector field on S^2 has at least one zero ("you can't comb the hair on a ball so that it is continuous and lies flat on the ball.")

For a vector field that never vanishes has index sum $= 0$.

Proof. We shall prove the theorem for an orientable M^2. Later on we show how to remove this slight restriction. The proof we give is due to Heinz Hopf who proved the n dimensional version of this theorem (the "Poincaré-Hopf" theorem).

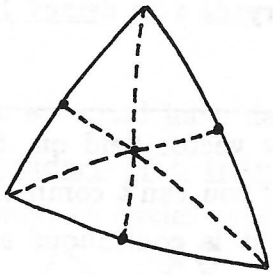
Let $\vec{v}$ and $\vec{w}$ be a pair of vector fields on M^2, each having a finite number of zeros z. Some zero of $\vec{v}$ may coincide with a zero of $\vec{w}$.

It is a fact that M^2 can be "triangulated," i.e.,



covered by "triangles" that are either disjoint or have a single common edge or have a single common vertex. While this result is not easy to prove it is usually very easy to exhibit such triangulations for a specific M^2.

By subdividing triangles



and by moving them around slightly we may assume that the triangles are so small that each of them lies in a coordinate patch and no triangle will have a zero of $\vec{v}$ or $\vec{w}$ on an edge or a vertex, and each triangle will contain at most one zero (of the fields $\vec{v}$ or $\vec{w}$) in its interior. Now with each triangle Δ_α of the triangulation we associate integers (using the coordinates (x_α, y_α))

$$j_{\vec{v}}(\Delta_\alpha) = \frac{1}{2\pi} \oint_{\partial\Delta_\alpha} d\theta_{\vec{v}}$$

$$j_{\vec{w}}(\Delta_\alpha) = \frac{1}{2\pi} \oint_{\partial\Delta_\alpha} d\theta_{\vec{w}}$$

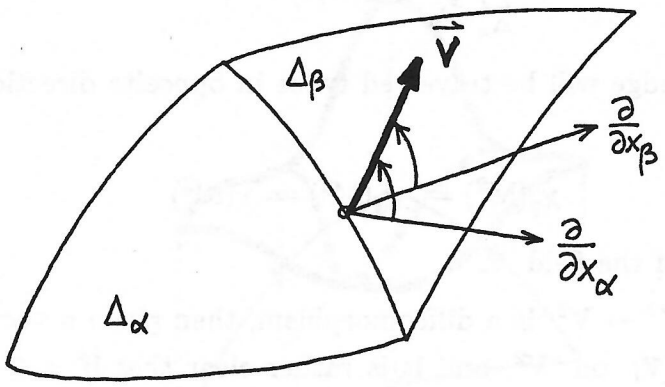
where $\theta_{\vec{v}}(x_\alpha,y_\alpha) = \measuredangle\left(\frac{\partial}{\partial x_\alpha},\vec{v}\right)$, etc. We have seen above that if Δ_α contains a zero of $\vec{v}$, then $j_{\vec{v}}(\Delta_\alpha)$ is the Kronecker index of $\vec{v}$ at this zero, while if Δ_α contains no zero of $\vec{v}$ then $j_{\vec{v}}(\Delta_\alpha) = 0$, etc. Thus we can write

$$\begin{cases} \sum\limits_{z\in\text{zeros of }\vec{v}} j_{\vec{v}}(z) = \sum\limits_{\Delta_\alpha} j_{\vec{v}}(\Delta_\alpha) \\ \sum\limits_{z\in\text{zeros of }\vec{w}} j_{\vec{w}}(z) = \sum\limits_{\Delta_\alpha} j_{\vec{w}}(\Delta_\alpha) \end{cases} \tag{28.4}$$

and so the difference of index sums is

$$\chi_{\vec{v}}(M^2) - \chi_{\vec{w}}(M^2) = \frac{1}{2\pi}\sum_{\Delta_\alpha}\left\{\oint_{\partial\Delta_\alpha} d\theta_{\vec{v}}(x_\alpha,y_\alpha) - \oint_{\partial\Delta_\alpha} d\theta_{\vec{w}}(x,y)\right\} . \tag{28.5}$$

Now while $\theta_{\vec{v}}(x_\alpha,y_\alpha)$ and $\theta_{\vec{w}}(x_\alpha,y_\alpha)$ depend strongly on the coordinate patch used, i.e., on Δ_α



$$d\theta_{\vec{v}}(x_\alpha,y_\alpha) - d\theta_{\vec{w}}(x_\alpha,y_\alpha) = d\,\measuredangle(\vec{w},\vec{v}))$$

is in fact <u>independent</u> of the patch used; it is simply the differential of the angle from $\vec{w}$ to $\vec{v}$ and is determined completely by the Riemannian metric on M^2.

$$\cos\measuredangle(\vec{w},\vec{v}) = \frac{1<\vec{w},\vec{v}>}{\|\vec{w}\|\,\|\vec{v}\|} .$$

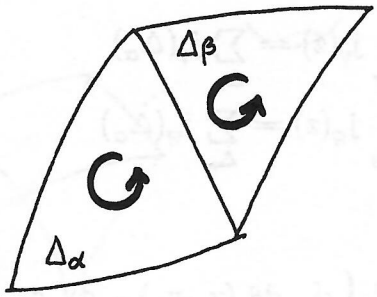
Thus

$$\chi_{\vec{v}}(M^2) - \chi_{\vec{w}}(M^2) = \frac{1}{2\pi}\sum_{\Delta_\alpha}\oint_{\partial\Delta_\alpha} d\,\measuredangle(\vec{w},\vec{v}) . \tag{28.6}$$

Now since M^2 is assumed orientable we may assume that the coordinate patch overlaps satisfy

$$\frac{\partial(x_\beta, y_\beta)}{\partial(x_\alpha, y_\alpha)} > 0$$

and thus the triangles can be oriented compatibly



Then

$$\sum_{\Delta_\alpha} \oint_{\partial\Delta_\alpha} d \not\perp(\vec{w}, \vec{v}) = 0$$

because each common edge will be traversed twice in opposite directions! We have now shown that

$$\chi_{\vec{v}}(M^2) = \chi_{\vec{w}}(M^2) = \chi(M^2)$$

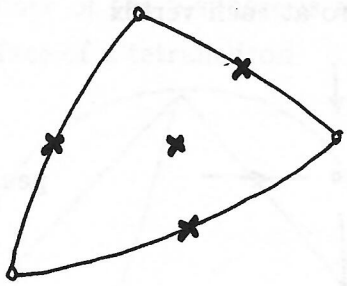
is in fact independent of the field $\vec{v}$.

Note that if $F : M^2 \to V^2$ is a diffeomorphism, then given a vector field $\vec{v}$ on M^2 we get another one $F_*(\vec{v})$ on V^2 and it is rather clear that if $z \in M^2$ is a zero of $\vec{v}$ of Kronecker index j then $F(z)$ will be a zero on V^2 of $F_*(\vec{v})$ of Kronecker index j also. Thus if F is a diffeomorphism then

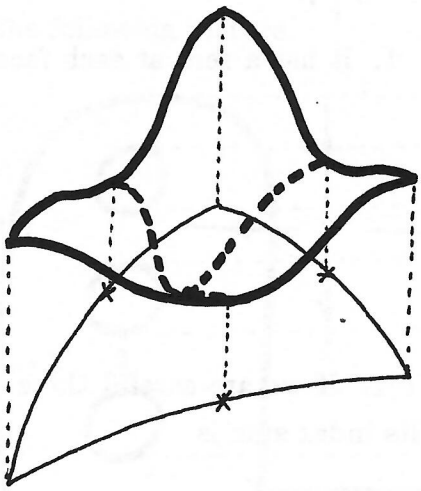
$$\chi(M^2) = \chi(V^2) \quad .$$

(2) $\chi(M)$ **and the Stiefel vector field.** To find $\chi(M)$ we need only define a vector field $\vec{v}$ on M for which we can compute $\sum j_{\vec{v}}$. For this purpose we look at the "Stiefel vector field" (Stiefel was a student of Hopf).

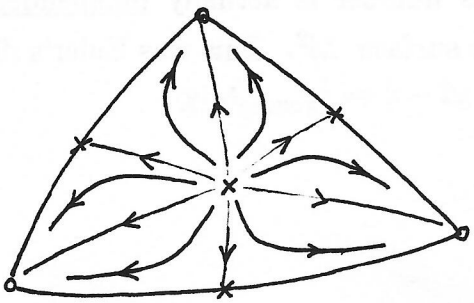
Consider a fixed triangulation of the surface M^2. Imagine the surface to be the surface of a planet. Consider a mountain range constructed as follows. Subdivide each triangle by introducing a "barycenter" on each edge and a "barycenter" for the face.



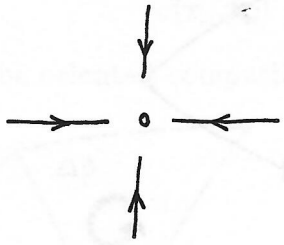
Now construct a mountain peak of height 2 at each original vertex, a mountain pass of height 1 at each edge barycenter and a pit of height 0 at each face barycenter.



This defines a function f on the surface, the altitude function. It has a maximum at each vertex, a minimum at each face barycenter and a saddle (minimax) at each edge barycenter. The gradient vector field ∇f is a vector field on M^2 of the following type (we have included the gradient lines):

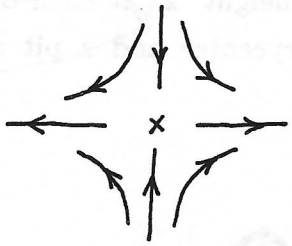


The vector field ∇f has a zero at each vertex



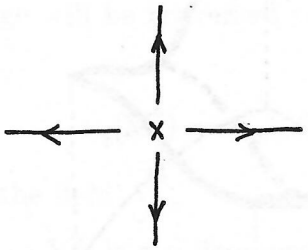
peak

with Kroncecker index j (peak) = + 1. It has a zero at each edge barycenter



pass

with Kronecker index − 1. It has a zero at each face barycenter



pit

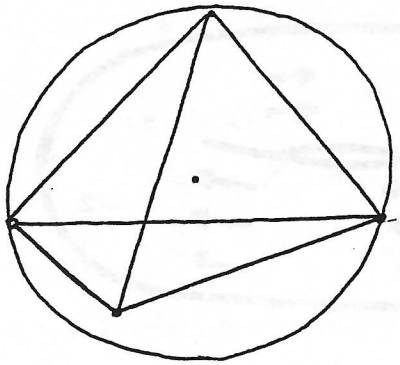
with Kronecker inder + 1. If we are careful there will be no other zeros. This is the Stiefel vector field. Its index sum is

$$\boxed{\begin{aligned}\chi(M^2) &= (\#\text{ pits}) - (\#\text{ passes}) + (\#\text{ peaks}) \\ &= (\#\text{ vertices}) - (\#\text{ edges}) + (\#\text{ faces})\end{aligned}} \tag{28.7}$$

$$\#\text{ vertices} - \#\text{ edges} + \#\text{ faces}$$

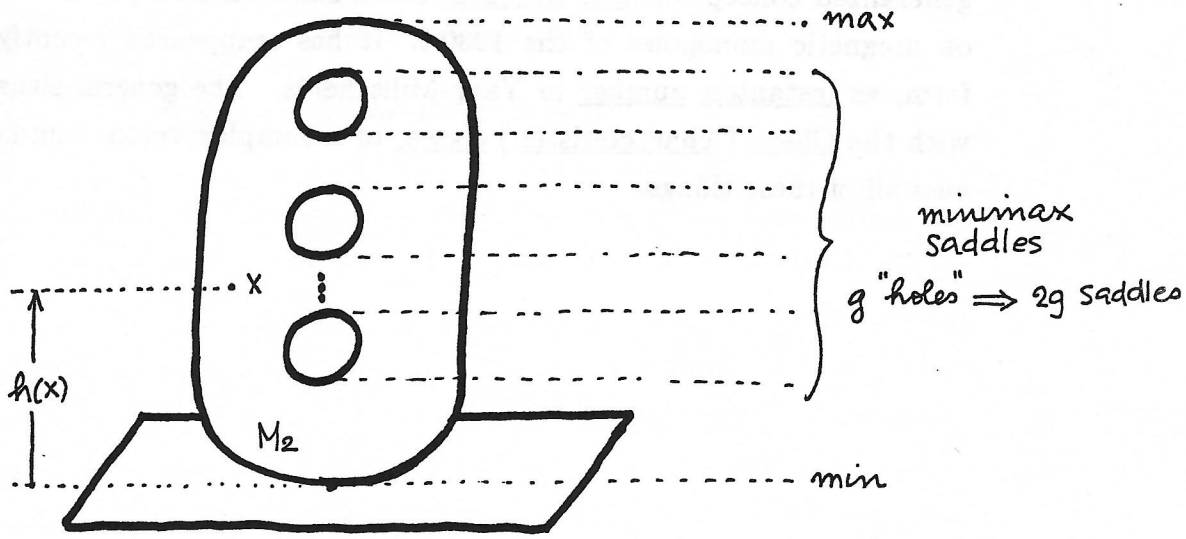
is called the Euler number or Euler characteristic of the triangulation. Poincaré's theorem shows us that this number is actually independent of the triangulation used; it depends only on the surface M^2. This was Euler's discovery.

The Euler number of the 2-sphere is easily computed because S^2 can be triangulated as the surface of a tetrahedron



(use central projection). In this case vertices – edges + faces = 4 – 6 + 4 = 2, agreeing with our examples following (28.3).

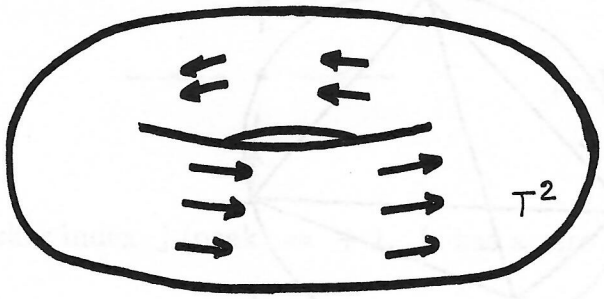
Of course one needn't use the specific Stiefel field. For example, for a surface of genus g consider the following picture.



If we look at the function h on such an M^2 defined by giving the height above a horizontal plane, we see that there is one minimum, where ∇ h has index + 1, one maximum, where ∇ h has index + 1, and 2g saddles where ∇ h has index – 1. Thus the surface of genus g has Euler characteristic

$$\chi(M^2_{\text{genus g}}) = 2 - 2g \quad .$$

For example, the torus has genus 1, $\chi(T^2) = 0$. It is then possible that the torus has a vector field with no zeros, and indeed it does.



(3) **Concluding remarks.** It is difficult for me to restrain my admiration for Poincaré's theorem. In future lectures we shall see the connection with the celebrated Gauss-Bonnet theorem, involving the area integral of the curvature of the surface. Curvature occurs in physics not only in general relativity but also in the generalized concept of field strength. Index sums entered physics with Dirac's work on magnetic monopoles of the 1930's. It has reappeared recently, in generalized form, as instanton number in Yang-Mills fields. The general situation is involved with the Chern (characteristic) classes of a complex vector bundle. We shall discuss all of these things.

More on Mechanics

A <u>symplectic manifold</u> is an even dimensional M^{2n} with a distinguished 2-form ω^2 such that

(i) ω is closed $d\omega = 0$
(ii) non degenerate $\det(\omega_{ij}) \neq 0$
or $\omega^n = \underbrace{\omega \wedge \dots \wedge \omega}_{n} \neq 0$

Best example $M^{2n} = T^*V^n$ ω = point carré two form

(i) M^{2n} is orientable $\omega \wedge \dots \wedge \omega$ defines positive orientation
(iii) <u>Thm of Darboux</u> there always exists locally a coord system $q^1, p_1, q^2, p_2, \dots q^n, p_n$ such that
$\omega = \sum_{i=1}^{n} dp_i \wedge dq^i$

If M^{2n} is a closed symplectic manifold then ω is never globally exact

M^{2n} H smooth equations

"hamilton's equations" $-dH = i_X \omega$

$$-\frac{\partial H}{\partial x^i} = X^j \omega_{ji} = -\omega_{ij} X^j$$

$$\omega_{ij} X^j = \frac{\partial H}{\partial x^i}$$

$$X^j = J^{ji} \frac{\partial H}{\partial x^i} \qquad J = \text{inverse matrix to } \omega$$

$$-dH(X) = i_X \omega(X) = \omega(X, X) = 0 = -\langle \nabla H, X \rangle$$

$\therefore \nabla H \perp$ to X

$$\omega^2(X, \nabla H) = -i_X(\nabla H) = -dH(\nabla H) = -\|\nabla H\|^2$$

$F: M^{2n} \to \mathbb{R}$

$$-dF = i_{X_F}\omega$$

$$-dG = i_{X_G}\omega$$

$$\mathcal{L}_X\omega = i_X d\omega + d(i_X\omega) = d(-dF) = 0$$

($d\omega = 0$)

$\therefore \omega$ is invariant

$$\mathcal{L}_Y\omega = 0 \Rightarrow i_Y d\omega + d(i_Y\omega) = 0$$

$$i_Y\omega = -dF \text{ locally}$$

$\therefore \omega$ is invariant under $X \Leftrightarrow i_X\omega = -dF$ for some F

$$-d\{F, G\} = i_{[X_F, X_G]}\omega$$

$$i_{[X_F, X_G]} = (\mathcal{L}_{X_F} \cdot i_{X_G} - i_{X_G}\mathcal{L}_{X_F})\omega$$

$$= \mathcal{L}_{X_F}(-dG) = -d X_F(G) = -d(\{F, G\})$$

$$\{F, G\} = X_F(G)$$

Maxwell's Eqs

$\vec{B}$ psendo vector $\quad \mathcal{B}^2 = i_{\vec{B}}\, vol_0$ true form

$\vec{E}$ vector $\quad \mathcal{E}^1 = E_1 dx^1 + E_2 dx^2 + E_3 dx^3$ (lower with metric)

covariant lorentz force law $\quad f^1 = e(\mathcal{E}^1 - i_{\vec{v}} \mathcal{B}^2)$

psedo forms $\begin{cases} \mathcal{H}^1 = B_1 dx^1 + B_2 dx^2 + B_3 dx^3 & \text{(lower with metric)} \\ \mathcal{D}^2 = i_{\vec{E}}\, vol_0 \end{cases}$

$\rho^3 = \rho\, vol_0$ = charge psedo 3 form

$j^2 = i_{\rho \vec{v}}\, vol_0$ = current psedo 2 form

psedo forms $\begin{cases} d\mathcal{D}^2 = 4\pi\rho^3 & \text{(Gauss law)} \\ d\mathcal{H}^1 = 4\pi j^2 + \frac{d\mathcal{D}^2}{dt} & \text{(Ampere's law)} \end{cases}$

true forms $\begin{cases} d\mathcal{B}^2 = 0 & \text{(no monopoles)} \\ d\mathcal{E}^1 = -\frac{d\mathcal{B}^2}{dt} & \text{(Faraday's law)} \end{cases}$