

THE DIFFERENTIAL GEOMETRY DIFFERENTIAL TOPOLOGY DUMP

ERNEST YEUNG [ERNESTYALUMNI@GMAIL.COM](mailto:ERNESTYALUMNI@GMAIL.COM)

CONTENTS		
<b>Part 1. Combinatorics, Probability Theory</b>		
1. Distributions and Densities		
<b>Part 2. Linear Algebra Review</b>		
<b>Part 3. Analytical Geometry</b>		
2. Applications of Vector Algebra to Analytic Geometry		
<b>Part 4. Manifolds</b>		
3. Examples		
4. Inverse Function Theorem		
5. Immersions		
6. Submersions; Rank Theorem		
7. Submanifolds; immersed submanifold, embedded submanifolds, regular submanifolds		
8. Curves, Integral Curves, and Flows		
9. Push-forward, Pull-back		
10. Tensors		
<b>Part 5. Lie Groups, Lie Algebra</b>		
11. Lie Groups		
12. Lie Groups and their Lie Algebras		
<b>Part 6. Cohomology; Stoke’s Theorem</b>		
13. Stoke’s Theorem		
<b>Part 7. Prástaro</b>		
<b>Part 8. Connections</b>		
14. Connections of Vector Bundles		
<b>Part 9. Holonomy</b>		
15. Path Groupoid of a smooth manifold; generalization of paths		
<b>Part 10. Complex Manifolds</b>		
<b>Part 11. Jets, Jet bundles, <math>h</math>-principle, <math>h</math>-Prinzipien</b>		
	<b>Part 12. Morse Theory</b>	41
	16. Morse Theory introduction from a physicist	41
2	17. Lagrange multipliers	42
3		
	<b>Part 13. Classical Mechanics</b>	42
4	18. Structure of Galilean Space-Time	43
	19. Kinematics	44
4	20. Fundamental Theorems of (Classical) Dynamics	46
4	21. Rigid Body Dynamics, Rigid Body Motion	50
	22. Waves	50
5	23. Fluid Mechanics, Fluid Flow	54
5	24. Viscous Fluids, Navier-Stokes Equation Applications	58
8	25. Thermodynamics	58
9		
10	<b>Part 14. Classical Mechanics applications</b>	59
13	26. Orbital Mechanics	59
13		
14	<b>Part 15. Quantum Field Theory</b>	64
15		
	<b>Part 16. The Standard Model</b>	64
17		
17	<b>Part 17. General Relativity</b>	65
30		
	<b>Part 18. WE Heraeus International Winter School on Gravity and Light</b>	65
30	Introduction (from EY)	65
30	27. Lecture 1: Topology	65
	Topology Tutorial Sheet	66
30	28. Lecture 2: Topological Manifolds	66
	Tutorial Topological manifolds	67
31	29. Lecture 3: Multilinear Algebra	68
31	30. Lecture 4: Differentiable Manifolds	70
	Tutorial 4 Differentiable Manifolds	71
35	31. Lecture 5: Tangent Spaces	72
39	Lecture 6: Fields	74
	Lecture 7: Connections	76
40	Lecture 8: Parallel Transport & Curvature (International Winter School on Gravity and Light 2015)	78
	Tutorial 8 Parallel transport & Curvature	80
40	Lecture 9: Newtonian spacetime is curved!	81

*Date:* 28 juillet 2016.  
*Key words and phrases.* Differential Geometry, Differential Topology.

32.	Lecture 10: Metric Manifolds
33.	Symmetry
34.	Integration
35.	Lecture 13: Relativistic spacetime
36.	Lecture 14: <a href="#">Matter</a>
37.	Lecture 15: Einstein gravity
	Tutorial 13 Schwarzschild Spacetime
38.	
39.	
40.	Canonical Formulation of GR I
41.	
42.	
43.	
44.	Lecture 22: <a href="#">Black Holes</a>
	References

ABSTRACT. Everything about Differential Geometry, Differential Topology

Part 1. Combinatorics, Probability Theory

**Theorem 1** (4.2. of Feller (1968) [1]). *Let  $r_1, \dots r_k \in \mathbb{Z}$ , s.t.  $r_1 + r_2 + \dots + r_k = n$ ;;  $r_i \geq 0$ .  
Let*

(1) 
$$\frac{N!}{r_1!r_2!\dots r_k!} =$$

*number of ways in which  $n$  elemnts can be divided into  $k$  ordered parts (partitioned into  $k$  subpopulations). cf. Eq. (4.7) of Feller (1968) [1].*  
*Note that the order of the subpopulations is essential in the sense that  $(r_1 = 2, r_2 = 3)$  and  $(r_1 = 3, r_2 = 2)$  represent different partitions. However, no attention is paid to the order within the groups.*

*Proof.*

(2) 
$$\binom{n}{r_1}\binom{n-r_1}{r_2}\binom{n-r_1-r_2}{r_3}\dots\binom{n-r_1-\dots-r_{k-2}}{r_{k-1}} = \frac{n!}{r_1!r_2!\dots r_k!}$$

i.e. in order to effect the desired partition, we have to select  $r_1$  elementsout of  $n$ , remaining  $n - r_1$  elements select a second group of size  $r_2$ , etc. After forming the  $(k - 1)$ st group there remains  $n - r_1 - r_2 - \dots - r_{k-1} = r_k$  elements, and these form the last group.  $\square$

cf. pp. 37 of Feller (1968) [1] Examples. (g) Bridge. 32 cards are partitioned into 4 equal groups  $\rightarrow 52!/(13!)^4$ .  
Probability each player has an ace (?).  
The 4 aces can be ordered in  $4! = 24$  ways, each order presents 1 possibility of giving 1 ace to each player.  
Remaining 48 cards distributed  $(48!)/(12!)^4$  ways.

$$\rightarrow p = 24 \frac{48!}{(12!)^4} / \frac{52!}{(13!)^4}$$

(h) A throw of 12 dice  $\rightarrow 6^{12}$  different outcomes total. Event each face appears twice can occur in as many ways as 12 dice can be arranged in 6 groups of 2 each.

$$\frac{12!}{(2!)^6} / \frac{52!}{(13!)^4}$$

0.0.1. *Application to Occupancy Problems; binomial coefficients.* cf. Sec. 5 Application to Occupancy Problems of Feller (1968) [1].  
Consider randomly placing  $r$  balls intos  $n$  cells.  
Let  $r_k =$  occupancy number = number of balls in  $k$ th cell.  
Every  $n$ -tuple of integers satisfying  $r_1 + r_2 + \dots + r_n = r$ ;  $r_k \geq 0$ . describes a possible configuration of occupancy numbers.  
With indistinguishable balls 2 distributions are distinguishable only if the corresponding  $n$ -tuples  $(r_1, \dots r_n)$  are not identical.  
(i) number of distinguishable distributions is  
(3) 
$$A_{r,n} = \binom{n+r-1}{r} = \binom{n+r-1}{n-1}$$
  
cf. Eq. (5.2) of Feller (1968) [1]  
(ii) number of distinguishable distributions in which no cell remains empty is  $\binom{r-1}{n-1}$ .  
*Proof.* Represent balls by stars, indicate  $n$  cells by  $n$  spaces between  $n + 1$  bars. e.g.  $r = 8$  balls .  
 $n = 6$  cells

$$\begin{array}{ccccccc} & 3 & & 10 & & 00 & & 4 \\ & | & * & * & * & | & * & | \\ & | & * & | & & | & & | \\ & | & * & * & * & * & | & \end{array}$$

Such a symbol necessarily starts and ends with a bar, but remaining  $n - 1$  bars and  $r$  starts appear in an arbitrary order. In this way, it becomes apparent that the number of distinguishable distributions equals the number of ways of selecting.  $r$  places out of  $n + r - 1$ ,  $\frac{(n+r-1)!}{(n-1)!r!} = \binom{n-1+r}{r}$   
 $|||| \dots ||$   $n + 1$  bars  
 $* * * \dots * *$   $r$  stars leave  $r - 1$  spaces

Condition that no cell be empty imposes the restriction that no 2 bars be adjacent.  $r$  stars leave  $r - 1$  spaces of which  $n - 1$  are to be occupied by bars. Thus  $\binom{r-1}{n-1}$  choices.  $\square$

Probability to obtain given occupancy numbers  $r_1, \dots r_n = \frac{r!}{r_1!r_2!\dots r_n!} / n^r$ , with  $r$  balls given by Thm. 4.2. of Feller (1968) [1], which is the Maxwell-Boltzmann distribution.

- (a) Bose-Einstein and Fermi-Dirac statistics. Consider  $r$  indistinguishable particles,  $n$  cells, each particle assigned to 1 cell.  
State of the system - random distribution of  $r$  particles in  $n$  cells.  
If  $n$  cells distinguishable,  $n^r$  arrangements equiprobable  $\rightarrow$  Maxwell-Boltzmann statistics.  
Bose-Einstein statistics: only distinguishable arrangements are considered, and each assigned probability  $\frac{1}{A_{r,n}}$

(4) 
$$A_{r,n} = \binom{n+r-1}{r} = \binom{n-1+r}{n-1}$$

cf. Eq. 5.2 of Feller (1968) [1]  
Fermi-Dirac statistics.  
(1) impossible for 2 or more particles to be in the same cell.  $\rightarrow r \leq n$ .  
(2) all distinguishable arrangements satisfying the first condition have equal probabilities.  
 $\rightarrow$  an arrangement is completely described by stating which of the  $n$  cells contain a particle  
 $r$  particles  $\rightarrow \binom{n}{r}$  ways  $r$  cells chosen.  
Fermi-Dirac statistics, there are  $\binom{n}{r}$  possible arrangements, prob.  $1/\binom{n}{r}$ .  
pp. 39. Feller (1968) [1]. Consider cells themselves indistinguishable! Disregard order among occupancy numbers.  
cf. Feller (1968) [1]

## 1. DISTRIBUTIONS AND DENSITIES

cf. Ch. 5 "Important Distributions and Densities", Grinstead and Snell (1997) [2]

**1.1. Discrete Uniform Distribution.**  $X$  := random variable representing outcome of an experiment where all outcomes equally likely, so  $X$  *uniformly distributed*.

If sample space  $S$  is size  $n$ , where  $0 < n < \infty$ , then distribution function

$$m(\omega) = \frac{1}{n} \quad \forall \omega \in S$$

Consider  $1 + \lfloor x \rfloor \in \{1 \dots n\}$ ,  
with probability  $\frac{1}{n}$  ( $\lfloor x \rfloor$  denotes greatest integer not exceeding  $x$ ).

If sample space is uncountable set with positive finite length, such as  $[0, 1]$ , then use cont. density function.

**1.2. Poisson Random Variable.** Sec. 5.2 "The Poisson Random Variable", Ross (2000) [3].

random variable  $X$ ,  $X \in \{0, 1, 2, \dots\}$  is **Poisson random variable** with parameter  $\lambda$ ,  $\lambda > 0$ , if its probability mass function is

$$(5) \quad P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, \dots$$

Since

$$\sum_{i=0}^{\infty} p(i) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = e^{-\lambda} e^{\lambda} = 1$$

then  $P\{X = i\} = e^{-\lambda} \frac{\lambda^i}{i!}$  defines a probability mass function.

*Moment generating function*  $\phi(t)$  for *Poisson random variable*

$$(6) \quad \phi(t) = E[e^{tX}] = \sum_{i=0}^{\infty} e^{ti} e^{-\lambda} \frac{\lambda^i}{i!} = e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda e^t)^i}{i!} = e^{-\lambda} e^{\lambda e^t} = \exp[\lambda(e^t - 1)]$$

Both *mean* and *variance* of Poisson random variable equal to parameter  $\lambda$ , since

$$\phi'(t) = \phi \cdot (\lambda e^t) = \lambda e^t \exp[\lambda(e^t - 1)]$$

$$\phi''(t) = \phi'(\lambda e^t) + \phi \lambda e^t = (\lambda e^t)(\lambda e^t + 1)\phi$$

$$E[X] = \phi'(0) = \lambda$$

$$\text{Var}(X) = \phi''(0) - (E[X])^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda$$

Poisson random variable when  $n$  large,  $p$  small, may be used as an approximation for binomial random variable. Suppose  $X$  binomial random variable with parameters  $(n, p)$ ,  $\lambda = np$

$$P\{X = i\} = \frac{n!}{(n-i)!i!} p^i (1-p)^{n-i} = \binom{n}{i} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} = \frac{n(n-1)\dots(n-i+1)}{n^i} \frac{\lambda^i}{i!} \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i}$$

For  $n$  large,  $p$  small,

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda} \quad \frac{n(n-1)\dots(n-i+1)}{n^i} \approx 1 \quad \left(1 - \frac{\lambda}{n}\right)^i \approx 1$$

Hence, for  $n$  large,  $p$  small

$$(7) \quad P\{X = i\} \approx e^{-\lambda} \frac{\lambda^i}{i!}$$

i.e. if  $n$  independent trials, each of which results in "success" with probability  $p \ll 1$ , then when  $n$  large ( $n \gg 1$ ) and  $p$  small ( $p \ll 1$ ),

number of successes occurring is approximately Poisson random variable with mean  $\lambda = np$ .

Examples of  $n$  large,  $p$  small for random variables that obey to good approximation the Poisson probability law:

- (1) number of misprints on page (or group of pages) of book  
Suppose there's small probability  $p$  each letter typed on page will be misprinted, so number of misprints on a given page will be approximately Poisson with mean  $\lambda = np$  where  $n$  is (presumably) the large number of letters on the page
- (2) number of people in community living to 100 years of age  
Suppose each person in a given community, independently, has a small probability  $p$  of reaching age 100, and so number of people that do will have approximately a Poisson distribution with mean  $np$  where  $n$  is large number of people in the community.
- (3) number of wrong telephone numbers that are dialed in a day
- (4) number of transistors that fail on first day of use
- (5) number of customers entering post office on given day
- (6) number of  $\alpha$ -particles discharged in fixed period of time from some radioactive particle

Example 5.2a (Ross (2000) [3]), Suppose average number of accidents occurring weekly on highway equals 3. Calculate probability that there's at least 1 accident this week.

**Solution Example 5.2a, Ross (2000) [3].** Let  $X$  denote number of accidents occurring on highway during this week. It's reasonable to suppose there's a large number of cars passing along stretch, each having small probability of being involved in an accident.

$$P\{X \geq 1\} = 1 - P\{X < 1\} = 1 - P\{X = 0\} = 1 - e^{-3} \frac{3^0}{0!} = 1 - e^{-3} = 0.9502$$

Example 5.2b. of Ross (2000) [3]. Suppose  $p$  = prob. item produced by certain machine is defective = 0.1. Find prob. sample of  $N = 10$  items will contain at most 1 defective item.

**Solution Example 5.2b, Ross (2000) [3].**

$$P(X \leq 1) = \sum_{j=0}^1 \binom{N}{j} p^j (1-p)^{N-j} = 0.7361 \text{ for } N = 10, p = 0.1$$

$$P(X \leq 1) \approx \sum_{j=0}^1 e^{-\lambda} \frac{\lambda^j}{j!} = e^{-1} \left( \frac{1^0}{0!} + \frac{1^1}{1!} \right) = 2e^{-1}$$

Example 5.2c of Ross (2000) [3]. Consider experiment counting number of  $\alpha$ -particles given off in 1-sec. interval. If on average 3.2  $\alpha$ -particles given off,

What's good approximation to prob. that no more than 2  $\alpha$ -particles will appear?

**Solution Example 5.2c, Ross (2000) [3].**

$$\lambda = Np = 3.2$$

$$P(X \leq 2) = \sum_{j=0}^2 e^{-\lambda} \frac{\lambda^j}{j!} = e^{-\lambda} \left( 1 + \sum_{j=1}^2 \frac{\lambda^j}{j!} \right)$$

Example 5.2d of Ross (2000) [3]. If average number of claims handled daily by insurance company is 5, what proportion of days have less than 3 claims?

What's prob. there'll be 4 claims in exactly 3 of the next 5 days?

Assume number of claims on different days is independent.

**Solution Example 5.2d, Ross (2000) [3].**

Suppose  $N$  = number of possible claims/clients large, so small prob. of making claim,

$$\lambda = 5$$

$$P(X < 3) = \sum_{j=0}^2 e^{-\lambda} \frac{\lambda^j}{j!} = e^{-\lambda} \sum_{j=0}^2 \frac{\lambda^j}{j!}$$

By law of large numbers (different days independent)  $P(X < 3) = 0.125$  days have fewer than 3 claims.

$$P(X = 4) = \text{prob. there'll be 4 claims on a day} = e^{-\lambda} \frac{\lambda^4}{4!}$$

$$P(Y = 3) = \text{prob. of 4 claims in exactly 3 of next 5 days} = \binom{5}{3} (P(X = 4))^3 (1 - P(X = 4))^2 = \binom{5}{3} (e^{-\lambda} \frac{\lambda^4}{4!})^3 (1 - e^{-\lambda} \frac{\lambda^4}{4!})^2$$

## Part 2. Linear Algebra Review

cf. *Change of Basis*, of Appendix B of John Lee (2012) [5].

**Exercise B.22.** Suppose  $V, W, X$  finite-dim. vector spaces

$S : V \rightarrow W$ ,  $T : W \rightarrow X$

- (a)  $\text{rank} S \leq \dim V$  with  $\text{rank} S = \dim V$  iff  $S$  injective
- (b)  $\text{rank} S \leq \dim W$  with  $\text{rank} S = \dim W$  iff  $S$  surjective
- (c) if  $\dim V = \dim W$  and  $S$  either injective or surjective, then  $S$  isomorphism
- (d)  $\text{rank} TS \leq \text{rank} S$   $\text{rank} TS = \text{rank} S$  iff  $\text{im} S \cap \ker T = 0$
- (e)  $\text{rank} TS \leq \text{rank} T$   $\text{rank} TS = \text{rank} T$  iff  $\text{im} S + \ker T = W$
- (f) if  $S$  isomorphism, then  $\text{rank} TS = \text{rank} T$
- (g) if  $T$  isomorphism, then  $\text{rank} TS = \text{rank} S$

EY : Exercise B.22(d) is useful for showing the chart and atlas of a Grassmannian manifold, found in the More examples, for smooth manifolds.

*Proof.* (a) Recall the **rank-nullity theorem**:

**Theorem 2** (Rank-Nullity Theorem).

$$(8) \quad \dim(\text{im}(S)) + \dim(\ker(S)) = \dim V$$

Now

$$\text{rank}(S) + \dim(\ker(S)) \equiv \dim(\text{im}(S)) + \dim(\ker(S)) = \dim V$$

$$\implies \boxed{\text{rank}(S) \leq \dim V}$$

If  $\text{rank}(S) = \dim V$ ,

then by rank-nullity theorem,  $\dim(\ker(S)) = 0$ , implying that  $\ker S = \{0\}$ .

Suppose  $v_1, v_2 \in V$  and that  $S(v_1) = S(v_2)$ . By linearity of  $S$ ,  $S(v_1) - S(v_2) = S(v_1 - v_2) = 0$ , which implies, since  $\ker S = \{0\}$ , that  $v_1 - v_2 = 0$ .

$\implies v_1 = v_2$ . Then by definition of injectivity,  $S$  injective.

If  $S$  injective, then  $S(v) = 0$  implies  $v = 0$ . Then  $\ker S = \{0\}$ . Then by rank-nullity theorem,  $\text{rank}(S) = \dim V$ .

(b)  $\forall w \in \text{im}(S)$ ,  $w \in W$ . Clearly  $\text{rank} S \leq \dim W$ .

If  $S$  surjective,  $\text{im}(S) = W$ . Then  $\dim(\text{im}(S)) = \text{rank} S = \dim W$ .

If  $\text{rank} S = \dim W = m$ , then  $\text{im}(S)$  has basis  $\{y_i\}_{i=1}^m$ ,  $y_i \in \text{im}(S)$ , so  $\exists x_i \in V$ ,  $i = 1 \dots m$  s.t.  $S(x_i) = y_i$ , with  $\{S(x_i)\}_{i=1}^m$  linearly independent.

Since  $\{S(x_i)\}_{i=1}^m$  linearly independent and  $\dim W = m$ ,  $\{S(x_i)\}_{i=1}^m$  basis for  $W$ .

$\forall w \in W$ ,  $w = \sum_{i=1}^m w^i S(x_i) = S(\sum_{i=1}^m w^i x_i)$ .  $\sum_{i=1}^m w^i x_i \in V$ .  $S$  surjective.

(c)

(d) Now

$$\dim V = \text{rank} TS + \text{nullity} TS$$

$$\dim V = \text{rank} S + \text{nullity} S$$

$\ker S \subseteq \ker TS$ , clearly, so  $\text{nullity} S \leq \text{nullity} TS$

$$\implies \boxed{\text{rank} TS \leq \text{rank} S}$$

If  $\text{rank} TS = \text{rank} S$ ,

then  $\text{nullity} S = \text{nullity} TS$

Suppose  $w \in \text{Im} S \cap \ker T$ ,  $w \neq 0$

Then  $\exists v \in S$ , s.t.  $w = S(v)$  and  $T(w) = 0$

Then  $T(w) = TS(v) = 0$ . So  $v \in \ker TS$

$v \notin \ker S$  since  $w = S(v) \neq 0$

This implies  $\text{nullity} TS > \text{nullity} S$ . Contradiction.

$$\implies \text{Im} S \cap \ker T = 0$$

If  $\text{Im} S \cap \ker T = 0$ ,

Consider  $v \in \ker TS$ . Then  $TS(v) = 0$ .

. Then  $S(v) \in \ker T$

$S(v) = 0$ ; otherwise,  $S(v) \in \text{Im} S$ , contradicting given  $\text{Im} S \cap \ker T = 0$

$v \in \ker S$

$\ker TS \subseteq \ker S$

$$\implies \ker TS = \ker S$$

So  $\text{nullity} TS = \text{nullity} S$

$$\implies \text{rank} TS = \text{rank} S$$

(e)

(f)

(g)

□

## Part 3. Analytical Geometry

### 2. APPLICATIONS OF VECTOR ALGEBRA TO ANALYTIC GEOMETRY

cf. Ch. 13 "Applications of Vector Algebra to Analytic Geometry", Apostol (1991) [45]

2.1. **Lines in  $n$ -space.** cf. Sec. 13.2 "Lines in  $n$ -space", Apostol (1991) [45].

**Definition 1** (Line). *Let given pt.  $P$ , given  $A$  nonzero vector.*

*Define line  $L(P; A)$*

$$(9) \quad L(P; A) := \{P + tA | t \in \mathbb{R}\} \equiv L(P; A) = \{P + tA\}$$

2.1.1. *Some simple properties of a straight line.* cf. Sec. 13.3 "Some simple properties of straight line", Apostol (1991) [45].

**Theorem 3** (Thm. 13.4 of Apostol (1991) [45]). *2 distinct pts. determine a line, i.e. if  $P \neq Q$ ,  $\exists!$  line containing both  $P$  and  $Q$ , namely*

$$(10) \quad \{P + t(P - Q)\}$$

*Proof.* Let  $L$  be line parallel through  $P$  to  $Q - P$ , i.e.  $L = L(P; Q - P) = \{P + t(Q - P)\}$ .

$P, Q \in L$  for  $t = 0, 1$ , respectively.

Suppose  $L'$  is another line containing  $P$  and  $Q$ . Since  $P \in L'$ ,  $L' = L'(P, A)$ ,  $Q \in L'$  so  $Q = P + At$ .

$Q - P = tA$ .  $t \neq 0$  since  $Q \neq P$ . So  $A = \frac{Q-P}{t}$ , i.e.  $A$  and  $Q - P$  are parallel.

By Thm. 13.1, since  $L(P; Q - P), L'(P, A)$  through same pt.  $P$  and direction vectors  $Q - P, A$  parallel, then  $L = L'$ .

□

**2.2. Planes in Euclidean  $n$ -plane.** cf. Sec. 13.6 ”Planes in Euclidean  $n$ -space”, Apostol (1991) [45].

**Definition 2.** Set  $M$  of pts. in  $V_n$  is called a plane if  $\exists$  pt.  $P$  and 2 linearly independent vectors  $A$  and  $B$  s.t.

$$(11) \quad M = \{P + sA + tB | s, t \in \mathbb{R}\}$$

EY 20211207: Consider the intersection of 2 planes:  $M \cap N$ :

$$M = \{P + sA + tB | s, t \in \mathbb{R}\}$$

$$N = \{Q + uC + vD\}$$

Let  $X \in M$  and  $X \in N$  so  $X = P + sA + tB = Q + uC + vD$ .

Let  $M \cap N \neq \emptyset$ . Let  $X_0 \in M \cap N$ ,  $X_0 = P + s_0A + t_0B = Q + u_0C + v_0D$ .

## Part 4. Manifolds

Schuller (2015) [37]

- (1) ~~(strikethrough)~~
- (2)
- (3) transitivity
- (4) totality:  $a \leq b$  or  $b \leq a$ ,  $\forall a, b$

**Definition 3** (upper bound). *Schuller, original, version 1:*

$u \in P$  **upper bound** to a subset  $T \subseteq P$  if

$\forall t \in T : t \leq u$  (EY: in  $t \leq u$ ,  $\leq$  is for partial ordered set)

*version 2, possible alternative (bottomline: not right):*

(of is it  $u \in P$  **upper bound** to a subset  $T \subseteq P$  if  $\nexists t \in T$  s.t.  $u \leq t$ )

## 3. EXAMPLES

### 3.1. Euclidean space.

3.1.1.  $\mathbb{R}^2$ . Consider  $\mathbb{R}^2$  as a smooth manifold.

Clearly,  $\mathbb{R}^2$  is covered by 1 chart, with so-called *Cartesian* coordinates (that also happen to be global coordinates, in this particular case):  $(\mathbb{R}^2, \mathcal{O}_{\text{std}}, \mathcal{A} = \{= \{(\mathbb{R}^2, \varphi_{\text{cart}})\})$ , with the chart (which is a homeomorphism, by definition)

$$\varphi_{\text{cart}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$\varphi_{\text{cart}} : p \equiv (x, y) \mapsto (x, y)$$

Clearly,  $\varphi_{\text{cart}} = 1_{\mathbb{R}^2}$  is a homeomorphism; in fact, it’s a smooth diffeomorphism.

$\mathcal{O}_{\text{std}}$  denotes the so-called “standard” topology, that open sets consists of Euclidean  $d$ -balls. I will assume this topology for this discussion of  $\mathbb{R}^d$ ’s.

Clearly, the above discussion generalizes to  $\mathbb{R}^d$ .

Consider the so-called *polar* coordinates

$$(12) \quad \begin{aligned} x &= r \cos \varphi \\ y &= r \sin \varphi \end{aligned}$$

where  $r \geq 0$  and  $\varphi \in \mathbb{R}$ .

Recall the definition of a coordinate char: a *coordinate chart* (or just chart) on  $M$  is a pair  $(U, \varphi)$  where  $U$  is an open subset of  $M$ , homeomorphism  $\varphi : U \rightarrow \widehat{U}$ , where open subset  $\widehat{U} = \varphi(U) \subseteq \mathbb{R}^n$ .

We want to construct a homeomorphism  $\varphi_{\text{pol}}^{-1}$  from an open subset  $V \subseteq \mathbb{R}^2$  to  $M = \mathbb{R}^2$  according to these polar coordinates:

$$\varphi_{\text{pol}}^{-1} : V \subseteq \mathbb{R}^2 \rightarrow \varphi_{\text{pol}}^{-1}(V) \subseteq M = \mathbb{R}^2$$

$$\varphi_{\text{pol}}^{-1} : (r, \varphi) \mapsto (x, y)$$

Immediately, from this choice of polar coordinates, we see problems with defining a homeomorphism, which is a bijection by definition. For instance, let  $(\varphi_{\text{pol}}^{-1})^{-1} \equiv \varphi_{\text{pol}}$  (defining this highly suggestive notation), and so

$$\varphi_{\text{pol}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$$

$$\varphi_{\text{pol}} : (x, y) \mapsto (r, \varphi)$$

The single point, the origin of  $\mathbb{R}^2$ , gets mapped to many possible values! In other words,

$$\varphi_{\text{pol}}(0, 0) = \{(r, \varphi) | r = 0 \text{ and } \varphi \in \mathbb{R}\}$$

Clearly this must be the case for any kind of coordinates involving cos’s, sin’s, etc. (i.e. transcendental functions). So the choice of these kinds of coordinates necessitate giving up on the possibility of a single chart covering the entire manifold of the Euclidean space.

Consider

$$V := \{(r, \varphi) | r > 0, \varphi \in (0, 2\pi)\} \subset \mathbb{R}^2$$

$$[0, \infty) \times [0, 2\pi) - V = \{(r, \varphi) | r = 0 \text{ or } \varphi = 0\}$$

Now  $\varphi_{\text{pol}}^{-1}([0, \infty) \times [0, 2\pi)) = \mathbb{R}^2$ , i.e.  $\varphi_{\text{pol}}^{-1}$  maps  $[0, \infty) \times [0, 2\pi)$  *onto*  $\mathbb{R}^2$ . But it maps  $\varphi_{\text{pol}}^{-1}(0, \varphi)$  onto  $(0, 0)$  for infinitely many values of  $\varphi \in [0, 2\pi)$ .

Now consider if  $\varphi = 0$ . Then  $x = r$  for  $r > 0$ . So  $x > 0$  and  $y = 0$ . If  $r = 0$ , then  $(x, y) = (0, 0)$ . Thus

$$y = 0$$

$$\varphi_{\text{pol}}^{-1}([0, \infty) \times [0, 2\pi) - V) = \{(x, y) | x \geq 0 \text{ and } y = 0\}$$

(If you have trouble seeing this, think of the union of the cases when  $r = 0$  or when  $\varphi = 0$ ). Thus

$$\varphi_{\text{pol}}^{-1}(V) = \{(x, y) | x < 0 \text{ or } y \neq 0\}$$

Thus, in summary, we can include this chart into atlas  $\mathcal{A}$ :

(13)

$$(U, \varphi_{\text{pol}}) \in \mathcal{A}$$

$$\varphi_{\text{pol}} : U \rightarrow \mathbb{R}^2$$

$$U = \{(x, y) | x < 0 \text{ or } y \neq 0\} \quad \varphi_{\text{pol}}(x, y) = (\sqrt{x^2 + y^2}, \begin{cases} \arccot\left(\frac{x}{y}\right) & \text{if } (x > 0 \text{ and } y \geq 0) \text{ or } (x > 0 \text{ and } y \leq 0) \\ \arccot\left(\frac{x}{y}\right) + \pi & \text{if } (x < 0 \text{ and } y \leq 0) \text{ or } (x \geq 0 \text{ and } y < 0) \end{cases})$$

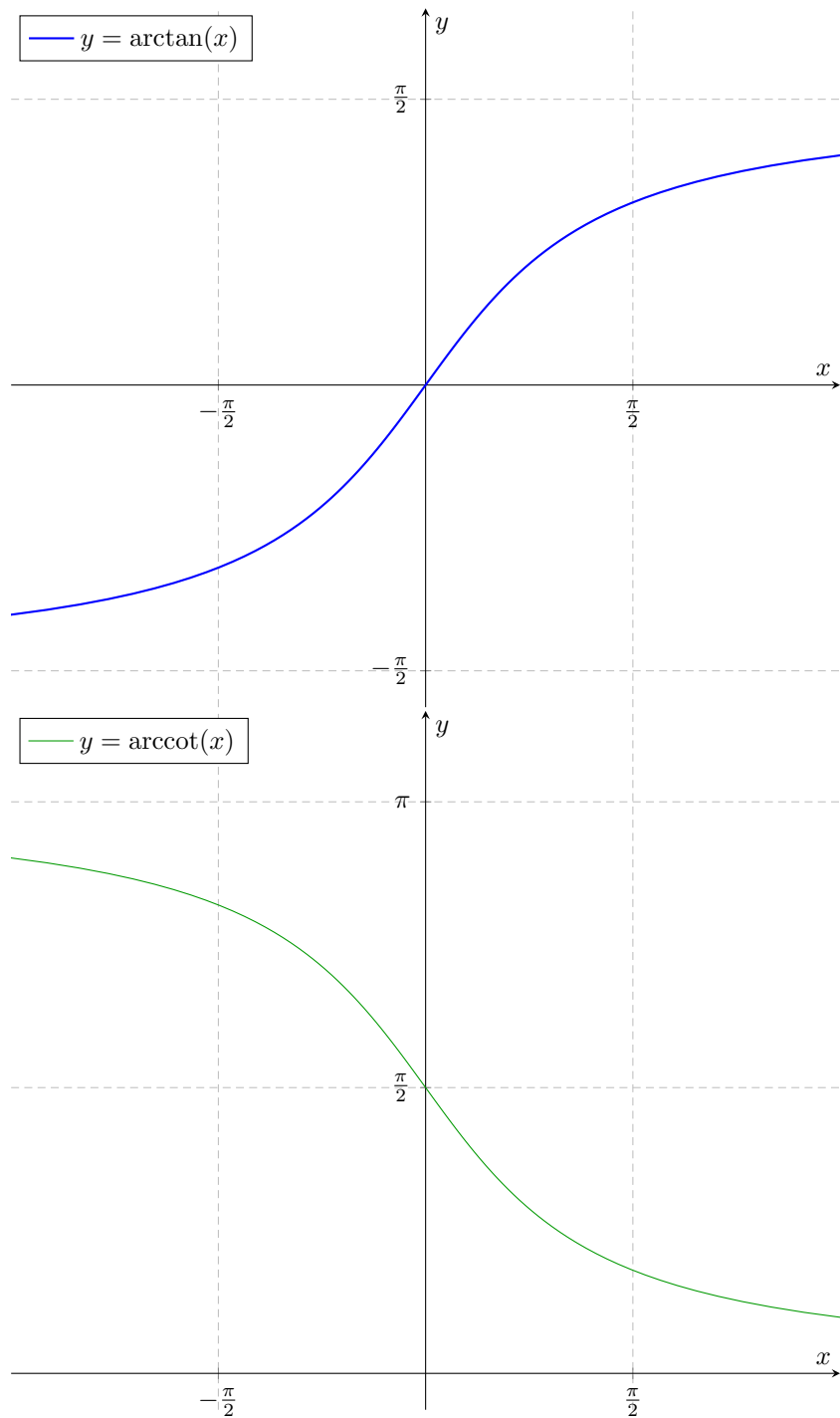
$$\varphi_{\text{pol}}(U) = \{(r, \varphi) | r > 0, \varphi \in (0, 2\pi)\}$$

$$\varphi_{\text{pol}}^{-1} : \varphi_{\text{pol}}(U) \rightarrow \mathbb{R}^2$$

$$\varphi_{\text{pol}}^{-1}(r, \varphi) = (r \cos \varphi, r \sin \varphi) = (x, y)$$

where  $0 < \arccot(x) < \pi$ , with limits  $\lim_{x \rightarrow -\infty} \arccot(x) = \pi$  and  $\lim_{x \rightarrow \infty} \arccot(x) = 0$  being well defined.

Recall the plots of arctan and arccot:



Let's try to cover  $\mathbb{R}^2$  as much as possible with these "polar coordinate"-type charts. Unfortunately, we'll find out that we can't cover the origin in a well-defined (meaning with one-to-one functions) manner, so we'll have to consider  $\mathbb{R}^2 \setminus 0 := \mathbb{R}^2 - \{0\}$  (the Euclidean space  $\mathbb{R}^2$  but without the origin).

Repeat some of the steps we did above. Consider open subset

$$U_+ = \{(x, y) | x > 0 \text{ or } y \neq 0\}$$

obtained by considering

$$V_+ = \{(r, \varphi) | r > 0, \varphi \in (-\pi, \pi)\} \in \mathbb{R}^2$$

$$[0, \infty) \times (-\pi, \pi] - V_+ = \{(r, \varphi) | r = 0 \text{ or } \varphi = \pi\}$$

Consider if  $\varphi = \pi$ . Then  $x = -r$  for  $r > 0$ , so  $x < 0$  and  $y = 0$ . If  $r = 0$ , then  $(x, y) = (0, 0)$ . So  $\varphi_{\text{pol},+}^{-1}(V_+) = \{(x, y) | x > 0 \text{ or } y \neq 0\}$

$$(U_+, \varphi_{\text{pol},+}) \in \mathcal{A}$$

$$U_+ = \{(x, y) | x > 0 \text{ or } y \neq 0\}$$

$$\varphi_{\text{pol},+}(U) = \{(r, \varphi) | r > 0, \varphi \in (-\pi, \pi)\}$$

$$\varphi_{\text{pol},+} : U_+ \rightarrow \mathbb{R}^2 \setminus 0$$

$$(14) \quad \varphi_{\text{pol},+}(x, y) = (\sqrt{x^2 + y^2}, \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } (x > 0 \text{ and } y \geq 0) \text{ or } (x > 0 \text{ and } y \leq 0) \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } (x \leq 0 \text{ and } y > 0) \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{if } (x \leq 0 \text{ and } y < 0) \end{cases})$$

$$\varphi_{\text{pol},+}^{-1} : \varphi_{\text{pol},+}(U) \rightarrow \mathbb{R}^2 \setminus 0$$

$$\varphi_{\text{pol},+}^{-1}(r, \varphi) = (r \cos \varphi, r \sin \varphi) = (x, y)$$

where  $\pi/2 < \arctan(x) < \pi/2$ , with limits  $\lim_{x \rightarrow -\infty} \arctan(x) = -\pi/2$  and  $\lim_{x \rightarrow \infty} \arctan(x) = \pi/2$  being well defined.

**3.2. Spheres.** cf. Example 1.4 (Spheres), Lee (2012) [5], pp. 5

$\forall n \geq 0, n \in \mathbb{Z}$ ,  $n$ -sphere  $\mathbb{S}^n$  is Hausdorff and second-countable because it's topological subspace of  $\mathbb{R}^{n+1}$ .  $\forall i = 1, \dots, n+1$ , let  $U_i^+ = \{(x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} : x^i > 0\}$ .

Similarly, let  $U_i^- = \{(x^1, \dots, x^{n+1}) \in \mathbb{R}^{n+1} : x^i < 0\}$

Let  $f : \mathbb{B}^n \rightarrow \mathbb{R}$  be  $f(u) = \sqrt{1 - |u|^2}$ .

$\forall i = 1, \dots, n+1$ ,  $U_i^+ \cap \mathbb{S}^n$  is graph of function

$$x^i = f(x^1 \dots \widehat{x}^i \dots x^{n+1})$$

where  $\widehat{x}^i$  indicates  $x^i$  is omitted, since

$$\sqrt{(x^1)^2 + (x^{n+1})^2} = 1^2 \text{ or } (x^i)^2 = 1 - \sum_{\substack{j \neq i \\ j=1}}^{n+1} (x^j)^2 \implies x^i = f(x^1, \dots, \widehat{x}^i, \dots, x^{n+1})$$

Similarly,  $U_i^- \cap \mathbb{S}^n$  is graph of

$$x^i = -f(x^1 \dots \widehat{x}^i \dots x^{n+1})$$

since  $x^i < 0$ .

Thus,  $\forall$  subset  $U_i^\pm \cap \mathbb{S}^n$  is locally Euclidean of dim.  $n$  and maps

$$\varphi_i^\pm : U_i^\pm \cap \mathbb{S}^n \rightarrow \mathbb{B}^n$$

$$\varphi_i^\pm(x^1 \dots x^{n+1}) = (x^1 \dots \widehat{x}^i \dots x^{n+1})$$

are graph coordinates for  $\mathbb{S}^n$ .



Since  $\forall$  pt.  $\in \mathbb{S}^n$  is in domain of at least 1 of these  $2n+2$  charts,  $\mathbb{S}^n$  is a topological  $n$ -manifold.

cf. Example 1.31 (Spheres), Lee (2012) [5], pp. 20,

$\forall i = 1 \dots n+1$ , let  $(U_i^\pm, \varphi_i^\pm)$  denote graph coordinate charts, s.t.

$$\begin{aligned} U_i^\pm &= \{(x^1 \dots x^{n+1}) \in \mathbb{R}^{n+1} : x^i \gtrless 0\} \\ \varphi_i^\pm : U_i^\pm \cap \mathbb{S}^n &\rightarrow \mathbb{B}^n \\ \varphi_i^\pm(x^1 \dots x^{n+1}) &= (x^1, \dots, \widehat{x}^i \dots x^{n+1}) \end{aligned}$$

$\forall$  distinct  $i, j$ , transition map  $\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}$  is computed; for  $i < j$ ,

$$\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}(u^1 \dots u^n) = (u^1, \dots, \widehat{u}^i \dots \pm \sqrt{1 - |u|^2}, \dots u^n)$$

with  $\sqrt{1 - |u|^2}$  in  $j$ th position; similarly for  $i = j$ .

Also  $\varphi_i^+ \circ (\varphi_i^-)^{-1}(x^1, \dots x^n) = \varphi_i^+(x^1 \dots - \sqrt{1 - \sum_i (x^i)^2} \dots x^n) = (x^1 \dots x^n)$ .

So  $\varphi_i^+ \circ (\varphi_i)^{-1} = \varphi_i^- \circ (\varphi_i^+)^{-1} = 1_{\mathbb{B}^n}$ .

$\{(U_i^\pm, \varphi_i^\pm)\}$  is a smooth atlas, defines smooth structure on  $\mathbb{S}^n$ . This is its **standard smooth structure**.

**Problem 1, Wald (1984) [36].**

(a) In Wald's notation,  $f_i^+ : O_i^+ \rightarrow D$ ; compare Lee's notation against Wald's, respectively:

$$\varphi_i^\pm : U_i^\pm \cap S^n \rightarrow \mathbb{B}^n \equiv f_i^\pm : O_i^\pm \rightarrow D^n$$

$$\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}(x^1 \dots x^n) = \varphi_i^\pm(x^1, \dots \pm \sqrt{1 - \sum_{k=1}^n (x^k)^2} \dots x^n) = (x^1 \dots \widehat{x}^i \dots \pm \sqrt{1 - \sum_{k=1}^n (x^k)^2}, \dots x^n)$$

Clearly,  $\varphi_i^\pm \circ (\varphi_j^\pm)^{-1}$  is  $C^\infty$  by taking  $\frac{\partial}{\partial x^k}$  to any order, since  $\mathbb{B}^n$  doesn't include its boundary.

(b) See *stereographic projection*:

**north pole**  $N = (0, \dots, 0, 1) \in S^n \subseteq \mathbb{R}^{n+1}$

**south pole**  $S = (0, \dots, 0, -1) \in S^n \subseteq \mathbb{R}^{n+1}$

**stereographic projection**

$$\begin{aligned} \sigma : \mathbb{S}^n \setminus \{N\} &\rightarrow \mathbb{R}^n \\ \sigma(x^1, \dots x^{n+1}) &= \frac{(x^1, \dots x^n)}{1 - x^{n+1}} \\ \tilde{\sigma} : \mathbb{S}^n \setminus \{S\} &\rightarrow \mathbb{R}^n \\ \tilde{\sigma}(x^1, \dots x^{n+1}) &= -\sigma(-x) = \frac{x^1, \dots x^n}{1 + x^{n+1}} \end{aligned}$$

**Problem 1-7, Lee (2012) [5], Stereographic projection of spheres.**

Let  $N$  denote the **north pole**  $(0, \dots, 0, 1) \in \mathbb{S}^n \subseteq \mathbb{R}^{n+1}$ , and let  $S$  denote the **south pole**  $(0, \dots, 0, -1)$ . Define the **stereographic projection**  $\sigma : \mathbb{S}^n \setminus \{N\} \rightarrow \mathbb{R}^n$  by

$$\sigma(x^1, \dots x^{n+1}) = \frac{(x^1, \dots x^n)}{1 - x^{n+1}}$$

Let  $\tilde{\sigma}(x) = -\sigma(-x)$  for  $x \in \mathbb{S}^n \setminus \{S\}$ .

(a) So given a

**north pole**  $N = (0, \dots, 0, 1) \in S^n \subseteq \mathbb{R}^{n+1}$

**south pole**  $S = (0, \dots, 0, -1) \in S^n \subseteq \mathbb{R}^{n+1}$

**stereographic projection**

$$\begin{aligned} \sigma : \mathbb{S}^n \setminus \{N\} &\rightarrow \mathbb{R}^n \\ \sigma(x^1, \dots x^{n+1}) &= \frac{(x^1, \dots x^n)}{1 - x^{n+1}} \\ \tilde{\sigma} : \mathbb{S}^n \setminus \{S\} &\rightarrow \mathbb{R}^n \\ \tilde{\sigma}(x^1, \dots x^{n+1}) &= -\sigma(-x) = \frac{x^1, \dots x^n}{1 + x^{n+1}} \end{aligned}$$

Consider  $\mathbf{x} \in \mathbb{S}^n \setminus \{N\} \rightarrow \mathbb{R}^n$ ,  $\mathbf{x} = (x^1 \dots x^{n+1})$ .

$$N + t(\mathbf{x} - N) = \mathbf{x}' \implies 1 + t(x^{n+1} - 1) = 0 \text{ or } t = \frac{1}{1 - x^{n+1}}$$

$$(x')^i = \frac{x^i}{1 - x^{n+1}} \quad \forall i \neq n+1 \implies \sigma(x^1 \dots x^{n+1}) = \frac{(x^1, \dots x^n)}{1 - x^{n+1}}$$

Thus we have a stereographic projection from the north pole.

$$S + t(\mathbf{x} - S) = \mathbf{x}' \implies 1 + t(x^{n+1} + 1) = 0 \text{ or } t = \frac{1}{1 + x^{n+1}}$$

$$\begin{aligned} (x')^i &= \frac{x^i}{1 + x^{n+1}} \quad \forall i \neq n+1 \\ \implies \tilde{\sigma}(x^1 \dots x^{n+1}) &= \frac{(x^1, \dots x^n)}{1 + x^{n+1}} \end{aligned}$$

**Thus we have a stereographic projection from the south pole.**

b

$$\begin{aligned} \sigma^{-1}(u^1 \dots u^n) &= \frac{(2u^1 \dots 2u^n, |u|^2 - 1)}{|u|^2 + 1} \\ \sigma \circ \sigma^{-1}(u^1 \dots u^n) &= \frac{(2u^1 \dots 2u^n)}{|u|^2 + 1} / \left(1 - \frac{|u|^2 - 1}{|u|^2 + 1}\right) = (u^1, \dots u^n) \\ \sigma^{-1} \circ \sigma(x^1, \dots x^{n+1}) &= \sigma^{-1} \frac{(x^1 \dots x^n)}{1 - x^{n+1}} = \frac{\left(\frac{2x^1}{1 - x^{n+1}}, \dots, \frac{2x^n}{1 - x^{n+1}}, \frac{(x^1)^2 + \dots + (x^n)^2}{(1 - x^{n+1})^2} - 1\right)}{\frac{(x^1)^2 + \dots + (x^n)^2}{(1 - x^{n+1})^2} + 1} = \\ &= \frac{\left(\frac{2x^1}{1 - x^{n+1}} \dots \frac{2x^n}{1 - x^{n+1}}, \frac{(x^1)^2 + \dots + (x^n)^2}{(1 - x^{n+1})^2} - \frac{1 - 2x^{n+1} + (x^{n+1})^2}{(1 - x^{n+1})^2}\right)}{\frac{2(1 - x^{n+1})}{(1 - x^{n+1})^2}} = \left(x^1, \dots x^n, \frac{1 - (x^{n+1})^2 - 1 + 2x^{n+1} - (x^{n+1})^2}{\frac{2}{1 - x^{n+1}}(1 - x^{n+1})^2}\right) = \\ &= (x^1, \dots x^n, \frac{2x^{n+1}(1 - x^{n+1})}{\frac{2}{1 - x^{n+1}}(1 - x^{n+1})^2}) = (x^1, \dots x^{n+1}) \end{aligned}$$

(c) Compute the transition map  $\tilde{\sigma} \circ \sigma^{-1}$ . (The coordinates defined by  $\sigma$  or  $\tilde{\sigma}$  are called **stereographic coordinates**).

$$\begin{aligned} \tilde{\sigma} \circ \sigma^{-1}(u^1 \dots u^n) &= \tilde{\sigma} \frac{(2u^1 \dots 2u^n, |u|^2 - 1)}{|u|^2 + 1} = \frac{\frac{2u^1}{|u|^2 + 1}, \dots, \frac{2u^n}{|u|^2 + 1}}{1 + \frac{|u|^2 - 1}{|u|^2 + 1}} = \frac{\frac{2u^1}{|u|^2 + 1}, \dots, \frac{2u^n}{|u|^2 + 1}}{\frac{2|u|^2}{|u|^2 + 1}} = \\ &= \boxed{\left(\frac{u^1}{|u|^2}, \dots, \frac{u^n}{|u|^2}\right)} \end{aligned}$$

3.2.1. *Polar coordinates.* Consider

$$\begin{aligned} \mathbf{x}(\theta, \varphi) &= (s\theta c\varphi, s\theta s\varphi, c\theta) \\ \theta &= \arccos z, \quad \theta \in [0, \pi], z \in [0, 1] \\ \varphi &= \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0, \varphi \in \left(\frac{-\pi}{2}, \frac{\pi}{2}\right) \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0 \text{ and } y > 0, \varphi \in \left(\frac{\pi}{2}, \pi\right) \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{if } x < 0 \text{ and } y < 0, \varphi \in \left(-\pi, \frac{-\pi}{2}\right) \\ \frac{\pi}{2} & \text{if } x = 0, y = 1 \\ \frac{-\pi}{2} & \text{if } x = 0, y = -1 \end{cases} \end{aligned}$$

#### 4. INVERSE FUNCTION THEOREM

Shastri (2011) had a thorough and lucid and explicit explanation of the Inverse Function Theorem [7]. I will recap it here. The following is also a blend of Wienhard's Handout 4 <https://web.math.princeton.edu/~wienhard/teaching/M327/handout4.pdf>

**Definition 4.** Let  $(X, a)$  metric space.

**contraction**  $\phi : X \rightarrow X$  if  $\exists$  constant  $0 < c < 1$  s.t.  $\forall x, y \in X$

$$d(\phi(x), \phi(y)) \leq cd(x, y)$$

**Theorem 4** (Contraction Mapping Principle). Let  $(X, d)$  complete metric space.

Then  $\forall$  contraction  $\phi : X \rightarrow X$ ,  $\exists ! y \in X$  s.t.  $\phi(y) = y$ ,  $y$  fixed pt.

*Proof.* Recall def. of complete metric space  $X$ ,  $X$  metric space s.t.  $\forall$  Cauchy sequence in  $X$  is convergent in  $X$  (i.e. has limit in  $X$ ).

$$x_1 = \phi(x_0)$$

$$x_2 = \phi(x_1)$$

$\forall x_0 \in X$ , Define  $\vdots$

$$x_j = \phi(x_{j-1})$$

$\vdots$

$$x_n = \phi(x_{n-1})$$

$$d(x_{n+1}, x_n) = d(\phi(x_n), \phi(x_{n-1})) \leq cd(x_n, x_{n-1}) \leq \cdots \leq c^n d(x_1, x_0)$$

for some  $0 < c < 1$ .

$$d(x_m, x_n) \leq d(x_n, x_{n-1}) + d(x_{n-1}, x_m) \leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_{m+1}, x_m) \leq \sum_{k=n-1}^m c^k d(x_1, x_0)$$

Thus,  $\forall \epsilon > 0$ ,  $\exists n_0 > 0$ , ( $n_0$  large enough) s.t.  $\forall m, n \in \mathbb{N}$  s.t.  $n_0 < n < m$ ,

$$d(x_m, x_n) \leq \sum_{k=n-1}^m c^k d(x_1, x_0) < \epsilon d(x_1, x_0)$$

Thus,  $\{x_n\}$  Cauchy sequence. Since  $X$  complete,  $\exists$  limit pt.  $y \in X$  of  $\{x_n\}$ .

$$\phi(y) = \phi(\lim_n x_n) = \lim_n \phi(x_n) = \lim_n x_{n+1} = y$$

Since by def. of  $y$  limit pt. of  $\{x_n\}$ ,  $\forall \epsilon > 0$ , then  $\{n \mid |x_n - y| \leq \epsilon, n \in \mathbb{N}\}$  is infinite.

Consider  $\delta > \mathbb{N}$ . Consider  $\{n \mid |x_n - y| \leq \delta, n \in \mathbb{N}\}$

$\exists N_\delta \in \mathbb{N}$  s.t.  $\forall n > N_\delta$ ,  $|x_n - y| < \delta$ ; otherwise,  $\forall N_\delta$ ,  $\exists n > N_\delta$  s.t.  $|x_n - y| \geq \delta$ . Then  $\{n \mid |x_n - y| \leq \delta, n \in \mathbb{N}\}$  finite.

Contradiction.

$\phi$  cont. so by def.  $\forall \epsilon > 0$ ,  $\exists \delta > 0$  s.t. if  $|x_n - y| < \delta$ , then  $|\phi(x_n) - \phi(y)| < \epsilon$ .

Pick  $N_\delta$  s.t.  $\forall n > N_\delta$ ,  $|x_n - y| < \delta$ , and so  $|\phi(x_n) - \phi(y)| < \epsilon$ . There are infinitely many  $\phi(x_n)$ 's that satisfy this, and so  $\phi(y)$  is a limit pt.

If  $\exists y_1, y_2 \in X$  s.t.  $\phi(y_1) = y_1$ , then

$$\phi(y_2) = y_2$$

$$d(y_1, y_2) = d(\phi(y_1), \phi(y_2)) \leq cd(y_1, y_2) \text{ with } c < 1$$

so  $c = 1$

□

**Theorem 5** (Inverse Function Theorem). Suppose open  $U \subset \mathbb{R}^n$ , let  $C^1 f : U \rightarrow \mathbb{R}^n$ ,  $x_0 \in U$  s.t.  $Df(x_0)$  invertible.

Then  $\exists$  open neighborhoods  $V \ni x_0$ ,  $W \ni f(x_0)$  s.t.  $V \subseteq U$  and  $W \subseteq \mathbb{R}^n$ , respectively, and s.t.

- (i)  $f : V \rightarrow W$  bijection
- (ii)  $g = f^{-1} : V \rightarrow U$  differentiable, i.e.  $g = f^{-1} : W \rightarrow V$  is  $C^1$
- (iii)  $D(f^{-1})$  cont. on  $W$ .
- (iv)  $Dg(y) = (Df(g(y)))^{-1} \quad \forall y \in W$

Also, notice that  $f(g(y)) = y \forall y \in W$ .

*Proof.* Consider  $\tilde{f}(x) = (Df(x_0))^{-1}(f(x + x_0) - f(x_0))$ . Then

$$\tilde{f}(0) = 0 \text{ and}$$

$$D\tilde{f} = (Df(x_0))^{-1}(Df(x + x_0) - 0)$$

$$D\tilde{f}(0) = (Df(x_0))^{-1}Df(x_0) = 1$$

So let  $\tilde{f} \rightarrow f$  (notation) and so assume, without loss of generality, that  $U \ni 0$ ,  $f(0) = 0$ ,  $Df(0) = 1$

Choose  $0 < \epsilon \leq \frac{1}{2}$ . Let  $0 < \delta < 1$  s.t. open ball  $V = B_\delta(0) \subseteq U$ , and  $\|Df(x) - 1\| < \epsilon$ .  $\forall x \in U$ , since  $Df$  cont. at 0.

Let  $W = f(V)$ .

$\forall y \in W$ , define  $\phi_y : V \rightarrow \mathbb{R}^n$

$$\phi_y(x) = x + (y - f(x))$$

$$D(\phi_y)(x) = 1 + -Df(x) \quad \forall x \in V$$

$$\|D(\phi_y)(x)\| = \|1 - Df(x)\| \leq \epsilon < 1$$

$\forall x_1, x_2 \in V$ , by mean value Thm. (not the equality that is only valid in 1-dim., but the inequality, that's valid for  $\mathbb{R}^d$ ,

$$\|\phi_y(x_1) - \phi_y(x_2)\| \leq \|D(\phi_y)(x')\| \|x_1 - x_2\|$$

for some  $x' = cx_2 + (1 - c)x_1$ ,  $c \in [0, 1]$ .  $V$  only needed to be convex set.

$$\implies \|\phi_y(x_1) - \phi_y(x_2)\| \leq \epsilon \|x_1 - x_2\|$$

Then  $\phi_y$  contraction mapping.

Suppose  $f(x_1) = f(x_2) = y$ ,  $x_1, x_2 \in V$ .

$$\phi_y(x_1) = x_1$$

$$\phi_y(x_2) = x_2$$

$$\|\phi_y(x_1) - \phi_y(x_2)\| = \|x_1 - x_2\| \leq \epsilon \|x_1 - x_2\| \quad \forall \epsilon > 0 \implies x_1 = x_2$$

$\implies f|_U$  injective.

$W = f(V)$ , so  $f : V \rightarrow W$  surjective.  $f$  bijective.



Fix  $y_0 \in W$ ,  $y_0 = f(x_0)$ ,  $x_0 \in V$ .

Let  $r > 0$  s.t.  $B_r(x_0) \subset V$ .

Consider  $B_{r\epsilon}(y_0)$ . If  $y \in B_{r\epsilon}(y_0)$ .

$$r\epsilon > \|y - y_0\| = \|y - f(x_0)\| = \|\phi_y(x_0) - x_0\| \text{ with}$$

$$\phi_y(x) = x + (y - f(x))$$

If  $x \in B_r(x_0)$ ,

$$\|\phi_y(x) - x_0\| \leq \|\phi_y(x) - \phi_y(x_0)\| + \|\phi_y(x_0) - x_0\| \leq \epsilon\|x - x_0\| + r\epsilon < 2r\epsilon = r$$

Thus  $\phi(B_r(x_0)) = B_{r\epsilon}(y_0)$ .

By contraction mapping principle,  $\exists a \in B_r(x_0)$ , s.t.  $\phi_y(a) = a$ . Then  $\phi_y(a) = a + (y - f(a)) = a \implies f(a) = y$ .

$y \in f(V) = W$ .

So  $B_{r\epsilon}(y_0) \subset W$ .  $W$  open.

Let  $\text{Mat}(n, n) \equiv$  space of all  $n \times n$  matrices;  $\text{Mat}(n, n) = \mathbb{R}^{n^2}$ .

There is a proof of the implicit function theorem and its various forms in Shastri (2011) [7], but I found Wienhard's Handout 4 for Math 327 to be clearer.<sup>1</sup>

**Theorem 6** (Implicit Function Theorem). *Let open  $U \subset \mathbb{R}^{m+n} \equiv \mathbb{R}^m \times \mathbb{R}^n$*

$$C^1 f : U \rightarrow \mathbb{R}^n$$

$$(a, b) \in U \text{ s.t. } f(a, b) = 0 \text{ and } D_y f|_{(a, b)} \text{ invertible.}$$

*Then  $\exists$  open  $V \ni (a, b)$ ,  $V \subset U$*

$$\exists \text{ open neighborhood } W \ni a, W \subseteq \mathbb{R}^m$$

$$\exists! C^1 g : W \rightarrow \mathbb{R}^n \text{ s.t.}$$

$$\{(x, y) \in V | f(x, y) = 0\} = \{(x, g(x)) | x \in W\}$$

Moreover,

$$dg_x = - (d_y f)^{-1}|_{(x, g(x))} d_x f|_{(x, g(x))}$$

and  $g$  smooth if  $f$ .

*Proof.* Define  $F : U \rightarrow \mathbb{R}^{m+n}$

$$F(x, y) = (x, f(x, y))$$

Then  $F(a, b) = (a, 0)$  (given), and

$$DF = \begin{bmatrix} 1 & \\ \frac{\partial f^i(x, y)}{\partial x^j} & \frac{\partial f^i(x, y)}{\partial y^j} \end{bmatrix} \equiv \begin{bmatrix} 1 & \\ D_x f & D_y f \end{bmatrix}$$

$DF(a, b)$  invertible.

By inverse function theorem, since  $DF(a, b)$  invertible at pt.  $(a, b)$ ,

$\exists$  open neighborhoods  $V \ni (a, b) \subseteq \mathbb{R}^m \times \mathbb{R}^n$  s.t.  $F$  diffeomorphism with  $F^{-1} : \widetilde{W} \rightarrow V$ .

$$\widetilde{W} \ni (a, 0) \subseteq \mathbb{R}^m \times \mathbb{R}^n$$

Set  $W = \{x \in \mathbb{R}^m | (x, 0) \in \widetilde{W}\}$ . Then  $\pi_1(\widetilde{W}) = W$  open in  $\mathbb{R}^m$ .

Define  $g : W \rightarrow \mathbb{R}^n$ ,

$$g(x) = \pi_2 \circ F^{-1}(x, 0) \text{ or}$$

$$F^{-1}(x, 0) = (h(x), g(x))$$

Now  $FF^{-1}(x, 0) = (x, 0) = (h(x), f(h(x), g(x)))$  so  $h(x) = x \forall x \in W$ ,  $0 = f(x, g(x))$ .

Then

$$\{(x, y) \in V | f(x, y) = 0\} = \{(x, y) \in V | F(x, y) = (x, 0)\} = \{(x, g(x)) | x \in W, 0 = f(x, g(x))\}$$

<sup>1</sup><https://web.math.princeton.edu/~wienhard/teaching/M327/handout4.pdf>

Since  $\pi$  smooth and  $F^{-1}$  is  $C^1$ ,  $g$  is  $C^1$ .

To reiterate,  $f(x, g(x)) = 0$  on  $W$ .

Using chain rule while differentiating  $f(x, g(x)) = 0$ ,

$$\begin{aligned} \partial_{x^j} f(x, g(x)) &= \frac{\partial f(x, g(x))}{\partial x^k} \frac{\partial x^k}{\partial x^j} + \frac{\partial f(x, g(x))}{\partial y^k} \frac{\partial g^k(x)}{\partial x^j} = D_x f|_{(x, g(x))} + (D_y f)|_{(x, g(x))} \cdot (Dg)_x = 0 \text{ or} \\ (Dg)_x &= - (D_y f)|_{x, g(x)} D_x f|_{(x, g(x))} \end{aligned}$$

□

## 5. IMMERSIONS

**Definition 5** (Immersion). *smooth  $f : M \rightarrow N$ , s.t.  $Df(p) : T_p M \rightarrow T_{f(p)} N$  injective. Then  $f$  **immersion** at  $p$ .*

Absil, Mahony, and Sepulchre [10] pointed out that another definition for a *immersion* can utilize the theorem that rank of  $Df \equiv DF = \dim T_p M$ . Indeed, recall these facts from linear algebra:  
for  $T : V \rightarrow W$ ,

It's always true that  $\text{rank} T \leq V$ , and

$$\text{rank} T \leq W$$

$\text{rank} T = \dim V$  iff  $T$  injective.

$\text{rank} T = \dim W$  iff  $T$  surjective.

$$\begin{array}{ccc} T_x M & \xrightarrow{DF(x)} & T_{F(x)} N = T_y N \\ \uparrow & & \uparrow \\ x \in M & \xrightarrow{F} & y = F(x) \in N \end{array}$$

$$M \xrightarrow{F} N$$

Now

$$\dim T_x M = \dim M$$

$$\dim T_{F(x)} N = \dim N$$

And

$$\text{rank}(DF(x)) \equiv \text{rank of } F$$

I know that the notation above is confusing, but this is what all Differential Geometry books apparently mean when they say "rank of  $F$ ".

Now

$$\text{rank}(DF(x)) = \dim(\text{im}(DF(x))) = \dim T_x M \text{ iff } DF(x) \text{ injective}$$

If  $\forall x \in M$ , this is the case, then  $F$  an **immersion**.

Apply the rank-nullity theorem in this case:

$$\begin{aligned} \text{rank}(DF(x)) + \dim \ker(DF(x)) &= \dim T_x M = \dim M \\ \implies \text{rank}(DF(x)) &= \dim M \leq \dim T_{F(x)} N = \dim N \text{ or } \dim M \leq \dim N \end{aligned}$$

Now

$$\text{rank}(DF(x)) = \dim T_{F(x)}N \text{ iff } DF(x) \text{ surjective}$$

If  $\forall x \in M$ , this is the case, then  $F$  an **submersion** .

$$\text{rank}(DF(x)) = \dim T_{F(x)}N = \dim N \leq \dim M$$

Shastri (2011) has this as the “Injective Form of Implicit Function Theorem”, Thm. 1.4.5, pp. 23 and Guillemin and Pollack (2010) has this as the “Local Immersion Theorem” on pp. 15, Section 3 “The Inverse Function Theorem and Immersions” [6].

**Theorem 7** (Local immersion Theorem i.e. Injective Form of Implicit Function Theorem). *Suppose  $f : M \rightarrow N$  immersion at  $p$ ,  $q = f(p)$ .*

*Then  $\exists$  local coordinates around  $p, q$ ,  $x, y$ , respectively s.t.  $f(x_1 \dots x_m) = (x_1 \dots x_m, 0 \dots 0)$ .*

*Proof.* Choose local parametrizations

$$\begin{array}{ccc} U \subseteq M & \xrightarrow{f} & N \supseteq V \\ \downarrow \phi & & \downarrow \psi \\ \phi(U) & \xrightarrow{f} & \psi(V) \end{array} \quad \begin{array}{l} \phi(p) = x \\ \psi(q) = y \end{array}$$

$$D(\psi f \varphi^{-1}) \equiv Df. \ Df(p) \text{ injective (given } f \text{ immersion). } Df(p) \in \text{Mat}(n, m)$$

By change of basis in  $\mathbb{R}^n$ , assume  $Df(p) = \begin{pmatrix} I_m \\ 0 \end{pmatrix}$ .

Now define  $G : \phi(U) \times \mathbb{R}^{n-m} \rightarrow \mathbb{R}^n$

$$G(x, z) = f(x) + (0, z)$$

Thus,  $DG(x, z) = 1$  and for open  $\phi(U) \times U_2$ ,  $G(\phi(U) \times U_2)$  open.

By inverse function theorem,  $G$  local diffeomorphism of  $\mathbb{R}^n$ , at 0.

Now  $f = G \circ \mathbf{i}$ , where  $\mathbf{i}$  is canonical immersion.

$$\begin{aligned} G(x, 0) &= f(x) \\ \implies G^{-1}G(x, 0) &= (x, 0) = G^{-1}f(x) \end{aligned}$$

Use  $\psi \circ G$  as the local parametrization of  $N$  around pt.  $q$ . Shrink  $U, V$  so that

$$\begin{array}{ccc} U \subseteq M & \xrightarrow{f} & N \supseteq V \\ \downarrow \phi & & \downarrow \psi \circ G \\ \phi(U) & \xrightarrow{\mathbf{i}} & \psi \circ G(V) \end{array}$$

**Theorem 8** (Implicit Function Thm.). *Let open subset  $U \subseteq \mathbb{R}^n \times \mathbb{R}^d$ ,  $(x, y) = (x^1 \dots x^n, y^1 \dots y^k)$  on  $U$ . Suppose smooth  $\Phi : U \rightarrow \mathbb{R}^k$ ,  $(a, b) \in U$ ,  $c = \Phi(a, b)$*

*If  $k \times k$  matrix  $\frac{\partial \Phi^i}{\partial y^j}(a, b)$  nonsingular, then  $\exists$  neighborhoods  $V_0 \subseteq \mathbb{R}^n$  of  $a$  and smooth  $F : V_0 \rightarrow W_0$  s.t.  $W_0 \subseteq \mathbb{R}^k$  of  $b$*

$\Phi^{-1}(c) \cap (V_0 \times W_0)$  is graph of  $F$ , i.e.  
 $\Phi(x, y) = c$  for  $(x, y) \in V_0 \times W_0$  iff  $y = F(x)$ .

6. SUBMERSIONS; RANK THEOREM

cf. pp. 20, Sec. 4 ”Submersions”, Ch. 1 of Guillemin and Pollack (2010) [6].  
Consider  $X, Y \in \mathbf{Man}$ , s.t.  $\dim X \geq \dim Y$ .

**Definition 6** (submersion). *If  $f : X \rightarrow Y$ , if  $Df_x \equiv df_x$  is surjective,  $f \equiv$  **submersion** at  $x$ .*

Recall that,

$$\begin{aligned} Df_x : T_x X &\rightarrow T_{f(x)} Y \\ \dim T_x X &\geq \dim T_{f(x)} Y \\ \text{rank } Df_x &\leq \dim T_{f(x)} Y, \text{ in general, while} \\ \text{rank } Df_x &= \dim T_{f(x)} Y \text{ iff } Df_x \text{ surjective} \end{aligned}$$

Canonical submersion is standard projection:

If  $\dim X = k$ ,  $k \geq l$ ,  
 $\dim Y = l$

$$(a_1 \dots a_k) \mapsto (a_1 \dots a_l)$$

**Theorem 9** (Local Submersion Theorem). *Suppose  $f : X \rightarrow Y$  submersion at  $x$ , and  $y = f(x)$ , Then  $\exists$  local coordinates around  $x, y$  s.t.*

$$f(x_1 \dots x_k) = (x_1 \dots x_l)$$

*i.e.  $f$  locally equivalent to canonical submersion near  $x$*

*Proof.* I’ll have a side-by-side comparison of my notation and the 1 used in Guillemin and Pollack (2010) [6] where I can.

For charts  $(U, \phi), (V, \psi)$  for  $X, Y$ , respectively,  $y = f(x)$  for  $x \in X$ ,

$$\begin{array}{ccc} U \subseteq X & \xrightarrow{f} & Y \supseteq V \\ \downarrow \phi & & \downarrow \psi \circ G \\ \mathbb{R}^k & \xrightarrow{\mathbf{i}} & \mathbb{R}^l \end{array} \quad \begin{array}{ccc} x & \xrightarrow{f} & f(x) = y \\ \downarrow \phi & & \downarrow \psi \\ \phi(x) = (a^1 \dots a^k) & \xrightarrow{g} & g(\phi(x)) = g(a^1 \dots a^k) = \psi(y) \end{array}$$

$Dg_x$  surjective, so assume it’s a  $l \times k$  matrix  $\begin{bmatrix} \mathbf{1}_l & 0 \end{bmatrix}$ .

Define

$$\begin{aligned} (15) \quad G : U \subset \mathbb{R}^k &\rightarrow \mathbb{R}^k \\ G(a) \equiv G(a^1 \dots a^k) &:= (g(a), a_{l+1}, \dots, a_k) \end{aligned}$$

Now

□

$$(16) \quad DG(a) = \begin{bmatrix} \mathbf{1}_l & 0 \\ & \mathbf{1}_{k-l} \end{bmatrix} = \mathbf{1}_k$$

so  $G$  local diffeomorphism (at 0).

So  $\exists G^{-1}$  as local diffeomorphism of some  $U'$  of  $a$  into  $U \subset \mathbb{R}^k$ .

By construction,

$$(17) \quad g = \mathbb{P}_l \circ G$$

where  $\mathbb{P}_l$  is the *canonical submersion*, the projection operator onto  $\mathbb{R}^l$ .

$$g \circ G^{-1} = \mathbb{P}_l$$

(since  $G$  diffeomorphism)

$$\begin{array}{ccc} U \subseteq X & \xrightarrow{f} & V \subseteq Y \\ \phi^{-1} \circ G^{-1} \uparrow & & \uparrow \psi^{-1} \\ \mathbb{R}^k & \xrightarrow{\mathbb{P}_l} & \mathbb{R}^l \end{array} \quad \text{for}$$

$$\begin{array}{ccc} \phi^{-1} \circ G^{-1}(a) \equiv \phi^{-1} \circ G^{-1}(a^1 \dots a^k) = x & \xrightarrow{f} & f(x) = y = \psi^{-1}(a^1 \dots a^l) \\ \phi^{-1} \circ G^{-1} \uparrow & & \uparrow \psi^{-1} \\ (a^1 \dots a^k) & \xrightarrow{\mathbb{P}_l} & (a^1 \dots a^l) \end{array}$$

$$\implies$$

□

”An obvious corollary worth noting is that if  $f$  is a submersion at  $x$ , then it is actually a submersion in a whole neighborhood of  $x$ .” Guillemin and Pollack (2010) [6]

Suppose  $f$  submersion at  $x \in f^{-1}(y)$ .

By local submersion theorem

$$f(x_1 \dots x_k) = (x_1 \dots x_l)$$

Choose  $y = (0, \dots, 0)$ .

Then, near  $x$ ,  $f^{-1}(y) = \{(0, \dots, 0, x_{l+1} \dots x_k)\}$  i.e. let  $V \ni x$  neighborhood of  $x$ , define  $(x_1 \dots x_k)$  on  $V$ .

Then  $f^{-1}(y) \cap V = \{(0 \dots 0, x_{l+1}, \dots x_k) | x_1 = 0, \dots x_l = 0\}$ .

Thus  $x_{l+1}, \dots x_k$  form a coordinate system on open set  $f^{-1}(y) \cap V \subseteq f^{-1}(y)$ .

Indeed,

$$\begin{array}{ccc} U \subseteq X & \xrightarrow{f} & V \subseteq Y \\ \downarrow \phi & & \downarrow \psi \\ \mathbb{R}^k & \xrightarrow{\mathbb{P}_l} & \mathbb{R}^l \end{array} \quad \begin{array}{ccc} x & \xrightarrow{f} & f(x) = y \\ \downarrow \phi & & \downarrow \psi \\ \phi(x) = (x^1 \dots x^k) & \xrightarrow{\mathbb{P}_l} & (x^1 \dots x^l) \end{array}$$

and now

$$\begin{array}{ccc} f^{-1}(y) & \xleftarrow{f^{-1}} & y \\ \uparrow \phi^{-1} & & \downarrow \psi \\ \{(0, \dots, 0, x^1 \dots x^k)\} & \xleftarrow{\mathbb{P}_l^{-1}} & (0 \dots 0) \end{array}$$

**6.1. Rank Theorem.** Lee (2012) [5] in pp. 85, Ch. 4 Submersions, Immersions, and Embeddings, combines Theorems 7, 9 (local immersion and local submersion theorems, respectively) into the ”Rank Theorem” (cf. Thm 4.12 ”Rank Theorem” of Lee (2012)):

**Theorem 10** (Rank Theorem). *Suppose smooth manifolds  $M, N$ ,  $\dim M = m$ ,  $\dim N = n$ , smooth map  $F : M \rightarrow N$ ,  $F$  has constant rank  $r$ .*

$\forall p \in M$ ,  $\exists$  smooth charts  $(U, \varphi)$  for  $M$ , centered at  $p$ ,  $(V, \psi)$  for  $N$ , centered at  $F(p)$ , s.t.

$$F(U) \subseteq V$$

in which  $F$  has coordinate representation of form

$$(18) \quad \widehat{F}(x^1 \dots x^r, x^{r+1} \dots x^m) = (x^1 \dots x^r, 0 \dots 0)$$

Particularly, if  $F$  smooth submersion,

$$\widehat{F}(x^1 \dots x^n, x^{n+1} \dots x^m) = (x^1 \dots x^n)$$

and if  $F$  smooth immersion

$$\widehat{F}(x^1 \dots x^m) = (x^1 \dots x^m, 0 \dots 0)$$

$$\begin{array}{ccc} \mathbb{R}^m & \xrightarrow{\psi F \varphi^{-1}} & \mathbb{R}^n \\ \uparrow \varphi & & \uparrow \psi \\ U \subset M & \xrightarrow{F} & V \subset N \end{array}$$

Also remember that  $DF(p) : T_p M \rightarrow T_{F(p)} N$

*Proof.*  $DF(p)$  has rank  $r$  (given). Then  $DF(p)$  is some  $r \times r$  submatrix of a  $n \times m$  matrix s.t.  $\det DF(p)$  nonzero.

By change of basis in  $\mathbb{R}^n$ , or reordering coordinates, assume  $DF(p)$  is upper left submatrix  $\left( \frac{\partial F^i}{\partial x^j} \right) \quad \forall i, j = 1, \dots r$ .

Relabel standard coordinate as

$$(x, y) = (x^1 \dots x^r, y^1 \dots y^{m-r}) \in \mathbb{R}^m$$

$$(v, w) = (v^1 \dots v^r, w^1 \dots w^{n-r}) \in \mathbb{R}^n$$

By initial translations of coordinates, assume without loss of generality  $p = (0, 0)$ ,  $F(p) = (0, 0)$

Suppose

$$F(x, y) = (Q(x, y), R(x, y))$$

for some smooth maps  $Q : U \rightarrow \mathbb{R}^r$ ,  $R : U \rightarrow \mathbb{R}^{n-r}$

Define

$$\varphi : U \rightarrow \mathbb{R}^m$$

$$\varphi(x, y) = (Q(x, y), y)$$

so

$$D\varphi(0, 0) = \begin{pmatrix} \frac{\partial Q^i}{\partial x^j}(0, 0) & \frac{\partial Q^i}{\partial y^j}(0, 0) \\ 0 & \delta_j^i \end{pmatrix}$$

$D\varphi(0, 0)$  nonsingular, since  $\det \frac{\partial Q^i}{\partial x^j} \neq 0$  (by hypothesis).

By inverse function thm.,  $\exists$  connected neighborhoods  $U_0$  of  $(0, 0)$ ,  $\tilde{U}_0$  of  $\varphi(0, 0) = (0, 0)$  s.t.

$$\varphi : U_0 \rightarrow \tilde{U}_0$$

is a diffeomorphism.

By shrinking  $U_0, \tilde{U}_0$ , assume  $\tilde{U}_0$  open cube.

Write  $\varphi^{-1}(x, y) = (A(x, y), B(x, y))$ , for some smooth functions  $A : \tilde{U}_0 \rightarrow \mathbb{R}^r$  ,  
 $B : \tilde{U}_0 \rightarrow \mathbb{R}^{m-r}$

$$(x, y) = \varphi(A(x, y), B(x, y)) = (Q(A(x, y), B(x, y)), B(x, y))$$

$$\begin{aligned} & B(x, y) = y \\ \implies & \varphi^{-1}(x, y) = (A(x, y), y) \end{aligned}$$

$$\varphi\varphi^{-1} = 1 \implies x = Q(A(x, y), y)$$

Recall that we had hypotehsized that

$$F(x, y) = (Q(x, y), R(x, y))$$

Then

$$F \circ \varphi^{-1}(x, y) = F(A(x, y), y) = (Q(A(x, y), y), R(A(x, y), y)) = (x, R(A(x, y), y))$$

and so

$$F \circ \varphi^{-1}(x, y) = (x, \tilde{R}(x, y))$$

where  $\tilde{R} : \tilde{U}_0 \rightarrow \mathbb{R}^{n-r}$

$$\tilde{R}(x, y) = R(A(x, y), y)$$

Compute

$$D(F \circ \varphi^{-1})(x, y) = \begin{pmatrix} \delta_j^i & 0 \\ \frac{\partial \tilde{R}^i}{\partial x^j}(x, y) & \frac{\partial \tilde{R}^i}{\partial y^j}(x, y) \end{pmatrix}$$

Since composing with a diffeomorphism doesn't change rank of map,  $D(F \circ \varphi^{-1})$  has rank  $r$  everywhere in  $\tilde{U}_0$ .

$\begin{pmatrix} \delta_j^i \\ \frac{\partial \tilde{R}^i}{\partial x^j}(x, y) \end{pmatrix} j = 1 \dots r$  are linearly independent, so  $\frac{\partial \tilde{R}^i}{\partial y^j}(x, y) = 0$  on  $\tilde{U}_0$ , so  $\tilde{R}^i$  independent of  $y^j$ .

Let  $S(x) = \tilde{R}(x, 0)$ , then

$$F \circ \varphi^{-1}(x, y) = (x, S(x))$$

Let open  $V_0 \subseteq V$ ,  $(0, 0) \in V$  be an open subset  $V_0 = \{(v, w) \in V : (v, 0) \in \tilde{U}_0\}$ .  
Then  $V_0$  is a neighborhood of  $(0, 0)$ .

Because  $\tilde{U}_0$  is a cube,  $F \circ \varphi^{-1}(x, y) = (x, S(x))$ ,

$$F \circ \varphi^{-1}(\tilde{U}_0) \subseteq V_0$$

so  $F(U_0) \subseteq V_0$ .

Define  $\psi : V_0 \rightarrow \mathbb{R}^n$

$$\psi(v, w) = (v, w - S(v))$$

Because  $\psi^{-1}(s, t) = (s, t + S(s))$ ,

it is a diffeomorphism.

Thus  $(V_0, \psi)$  is a smooth chart.

$$\psi \circ F(\varphi^{-1}(x, y)) = \psi(x, S(x)) = (x, S(x) - S(x)) = (x, 0)$$

**Definition 7** (regular value). *For smooth  $f : X \rightarrow Y$ ,  $X, Y \in \mathbf{Man}$ ,*

*$y \in Y$  is a **regular value** for  $f$  if  $Df_x : T_x X \rightarrow T_y Y$  surjective  $\forall x$  s.t.  $f(x) = y$ .*

*$y \in Y$  **critical value** if  $y$  not a regular value of  $f$ .*

Absil, Mahony, and Sepulchre [10] pointed out that another definition for a *regular value* can utilize the theorem that rank of  $Df \equiv DF = \dim T_p N = \dim N$ , iff  $DF(p)$  surjective, for  $p \in M$ ,  $F : M \rightarrow N$ . Then

**regular value**  $y \in N$ , of  $F$ , if rank of  $F \equiv \text{rank}(DF(x)) = \dim N$ ,  $\forall x \in F^{-1}(y)$ , for  $F : M \rightarrow N$ .

**Theorem 11** (Preimage theorem). *If  $y$  regular value of  $f : X \rightarrow Y$ ,  $f^{-1}(y)$  is a submanifold of  $X$ , with  $\dim f^{-1}(y) = \dim X - \dim Y$*

*Proof.* Given  $y$  is a regular value of  $f : X \rightarrow Y$ ,

$\forall x \in f^{-1}(y)$ ,  $Df_x : T_x X \rightarrow T_y Y$  is surjective. By local submersion theorem,

$$f(x^1 \dots x^k) = (x^1 \dots x^l) = y$$

Since  $x \in f^{-1}(y)$ ,  $(x^1 \dots x^k) = (y^1 \dots y^l, x^{l+1} \dots x^k)$ .

For this chart for  $(U, \varphi)$ ,  $U \ni x$ , consider  $(U \cap f^{-1}(y), \psi)$  with  $\psi(x) = (x^{l+1} \dots x^k) \quad \forall x \in U \cap f^{-1}(y)$ .

$\forall f^{-1}(y)$  submanifold with  $\dim f^{-1}(y) = k - l = \dim X - \dim Y$ . □

*Examples for emphasis*

If  $\dim X > \dim Y$ ,

if  $y \in Y$ , regular value of  $f : X \rightarrow Y$ ,

$f$  submersion,  $\forall x \in f^{-1}(y)$

If  $\dim X = \dim Y$ ,

$f$  local diffeomorphism  $\forall x \in f^{-1}(y)$

If  $\dim X < \dim Y$ ,  $\forall y \in f(X)$  is a critical value.

**Example:**  $O(n)$  as a submanifold of  $\text{Mat}(n, n)$

Given  $\text{Mat}(n, n) \equiv M(n) = \{n \times n \text{ matrices}\}$  is a manifold; in fact  $\text{Mat}(n, n) \cong \mathbb{R}^{n^2}$ ,

Consider  $O(n) = \{A \in \text{Mat}(n, n) | AA^T = 1\}$ .

$$(19) \quad AA^T \in \text{Sym}(n) \equiv S(n) = \{S \in \text{Mat}(n, n) | S^T = S\} = \{ \text{symmetric } n \times n \text{ matrices} \}$$

$\text{Sym}(n)$  submanifold of  $\text{Mat}(n, n)$ ,  $\text{Sym}(n)$  diffeomorphic to  $\mathbb{R}^k$  (i.e.  $\text{Sym}(n) \cong \mathbb{R}^k$ ),  $k := \frac{n(n+1)}{2}$ .

$$f : \text{Mat}(n, n) \rightarrow \text{Sym}(n)$$

$$f(A) = AA^T$$

Notice  $f$  is smooth,

$$f^{-1}(1) = O(n)$$

$$Df_A(B) = \lim_{s \rightarrow 0} \frac{f(A + sB) - f(A)}{s} = \lim_{s \rightarrow 0} \frac{(A + sB)(A^T + sB^T) - AA^T}{s} = AB^T + BA^T$$

If  $Df_A : T_A \text{Mat}(n, n) \rightarrow T_{f(A)} \text{Sym}(n)$  surjective when  $A \in f^{-1}(1) = O(n)$  (???)

**Proposition 1.** *If smooth  $g_1 \dots g_l \in C^\infty(X)$  on  $X$  are independent  $\forall x \in X$ , s.t.  $g_i(x) = 0, \forall i = 1 \dots l$ , then  $Z = \{x \in X | g_1(x) = \dots = g_l(x) = 0\}$  is a submanifold of  $X$  s.t.  $\dim Z = \dim X - l$ .*

*Take note that  $g_1 \dots g_l$  are independent at  $x$  means, really, that  $D(g_1)_x \dots D(g_l)_x$  are linearly independent on  $T_x X$ .*

*Proof.* Suppose smooth  $g_1 \dots g_l \in C^\infty(X)$  on manifold  $X$  s.t.  $\dim X = k \geq l$ .

Consider  $g = (g_1 \dots g_l) : X \rightarrow \mathbb{R}^l$ ,  $Z \equiv g^{-1}(0)$ .

Since  $\forall g_i$  smooth,  $D(g_i)_x : T_x X \rightarrow \mathbb{R}$  linear.

Now for

$$Dg_x = (D(g_1)_x \dots D(g_l)_x) : T_x X \rightarrow \mathbb{R}^l$$

□

By rank-nullity theorem (linear algebra),  $Dg_x$  surjective iff  $\text{rank } Dg_x = l$  i.e.  $l$  functionals  $D(g_1)_x \dots D(g_l)_x$  are linearly independent on  $T_x X$ .

”We express this condition by saying the  $l$  functions  $g_1 \dots g_l$  are independent at  $x$ .” (Guillemin and Pollack (2010) [6])

□

## 7. SUBMANIFOLDS; IMMERSED SUBMANIFOLD, EMBEDDED SUBMANIFOLDS, REGULAR SUBMANIFOLDS

**Definition 8** (Embedded Submanifold).

Recall immersion:

$F : M \rightarrow N$  immersion iff  $DF$  injective, i.e. iff  $\text{rank} DF = \dim M$ .

Consider manifolds  $M \subseteq N$ .

Consider inclusion map  $i : M \rightarrow N$ .

$$i : x \mapsto x$$

If  $i$  immersion,  $Di(x) = \frac{\partial y^i}{\partial x^j} = \delta_j^i$  if  $y^i = x^i$ ,  $\forall i = 1, \dots, \dim M$ .

**Definition 9** (immersed submanifold). ***immersed submanifold**  $M \subseteq N$  if inclusion  $i : M \rightarrow N$  is an immersion.*

cf. 3.3 Embedded Submanifolds of Absil, Mahony, and Sepulchre [10], also Ch. 5 Submanifolds, pp. 108, **Immersed Submanifolds** of John Lee (2012) [5].

Immersed submanifolds often arise as images of immersions.

**Proposition 2** (Images of Immersions as submanifolds). *Suppose smooth manifold  $M$ , smooth manifold with or without boundaries  $N$ ,*

*injective, smooth immersion  $F : M \rightarrow N$  ( $F$  injective itself, not just immersion)*

*Let  $S = F(M)$ .*

*Then  $S$  has unique topology and smooth structure of smooth submanifolds of  $N$  s.t.  $F : M \rightarrow S$  diffeomorphism.*

cf. Prop. 5.18 of John Lee (2012) [5].

*Proof.* Define topology of  $S$ : set  $U \subseteq S$  open iff  $F^{-1}(U) \subseteq M$  open ( $F^{-1}(U \cap V) = F^{-1}(U) \cap F^{-1}(V)$ ,  $F^{-1}(U \cup V) = F^{-1}(U) \cup F^{-1}(V)$ ).

Define smooth structure of  $S$ :  $\{F(U), \varphi \circ F^{-1}|(U, \varphi) \in \text{atlas for } M, \text{ i.e. } (U, \varphi) \text{ any smooth chart of } M\}$ .

”smooth compatibility condition”:

$$(\varphi_2 \circ F^{-1})(\varphi_1 F^{-1})^{-1} = \varphi_2 \circ F^{-1} F \varphi_1^{-1} = \varphi_2 \varphi_1^{-1}$$

since  $\varphi_2 \varphi_1^{-1}$  diffeomorphism ( $\varphi_2 \varphi_1^{-1}$  bijection and it and inverse is differentiable)

$F$  diffeomorphism onto  $F(M)$ .

and these are the only topology and smooth structure on  $S$  with this property:

$$S \xrightarrow{F^{-1}} M \xrightarrow{F} N = S \hookrightarrow M$$

and  $F^{-1}$  diffeomorphism,  $F$  smooth immersion, so  $i : S \rightarrow M$  smooth immersion.

## 8. CURVES, INTEGRAL CURVES, AND FLOWS

cf. John Lee (2012) [5], Ch. 9, deals with time-dependent vector fields and I don’t see other texts or references handling such an important, but overlooked, case.

**8.1. Curves in Euclidean space.** cf. 4.1 Jeffrey Lee (2009) [4]

If  $C$  is 1-dim. submanifold of  $\mathbb{R}^n$ ,  $p \in C$ ,  $\exists$  chart  $(V, y)$  of  $C$ ,  $p \in V$  s.t.  $y(V)$  is a connected open interval  $I \subset \mathbb{R}$ , inverse map  $y^{-1} : I \rightarrow V \subset M$  is a local parametrization.

idea is to extract information that’s appropriately independent of parametrization.

If  $\gamma : I \rightarrow \mathbb{R}^n$ ,  $c : J \rightarrow \mathbb{R}^n$  curves with same image,

$c$  is a **positive reparametrization** of  $\gamma$  if  $\exists$  smooth  $h : J \rightarrow I$  with  $h' > 0$  s.t.  $c = \gamma \circ h$ ,

in this case,  $\gamma, c$  have same **sense** and same **orientation**.

Assume  $\gamma : I \rightarrow \mathbb{R}^n$  has  $\|\gamma'\| > 0$ ,

Such a curve is **regular**, i.e. curve is an immersion (Recall, an immersion would be smooth  $\gamma : I \rightarrow \mathbb{R}^n$ , s.t.  $D\gamma(p) : T_p I \rightarrow T_{\gamma(p)} \mathbb{R}^n$  injective).

**Definition 10** (unit tangent field in Euclidean space, 4.3 Lee (2009) [4]). *If  $\gamma : I \rightarrow \mathbb{R}^n$  regular curve, then  $\mathbf{T}(t) := \frac{\gamma'(t)}{\|\gamma'(t)\|}$  defines **unit tangent field** along  $\gamma$  ( $\|T\| = 1$ )*

**length of a curve** defined on closed interval  $\gamma : [t_1, t_2] \rightarrow \mathbb{R}^n$ ,

$$L = \int_{t_1}^{t_2} \|\gamma'(t)\| dt$$

Define **arc length function** for curve  $\gamma : I \rightarrow \mathbb{R}^n$  by choosing  $t_0 \in I$ ,

$$s = h(t) := \int_{t_0}^t \|\gamma'(\tau)\| d\tau$$

If curve is smooth and regular, then  $h' = \|\gamma'(\tau)\| > 0$ , so by inverse function theorem,  $\exists$  smooth  $h^{-1}$  (since  $h'$  invertible (with  $1/h'$ )).

If  $c(s) = \gamma h^{-1}(s)$ , then  $\|c'(s)\| := \|c'(s)\| = 1$ ,  $\forall s$

$$c'(s) = (\gamma(h^{-1}(s)))' = \gamma'(h^{-1}(s)) \cdot \frac{dh^{-1}}{ds}(s)$$

$$\|c'(s)\| = \|\gamma'(t)\| \left\| \frac{dh^{-1}}{ds}(s) \right\| = h' \cdot \frac{1}{h'} = 1$$

Curves parametrized by arc length are **unit speed curves**.

For a unit speed curve,  $\frac{dc}{ds}(s) = \mathbf{T}(s)$ .

**Definition 11** (curvature vector, curvature function, principal normal in Euclidean space, 4.4 Lee (2009) [4]). *Let  $c : I \rightarrow \mathbb{R}^n$  be a unit speed curve. vector valued function*

$$\kappa(s) := \frac{d\mathbf{T}}{ds}(s)$$

*is called the **curvature vector**.*

$$\kappa(s) := \|\kappa(s)\| = \left\| \frac{d\mathbf{T}}{ds}(s) \right\|$$

*is called the **curvature function**.*

□ *If  $\kappa(s) > 0$ , then define principal normal*

$$\mathbf{N}(s) = \left\| \frac{d\mathbf{T}}{ds}(s) \right\|^{-1} \frac{d\mathbf{T}}{ds}(s)$$

s.t.  $\frac{d\mathbf{T}}{ds} = \kappa \mathbf{N}$

**8.2. Integral Curves.** cf. Ch. 9 of Lee (2012) [5]

integral curves - smooth curves whose velocity at each point is equal to the value of the smooth vector field there.

Let smooth manifold  $M$  with or without boundary.

If smooth curve  $\gamma : J \rightarrow M$ , then  $\forall t \in J$ , velocity vector  $\gamma'(t)$  is a vector in  $T_{\gamma(t)} M$ .

**Definition 12** (Integral Curve  $\Gamma$ ). *Integral curve:  $\mathbf{v}, \gamma$ .*

*If vector field  $\mathbf{v}$  on  $M$ ,*

***integral curve** of  $\mathbf{v}$  is differentiable curve  $\gamma : J \rightarrow M$ , whose velocity  $\forall p \in M$ ,  $V_{\gamma(t)}$*

$$\gamma'(t) = V_{\gamma(t)}, \forall t \in J$$

Finding integral curves boils downs to solving a system of ordinary differential equations in a smooth chart.

Suppose smooth vector field  $V$  on  $M$ , smooth curve  $\gamma : J \rightarrow M$

On smooth coordinate domain  $U \subseteq M$ ,

Write  $\gamma$  in local coordinates as

$$\begin{aligned}\gamma(t) &= (\gamma^1(t), \dots, \gamma^n(t)) \\ \gamma^i(t) &= x^i \circ \gamma(t) \text{ where } x^i : U \subseteq M \rightarrow \mathbb{R}\end{aligned}$$

Then condition for  $\gamma$  to be an integral curve.

$$\gamma'(t) \equiv \dot{\gamma}(t) = V_{\gamma(t)}$$

can be written as

$$\dot{\gamma}^i(t) \left. \frac{\partial}{\partial x^i} \right|_{\gamma(t)} = V^i(\gamma(t)) \left. \frac{\partial}{\partial x^i} \right|_{\gamma(t)}$$

which reduces to an autonomous system of ordinary differential equations (ODEs)

$$\begin{aligned}\dot{\gamma}^1(t) &= V^1(\gamma^1(t), \dots, \gamma^n(t)) \\ &\vdots \\ \dot{\gamma}^n(t) &= V^n(\gamma^1(t), \dots, \gamma^n(t))\end{aligned}\tag{20}$$

20200306 EY: I use this notation which is more useful for me, e.g. taking the double derivative of the curve  $\gamma$ :

$$\begin{aligned}\dot{\gamma}(t) &= v(\gamma(t)) \in T_{\gamma(t)}U \\ \dot{\gamma}(t) &= v^i(\gamma^1(t), \dots, \gamma^n(t)) \\ \frac{d}{dt} \dot{\gamma}^i(t) &= \frac{d}{dt} v^i(\gamma^1(t), \dots, \gamma^n(t)) = \frac{d}{dt} v^i(x^1, \dots, x^n) \\ &= \frac{\partial v^i}{\partial x^j}(x) \dot{x}^j = \dot{\gamma}^j \frac{\partial \dot{\gamma}^i}{\partial x^j}\end{aligned}$$

## 9. PUSH-FORWARD, PULL-BACK

9.1. **Push-forward in coordinates**,  $f_* : T_p M \rightarrow T_{f(p)} N$ . Using Nakahara (2003) [9]’s notation.

Consider a smooth mapping  $f : M \rightarrow N$ ,  $\dim M = m, \dim N = n$ .

$$f_* : T_p M \rightarrow T_{f(p)} N$$

With

$$\begin{aligned}(U, \varphi) \subset M & \quad \varphi : U \rightarrow \mathbb{R}^m & \varphi^i \equiv x^i : U \rightarrow \mathbb{R} \\ (V, \psi) \subset N & \quad \psi : V \rightarrow \mathbb{R}^n & \psi^j \equiv y^j : V \rightarrow \mathbb{R}\end{aligned}$$

$$\begin{aligned}\text{Now } X \in T_p M \text{ , so we can write } X &= X^i \frac{\partial}{\partial x^i}, \text{ and } f_* X \in T_{f(p)} N, \text{ so we can write } f_* X = Y^j \frac{\partial}{\partial y^j}. \\ Y \in T_{f(p)} N & \quad Y = Y^j \frac{\partial}{\partial y^j}\end{aligned}$$

Let  $g \in \mathcal{F}(N)$

$$g : N \rightarrow \mathbb{R}$$

Now  $g \circ \psi^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}$  or  $g \circ \psi^{-1} : \psi(N) \rightarrow \mathbb{R}$ .

For notation, let  $g \circ \psi^{-1}(y) \equiv g(y)$ , in that we treat  $y$  to be a point with coordinates in  $\mathbb{R}^n$  notation-wise, and not as  $y^j : V \rightarrow \mathbb{R}$ .

Since  $g \circ f : M \rightarrow \mathbb{R}$ , so  $g \circ f \in \mathcal{F}(M)$ .

$$(f_* X)[g] \equiv X[g \circ f]$$

$$\begin{aligned}(21) \quad (f_* X)[g] &= (Y^j \frac{\partial}{\partial y^j})[g] = Y^j \frac{\partial g}{\partial y^j} = X^i \frac{\partial}{\partial x^i} [g \circ f] = X^i \frac{\partial}{\partial x^i} g(f^j(x)) = X^i \frac{\partial g}{\partial y^j}(y) \frac{\partial f^j}{\partial x^i}(x) = X^i \frac{\partial f^j}{\partial x^i}(x) \frac{\partial}{\partial y^j} g \\ Y^j &= X^i \frac{\partial f^j}{\partial x^i} = X^i \frac{\partial y^j}{\partial x^i} = F^j_i(x) X^i\end{aligned}$$

where  $f : M \rightarrow N$ , but  $\psi \circ f \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ , but write this notation as  $f^j(x) = y^j$ .

Also, notice that for Eq. [21](#),

$$\begin{aligned}g \circ f &: M \rightarrow \mathbb{R} \\ g \circ \psi^{-1} \circ \psi \circ f \circ \varphi^{-1} \text{ and } g \circ f \circ \varphi^{-1} : \varphi(U) \subset \mathbb{R}^m &\rightarrow \mathbb{R} \\ (g \circ \psi^{-1})(\psi \circ f \circ \varphi^{-1})(x) &\equiv g(y(x))\end{aligned}$$

9.2. **Pullback in coordinates**,  $f^* : T_{f(p)}^* N \rightarrow T_p^* M$ . For pullback  $f^* : T_{f(p)}^* N \rightarrow T_p^* M$ ,

since  $\omega \in T_{f(p)}^* N$ , then write as  $\omega_j dy^j$

since  $f^* \omega \in T_p^* M$ , so  $f^* \omega = \psi_i dx^i$

$$\begin{aligned}\langle f^* \omega, X \rangle &\equiv f^* \omega(X) = \langle \psi_i dx^i, X^j \frac{\partial}{\partial x^j} \rangle \equiv \psi_i dx^i (X^j \frac{\partial}{\partial x^j}) = \psi_i X^i = \langle \omega, f_* X \rangle \equiv \omega(f_* X) = \\ &= \langle \omega_j dy^j, Y^i \frac{\partial}{\partial y^i} \rangle = \omega_j dy^j (Y^i \frac{\partial}{\partial y^i}) = \omega_j dy^j \left( X^i \frac{\partial y^k}{\partial x^i} \frac{\partial}{\partial y^k} \right) = \omega_j \frac{\partial y^i}{\partial x^i} X^i \\ &\implies \psi_i = \omega_j \frac{\partial y^j}{\partial x^i}\end{aligned}$$

Thus

**Theorem 12** (Push-forward in coordinates).

$$\text{For smooth } f : M \rightarrow N, \quad \begin{aligned} \dim M &= m \\ \dim N &= n \end{aligned}$$

$$\begin{aligned}X &= X^i \frac{\partial}{\partial x^i} \in T_p M \\ Y &= Y^j \frac{\partial}{\partial y^j} \in T_{f(p)} N\end{aligned} \quad \text{then}$$

$$(22) \quad \boxed{\begin{aligned} f_* : T_p M &\rightarrow T_{f(p)} N \\ f_*(X) &= Y \\ Y^j &= X^i \frac{\partial y^j}{\partial x^i} \end{aligned} \quad s.t.}$$

where  $f_*$  is the **push-forward**, and where colloquially,  $y^j = y^j(x)$ , but actually  $y^j \equiv \psi \circ f \circ \varphi^{-1}$

And

**Theorem 13** (Pullback in coordinates). *For smooth  $f : M \rightarrow N$ ,*

$$\begin{aligned} \dim M &= m \\ \dim N &= n \end{aligned}$$

$\omega \in T_{f(p)}^* N$ ,  $\psi = \psi_i dx^i \in T_p^* M$ , then



$$(23) \quad \boxed{\begin{array}{l} f^* : T_{f(p)}^* N \rightarrow T_p^* M \\ f^* \omega = \psi_i dx^i \\ \psi_i = \omega_j \frac{\partial y^j}{\partial x^i} \end{array}} \quad \text{where}$$

## 10. TENSORS

I'll go through Ch.7 *Tensors* of Jeffrey Lee (2009) [4].

**Definition 13** (7.1[4]). *Let  $V, W$  be modules over commutative ring  $R$ , with unity. Then, algebraic  $W$ -valued tensor on  $V$  is multilinear map.*

$$(24) \quad \tau : V_1 \times V_2 \times \cdots \times V_m \rightarrow W$$

where  $V_i = \{V, V^*\} \quad \forall i = 1, 2, \dots, m$ .

If for  $r, s$  s.t.  $r + s = m$ , there are  $r \quad V_i = V^*, s \quad V_i = V$ , tensor is  $r$ -contravariant,  $s$ -covariant; also say tensor of total type  $\binom{r}{s}$ .

EY : 20170404 Note that

$$\begin{aligned} &(\tau_\beta^{i\alpha} \frac{\partial}{\partial x^i} \text{ or } \tau_\beta^{i\alpha} e_i)(\omega_j dx^j \text{ or } \omega_j e^j \in V^*) \\ &(\tau_{i\alpha}^\beta dx^i \text{ or } \tau_{i\alpha}^\beta e^i)(X^j \frac{\partial}{\partial x^j} \text{ or } X^j e_j \in V) \end{aligned}$$

$\exists$  natural map  $V \rightarrow V^{**}$ ,  $\tilde{v} : \alpha \mapsto \alpha(v)$ . If this map is an isomorphism,  $V$  is **reflexive** module, and identify  $V$  with  $V^{**}$ .

**Exercise 7.5.** Given vector bundle  $\pi : E \rightarrow M$ , open  $U \subset M$ , consider sections of  $\pi$  on  $U$ , i.e. cont.  $s : U \rightarrow E$ , where  $(\pi \circ s)(u) = u, \quad \forall u \in U$ .

Consider  $E^* \ni \omega = \omega_i e^i$ .

$\forall s \in \Gamma(E), \omega(s) = \omega_i(s(x))^i, \quad \forall x \in U \subset M$ . So define  $\tilde{s} : \omega, x \mapsto \omega(s(x)), \quad \forall x \in U$ .

If  $\tilde{s} = 0, \tilde{s}(\omega, x) = \omega(s(x)) = 0 \quad \forall \omega \in E^*, \forall x \in U$ , and so  $s = 0$ . (Let  $\omega_i = \delta_{iJ}$  for some  $J$ , and so  $s^J(x) = 0 \quad \forall J$ ).

$s = 0$ . So  $\ker(s \mapsto \tilde{s}) = \{0\}$  (so condition for injectivity is fulfilled).

Since  $\tilde{s} : \omega, x \mapsto \omega(s(x)), \forall \omega \in E^*, \forall x \in U, s \mapsto \tilde{s}$  is surjective.

$s \mapsto \tilde{s}$  is an isomorphism so  $\Gamma(E)$  is a *reflexive* module.

**Proposition 3.** *For  $R$  a ring (special case),  $\exists$  module homomorphism:*

*tensor product space  $\rightarrow$  tensor, as a multilinear map, i.e.  $\exists$*

$$(25) \quad (\otimes_{i=1}^r V) \otimes (\otimes_{j=1}^s V^*) \rightarrow T_s^r(V; R) \\ u_1 \otimes \cdots \otimes u_r \otimes \beta^1 \otimes \cdots \otimes \beta^s \in (\otimes^r V) \otimes (\otimes^s V^*) \mapsto (\alpha^1 \dots \alpha^r, v_1 \dots v_s) \mapsto \alpha^1(u_1) \dots \alpha^r(u_r) \beta^1(v_1) \dots \beta^s(v_s)$$

Indeed, consider

$$(\alpha^1 \dots \alpha^r, v_1 \dots v_s) \in \underbrace{V^* \times \cdots \times V^*}_r \times \underbrace{V \times \cdots \times V}_s \mapsto \alpha^1(u_1) \dots \alpha^r(u_r) \beta^1(v_1) \dots \beta^s(v_s)$$

and so for

$$\begin{aligned} \alpha^i &= \alpha_\mu^i e^\mu, & i &= 1, 2, \dots, r, \mu = 1, 2, \dots, \dim V^* & \alpha^i(u_i) &= \alpha_\mu^i u_i^\mu \\ v_i &= v_\mu^i e_\mu, & i &= 1, 2, \dots, s, \mu = 1, 2, \dots, \dim V & \beta^i(v_i) &= \beta_\mu^i v_\mu^i \end{aligned}$$

So that

$$\begin{aligned} \alpha^1(u_1) \dots \alpha^r(u_r) \beta^1(v_1) \dots \beta^s(v_s) &= \alpha_{\alpha_1}^1 u_1^{\alpha_1} \dots \alpha_{\alpha_r}^r u_r^{\alpha_r} \beta_{\mu_1}^1 v_1^{\mu_1} \dots \beta_{\mu_s}^s v_s^{\mu_s} = \\ &= (u_1^{\alpha_1} \dots u_r^{\alpha_r} \beta_{\mu_1}^1 \dots \beta_{\mu_s}^s)(\alpha_{\alpha_1}^1 \dots \alpha_{\alpha_r}^r v_1^{\mu_1} \dots v_s^{\mu_s}) \end{aligned}$$

Identify  $u_1 \otimes \cdots \otimes u_r \otimes \beta^1 \otimes \cdots \otimes \beta^s$  with this multilinear map.

**Proposition 4.** *If  $V$  is finite-dim. vector space, or if  $V = \Gamma(E)$ , for vector bundle  $E \rightarrow M$ , map*

$$(26) \quad (\otimes_{i=1}^r V) \otimes (\otimes_{j=1}^s V^*) \rightarrow T_s^r(V; R)$$

*is an isomorphism.*

**Definition 14.** *tensor that can be written as*

$$(27) \quad u_1 \otimes \cdots \otimes u_r \otimes \beta^1 \otimes \cdots \otimes \beta^s \equiv u_1 \otimes \cdots \otimes \beta^s$$

is **simple** or **decomposable**.

Now well that not *all* tensors are simple.

**Definition 15** (7.7[4], tensor product).  $\forall S \in T_{s_1}^{r_1}(V), \forall T \in T_{s_2}^{r_2}(V)$ , define tensor product

$$(28) \quad S \otimes T \in T_{s_1+s_2}^{r_1+r_2}(V) \\ S \otimes T(\theta^1 \dots \theta^{r_1+r_2}, v_1 \dots v_{s_1+s_2}) := S(\theta^1 \dots \theta^{r_1}, v_1 \dots v_{s_1}) T(\theta^{r_1+1} \dots \theta^{r_1+r_2}, v_{s_1+1} \dots v_{s_1+s_2})$$

**Proposition 5** (7.8[4]).

$$\tau^{i_1 \dots i_r}_{j_1 \dots j_s} e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} = \tau(e^{i_1} \dots e^{i_r}, e_{j_1} \dots e_{j_s}) e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} = \tau$$

So  $\{e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} | i_1 \dots i_r, j_1 \dots j_s \in 1 \dots n\}$  spans  $T_s^r(V; R)$

**Exercise 7.11.** Let basis for  $V \quad e_1 \dots e_n$ , corresponding dual basis for  $V^* \quad e^1 \dots e^n$

Let basis for  $V \quad \bar{e}_1 \dots \bar{e}_n$ , corresponding dual basis for  $V^* \quad \bar{e}^1 \dots \bar{e}^n$

s.t.

$$\begin{aligned} \bar{e}_i &= C_i^k e_k \\ \bar{e}^i &= (C^{-1})^i_k e^k \end{aligned}$$

EY:20170404, keep in mind that

$$\begin{aligned} Ax &= e_i A_i^k e^k (x^j e_j) = e_i A_i^j x^j = A_i^j x^j e_i \\ Ae_j &= e_k A_i^k e^i (e_j) = A_j^k e_k = \bar{e}_j \\ \bar{\tau}_{jk}^i \bar{e}_i \otimes \bar{e}^j \otimes \bar{e}^k &= \bar{\tau}_{jk}^i C_i^l e_l (C^{-1})_m^j e^m (C^{-1})_n^k e^n = \bar{\tau}_{jk}^i C_i^l (C^{-1})_m^j (C^{-1})_n^k = \tau_{mn}^l \\ \bar{\tau}_{jk}^i &= C_k^c C_j^b (C^{-1})_a^i \tau_{bc}^a \end{aligned}$$

On Remark 7.13 of Jeffrey Lee (2009) [4]: first, egregious typo for  $L(V, V)$ ; it should be  $L(V, W)$ . Onward, for  $L(V, W)$ , consider  $W \otimes V^* \ni w \otimes \alpha$  s.t.

$$(w \otimes \alpha)(v) = \alpha(v)w \in W, \forall v \in V, \text{ so } w \otimes \alpha \in L(V, W)$$

Now consider (category of) left  $R$ -module,

$$(29) \quad {}_R \mathbf{Mod} \ni {}_{\text{Mat}_{\mathbb{K}}(N, M)} \mathbb{K}^N$$

where

$$\begin{aligned} V &= \mathbb{K}^N \\ W &= \mathbb{K}^M \end{aligned}$$

For  $A \in \text{Mat}_{\mathbb{K}}(N, M)$ ,  $x \in \mathbb{K}^N$ ,

$$e_i A^i_{\phantom{i}\mu} e^\mu (x^\nu e_\nu) = Ax = e_i A^i_\mu x^\mu, \quad i = 1, 2, \dots M, \mu = 1, 2, \dots N$$

$$A \in \text{Mat}_{\mathbb{K}}(N, M) \cong W \otimes V^* \cong L(V, W)$$

Consider

$$\begin{aligned} \alpha &\in (\mathbb{K}^N)^* = V^* & \alpha &= \alpha_\mu e^\mu \\ w &\in \mathbb{K}^M = W & w &= w^i e_i \\ \alpha \otimes w &= w \otimes \alpha = w^i \alpha_\mu e_i \otimes e^\mu \end{aligned}$$

(remember, isomoprhism between  $\text{Mat}_{\mathbb{K}}(N, M)$  and  $W \otimes V^*$  guaranteed, if  $V, W$  are free  $R$ -modules,  $R = \mathbb{K}$ ).

Let  $V, W$  be left  $R$ -modules, i.e.  $V, W \in {}_R\mathbf{Mod}$ .

$$V^* \in \mathbf{Mod}_R$$

For  $V^* \otimes W \in \mathbf{Mod}_R \otimes {}_R\mathbf{Mod}$

$$\begin{aligned} \alpha &\in V^*, w \in W \\ (\alpha \otimes w)(v) &= \alpha(v)w, \text{ for } v \in V \in {}_R\mathbf{Mod} \end{aligned}$$

But  $(w \otimes \alpha)(v) = w\alpha(v)$ .

Note  $\alpha(v) \in R$ .

Let  $V, W$  be right  $R$ -modules, i.e.  $V, W \in \mathbf{Mod}_R$ .

$$V^* \in {}_R\mathbf{Mod}$$

For  $W \otimes V^* \in \mathbf{Mod}_R \otimes {}_R\mathbf{Mod}$ .

$$\begin{aligned} \alpha &\in V^*, w \in W \\ (v)(w \otimes \alpha) &= w\alpha(v), \text{ with } \alpha(v) \in R, v \in V \end{aligned}$$

So  $W \otimes V^* \cong L(V, W)$ , for  $V, W \in \mathbf{Mod}_R$

**Definition 16** (7.20[4], **contraction**). *Let  $(e_1, \dots e_n)$  basis for  $V$ ,  $(e^1 \dots e^n)$  dual basis.*

*If  $\tau \in T^r_s(V)$ , then for  $k \leq r$ ,  $l \leq s$ , define*

$$\begin{aligned} C_l^k \tau &\in T^{r-1}_{s-1}(V) \\ C_l^k \tau(\theta^1 \dots \theta^{r-1}, w_1 \dots w_{s-1}) &:= \\ (30) \quad \sum_{a=1}^n \tau(\theta^1 \dots \underbrace{e^a}_{kth \ position} \dots \theta^{r-1}, w_1 \dots \underbrace{e_a}_{ith \ position} \dots w_{s-1}) \end{aligned}$$

$C_l^k$  is called **contraction**, for some single  $1 \leq k \leq r$ , some single  $1 \leq l \leq s$ ,

$$C_l^k : T^r_s(V) \rightarrow T^{r-1}_{s-1}(V)$$

s.t.

$$(C_l^k \tau)^{i_1 \dots \widehat{i_k} \dots i_r}_{j_1 \dots \widehat{j_l} \dots j_s} := \tau^{i_1 \dots a \dots i_r}_{j_1 \dots a \dots j_s}$$

Universal mapping properties can be invoked to give a basis free definition of contraction (EY : 20170405???)

IN general,

$$\forall v_1 \dots v_s \in V, \forall \alpha^1 \dots \alpha^r \in V^*$$

so that

$$\begin{aligned} v_j &= v_j^\mu e_\mu & j &= 1 \dots s, & \mu &= 1, \dots \dim V \\ \alpha^i &= \alpha^i_\mu e^\mu & i &= 1 \dots r, & \mu &= 1 \dots \dim V^* \end{aligned}$$

then  $\forall \tau \in T^r_s(V)$ ,

$$\begin{aligned} \tau(\alpha^1 \dots \alpha^r, v_1 \dots v_s) &= \tau(\alpha_{\mu_1}^1 e^{\mu_1} \dots \alpha_{\mu_r}^r e^{\mu_r}, v_1^{\nu_1} e_{\nu_1} \dots v_s^{\nu_s} e_{\nu_s}) = \\ &= \alpha_{\mu_1}^1 \dots \alpha_{\mu_r}^r v_1^{\nu_1} \dots v_s^{\nu_s} \tau(e^{\mu_1} \dots e^{\mu_r}, e_{\nu_1} \dots e_{\nu_s}) = \alpha_{\mu_1}^1 \dots \alpha_{\mu_r}^r v_1^{\nu_1} \dots v_s^{\nu_s} \tau^{\mu_1 \dots \mu_r}_{\nu_1 \dots \nu_s} \end{aligned}$$

which is equivalent to

$$\begin{array}{ccc} \tau \in T^r_s(V) & \xrightarrow{\alpha^1 \otimes \dots \otimes \alpha^r \otimes v_1 \otimes \dots \otimes v_s \otimes} & \alpha^1 \otimes \dots \otimes \alpha^r \otimes v_1 \otimes \dots \otimes v_s \otimes \tau \\ & & \downarrow C^1_{s+1} C^2_{s+2} \dots C^r_{r+s} C^r_1 C^{r+1}_2 \dots C^{r+s}_s \\ & & \tau(\alpha^1 \dots \alpha^r, v_1 \dots v_s) \in R \end{array}$$

where I've tried to express the right- $R$ -module, "right action" on  $\alpha^1 \otimes \dots \otimes \alpha^r \otimes v_1 \otimes \dots \otimes v_s \in V^* \otimes \dots \otimes V$ .

Conlon (2008) [19]

### 10.1. Metric.

10.1.1. *Induced Metric.* cf. 7.1.2. Induced metric, Nakahara (2003) [9]. Let  $M$  be  $m$ -dim. submanifold of  $n$ -dim. Riemannian manifold  $N$  with metric  $g_N$ .

If  $f : M \rightarrow N$  embedding which induced the submanifold structure of  $M$ . (Sec. 5.2 of Nakahara (2003) [9]).

(Recall, smooth  $f : M \rightarrow N$ ,  $\dim M \leq N$ ,  $f$  immersion if  $f_* : T_p M \rightarrow T_{f(p)} N$  injection, so  $\text{rank } f_* = \dim M$ .  $f$  embedding if  $f$  immersion and  $f$  injection. Also  $f(M)$  submanifold of  $N$ ).

**Proposition 6.** *pullback  $f^*$  induced a natural metric  $g_M = f^* g_N$  on  $M$  is given by*

$$(31) \quad g_{M\mu\nu}(x) = g_{N\alpha\beta} \frac{\partial f^\alpha}{\partial x^\mu} \frac{\partial f^\alpha}{\partial x^\nu}$$

Recall that

$$\begin{aligned} f^* : T^*_{f(p)} N &\rightarrow T^*_p M & \langle f^* \omega, V \rangle &= \langle \omega, f_* V \rangle \\ \omega &\in T^*_{f(p)} N & \text{and} \\ V &\in T_p M & f_* V[g] &= V[gf], g \in \mathcal{F}(N) \end{aligned}$$

Using our notation, consider

$$\begin{aligned} (U, \varphi), U \subset M, \varphi : U \subset M &\rightarrow \mathbb{R}^m \\ (V, \psi), V \subset N, \psi : V \subset N &\rightarrow \mathbb{R}^n \end{aligned}$$

Colloquially, say  $y^j = y^j(x)$ , but actually  $y^j = \psi \circ f \circ \varphi^{-1}$ , so that

$$f_*(X) = Y, \text{ where } X = X^i \frac{\partial}{\partial x^i}, Y = Y^j \frac{\partial}{\partial y^j} = X^i \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}$$

We also observe that

$$f_* V[g] = V[g \circ f] = V^i \frac{\partial}{\partial x^i} (g \circ f) = V^i \frac{\partial g}{\partial y^i} \frac{\partial f^j}{\partial x^i} = V^i \frac{\partial f^j}{\partial x^i} \frac{\partial g}{\partial y^j}$$

where notation-wise,  $X \equiv V$ ,  $y \equiv f$  with respect to ours and Nakahara's notation.

Now  $g_N : T_{f(p)} N \otimes T_{f(p)} N \rightarrow \mathbb{R}$ , then if we just input one vector into  $g_N$ , it becomes a covector  $g_N(U, ) \in T^*_{f(p)} N$ .

So then

$$\begin{aligned} \langle f^* g_N(U, ), V \rangle &= \langle g_N(U, ), f_* V \rangle = \langle g_N(U, ), V^i \frac{\partial f^j}{\partial x^i} \frac{\partial}{\partial y^j} \rangle = \langle g_N(U, ), V^i \frac{\partial}{\partial y^j} \rangle \frac{\partial f^j}{\partial x^i} \\ g_N(U, V^i \frac{\partial}{\partial y^i}) \frac{\partial f^j}{\partial x^i} &= g_N(U^k \frac{\partial}{\partial y^k}, V^i \frac{\partial}{\partial y^j}) \frac{\partial f^j}{\partial x^i} = (g_N)_{aj} V^i \frac{\partial f^j}{\partial x^i} \end{aligned}$$

where we let  $U^k = \delta^{ak}$  in the last step.

Let  $V = \delta^{bi} \frac{\partial}{\partial x^i}$ , so  $(g_N)_{aj} \frac{\partial f^j}{\partial x^b} = \langle f^* g_N \left( \frac{\partial}{\partial y^a}, \right), \frac{\partial}{\partial x^b} \rangle$ .

By symmetry,

(32)

$$(g_M)_{ab} = (g_N)_{ij} \frac{\partial f^i}{\partial x^a} \frac{\partial f^j}{\partial x^b}$$

where  $f : M \rightarrow N$ .

**Problem 8, pp. 28, Wald (1984) [36].**

(a) The metric of flat, 3-dim. Euclidean space is

$$ds^2 = dx^2 + dy^2 + dz^2$$

Show that the metric components  $g_{\mu\nu}$  in spherical polar coordinates  $r, \theta, \phi$  defined by

$$r = (x^2 + y^2 + z^2)^{1/2}$$

$$\cos \theta = z/r$$

$$\tan \phi = y/x$$

is given by

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

(b) The spacetime metric of special relativity is

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$$

Find the components,  $g_{\mu\nu}$  and  $g^{\mu\nu}$ , of the metric and inverse metric in ”rotating coordinates,” defined by

$$t' = t$$

$$x' = (x^2 + y^2)^{1/2} \cos (\phi - \omega t)$$

$$y' = (x^2 + y^2)^{1/2} \sin (\phi - \omega t)$$

$$z' = z$$

where  $\tan \phi = y/x$ .

**Solution 8, pp. 28, Wald (1984) [36].**

(a) Consider  $(r, \theta, \varphi) \rightarrow (x, y, z)$  s.t.

$$\begin{array}{l} x = r s \theta c \varphi \\ y = r s \theta s \varphi \\ z = r c \theta \end{array} \text{ so consider } \left| \begin{array}{ccc} s \theta c \varphi & r c \theta c \varphi & -r s \theta s \varphi \\ s \theta s \varphi & r c \theta s \varphi & r s \theta c \varphi \\ c \theta & -r s \theta & 0 \end{array} \right|$$

Since  $(g_N)_{ij} = \delta_{ij}$  in our case,

$$(g_M)_{rr} = 1$$

$$(g_M)_{\theta\theta} = r^2$$

$$(g_M)_{\varphi\varphi} = r^2 \sin^2 \theta$$

By summing terms across columns,  $(g_M)_{ab} = 0$  if  $a \neq b$  and  $(g_M)_{aa}$  obtained by squaring and summing terms across a single column.

Thus, some physicists write as a (sometimes confusing) shorthand for the metric tensor

$$g_M = dr \otimes dr + r^2 d\theta \otimes d\theta + r^2 \sin^2 \theta d\varphi \otimes d\varphi$$

to be

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

(b) Consider  $(t', x', y', z') \rightarrow (t, x, y, z)$ . Let  $\rho = (x^2 + y^2)^{1/2}$ . Consider the Jacobian:

$$\left| \begin{array}{ccc} 1 & & \\ \omega \rho s(\phi - \omega t) & \frac{x}{\rho} c(\phi - \omega t) & \frac{y}{\rho} c(\phi - \omega t) \\ -\omega \rho c(\phi - \omega t) & \frac{x}{\rho} s(\phi - \omega t) & \frac{y}{\rho} s(\phi - \omega t) \\ 0 & & 1 \end{array} \right|$$

The terms with the same components, e.g.  $g_{t't'}$ ,  $g_{x'x'}$ ,  $\dots$  are  $(-1 + \omega^2 \rho^2)$ ,  $\frac{x^2}{\rho^2}$ ,  $\frac{y^2}{\rho^2}$ , 1, and the cross terms are  $\frac{xy}{\rho^2} dx \otimes dy$ .

Part 5. Lie Groups, Lie Algebra

11. LIE GROUPS

- : Lie Groups
- : Groups
- : Ring
- : group algebra
- : Group Ring
- : Representation Theory
- : Modules
- :  $kG$ -modules

From Sec. 8.1 “Noncommutative Rings” of Rotman (2010) [12]:

**Definition 17.** *ring  $R$  - additive abelian group equipped with multiplication  $R \times R \rightarrow R$  s.t.  $\forall a, b \in R$*   
 $(a, b) \mapsto ab$

- (i)  $a(bc) = (ab)c$
- (ii)  $a(b + c) = ab + ac$ ,  $(b + c)a = ba + ca$
- (iii)  $\exists 1 \in R$  s.t.  $\forall a \in R$ ,  $1a = a = a1$

Example 8.1[12]

- (ii) group algebra  $kG$ ,  $k$  commutative ring,  $G$  group, “its additive abelian group is free  $k$ -module having basis labeled by elements of  $G$ ,  
i.e.  $\forall a \in kG$ ,  $a = \sum_{g \in G} a_g g$ ,  $a_g \in k$ ,  $\forall g \in G$ ,  $a_g \neq 0$  for only finitely many  $g \in G$ .

define (ring) multiplication  $kG \times kG \rightarrow kG$   $\forall a, b \in kG$ ,  
 $ab = ab$   
 $a = \sum_{g \in G} a_g g$  to be  
 $b = \sum_{h \in G} b_h g$

$$\left( \sum_{g \in G} a_g g \right) \left( \sum_{h \in G} b_h h \right) = \sum_{z \in G} \left( \sum_{gh=z} a_g b_h \right) z$$

**Definition 18.** *Given  $R$  ring, left  $R$ -module is (additive) abelian group  $M$  equipped with*

*scalar multiplication  $R \times M \rightarrow M$  s.t.  $\forall m, m' \in M$ ,  $\forall r, r', 1 \in R$*   
 $(r, m) \mapsto rm$

- (i)  $r(m + m') = rm + rm'$
- (ii)  $(r + r')m = rm + r'm$
- (iii)  $(rr')m = r(r'm)$

(iv)  $1m = m$

EY : 20150922 Example : for  $kG$ -module  $V^\sigma$ , for  $r \in kG$ , so  $r = \sum_{g \in G} a_g g$

$$\begin{array}{ccc} R \times M & \rightarrow & M \\ (r, m) & \mapsto & rm \end{array} \implies \begin{array}{ccc} kG \times V & \rightarrow & V \\ (r, v) & \mapsto & tv \end{array}$$

For some representation  $\sigma : G \rightarrow GL(V)$ ,

$$rv = \sum_{g \in G} a_g g \cdot v = \sum_{g \in G} a_g \sigma_g(v)$$

So a  $kG$ -module needs to be associated with some chosen representation.

Note for  $V$  as an additive abelian group,  $\forall u, v, w \in V$ ,

$$v + w = w + v, (u + v) + w = u + (v + w)$$

$$v + 0 = v \quad \forall v \in V \text{ for } 0 \in V$$

$$v + (-v) = 0 \quad \forall v \in V$$

So a vector space can be an additive abelian group.

Note that

$$r(v + w) = \left( \sum_{g \in G} a_g g \right) (v + w) = \left( \sum_{g \in G} a_g \sigma_g \right) (v + w) = \sum_{g \in G} a_g \sigma_g(v) + \sum_{g \in G} a_g \sigma_g(w) = rv + rw$$

$$(r + r')v = \left( \sum_{g \in G} a_g g + b_g g \right) v = \sum_{g \in G} (a_g \sigma_g + b_g \sigma_g) v = \sum_{g \in G} a_g \sigma_g(v) + \sum_{g \in G} b_g \sigma_g(v) = rv + r'v$$

$$(rr')v = \left( \sum_{g \in G} a_g g \sum_{h \in G} b_h h \right) v = \left( \sum_{z \in G} \sum_{gh=z} a_g b_h z \right) v = \sum_{z \in G} \sum_{gh \in z} a_g b_h \sigma_z(v) = \sum_{g \in G} \sum_{h \in G} a_g b_h \sigma_g \sigma_h(v)$$

since  $\sigma(gh) = \sigma(g)\sigma(h) = \sigma_g \sigma_h = \sigma_{gh}$  ( $\sigma$  homomorphism)

$$1v = \sigma(1)v = 1v = v$$

From Sec. 8.3 “Semisimple Ring” of Rotman (2010) [12]:

**Definition 19.**  $k$ -representation of group  $G$  is homomorphism

$$\sigma : G \rightarrow GL(V)$$

where  $V$  is vector field over field  $k$

**Proposition 7** (8.37 Rotman (2010)[12]).  $\forall k$ -representation  $\sigma : G \rightarrow GL(V)$  equips  $V$  with structure of left  $kG$ -module, denote module by  $V^\sigma$ .

Conversely,  $\forall$  left  $kG$ -module  $V$  determines  $k$ -representation  $\sigma : G \rightarrow GL(V)$

*Proof.* Given  $\sigma : G \rightarrow GL(V)$ ,

$$\sigma_g =: \sigma(g) : V \rightarrow V$$

define

$$\begin{array}{ccc} kG \times V & \rightarrow & V \\ \left( \sum_{g \in G} a_g g \right) & & v = \sum_{g \in G} a_g \sigma_g(v) \end{array}$$

$$v, w \in V$$

Let  $r, r', 1 \in kG$

$$r = \sum_{g \in G} a_g g$$

$$r(v + w) = \left( \sum_{g \in G} a_g g \right) (v + w) = \left( \sum_{g \in G} a_g \sigma_g \right) (v + w) = \sum_{g \in G} a_g \sigma_g(v) + \sum_{g \in G} a_g \sigma_g(w) = rv + rw$$

$$(r + r')v = \left( \sum_{g \in G} a_g g + b_g g \right) v = \sum_{g \in G} (a_g \sigma_g + b_g \sigma_g) v = \sum_{g \in G} a_g \sigma_g(v) + \sum_{g \in G} b_g \sigma_g(v) = rv + r'v$$

$$(rr')v = \left( \sum_{g \in G} a_g g \sum_{h \in G} b_h h \right) v = \left( \sum_{z \in G} \sum_{gh=z} a_g b_h z \right) v = \sum_{z \in G} \sum_{gh \in z} a_g b_h \sigma_z(v) = \sum_{g \in G} \sum_{h \in G} a_g b_h \sigma_g \sigma_h(v)$$

since  $\sigma(gh) = \sigma(g)\sigma(h) = \sigma_g \sigma_h = \sigma_{gh}$  ( $\sigma$  homomorphism)

$$1v = \sigma(1)v = 1v = v$$

Conversely, assume  $V$  left  $kG$ -module.

If  $g \in G$ , then  $v \mapsto gv$  defines  $T_g : V \rightarrow V$ .  $T_g$  nonsingular since  $\exists T_g^{-1} = T_{g^{-1}}$

Define  $\sigma : G \rightarrow GL(V)$

$$\sigma : g \mapsto T_g$$

$\sigma$   $k$ -representation

$$\sigma(gh) = T_{gh} = T_g T_h = \sigma(g)\sigma(h)$$

$$\sigma(gh)(v) = T_{gh}v = ghv = T_g T_h v = \sigma(g)\sigma(h)v \quad \forall v \in V$$

□

**Proposition 8.** Let group  $G$ , let  $\sigma, \tau : G \rightarrow GL(V)$  be  $k$ -representations, field  $k$ .

If  $V^\sigma, V^\tau$  corresponding  $kG$ -modules in Prop. 7 (Prop. 8.37 in Rotman (2010) [12]), then

$V^\sigma \simeq V^\tau$  as  $kG$ -modules iff  $\exists$  nonsingular  $\varphi : V \rightarrow V$  s.t.

$$\varphi\tau(g) = \sigma(g)\varphi \quad \forall g \in G$$

*Proof.* If  $\varphi : V^\tau \rightarrow V^\sigma$   $kG$ -isomorphism, then  $\varphi : V \rightarrow V$  isomorphism s.t.

$$\varphi(\sum a_g gv) = (\sum a_g g)\varphi(v) \quad \forall v \in V, \forall g \in G$$

in  $V^\tau$ ,  $kG \times V \rightarrow V$

$$gv = \tau(g)(v)$$

in  $V^\sigma$ ,  $kG \times V \rightarrow V$  scalar multiplication

$$gv = \sigma(g)(v)$$

$$\implies \forall g \in G, v \in V, \quad \varphi(\tau(g)(v)) = \sigma(g)(\varphi(v))$$

I think

$$\varphi(gv) = \varphi(\tau(g)(v)) = g\varphi(v) = \sigma(g)\varphi(v)$$

$$\implies \varphi\tau(g) = \sigma(g)\varphi \quad \forall g \in G$$

Conversely, if  $\exists$  nonsingular  $\varphi : V \rightarrow V$  s.t.  $\varphi\tau(g) = \sigma(g)\varphi \quad \forall g \in G$

$$\varphi\tau(g)v = \varphi(\tau(g)v) = \sigma(g)\varphi(v) \quad \forall g \in G, \forall v \in V$$

Consider scalar multiplication

$$\begin{aligned} kG \times V &\rightarrow V \\ \sum_{g \in G} a_g g(v) &= \sum_{g \in G} a_g \tau_g(v) \\ \varphi \left( \sum_{g \in G} a_g \tau_g(v) \right) &= \varphi \left( \sum_{g \in G} a_g \tau(g)v \right) = \sum_{g \in G} a_g \sigma(g) \sigma(g) \varphi(v) = \left( \sum_{g \in G} a_g g \right) \varphi(v) \end{aligned}$$

□

Admittedly, after this exposition from Rotman (2010) [12], I still didn't understand how  $kG$ -modules relate to representation theory and group rings. I turned to Baker (2011) [13], which we'll do right now. Note that I found a lot of links to online resources on representation theory from Khovanov's webpage <http://www.math.columbia.edu/~khovanov/resource/>.

Note,

**Definition 20.** *vector subspace  $W \subseteq V$  is called a*

*$G$ -submodule,  $G$ -subspace, EY : 20150922 “invariant” subspace?*

*if  $\forall g \in G$ , for representation  $\rho : G \rightarrow GL_k(V)$ ,  $\rho_g(w) \in W$ ,  $\forall w \in W$ ,  $\forall g \in G$  i.e. closed under “action of elements of  $G$ ” with  $\rho_g =: \rho(g) : V \rightarrow V$*

Given basis  $\mathbf{v} = \{v_1 \dots v_n\}$  for  $V$ ,  $\dim_k V = n$ ,  $\forall g \in G$ ,

$$\rho_g v_j = \rho(g) v_j = r_{kj}(g) v_k$$

for, indeed,

$$\rho_g x^j v_j = \rho(g) x^j v_j = x^j \rho(g) v_j = x^j r_{kj}(g) v_k = r_{kj} x^j v_k$$

so that

$$\begin{aligned} \rho : G &\rightarrow GL_k(V) \\ \rho(g) &= [r_{ij}(g)] \end{aligned}$$

Example 2.1 (Baker (2011) [13]): Let  $\rho : G \rightarrow GL_k(V)$  where  $\dim_k V = 1$

$$\forall v \in V, v \neq 0, \forall g \in G, \lambda_g \in k \text{ s.t. } g \cdot v = \rho_g(v) = \lambda_g v$$

$$\rho(hg)v = \rho_h \rho_g v = \lambda_{hg} v = \lambda_h \lambda_g v \implies \lambda_{hg} = \lambda_h \lambda_g$$

$\implies \exists$  homomorphism  $\Lambda : G \rightarrow k^\times$

$$\Lambda(g) = \lambda_g$$

From Sec. 2.2 “ $G$ -homomorphisms and irreducible representations” of Baker (2011) [13], suppose  $\rho : G \rightarrow GL_k(V)$  are 2 representations  $\sigma : G \rightarrow GL_k(W)$

Many names for the same thing:  $G$ -equivalent,  $G$ -linear,  $G$ -homomorphism, EY : 20150922  $kG$ -isomorphic?

If  $\forall g \in G$ ,

$$\begin{array}{ccc} V & \xrightarrow{\varphi} & V \\ \tau_g \downarrow & & \downarrow \sigma_g \\ V & \xrightarrow{\varphi} & V \end{array} \iff V^\tau \cong V^\sigma$$

Indeed, define

$$\varphi : V^\tau \rightarrow V^\sigma$$

$$\varphi(v + w) = \varphi(v) + \varphi(w)$$

$$\varphi(rv) = \varphi\left(\sum_{g \in G} a_g g \cdot v\right) = \varphi\left(\sum_{g \in G} a_g \tau_g(v)\right) = \sum_{g \in G} a_g \varphi(\tau_g(v)) = \sum_{g \in G} a_g \sigma_g \cdot \varphi(v) = r \varphi(v)$$

EY : 20150922 So  $\varphi$  is a  $kG$ -isomorphism between left  $kG$  modules  $V^\tau$  and  $V^\sigma$  if it's bijective and is “linear” in “scalars”  $r \in kG$ , i.e.  $\varphi(rv) = r\varphi(v)$ .

Define action of  $G$  on  $\text{Hom}_k(V, W)$  ( $\text{Hom}_k(V, W)$  is the vector space of  $k$ -linear transformations  $V \rightarrow W$ )

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ v & \xmapsto{f} & f(v) \end{array} \quad \forall f \in \text{Hom}_k(V, W), f : V \rightarrow W$$

Consider

$$G \times \text{Hom}_k(V, W) \rightarrow \text{Hom}_k(V, W)$$

$$(g \cdot f) \mapsto (\sigma_g f) \circ \rho_{g^{-1}} \text{ i.e. } (g \cdot f)(v) = \sigma_g f(\rho_{g^{-1}} v) \quad (f \in \text{Hom}_k(V, W))$$

Let  $g, h \in G$ ,

$$(gh \cdot f)(v) = g \cdot \sigma_h f(\rho_{h^{-1}} v) = \sigma_g \sigma_h f \rho_{h^{-1}} \rho_{g^{-1}}(v) = (\sigma_{gh} f \rho_{(gh)^{-1}})(v)$$

Thus,  $G \times \text{Hom}_k(V, W) \rightarrow \text{Hom}_k(V, W)$  is thus another  $G$ -representation of  $G$ .

$$(g \cdot f) \mapsto (\sigma_g f) \circ \rho_{g^{-1}}$$

For  $k$ -representation  $\rho$ , if the only  $G$ -subspaces of  $V$  are  $\{0\}$ ,  $V$ ,  $\rho$  **irreducible** or **simple**.

$$\rho_g(\{0\}) = \{0\}$$

$$\rho_g(V) = V$$

given subrepresentation  $W \subseteq V$ ,  $V/W$  admits linear action of  $G$ ,  $\bar{\rho}_W : G \rightarrow GL_k(V/W)$  quotient representation

$$\bar{\rho}_W(g)(v + W) = \rho(g)(v) + W$$

if  $v' - v \in W$

$$\rho(g)(v') + W = \rho(g)(v + (v' - v)) + W = (\rho(g)(v) + \rho(g)(v' - v)) + W = \rho(g)(v) + W$$

**Proposition 9** (2.7 Baker (2011)[13]). *if  $f : V \rightarrow W$   $G$ -homomorphism, then*

(a)  *$\ker f$  is  $G$ -subspace of  $V$*

(b)  *$\text{im} f$  is  $G$ -subspace of  $W$*

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \rho_g \downarrow & & \downarrow \sigma_g \\ V & \xrightarrow{f} & W \end{array}$$

*Proof.* Recall

(a) Let  $v \in \ker f$ . Then  $\forall g \in G$ ,

$$f(\rho_g v) = \sigma_g f(v) = 0$$

so  $\rho_g v \in \ker f$ ,  $\forall g \in G$ . So  $\ker f$  is  $G$ -subspace of  $V$

- (b) Let  $w \in \text{im} f$ . So  $w = f(u)$  for some  $u \in V$

$$\sigma_g w = \sigma_g f(u) = f(\rho_g u) \in \text{im} f$$

So  $\text{im} f$  is  $G$ -subspace of  $W$

**Theorem 14** (Schur's Lemma). *Let  $\rho : G \rightarrow GL_{\mathbb{C}}(V)$  be irreducible representations of  $G$  over field  $k = \mathbb{C}$ ; let  $f : V \rightarrow W$  be  $\sigma : G \rightarrow GL_{\mathbb{C}}(W)$*

$G$ -linear map.

- (a) if  $f \neq 0$ ,  $f$  isomorphism. True  $\forall k$  field, not just  $\mathbb{C}$   
 (b) if  $V = W$ ,  $\rho = \sigma$ , then for some  $\lambda \in \mathbb{C}$ ,  $f$  given by  $f(v) = \lambda v$  ( $v \in V$ ) (true for algebraically closed fields)

*Proof.* (a) By Prop. 9,  $\ker f \subseteq V$ ,  $\text{im} f \subseteq W$  are  $G$ -subspaces.  
 For  $\rho$ , only  $G$ -subspaces are 0 or  $V$ , so if  $\ker f = V$ ,  $f = 0$ . If  $\ker f = 0$ ,  $f$  injective.  
 For  $\sigma$ , only  $G$ -subspaces are 0 or  $V$ , so  $\text{im} f = 0$ ,  $f = 0$ . If  $\text{im} f = V$ ,  $f$  surjective.  
 $\implies f$  isomorphism.  
 (b) Let  $\lambda \in \mathbb{C}$  be an eigenvalue of  $f$ ,  $f(v_0) = \lambda v_0$  eigenvector,  $v_0 \neq 0$ .  
 Let linear  $f_\lambda : V \rightarrow V$  s.t.

$$f_\lambda(v) = f(v) - \lambda v \quad (v \in V)$$

$$\forall g \in G$$

$$\rho_g f_\lambda(v) = \rho_g f(v) - \rho_g \lambda v = f(\rho_g v) - \lambda \rho_g v = f_\lambda(\rho_g v)$$

So  $f_\lambda$  is  $G$ -linear, for

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \rho_g \downarrow & & \downarrow \rho_g \\ V & \xrightarrow{f} & V \end{array}$$

Since  $f_\lambda(v_0) = 0$ , by Prop. 9,  $\ker f_\lambda = V$ , (for  $\ker f_\lambda \neq 0$  and so  $\ker f_\lambda = V$ )  
 By rank-nullity theorem,  $\dim V = \dim \ker f_\lambda + \dim \text{im} f_\lambda$ .  
 So  $\text{im} f_\lambda = 0$ , and so  $f_\lambda(v) = 0$  ( $\forall v \in V$ )  $\implies f(v) = \lambda v$

Schur's lemma, at least the first part, implies that the left  $kG$ -modules associated with representations  $\rho, \sigma$  are  $kG$ -isomorphic, i.e.

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \rho_g \downarrow & & \downarrow \sigma_g \\ V & \xrightarrow{f} & W \end{array} \iff V^\rho \stackrel{f}{\cong} V^\sigma$$

with  $f$  being an isomorphism between  $V^\rho$  and  $V^\sigma$  s.t.

$$\begin{aligned} f(v+w) &= f(v) + f(w) \quad \forall v, w \in (V^\sigma, +) \\ f(rv) &= rf(v) \quad \forall r = \sum_{g \in G} a_g g \in kG \end{aligned}$$

Kosmann-Schwarzbach's **Groups and Symmetries**[14] is a very lucid text that's mathematically rigorous enough and practical for physicists. It's really good and very clear. Let's follow its development for  $SU(2), SO(3), SL(2, \mathbb{C})$  and corresponding Lie algebras  $\mathfrak{su}(2), \mathfrak{so}(3), \mathfrak{sl}(2, \mathbb{C})$ .

From Chapter 2 "Representations of Finite Groups" of Kosmann-Schwarzbach (2010) [14]

□ **Definition 21** (2.1 Kosmann-Schwarzbach (2010)[14]). *On  $L^2(G)$ , scalar product defined by*

$$\langle f_1 | f_2 \rangle = \frac{1}{|G|} \sum_{g \in G} \overline{f_1(g)} f_2(g)$$

$f_1, f_2 \in \mathcal{F}(G) \equiv \mathbb{C}[G]$  vector space of functions on  $G$  taking values on  $\mathbb{C}$

**Definition 22** (2.3 Kosmann-Schwarzbach (2010)[14]). *Let  $(E, \rho)$  be representation of  $G$*

character of  $\rho \equiv \chi_\rho : G \rightarrow \mathbb{C}$

$$\chi_\rho(g) = \text{tr}(\rho(g)) = \sum_{i=1}^n (\rho(g))_{ii}$$

*Note: equivalent representations have same character  
 each conjugacy class of  $G$ , function  $\chi_p$  is constant*

Looking at Def. 21

$$\langle \chi_{\rho_1} | \chi_{\rho_2} \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_{\rho_1}(g^{-1}) \chi_{\rho_2}(g)$$

since  $\overline{\chi_{\rho_1}(g)} = \chi_{\rho_1}(g^{-1})$  by unitarity of representation with respect to scalar product  $\langle, \rangle$

**Proposition 10** (2.7 Kosmann-Schwarzbach (2010)[14]). *Let  $(E_1, \rho_1)$  be representations of  $G$ , let linear  $u : E_1 \rightarrow E_2$ .  
 $(E_2, \rho_2)$*

*Then  $\exists$  linear  $T_u$  s.t.*

$$(33) \quad \begin{aligned} T_u &: E_1 \rightarrow E_2 \\ T_u &= \frac{1}{|G|} \sum_{g \in G} \rho_2(g) u \rho_1(g)^{-1} \end{aligned}$$

□

*so that  $\rho_2(g) T_u = T_u \rho_1(g) \quad \forall g \in G$*

*Proof.*

$$\rho_2(g) T_u = \frac{1}{|G|} \sum_{h \in G} \rho_2(gh) u \rho_1(h^{-1}) = \frac{1}{|G|} \sum_{k \in G} \rho_2(k) u \rho_1(k^{-1}g) = T_u \rho_1(g)$$

□

Thus, diagrammatically, we have that

$$E_1 \xrightarrow{u} E_2 \implies \begin{array}{ccc} E_1 & \xrightarrow{T_u} & E_2 \\ \rho_1(g) \downarrow & & \downarrow \rho_2(g) \\ E_1 & \xrightarrow{T_u} & E_2 \end{array}$$



From Definition 1.12 of Kosmann-Schwarzbach [14], “representations  $\rho_1$  and  $\rho_2$  are called **equivalent** if there is a bijective intertwining operator for  $\rho_1$  and  $\rho_2$ .” So I will interpret this as if an intertwining operator is not bijective, then the representations  $\rho_1, \rho_2$  are not equivalent.

**Proposition 11** (2.8 Kosmann-Schwarzbach (2010)[14]). *Let  $(E_1, \rho_1)$  be irreducible representations of  $G$ , let linear  $u : E_1 \rightarrow E_2$ ,  $(E_2, \rho_2)$*

*define  $T_u$  by  $T_u = \frac{1}{|G|} \sum_{g \in G} \rho_2(g) u \rho_1(g)^{-1}$  by Eq. 33.*

- (i) *If  $\rho_1, \rho_2$  inequivalent, then  $T_u = 0$*
- (ii) *If  $E_1 = E_2 = E$  and  $\rho_1 = \rho_2 = \rho$ , then*

$$T_u = \frac{\text{tr}(u)}{\dim E} 1_E$$

*Proof.* (i) if  $\rho_1, \rho_2$  are inequivalent, by definition,  $T_u$  is not isomorphic. Then by Schur’s lemma (first part),  $T_u = 0$   
(ii) By Schur’s lemma,  $T_u(v) = \lambda v \quad \forall v \in E = E_1 = E_2$ . So  $T_u = \lambda 1_E$ .  $\text{tr} T_u = \lambda \dim E$  or  $\lambda = \frac{\text{tr} T_u}{\dim E}$ . Thus,  $T_u = \frac{\text{tr} T_u}{\dim E} 1_E$   $\square$

Let  $(e_1 \dots e_n)$  basis of  $E$

$(f_1 \dots f_p)$  basis of  $F$

$$\forall u \in \mathcal{L}(E, F), \quad \begin{aligned} u : E &\rightarrow F \\ u(x) &= u(x^j e_j) = x^j u(e_j) = x^j u^i_j f_i \quad \text{for } x = x^j e_j \in E \\ u &= u^i_j e^j \otimes f_i \quad y = y^i f_i \in F \end{aligned}$$

For

$$\begin{aligned} T : E^* \otimes F &\rightarrow \mathcal{L}(E, F) \\ T(\xi \otimes y) &= u^i_j e^j \otimes f_i \text{ i.e. set } T(\xi \otimes y) \text{ to this } u \\ T(\xi \otimes y) &= T(\xi e^l \otimes y^k f_k) = \xi_l y^k T(e^l \otimes f_k) = (\xi_l y^k T_{kj}^{li}) e^j \otimes f_i \implies \xi_l y^k T_{kj}^{li} = u^i_j \end{aligned}$$

*Exercises.* Exercises of Ch. 2 Representations of Finite Groups [14]

**Exercise 2.6.** [14] *The dual representation.*

Let  $(E, \pi)$  representation of group  $G$ .

$\forall g \in G, \xi \in E^*, x \in E$ , set  $\langle \pi^*(g)(\xi), x \rangle = \langle \xi, \pi(g^{-1})(x) \rangle$

- (a) *dual* (or *contragredient*) of  $\pi, \pi^* : G \rightarrow \text{End}(E^*)$ ,  $\pi^*$  is a representation, since

$$\begin{aligned} \langle \pi^*(gh)(\xi), x \rangle &= \langle \xi, \pi((gh)^{-1})(x) \rangle = \langle \xi, \pi(h^{-1}g^{-1})(x) \rangle = \langle \xi, \pi(h^{-1})\pi(g^{-1})(x) \rangle = \langle \xi, \pi(h^{-1})(\pi(g^{-1})(x)) \rangle = \\ &= \langle \pi^*(h)(\xi), \pi(g^{-1})(x) \rangle = \langle \pi^*(g)\pi^*(h)(\xi), x \rangle \end{aligned}$$

since this is true,  $\forall x \in E, \forall \xi \in E^*, \pi^*(gh) = \pi^*(g)\pi^*(h)$ .

dual  $\pi^*$  of  $\pi$  is a representation.

- (b) Consider  $G \times \mathcal{L}(E, F) \rightarrow \mathcal{L}(E, F)$ .

$$g \cdot u = \rho(g) \circ u \circ \pi(g^{-1})$$

Define

$$\begin{aligned} \sigma : G &\rightarrow \text{End}(\mathcal{L}(E, F)) \\ \sigma(g) : \mathcal{L}(E, F) &\rightarrow \mathcal{L}(E, F) \\ \sigma(g)(u) &= \rho(g) \circ u \circ \pi(g^{-1}) \end{aligned}$$

Let  $(e_1 \dots e_n)$  be a basis of  $E$ . Let  $\xi = \xi_i e^i \in E^*, x = x^j e_j \in E$ .

Consider the isomorphism  $T : E^* \otimes F \rightarrow \mathcal{L}(E, F)$  defined as<sup>2</sup>

$$T : E^* \otimes F \rightarrow \mathcal{L}(E, F) = \text{Hom}(E, F)$$

$$\xi \otimes y \mapsto (x \mapsto \xi(x)y)$$

Choose bases  $(e_1 \dots e_n)$  of  $E$   
 $(e^1 \dots e^n)$  of  $E^*$ . Then  
 $(f_1 \dots f_p)$  of  $F$

$$T(e^j \otimes f_i)(x) = T(e^j \otimes f_i)(x^k e_k) = \delta_k^j x^k f_i = x^j f_i$$

$$T(e^j \otimes f_i)(e_k) = \delta_k^j f_i$$

Consider

$$u \in \mathcal{L}(E, F)$$

$$u : E \rightarrow F$$

$$u(x) = u(x^j e_j) = x^j u(e_j) = x^j u^i_j f_i$$

$$u(e_j) = u^i_j f_i \text{ i.e. } u : e_j \rightarrow u^i_j f_i$$

Then  $\forall u \in \mathcal{L}(E, F)$ ,

$$T(u^i_j e^j \otimes f_i)(e_k) = u^i_j \delta_k^j f_i = u^i_k f_i = u(e_k) \implies u = T(u^i_j e^j \otimes f_i)$$

so  $T$  is surjective.

With  $T(\xi \otimes y) = T(\xi' \otimes y')$ ,

$$T(\xi \otimes y)(x) = T(\xi' \otimes y')(x)$$

$$\xi(x)y = \xi'(x)y' \implies \xi(x)y - \xi'(x)y' = 0$$

which implies that  $\xi \otimes y = \xi' \otimes y'$ . So  $T$  is injective. Or, one could consider that  $T^{-1} : \mathcal{L}(E, F) \rightarrow E^* \otimes F$ ,  $T^{-1} : u \mapsto u^i_j e^j \otimes f_i$ , which is the inverse of  $T$ .

**Remark 1.**

$$E^* \otimes F \xrightarrow{T} \mathcal{L}(E, F) = \text{Hom}(E, F)$$

$$(\xi, y) \mapsto (x \mapsto \xi(x)y)$$

and so  $(e^j \otimes f_i) \mapsto (x \mapsto e^j(x)f_i = x^j f_i)$

So  $E^* \otimes F$  is isomorphic to  $\mathcal{L}(E, F) = \text{Hom}(E, F)$

For representation  $\pi$ ,

$$\pi : G \rightarrow \text{End}(E)$$

$$\pi(g) : E \rightarrow E$$

$$\pi(g)(x) = \pi(g)(x^j e_j) = x^j \pi(g)(e_j) = x^j \pi(g)^i_j e_i = (\pi(g)^i_j x^j e_i$$

Consider this matrix formulation:

$$\begin{aligned} \pi^*(g)(\xi) &= \pi^*(g)(\xi_i e^i) = \xi_i \pi^*(g)(e^i) = \xi_i (\pi^*(g))^i_j e^j \\ &\implies \langle \pi^*(g)(\xi), x \rangle = \xi_i (\pi^*(g))^i_j x^j \end{aligned}$$

and

$$\langle \xi, \pi(g^{-1})(x) \rangle = \xi_i \pi(g^{-1})^i_j x^j$$

<sup>2</sup>Mathematics stackexchange Isomorphism between Hom and tensor product [duplicate] <http://math.stackexchange.com/questions/428185/isomorphism-between-hom-and-tensor-product>  
<http://math.stackexchange.com/questions/57189/understanding-isomorphic-equivalences-of-tensor-product>

so that

$$\langle \pi^*(g)(\xi), x \rangle = \langle \xi, \pi(g^{-1})(x) \rangle \implies \pi(g^{-1})^i_j = (\pi^*(g))^i_j$$

Thus, given a choice of basis for  $E$ , the *dual* of  $\pi$ ,  $\pi^*(g)^i_j$ , and  $\pi(g^{-1})^i_j$  are formally equal.

So for a choice of basis of  $E$  and of  $F$ ,

$$(\pi^* \otimes \rho)(g)(\xi, y) = (\pi^*(g) \otimes \rho(g))(\xi, y) = \pi^*(g)\xi \otimes \rho(g)y = \xi_l \pi(g^{-1})^l_j e^j \otimes \rho(g)^i_k y^k f_i = \rho(g)^i_k y^k \xi_l \pi(g^{-1})^l_j e^j \otimes f_i$$

Applying  $T$ ,

$$T(\pi^* \otimes \rho)(g)(\xi, \rho) = \rho(g)^i_k y^k \xi_l \pi(g^{-1})^l_j = \rho(g)T(\xi, y)\pi(g^{-1})$$

$$\begin{array}{ccc} E^* \otimes F & \xrightarrow{T} & \mathcal{L}(E, F) \\ \downarrow (\pi^* \otimes \rho)(g) & & \downarrow \sigma(g) \\ E^* \otimes F & \xrightarrow{T} & \mathcal{L}(E, F) \end{array} \quad \begin{array}{ccc} (\xi, y) & \xrightarrow{T} & (x \mapsto \xi(x)y) = y^i \xi_j \\ \downarrow (\pi^* \otimes \rho)(g) & & \downarrow \sigma(g) \\ \pi^*(g)(\xi) \otimes \rho(g)y & \xrightarrow{T} & \rho(g)y^i \xi_j \pi(g^{-1}) = \rho(g)T(\xi, y)\pi(g^{-1}) \end{array}$$

Thus

Thus, representation  $\sigma(g)$  is equivalent to representation  $(\pi^* \otimes \rho)$ , a tensor product of representations.

**Exercise 2.15.** *Representation of  $GL(2, \mathbb{C})$  on the polynomials of degree 2*

Let group  $G$ , let representation  $\rho$  of  $G$  on  $V = \mathbb{C}^n$ , i.e.  $\rho : G \rightarrow \text{End}(V)$

Let  $P^{(k)}(V)$  vector space of complex polynomials on  $V$  that are homogeneous of degree  $k$ .

For  $f \in P^{(k)}(V)$ , the general form is

$$f = \sum_{\substack{i_1+i_2+\dots+i_n=k \\ 0 \leq i_j \leq k}} a_{i_1 i_2 \dots i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$$

Given

$$\binom{n+k}{k} = \binom{k-1}{k-1} + \binom{k}{k-1} + \dots + \binom{n+k-1}{k-1} = \sum_{i=0}^n \binom{k-1+i}{k-1}$$

$\binom{k+n-1}{n-1}$  is number of monomials of degree  $k$ .

So  $\dim P^{(k)}(V) = \binom{k+n-1}{n-1}$ . This is a very lucid and elementary exposition on the basics of polynomials which I found was useful for the basic facts I forgot<sup>3</sup>.

So we have the graded algebra

$$P(V) = \bigoplus_{k=0}^{\infty} P^{(k)}(V)$$

$$\rho^{(k)} : G \rightarrow \text{End}(P^{(k)}(V))$$

$$\rho^{(k)}(g) : P^{(k)}(V) \rightarrow P^{(k)}(V)$$

$$\rho^{(k)}(g)(f) = f \circ \rho(g^{-1})$$

This is a representation of  $G$  since

(a)

$$\begin{aligned} \rho^{(k)}(gh)(f) &= f \circ \rho(gh)^{-1} = f \circ \rho(h^{-1}g^{-1}) = f \circ \rho(h^{-1})\rho(g^{-1}) \\ \rho^{(k)}(g)\rho^{(k)}(h)(f) &= \rho^{(k)}(g)(f \circ \rho(h^{-1})) = f \circ \rho(h^{-1}) \circ \rho(g^{-1}) \end{aligned} \implies \rho^{(k)}(gh) = \rho^{(k)}(g)\rho^{(k)}(h)$$

(b) Choose basis  $(e_1 \dots e_n)$  of  $V$ ,  $x = x^j e_j \in V$ ,  $\rho : G \rightarrow \text{End}(V)$ , and so  $\rho(g)(x) = \rho(g)(x^j e_j) = x^j \rho(g)(e_j) = x^j (\rho(g))^i_j e_i$ .

With  $\xi(e_i) = \xi_i \implies \langle \xi, \rho(g^{-1})x \rangle = \xi_i x^j (\rho(g^{-1}))^i_j$

$\forall \xi \in V^*$ ,  $\xi = \xi_i e^i$ ,

$$\rho^*(g)(\xi) = \rho^*(g)(\xi_i e^i) = \xi_i \rho^*(g)^i_j e^j$$

$$\implies \langle \rho^*(g)(\xi), x \rangle = \xi_i x^j (\rho^*(g))^i_j \implies (\rho^*(g))^i_j = (\rho(g^{-1}))^i_j$$

So  $\forall f \in P^{(1)}(V)$ ,  $x \in V$ ,  $\rho(g^{-1})x = x^j (\rho(g^{-1}))^i_j e_i$ . So  $f \circ \rho(g^{-1})(x) = \sum_{i=1}^n a_i (\rho(g^{-1}))^i_j x^j = \sum_{i=1}^n a_i (\rho^*(g))^i_j x^j$

$$\implies \rho^{(1)}(g)(f) = f \circ \rho^*(g)$$

(c) Suppose  $G = GL(2, \mathbb{C})$ ,  $V = \mathbb{C}^2$ ,  $\rho$  fundamental representation  $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ,  $g^{-1} = \frac{1}{\det g} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$  for  $\det g = ad - bc$ .

Let  $k = 2$ ,  $\dim P^{(2)}(\mathbb{C}^2) = \binom{2+2-1}{2-1} = \binom{3}{1} = 3$

$\forall f \in P^{(2)}(\mathbb{C}^2)$ ,  $f(x, y) = Ax^2 + 2Bxy + Cy^2$

Let

$$P^{(2)}(\mathbb{C}^2) \rightarrow \mathbb{C}^3$$

$$f(x, y) = Ax^2 + 2Bxy + Cy^2 \mapsto \begin{pmatrix} A \\ B \\ C \end{pmatrix} \in \mathbb{C}^3$$

Call this transformation  $T$ ,  $T : P^{(2)}(\mathbb{C}^2) \rightarrow \mathbb{C}^3$ .

$\forall \begin{pmatrix} A \\ B \\ C \end{pmatrix} \in \mathbb{C}^3$ ,  $f(x, y) = Ax^2 + 2Bxy + Cy^2$  and  $Tf(x, y) = \begin{pmatrix} A \\ B \\ C \end{pmatrix}$ .  $T$  surjective.

Suppose  $Tf(x, y) = Tf'(x, y)$ ,

$$\implies Ax^2 + 2Bxy + Cy^2 = A'x^2 + 2B'xy + C'y^2$$

$$\implies (A - A')x^2 + 2(B - B')xy + (C - C')y^2 = 0$$

Then since the monomials form a basis, and its basis elements are independent (by definition), then  $A = A'$ ,  $B = B'$ ,  $C = C'$ .  $T$  injective. So  $T$  is bijective, an isomorphism.

(This is all in `groups.sage`)

```
sage: P2CC.<x,y> = PolynomialRing(CC,2) # this declares a PolynomialRing of field of complex numbers,
# of order 2 (i.e. only 2 variables for a polynomial, such as x, y)
sage: A = var('A')
sage: assume(A, 'complex')
sage: B = var('B')
sage: assume(B, 'complex')
sage: C = var('C')
sage: assume(C, 'complex')
sage: f(x,y) = A*x**2 + 2*B*x*y + C*y**2
```

```
sage: a = var('a')
sage: assume(a, 'complex')
sage: b = var('b')
sage: assume(b, 'complex')
sage: c = var('c')
sage: assume(c, 'complex')
sage: d = var('d')
sage: assume(d, 'complex')
sage: g = Matrix([[a,b],[c,d]])
sage: X = Matrix([[x],[y]])
sage: f((g.inverse()*X)[0,0], (g.inverse()*X)[1,0]).expand()
sage: f((g.inverse()*X)[0,0], (g.inverse()*X)[1,0]).expand().coefficient(x^2).full_simplify()
(C*c^2 - 2*B*c*d + A*d^2)/(b^2*c^2 - 2*a*b*c*d + a^2*d^2)
sage: f((g.inverse()*X)[0,0], (g.inverse()*X)[1,0]).expand().coefficient(x*y).full_simplify()
-2*(C*a*c + A*b*d - (b*c + a*d)*B)/(b^2*c^2 - 2*a*b*c*d + a^2*d^2)
```

<sup>3</sup>Polynomials. Math 4800/6080 Project Course <http://www.math.utah.edu/~bertram/4800/PolyIntroduction.pdf>

```
sage: f((g.inverse()*X)[0,0],(g.inverse()*X)[1,0]).expand().coefficient(y^2).full_simplify()
(C*a^2 - 2*B*a*b + A*b^2)/(b^2*c^2 - 2*a*b*c*d + a^2*d^2)
```

So

$$\begin{aligned} \rho^{(2)}(g)(f)(x,y) &= f \circ \rho(g^{-1})(x,y) = \\ &= \frac{Cc^2 - 2Bcd + Ad^2}{(ad - bc)^2}x^2 + -2\frac{(Cac + Abd - (bc + ad)B)}{(ad - bc)^2}xy + \frac{Ca^2 - 2Bab + Ab^2}{(ad - bc)^2}y^2 \end{aligned}$$

So define  $\tilde{\rho} : G \rightarrow \text{End}(\mathbb{C}^3)$ .  $\tilde{\rho}$  is a representation, for

$$\forall v = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \in \mathbb{C}^3, \quad \tilde{\rho}(gh)(v) = T \circ f \circ \rho((gh)^{-1}) = T \circ f \circ \rho(h^{-1}g^{-1}) = T \circ f \circ \rho(h^{-1})\rho(g^{-1})$$

$$\text{Now } \tilde{\rho}(h)(v) = T \circ f \circ \rho(h^{-1})$$

$$\implies \tilde{\rho}(g)\tilde{\rho}(h)(v) = T \circ (f \circ \rho(h^{-1})) \circ \rho(g^{-1}) = T \circ f \circ \rho(h^{-1})\rho(g^{-1}) \text{ and so}$$

$$\tilde{\rho}(gh) = \tilde{\rho}(g)\tilde{\rho}(h)$$

And so

$$\tilde{\rho}^*(g)(v) = Tf\rho(g^{-1})$$

and consider this commutation diagram, that (helped me at least and) clarifies the relationships:

$$\begin{array}{ccc} P^{(2)}(\mathbb{C}^2) & \xrightarrow{T} & \mathbb{C}^3 \\ \rho^{(2)}(g) \downarrow & & \downarrow \tilde{\rho}(g) \\ P^{(2)}(\mathbb{C}^2) & \xrightarrow{T} & \mathbb{C}^3 \end{array} \quad \begin{array}{ccc} f & \xrightarrow{T} & \begin{pmatrix} A \\ B \\ C \end{pmatrix} \\ \rho^{(2)}(g) \downarrow & & \downarrow \tilde{\rho}(g) \\ f \circ \rho(g^{-1}) & \xrightarrow{T} & \begin{pmatrix} D \\ E \\ F \end{pmatrix} \end{array}$$

with

$$\begin{pmatrix} D \\ E \\ F \end{pmatrix} = \begin{pmatrix} \frac{Cc^2 - 2Bcd + Ad^2}{(ad - bc)^2} \\ -2\frac{(Cac + Abd - (bc + ad)B)}{(ad - bc)^2} \\ \frac{Ca^2 - 2Bab + Ab^2}{(ad - bc)^2} \end{pmatrix}$$

Now define the dual  $\tilde{\rho}^*$  as such:

$$\tilde{\rho}^*(g) : (\mathbb{C}^3)^* \rightarrow (\mathbb{C}^3)^*$$

$$\tilde{\rho}^*(g) = \tilde{\rho}(g^{-1})$$

$$\forall \xi \in (\mathbb{C}^3)^*$$

$$\tilde{\rho}^*(g)\xi = \xi_i(\tilde{\rho}^*(g))^i_j e^j = \xi_i(\tilde{\rho}(g^{-1}))^i_j e^j$$

$$\text{So for } v = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \in \mathbb{C}^3, f = T^{-1}v = Ax^2 + 2Bxy + Cy^2 \in P^2(\mathbb{C}^2),$$

$$\tilde{\rho}(g^{-1})(v) = T \circ (f\rho(g)) = \begin{bmatrix} Aa^2 + 2Bac + Cc^2 \\ Aab + Bbc + Bad + Ccd \\ Ab^2 + 2Bbd + Cd^2 \end{bmatrix}$$

which was found using Sage Math:

```
sage: f((g*X)[0,0],(g*X)[1,0])
(a*x + b*y)^2*A + 2*(a*x + b*y)*(c*x + d*y)*B + (c*x + d*y)^2*C
sage: f((g*X)[0,0],(g*X)[1,0]).expand()
A*a^2*x^2 + 2*B*a*c*x^2 + C*c^2*x^2 + 2*A*a*b*x*y + 2*B*b*c*x*y + 2*B*a*d*x*y + 2*C*c*d*x*y + A*b^2*y^2 + 2*B*b*d*y^2 + C*d^2*y^2
sage: f((g*X)[0,0],(g*X)[1,0]).expand().coefficient(x^2)
A*a^2 + 2*B*a*c + C*c^2
sage: f((g*X)[0,0],(g*X)[1,0]).expand().coefficient(x*y)
2*A*a*b + 2*B*b*c + 2*B*a*d + 2*C*c*d
sage: f((g*X)[0,0],(g*X)[1,0]).expand().coefficient(y^2)
A*b^2 + 2*B*b*d + C*d^2
```

or

```
sage: T(f((g*X)[0,0],(g*X)[1,0]).expand())
[A*a^2 + 2*B*a*c + C*c^2,
 2*A*a*b + 2*B*b*c + 2*B*a*d + 2*C*c*d,
 A*b^2 + 2*B*b*d + C*d^2]
```

So then

$$\tilde{\rho}(g^{-1}) = \begin{bmatrix} a^2 & 2ac & c^2 \\ 2ab & 2(ad + bc) & 2cd \\ b^2 & 2bd & d^2 \end{bmatrix}$$

So then

$$\tilde{\rho}^*(g) = \begin{bmatrix} a^2 & 2ac & c^2 \\ 2ab & 2(ad + bc) & 2cd \\ b^2 & 2bd & d^2 \end{bmatrix}$$

and operate on row vectors  $\xi \in (\mathbb{C}^3)^*$  with  $\tilde{\rho}^*(g)$  from the row vector's right.

More: Let  $G = SU(2)$ . Then  $U = e^{i\phi} \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$

$$\tilde{\rho} : SU(2) \rightarrow \text{End}(\mathbb{C}^3)$$

$$\tilde{\rho}(U) : \mathbb{C}^3 \rightarrow \mathbb{C}^3$$

$$\tilde{\rho}(U)(v) = e^{-2i\varphi} \begin{bmatrix} A\bar{a}^2 + 2B\bar{a}\bar{b} + C\bar{b}^2 \\ -A\bar{a}\bar{b} + B + C\bar{a}\bar{b} \\ Ab^2 - 2Bab + Ca^2 \end{bmatrix}$$

$$\implies \tilde{\rho}(U) = e^{-2i\varphi} \begin{bmatrix} -\bar{a}^2 & 2\bar{a}\bar{b} & \bar{b}^2 \\ -\bar{a}\bar{b} & 1 & \bar{a}\bar{b} \\ b^2 & -2ab & a^2 \end{bmatrix}$$

cf. Ch. 5 Lie Groups of Jeffrey Lee (2009) [4]

**Definition 23** (Lie Group). ***Lie Group**  $G$  := smooth manifold  $G$  is a **Lie Group** if  $G$  is a group (abstract group), s.t.*

*multiplication map  $\mu : G \times G \rightarrow G$*

$$\mu(g,h) = gh$$

*inverse map  $\text{inv} : G \rightarrow G$*

$$\text{inv}(g) = g^{-1}$$

*are  $C^\infty$  maps.*

*If group is abelian, use additive notation  $g + h$  for group operation.*

**Definition 24** ( $GL(n, \mathbb{R})$ ).  *$GL(n, \mathbb{R})$  := group of all invertible real  $n \times n$  matrices.*

*global chart on  $GL(n, \mathbb{R}) = \{x_j^i\}$ ,  $n^2$  functions  $x_j^i$ , where if  $A \in GL(n, \mathbb{R})$ , then  $x_j^i(A)$  is  $ij$ th entry of  $A$ .*

*Claim:  $GL(n, \mathbb{R})$  is a Lie group.*

*Proof.* multiplication is clearly smooth:  $(AB)_{ij} = A_{ik}B_{kj}$ ,

$$\frac{\partial}{\partial x_m^i}(x_k^i(A)x_j^k(B)) = \delta_l^i \delta_k^m x_j^k(B) + x_k^i(A) \delta_l^k \delta_j^m$$

inversion map; appeal to formula for  $A^{-1}$ ,  $A^{-1} = \text{adj}(A)/\det(A)$ ,  $\text{adj}(A) \equiv$  adjoint matrix (whose entries are cofactors).  
 $\implies A^{-1}$  depends smoothly on entries of  $A$ .

Similarly,  $GL(n, \mathbb{C})$ , group of invertible  $n \times n$  complex matrices, is a Lie group.

**Exercise 5.5.** Let subgroup  $H$  of  $G$ , consider cosets  $gH$ ,  $g \in G$ .

Recall  $G$  is disjoint union of cosets of  $H$ .

*Claim:* if  $H$  open, so are all its cosets. And  $H$  closed.

*Proof.* cf. [stackexchange: Open subgroups of a topological group are closed](#)

$gH = \{gh|h \in H\}$  is an open neighborhood of  $g$  (since  $1 \in H$ , and mapping  $h \mapsto gh$  sends open sets to open sets, since its inverse,  $gh \mapsto h$ , is  $C^\infty$  (so continuous)).

$$\begin{array}{ll} gH \rightarrow H & H \rightarrow gH \\ gh \xrightarrow{g^{-1}} h = \mu(g^{-1}, gh) & h \xrightarrow{g} gh = \mu(g, h) \end{array}$$

Then  $\forall$  coset  $gH$ ,  $gH$  is open.

Suppose  $g' \in H^c \equiv G - H \equiv G \setminus H$ .

Consider  $h \in H$ , if  $g'h \in H$ , then  $g' = (g'h)h^{-1} \in H$  (recall  $h^{-1} \in H$ , and  $H$  is a subgroup).

Contradiction.

$\implies \forall g' \in H^c$ ,  $\exists$  open neighborhood  $g'H \subset H^c$ , so  $H^c$  open (by definition). Then  $H$  closed.

cf. Thm. 5.6 in Jeffrey Lee (2009) [4].

**Theorem 15.** *If  $G$  connected Lie group,  $U$  neighborhood of identity element  $e$ , then  $U$  generates the group, i.e.  $\forall g \in G$ ,  $g$  is a product of elements of  $U$ .*

*Proof.* Note  $V = \text{inv}(U) \cap U$  is an open neighborhood of  $e$ . Note  $\text{inv}(V) = V$ .  $\text{inv}(V) \equiv V^{-1} = \{V^{-1}|v \in V\}$ . We say that  $V$  is *symmetric*.

*Claim:*  $V$  generates  $G$ .

$\forall$  open  $W_1$ , open  $W_2 \subset G$ ,

$W_1W_2 = \{w_1w_2|w_1 \in W_1, w_2 \in W_2\}$  is an open set being a union of open sets  $\bigcup_{g \in W_1} gW_2$ .

Thus, inductively defined sets

$$V^n = VV^{n-1}, \quad n = 1, 2, 3, \dots$$

are open.

$$e \in V \subset V^2 \subset \dots V^n \subset \dots$$

It's easy to check that each  $V^n$  is symmetric.

$$\text{inv}(V) = V$$

$$\text{inv}(V^2) = \text{inv}\left(\bigcup_{v \in V} vV\right) = V\text{inv}(V) = V = V^2$$

$$\text{inv}(V^{n+1}) = \text{inv}\left(\bigcup_{v \in V} vV^n\right) = V\text{inv}(V^n) = VV^n = V^{n+1}$$

so  $V^\infty := \bigcup_{n=1}^\infty V^n$  is symmetric.

$V^\infty$  closed under inversion, also multiplication. Thus  $V^\infty$  is an open subgroup.

From Exercise 5.5, Jeffrey Lee (2009) [4], i.e. Exercise **11**,  $V^\infty$  also closed, since  $G$  is connected,  $V^\infty = G$ . (a topological space  $X$  is **connected** iff the only open and closed (clopen) sets are  $\emptyset$  and  $X$ ).

□

**Definition 25.** *Identity component of  $G, G_0$ .*

$G_0 :=$  connected component of Lie group  $G$  that contains identity;

□  $G_0$  is a Lie group, and is generated by any open neighborhood of the identity.

**Definition 26.** *For Lie group  $G$ , fixed element  $g \in G$ ,*

*left translation (by  $g$ )  $L_g : G \rightarrow G$ ,  $L_g x = gx$ ,*  $\forall x \in G$

*right translation (by  $g$ )  $R_g : G \rightarrow G$ ,  $R_g x = xg$ ,*  $\forall x \in G$

$L_g, R_g$  are diffeomorphisms with  $L_g^{-1} = L_{g^{-1}}, R_g^{-1} = R_{g^{-1}}$ .

**Definition 27** (Product Lie group). *If  $G, H$  are Lie groups, then product manifold  $G \times H$  is a Lie group, where multiplication*

$$(g_1, h_1) \cdot (g_2, h_2) = (g_1g_2, h_1h_2)$$

*Lie group  $G \times H$  is called **product Lie group***

e.g. product group  $S^1 \times S^1 \equiv$  2-torus group.

Generally, higher torus groups  $T^n = S^1 \times \dots \times S^1$  ( $n$  factors).

**Definition 28** (Lie subgroup of  $G, H$ ). *Let  $H$  be an abstract subgroup of Lie group  $G$ .*

*If  $H$  is a Lie group s.t. inclusion map  $i : H \rightarrow G \equiv H \hookrightarrow G$  is an immersion, then  $H$  is a*

□ ***Lie subgroup*** of  $G$ .

Recall  $i : H \rightarrow G$  immersion iff  $Di$  injective, i.e. iff  $\text{rank} Di = \dim H$

cf. Prop. 5.9 in Jeffrey Lee (2009) [4].

**Proposition 12.** *If  $H$  abstract subgroup of Lie group  $G$ , that's also a regular submanifold  $\equiv$  embedded submanifold, then  $H$  closed Lie subgroup.*

*Recall that*

$$\text{embedded submanifold} \equiv \text{regular submanifold}$$

Each name is used frequently and we shouldn't be biased against one or the other; we'll have to refer to both, to emphasize they're *exactly the same*.

embedded submanifold  $\equiv$  regular submanifold is an immersed submanifold s.t. inclusion map  $i$  is a topological embedding,

i.e. embedded submanifold  $\equiv$  regular submanifold  $S \subset M$ ,

immersed submanifold  $S$  if  $i : S \rightarrow M \equiv S \hookrightarrow M$  is an immersion, i.e.  $Di$  injective, i.e.  $\text{rank} Di \equiv \dim S$ .

topological embedding  $:=$  homeomorphism onto its image, i.e.

injective cont. map  $f : X \rightarrow Y$ ,  $X, Y$  topological spaces, is a **topological embedding**

if  $f$  is a homeomorphism between  $X$  and  $f(X)$ .

$f$  homeomorphism is a bijection, continuous, and  $f^{-1}$  continuous.

e.g.  $\forall$  embedding  $f : M \rightarrow N$ ,  $f(M) \subset N$  naturally has the structure of an embedding submanifold  $\equiv$  regular submanifold.

*Useful, intrinsic definition* of **embedded submanifold**  $\equiv$  regular submanifold.

Let manifold  $M$ ,  $\dim M = n$ , let  $k \in \mathbb{Z}^+$ , s.t.  $0 \leq k \leq n$ .

A  $k$ -dim. embedded submanifold  $\equiv$  regular submanifold  $S$  is subset  $S \subset M$  s.t.  $\forall p \in S$ ,  $\exists$  chart  $(U \subset M, \varphi : U \rightarrow \mathbb{R}^n \ni 0)$ , s.t.  $\varphi(S \cap U)$  is the intersection of a  $k$ -dim. plane with  $\varphi(U)$ .

(pairs  $(S \cap U, \varphi|_{S \cap U})$  form an atlas for differential structure on  $S$ ).

Proof 1:

*Proof.*  $H$  subgroup of  $G$ , so  
multiplication map  $H \times H \rightarrow H$   
inversion map  $H \rightarrow H$   
are restrictions of multiplication and inversion maps on  $G$ .

Since  $H$  regular submanifold, maps are smooth.  
Recall  $H$  regular submanifold iff  $H$  immersive submanifold (i.e.  $H \hookrightarrow G$  is an immersion) and  $H$  topological subspace of  $G$ , i.e. submanifold topology on  $H$  is same as subspace topology.

Claim:  $H$  closed.  
Let  $x_0 \in \overline{H}$   
Let  $(U, x)$  be a chart adapted to  $H$ , whose domain contains  $e$ .  
Let

$$\delta : G \times G \rightarrow G$$

$$\delta(g_1, g_2) = g_1^{-1}g_2$$

Choose open set  $V$  s.t.  $e \in V \subset \overline{V} \subset U$ .  
By continuity map  $\delta$ , find open neighborhood  $O$  of identity  $e$  s.t.  $O \times O \subset \delta^{-1}(V)$   
If  $\{h_i\}$  sequence in  $H$  converging to  $x_0 \in \overline{H}$ , then  $x_0^{-1}h_i \rightarrow e$  and  $x_0^{-1}h_i \in O$  for all sufficiently large  $i$ .  
Since  $h_j^{-1}h_i = (x_0^{-1}h_j)^{-1}x_0^{-1}h_i$ ,  $h_j^{-1}h_i \in V$  for sufficiently large  $i, j$ .

For any sufficiently large fixed  $j$ ,

$$\lim_{i \rightarrow 0} h_j^{-1}h_i = h_j^{-1}x_0 \in \overline{V} \subset U$$

Since  $U$  is domain of a single-slice chart,  $U \cap H$  closed in  $U$ .

Thus, since  $\forall h_j^{-1}h_i \in U \cap H$ ,  $h_j^{-1}x_0 \in U \cap H \subset H$ ,  $\forall$  sufficiently large  $j$ .  
 $\implies x_0 \in H$ , and since  $x_0$  arbitrary, done.

Proof 2:  
cf. 9.2 The Closed Subgroup Theorem I of 427 Notes<sup>4</sup>

*Proof.* **Claim:** Since  $H$  is an embedded submanifold  $\equiv$  regular submanifold,  $\exists$  neighborhood  $U$  of  $1$ ,  $1 \in G$ , s.t.  $U \cap H$  closed in  $U$ .

Let  $x_0 \in \overline{H}$ ,  $\overline{H} \equiv$  closure of  $x_0$ .  
Then  $x_0U^{-1} \subseteq G$  is a neighborhood of  $x_0$  in  $G$  (since  $1 \in U^{-1}$ ,  $x_01 = x_0 \in x_0U^{-1}$ )

$$\implies x_0U^{-1} \cap H \neq \emptyset$$

$\forall x \in x_0U^{-1} \cap H$ ,  $x = x_0U^{-1}$  for some  $u \in U$ . Thus,  $x^{-1}x_0 = u \in U$ .

Now  
 $L_{x^{-1}} : G \rightarrow G$  is a homeomorphism, so  $L_{x^{-1}}(H) = H$ . By continuity,  $L_{x^{-1}}(\overline{H}) = \overline{H}$ . Thus  $x^{-1}x_0 \in \overline{H}$ .

**Claim:**  $x^{-1}x_0 \in H \cap U$ .  
Since  $x^{-1}x_0 \in \overline{H} \cap U$ ,  $\exists$  sequence  $\{h_i\} \subset H \cap U$  s.t.  $h_0 \rightarrow x^{-1}x_0$ .  
But recall  $H \cap U$  closed in  $U$ , so  $x^{-1}x_0 \in H \cap U$ .

$$\implies x_0 \in xH = H, \quad \overline{H} \subseteq H$$

Thus  $H$  closed.

<sup>4</sup><https://faculty.math.illinois.edu/~lerman/519/s12/427notes.pdf>

**Claim:** If  $H$  abstract subgroup of Lie group  $G$ , that's also an embedded submanifold  $\equiv$  regular submanifold, then  $H$  is a Lie subgroup.

Recall that by definition, Lie group has group multiplication and inverse map to be  $C^\infty$ . Then, just show group multiplication is  $C^\infty$ , first.

Since  $G$  is a Lie group, then

$$\mu : G \times G \rightarrow G$$

$$\mu(x, y) = xy$$

is  $C^\infty$  (by definition).

Then  $\mu : G \times G \rightarrow G$  cont.  
Consider subgroup  $H \subseteq G$  and  $\mu : H \times H \rightarrow H$ .  
Since  $H \times H \subseteq G \times G$ ,  $\forall (x, y) \in H \times H$  (fix  $(x, y) \in H \times H$ ),  $\forall$  neighborhood  $V$  of  $\mu(x, y) = xy$ ,  $V \subset G$ ,  $\exists$  neighborhood  $U$  of  $(x, y)$  s.t.  $\mu(U) \subseteq V$  (by  $\mu : G \times G \rightarrow G$  cont.).  
Since  $H$  embedded submanifold  $\equiv$  regular submanifold of  $G$ ,  
 $\exists$  neighborhood  $V' \subseteq V$  of  $xy \in G$ , coordinate map  $\varphi : V' \rightarrow \mathbb{R}^n$  ( $n = \dim G$ ) s.t.

$$\varphi(H \cap V') = \varphi(V') \cap (\mathbb{R}^k \times \{0\})$$

where  $k = \dim H$   
(since  $H$  is a  $k$ -dim. embedded submanifold  $\equiv$  regular submanifold,  $H \subseteq G$ , s.t.  $\forall p \in H$ ,  $\exists$  chart  $(V \subset G, \varphi : U \rightarrow \mathbb{R}^n \ni 0)$ , s.t.  $\varphi(U \cap V) = \varphi(V) \cap (\mathbb{R}^k \times \{0\})$ ).  
Now

$$\varphi \circ \mu : \mu^{-1}(V') \cap U \rightarrow \mathbb{R}^n \text{ is } C^\infty, \text{ and } \varphi \circ \mu(\mu^{-1}(V') \cap U) \subseteq \mathbb{R}^k \times \{0\}$$

Let projection  $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^k$  be the standard projection,

$$\pi \circ \varphi \circ \mu : \mu^{-1}(V') \cap U \rightarrow \mathbb{R}^k \text{ is } C^\infty$$

$\implies \mu$  is  $C^\infty$

From Chapter 4 “Lie Groups and Lie Algebras” of Kosmann-Schwarzbach (2010) [14]  
While Proposition 2.6 of Kosmann-Schwarzbach (2010) [14] states that

$$\det(\exp(X)) = \exp(\operatorname{tr} X)$$

here are some other resources online that gave further discussion on the characteristic polynomial,  $\det(A - \lambda 1)$  and the different terms of it, called Newton identities:

- [http://scipp.ucsc.edu/~haber/ph116A/charpoly\\_11.pdf](http://scipp.ucsc.edu/~haber/ph116A/charpoly_11.pdf)
- <http://math.stackexchange.com/questions/1126114/how-to-find-this-lie-algebra-proof-that-mathfraksl-is-tr>
- <http://mathoverflow.net/questions/131746/derivative-of-a-determinant-of-a-matrix-field>

**Theorem 16** (5.1 [14]). *Consider  $\mathfrak{g} = \{X = \gamma'(0) | \gamma : 1 \rightarrow G \text{ of class } C^1, \gamma(0) = 1\}$   
Let Lie group  $G$*

- $\mathfrak{g}$  vector subspace of  $\mathfrak{gl}(n, \mathbb{R})$
- $X \in \mathfrak{g}$  iff  $\forall t \in \mathbb{R}, \exp(tX) \in G$
- if  $X \in \mathfrak{g}$ , if  $g \in G$ , then  $gXg^{-1} \in \mathfrak{g}$
- $\mathfrak{g}$  closed under matrix commutator, i.e. if  $X, Y \in \mathfrak{g}$ ,  $[X, Y] \in \mathfrak{g}$

*Proof.* (i)



- (ii) If  $\exp(tX) \in G$ , then  $X \frac{d}{dt} \exp(tX) \Big|_{t=0} \in \mathfrak{g}$  (by def.)  
 If  $X \in \mathfrak{g}$ , then by def.,  $X = \frac{d}{dt} \gamma(t) \Big|_{t=0}$  with  $\gamma(t) \in G$ .  
 Now Taylor expand;  $\forall k \in \mathbb{Z}^+$

$$\begin{aligned} \gamma\left(\frac{t}{k}\right) &= 1 + \frac{t}{k}X + O\left(\frac{1}{k^2}\right) = \exp\left(\frac{t}{k}X + O\left(\frac{1}{k^2}\right)\right) \\ &\implies \left(\gamma\left(\frac{t}{k}\right)\right)^k = \exp(tX) \\ \gamma\left(\frac{t}{k}\right) &\in G \quad \forall k \in \mathbb{Z}^+ \end{aligned}$$

$G$  closed subgroup, so  $\lim_{k \rightarrow \infty} (\gamma(\frac{t}{k}))^k = \exp(tX) \in G$

- (iii)  
(iv)

**Definition 29** (Lie algebra). *Lie algebra  $\mathfrak{g}$ , tangent space to  $G$  at 1, i.e.  $\mathfrak{g} := T_1G$  is called Lie algebra of Lie group  $G$ .*

$$\mathfrak{g} := \{X = \gamma'(0) | \gamma : 1 \rightarrow G \text{ of class } C^1, \gamma(0) = 1\} = T_1G$$

This is based on Proposition 5.3 of Kosmann-Schwarzbach (2010) [14].

For Lie group

$$U(n) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = 1\}$$

If  $X \in \mathfrak{u}(n)$ , then  $\exp(tX) \in U(n)$ . Then

$$\exp(tX) \exp(tX)^\dagger = (1 + tX + O(t^2))(1 + tX^\dagger + O(t^2)) = 1 + t(X + X^\dagger) + O(t^2) = 1 \forall t \in \mathbb{R} \implies X + X^\dagger = 0$$

i.e.  $X \in \mathfrak{u}(n)$  is an anti-Hermitian complex  $n \times n$  matrix.

$$\mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) | X + X^\dagger = 0\}$$

*Physicists:*  $X = iA$  and so  $A - A^\dagger$ .  $A \in \mathfrak{u}(n)$  is a Hermitian complex  $n \times n$  matrix.

$$\mathfrak{u}(n) = \{A \in \mathfrak{gl}(n, \mathbb{C}) | A - A^\dagger = 0\}$$

Regardless,  $\dim_{\mathbb{R}} \mathfrak{u}(n) = n^2 = 2n^2 - n^2$

For Lie group

$$SU(n) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = 1, \det U = 1\}$$

Then

$$\mathfrak{su}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) | X + X^\dagger = 1, \text{tr} X = 0\}$$

is the Lie algebra of traceless anti-Hermitian complex  $n \times n$  matrices, and that

$$\dim_{\mathbb{R}} \mathfrak{su}(n) = n^2 - 1$$

In summary,

$$\begin{array}{ccc} \mathfrak{u}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) | X + X^\dagger = 0\} & & \mathfrak{su}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) | X + X^\dagger = 0, \text{tr} X = 0\} \\ \exp(tX) \downarrow & & \exp(tX) \downarrow \\ U(n) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = 1\} & & SU(n) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = 1, \det U = 1\} \end{array}$$

$$\dim_{\mathbb{R}} \mathfrak{u}(n) = n^2 \quad \dim_{\mathbb{R}} \mathfrak{su}(n) = n^2 - 1$$

From Chapter 5 “Lie Groups  $SU(2)$  and  $SO(3)$ ” of Kosmann-Schwarzbach (2010) [14],

11.0.1. *Bases of  $\mathfrak{su}(2)$ , Subsection 1.1 of Chapter 5 of Kosmann-Schwarzbach (2010) [14].* Recall that

$$\mathfrak{su}(n) = \{X \in \mathfrak{gl}(n, \mathbb{C}) | X + X^\dagger = 0, \text{tr} X = 0\}$$

$$\exp(tX) \downarrow$$

$$SU(n) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = 1, \det U = 1\}$$

$$\dim_{\mathbb{R}} \mathfrak{su}(n) = n^2 - 1$$

and so

$$\mathfrak{su}(2) = \{X \in \mathfrak{gl}(2, \mathbb{C}) | X + X^\dagger = 0, \text{tr} X = 0\}$$

$$\exp(tX) \downarrow$$

$$SU(2) = \{U \in GL(n, \mathbb{C}) | UU^\dagger = 1, \det U = 1\}$$

$$\dim_{\mathbb{R}} \mathfrak{su}(2) = 3$$

Also, recall that  $\mathfrak{g} \subseteq \mathfrak{gl}(n, \mathbb{C})$  is a vector subspace (16) and that  $X \in \mathfrak{g}$  iff  $\forall t \in \mathbb{R}, \exp(tX) \in G$ .

if  $X \in \mathfrak{g}$ , if  $g \in G$ , then  $gXg^{-1} \in \mathfrak{g}$

$\mathfrak{g}$  closed under  $\mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$

$$(X, Y) \mapsto [X, Y]$$

and so with  $\mathfrak{g}$  as a vector space, we can have a choice of bases.

$$\begin{aligned} \xi_1 &= \frac{i}{2} \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \\ \text{(a) } \xi_2 &= \frac{1}{2} \begin{pmatrix} & -1 \\ 1 & \end{pmatrix} \\ \xi_3 &= \frac{i}{2} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} \\ &\text{satisfying} \end{aligned}$$

$$[\xi_k, \xi_l] = \epsilon_{klm} \xi_m$$

(b) *Physics*



$$\sigma_1 = -2i\xi_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$$
$$\sigma_2 = 2i\xi_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}$$
$$\sigma_3 = -2i\xi_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$$

satisfying

$$[\sigma_k, \sigma_l] = 2i\epsilon_{klm}\sigma_m$$

EY : 20151001 Sage Math 6.8 doesn’t run on Mac OSX El Capitan: I suspect that it’s because in Mac OSX El Capitan, /usr cannot be modified anymore, even in an Administrator account. The TUG group for MacTeX had a clear, thorough, and useful (i.e. copy UNIX commands, paste, and run examples) explanation of what was going on:

<http://tug.org/mactex/elcapitan.html>

So keep in mind that my code for Sage Math is for Sage Math 6.8 that doesn’t run on Mac OSX El Capitan. I’ll also use sympy in Python as an alternative and in parallel.

One can check in sympy the traceless anti-Hermitian (or Hermitian) property of the bases and Pauli matrices, and the commutation relations (see groups.py):

```
import itertools
from itertools import product, permutations

import sympy
from sympy import I, LeviCivita
from sympy import Rational as Rat

from sympy.physics.matrices import msigma # <class 'sympy.matrices.dense.MutableDenseMatrix'>

def commute(A,B):
    """
    commute = commute(A,B)
    commute takes the commutator of A and B
    """
    return (A*B - B*A)

def xi(i):
    """
    xi = xi(i)
    xi is a function that returns the independent basis for
    Lie algebra su(2)\equiv su(2,\mathbb{C}) of Lie group SU(2) of
    traceless anti-Hermitian matrices, based on msigma of sympy
    cf. http://docs.sympy.org/dev/_modules/sympy/physics/matrices.html#msigma
    """
    if i not in [1,2,3]:
        raise IndexError("Invalid Pauli index")
    elif i==1:
        return I/Rat(2)*msigma(1)
    elif i==2:
        return -I/Rat(2)*msigma(2)
    elif i==3:
        return I/Rat(2)*msigma(3)
```

```
## check anti-Hermitian property and commutation relations with xi
# xi is indeed anti-Hermitian
xi(1) == -xi(1).adjoint() # True
xi(2) == -xi(2).adjoint() # True
xi(3) == -xi(3).adjoint() # True
```

```
# xi obeys the commutation relations
```

```
for i,j in product([1,2,3],repeat=2): print i,j

for i,j in product([1,2,3],repeat=2): print i,j, "\tCommutator:", commute(xi(i),xi(j))

## check traceless Hermitian property and commutation relations with Pauli matrices
# Pauli matrices i.e. msigam is indeed traceless Hermitian

msigma(1) == msigma(1).adjoint() # True
msigma(2) == msigma(2).adjoint() # True
msigma(3) == msigma(3).adjoint() # True

msigma(1).trace() == 0 # True
msigma(2).trace() == 0 # True
msigma(3).trace() == 0 # True

# Pauli matrices obey commutation relation
print "For Pauli matrices, the commutation relations are:\n"
for i,j in product([1,2,3],repeat=2): print i,j, "\tCommutator:", commute(msigma(i),msigma(j))

for i,j,k in permutations([1,2,3],3): print "Commute:", i,j,k, msigma(i), msigma(j), \ \
":\tand is 2*i of", msigma(k), commute(msigma(i),msigma(j)) == 2*I*msigma(k)*LeviCivita(i,j,k)
```

And finally the traceless property of the Pauli matrices:

```
>>> msigma(1).trace()
0
>>> msigma(2).trace()
0
>>> msigma(3).trace()
0
```

11.1. **Rotation Group  $SO(N)$  and Lie Algebra  $\mathfrak{so}(3)$ .** cf. Kambe (2009) [24], Appendix C  
 $g \in SO(N)$  represented by  $N \times N$  orthogonal matrix (i.e.  $gg^T = 1$ ) s.t.  $\det g = 1$ .

Let curve  $\xi(t)$  on  $SO(N)$ , initially from identity 1, with tangent vector  $\mathbf{a}$  at 1.  
 $\implies \xi(t) = 1 + t\mathbf{a} + O(t^2)$  for an infinitesimal parameter  $t$ .  
 $\mathbf{a} = \xi'(0)$  is an element of tangent space  $T_1SO(N)$  i.e. the Lie algebra  $\mathfrak{so}(N)$ , i.e.  $T_1SO(N) = \mathfrak{so}(N)$  by orthogonality condition  $\xi(t)\xi^T(t) = 1$ , then

(34)  $(1 + t\mathbf{a} + \dots)(1 + t\mathbf{a}^T + \dots) = 1 \implies \mathbf{a} + \mathbf{a}^T = 0$  or  $\mathbf{a} = \mathbf{a}^T$

$\implies \mathbf{a}$  skew-symmetric.

11.1.1.  **$\mathfrak{so}(3)$ .**  $\dim(\mathfrak{so}(3)) = 3$ , vector space of Lie algebra  $\mathfrak{so}(3)$  represented by skew-symmetric basis  $(E_1, E_2, E_3)$

$$E_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$[E_i, E_j] = \epsilon_{ijk}E_k \text{ where } [E_i, E_j] = E_iE_j - E_jE_i$$

11.2. **Spin.** Let’s follow the development by Baez and Muniain (1994) on pp. 175 of the Section II.1 “Lie Groups”, the second (II) chapter on “Symmetry” [11].

Let  $V = \mathbb{C}^2$ ,  $G = SU(2)$ . Then consider the graded algebra of polynomials on  $V = \mathbb{C}^2 \ni (x, y)$

$$P(V) = \bigoplus_{k=0}^{\infty} P^{(k)}(V) = \bigoplus_{\substack{j=0 \\ 2j \in \mathbb{Z}}}^{\infty} P^{(2j)}(V) = \bigoplus_{\substack{j=0 \\ j \in \mathbb{Z}}}^{\infty} P^{(2j)}(V) \oplus \bigoplus_{\substack{j=1/2 \\ 2j \text{ odd}}}^{\infty} P^{(2j)}(V)$$
$$P^{(2j)}(V) \equiv \text{vector space of complex polynomials of degree } 2j$$

and recall this representation on  $P^{(2j)}(V)$

$$\begin{aligned}\rho^{(2j)} : G &\rightarrow \text{End}(P^{(2j)}(V)) \\ \rho^{(2j)} : P^{(2j)}(V) &\rightarrow P^{(2j)}(V) \\ \rho^{(2j)}(g)(f) &= f \circ \rho(g^{-1}) \text{ where } \rho \text{ is the fundamental representation of } G = SU(2) \\ \rho^{(2j)}(g)(f)(v) &= f \circ \rho(g^{-1})(v) \quad \forall f \in P^{(2j)}(V), \forall v \in V = \mathbb{C}^2\end{aligned}$$

Note,  
 $\dim P^{(2j)} = \binom{2j+2-1}{2-1} = 2j + 1$

**Exercise 21.** [11] *spin-0* Consider the trivial representation  $\tau$ :

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{T} & P^{(0)}(V) \\ \tau(g) \downarrow & & \downarrow \rho^{(0)}(g) \\ \mathbb{C} & \xrightarrow{T} & P^{(0)}(V) \end{array}$$
$$\begin{aligned}\tau : G &\rightarrow \text{End}(\mathbb{C}) \\ \tau(g) : \mathbb{C} &\rightarrow \mathbb{C} \\ \tau(g) &= 1_{\mathbb{C}}\end{aligned}$$

Clearly,  $P^{(0)}(V) = \mathbb{C}$ , since  $P^{(0)}(V)$  consists of polynomials of constants in  $\mathbb{C}$ .  
Consider  $c_0 \in \mathbb{C}$ ,  $f = k_0 \in P^{(0)}(V)$   
 $\rho^{(0)}(g)(f) = f \circ \rho(g^{-1}) = k_0$   
 $\implies \rho^0(g)T(c_0) = T \circ \tau(g)c_0 = T(c_0)$ . Let  $T = 1_{\mathbb{C}} = 1_{P^0(V)}$   
So  $\rho^{(0)}(g) = \tau(g) = 1$ .  $T = 1$ . So representations  $\rho^{(0)}$  and trivial representation  $\tau$  on  $G$  are equivalent.

**Exercise 22.** [11] *spin- $\frac{1}{2}$*  For spin- $\frac{1}{2}$ ,  $j = \frac{1}{2}$ ,  $2j = 1$ .

$\forall f \in P^{(1)}(V)$ ,  $V = \mathbb{C}^2$ . So in general form,  $f(x, y) = ax + by \in P^{(1)}(V)$ ,  $\begin{pmatrix} x \\ y \end{pmatrix} \in V = \mathbb{C}^2$

Recall the fundamental representation

$$\begin{aligned}\rho : G &\rightarrow GL(2, \mathbb{C}) \equiv GL(\mathbb{C}^2) \\ \rho(g) : \mathbb{C}^2 &\rightarrow \mathbb{C}^2 \\ \rho(g) &= g\end{aligned}$$

So consider  $T$  such that

$$\begin{array}{ccc} \mathbb{C}^2 & \xrightarrow{T} & P^{(1)}(V) \\ \rho(g) \downarrow & & \downarrow \rho^{(1)}(g) \\ \mathbb{C}^2 & \xrightarrow{T} & P^{(1)}(V) \end{array}$$

Consider  $\forall v \in \mathbb{C}^2$ ,  $v = \begin{pmatrix} x \\ y \end{pmatrix}$ , then

$$\rho(g)v = gv = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix}$$

```
sage: g*X
[a*x + b*y]
[c*x + d*y]
```

For notation, let  $U \in G = SU(2)$  s.t.  $UU^\dagger = 1$ .  
Consider  $(\rho^{(2j)}(U)(f))(x) = f(U^{-1}x)$ ,  $\forall x \in \mathbb{C}^2$ .

Choose  $f(x, y) = x$ . So for  $f(x, y) = Ax + By$ ,  $A = 1, B = 0$ . Choose  $U = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$  so  $U^{-1} = \begin{pmatrix} \bar{a} & -b \\ \bar{b} & a \end{pmatrix}$ . Then

$$U^{-1}x = \begin{pmatrix} \bar{a}x - by \\ \bar{b}x + ay \end{pmatrix}$$

So

$$\begin{aligned}(\rho^{(1)}(U)(f))(x) &= f(U^{-1}x) = \bar{a}x - by \\ (\rho^{(1)}(U)(f))(x) &= f(U^{-1}x) = \bar{b}x + ay \text{ for } f(x, y) = y\end{aligned}$$

Let  $f(x, y) = Ax + By$

$$(\rho^{(1)}(U)(f))(x) = f(U^{-1}x) = (A\bar{a} + B\bar{b})x + (Ba - Ab)y = (\bar{a}x - by)A + (\bar{b}x + ay)B = (A\bar{a} + B\bar{b})x + (Ba - Ab)y$$

which was calculated with the assistance of Sage Math:

```
sage: U_try1 = Matrix( [[a.conjugate(), -b],[b.conjugate(), a] ] )
sage: f1( U_try1*X).coefficient(x)
A*conjugate(a) + B*conjugate(b)
sage: f1( U_try1*X).coefficient(y)
B*a - A*b
```

Treating  $P^{(1)}(\mathbb{C}^2)$  as a vector space, in its matrix formulation, then  $f(x, y) = Ax + By \in P^{(1)}(\mathbb{C}^2)$  is treated as  $\begin{bmatrix} A \\ B \end{bmatrix}$ , then  $(\rho^{(1)}(U)f)$  is

$$\implies \begin{bmatrix} \bar{a} & \bar{b} \\ -b & a \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} A\bar{a} + B\bar{b} \\ -Ab + Ba \end{bmatrix}$$

so conclude in general that  $\rho^{(1)}(U) = (U^\dagger)^T$ .

Now, as Kosmann-Schwarzbach (2010) [14] says, on pp. 13, Chapter 2 Representations of Finite Groups, “Two representations  $(E_1, \rho_1)$  and  $(E_2, \rho_2)$  are equivalent if and only if there is a basis  $B_1$  of  $E_1$  and a basis  $B_2$  of  $E_2$  such that for every  $g \in G$ , the matrix of  $\rho_1(g)$  in the basis  $B_1$  is equal to the matrix of  $\rho_2(g)$  in the basis  $B_2$ . In particular, if the representations  $(E_1, \rho_1)$  and  $(E_2, \rho_2)$  are equivalent, then  $E_1$  is isomorphic to  $E_2$ .” So we need a change of basis between  $\rho(U) = U$  and  $\rho^{(1)}(U)$ . What’s the linear transformation  $T$  s.t.

$$T^{-1}\rho^{(1)}(U)T = U?$$

By intuition,

$$T = \sigma_x \sigma_z \equiv \sigma_1 \sigma_3$$

where  $\sigma_i$ ’s are Pauli matrices.

Indeed,

```
sage: Paulimat[3] * Paulimat[1]*U_try*Paulimat[1] * Paulimat[3]
[conjugate(a) conjugate(b)]
[          -b          a]
```

Then  $\rho^{(1)}(U) \circ T = TU$ , so this  $T = \sigma_1 \sigma_3$  is an “intertwining operator” between  $\rho^{(1)}(U)$  and fundamental representation  $\rho(U) = U$ , with  $T = \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}$ , and  $T^{-1} = \begin{bmatrix} & 1 \\ -1 & \end{bmatrix}$ .  
 $T$  is an isomorphism between  $\mathbb{C}^2$  and  $P^{(1)}(\mathbb{C}^2)$ . So fundamental representation  $\rho$  of  $G = SU(2)$  is equivalent to  $\rho^{(1)}(U)$  on  $P^{(1)}(\mathbb{C}^2)$ .

**Exercise 23.** [11] (Also from Exercise 2.6 of Kosmann-Schwarzbach (201) [14])

Let  $(E, \pi)$  representation of group  $G$ .  
 $\forall g \in G$ ,  $\xi \in E^*$ ,  $x \in E$ , set  $\langle \pi^*(g)(\xi), x \rangle = \langle \xi, \pi(g^{-1})(x) \rangle$

*dual* (or *contragredient*) of  $\pi$ ,  $\pi^* : G \rightarrow \text{End}(E^*)$ ,  $\pi^*$  is a representation, since

$$\begin{aligned} \langle \pi^*(gh)(\xi), x \rangle &= \langle \xi, \pi((gh)^{-1})(x) \rangle = \langle \xi, \pi(h^{-1}g^{-1})(x) \rangle = \langle \xi, \pi(h^{-1})\pi(g^{-1})(x) \rangle = \langle \xi, \pi(h^{-1})(\pi(g^{-1})(x)) \rangle = \\ &= \langle \pi^*(h)(\xi), \pi(g^{-1})(x) \rangle = \langle \pi^*(g)\pi^*(h)(\xi), x \rangle \end{aligned}$$

since this is true,  $\forall x \in E, \forall \xi \in E^*, \pi^*(gh) = \pi^*(g)\pi^*(h)$ .

dual  $\pi^*$  of  $\pi$  is a representation.

**11.3. Adjoint Representation.** I will first follow Sec. 7.3 The Adjoint Representation of Ch. 4 Lie Groups and Lie Algebras of Kosmann-Schwarzbach (201) [14]).

The *conjugation action*  $\mathcal{C}_g : G \rightarrow G$  is defined as

$$\begin{aligned} \mathcal{C}_g : G &\rightarrow G \\ \mathcal{C}_g : h &\mapsto ghg^{-1} \end{aligned}$$

So

$$\begin{aligned} \mathcal{C} : G &\rightarrow \text{Aut}(G) \\ \mathcal{C}g &= \mathcal{C}_g \end{aligned}$$

Now define the *adjoint action* of  $g$  as the differential or push forward of  $\mathcal{C}_g$ :

$$\text{Ad}_g := D_1\mathcal{C}_g \equiv (\mathcal{C}_g)_{*1} \equiv (\mathcal{C}_g)_*|_{g=1} \quad (\text{adjoint action of } g)$$

Now  $\text{Ad}_g : \mathfrak{g} \rightarrow \mathfrak{g}$ , so  $\text{Ad} : G \rightarrow \text{End}(\mathfrak{g})$

Note  $\mathcal{C}_{gg'} = \mathcal{C}_g\mathcal{C}_{g'} \equiv \mathcal{C}(gg') = \mathcal{C}(g) \circ \mathcal{C}(g')$  and so

$$\xrightarrow{D_1} \text{Ad}_{gg'} = \text{Ad}_g \circ \text{Ad}_{g'}$$

Kosmann-Schwarzbach (201) [14]) claims, because  $\text{Ad}_g = 1_{\mathfrak{g}}$  when  $g = 1$ ,

$\text{Ad} : G \rightarrow GL(\mathfrak{g})$  is a representation of  $G$  on  $\mathfrak{g}$ . (EY : 20160505 ???)

$\text{Ad} : g \mapsto \text{Ad}_g$

**Definition 30.** *representation*  $Ad$  of  $G$  on  $V = \mathfrak{g}$  is called *adjoint representation of Lie group*  $G$ .

Denote adjoint representation of Lie algebra  $\mathfrak{g}$ ,  $\text{ad}$ .

By definition,  $\text{Ad}_{\exp(tX)} = \exp(t\text{ad}_X)$

cf. Prop. 7.8 of Kosmann-Schwarzbach (201) [14])

**Proposition 13.** (1) *Let  $A$  invertible matrix,  $A \in \text{Lie group } G$ .*

*Let  $X$  matrix s.t.  $X \in \mathfrak{g}$ . Then*

$$\text{Ad}_A(X) = AXA^{-1}$$

(2) *Let  $X, Y \in \mathfrak{g}$ . Then*

$$\text{ad}_X(Y) = [X, Y]$$

(3) *Let  $X, Y \in \mathfrak{g}$ . Then*

$$\text{ad}_{[X, Y]} = [\text{ad}_X, \text{ad}_Y]$$

*Proof.* (1) By def.,  $\forall B \in G, \mathcal{C}_A(B) = ABA^{-1}$ , and thus

$$\text{Ad}_A(X) = \left. \frac{d}{dt} A \exp(tX) A^{-1} \right|_{t=0} = AXA^{-1}$$

(2)

$$\begin{aligned} \text{ad}_X(Y) &= \left. \frac{d}{dt} \text{Ad}_{\exp(tX)}(Y) \right|_{t=0} = \left. \frac{d}{dt} \exp(tX) Y \exp(tX) \right|_{t=0} = \\ &= XY - YX = [X, Y] \end{aligned}$$

(3) Use Jacobi identity:

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \text{ or}$$

$$[[A, B], C] = [A, [B, C]] - [B, [A, C]]$$

$\text{ad}_{[X, Y]}C = [[X, Y], C] = [X, [Y, C]] - [Y, [X, C]] = [X, \text{ad}_Y C] - [Y, \text{ad}_X C]$  and that

$$\text{ad}_X \text{ad}_Y C = [X, [Y, C]] \implies \text{ad}_{[X, Y]} C = [\text{ad}_X, \text{ad}_Y] C$$

□

**11.4. Invariant Vector Fields, Left-invariant Vector Field, Right-invariant Vector Field, Invariant Vector Fields.**

cf. pp. 26, Sec. 1.7 Lie Group and Invariant Vector Fields, Kambe (2009) [24]

Consider group  $G$  of smooth transformations (maps) of a manifold  $M$  into itself, i.e. Lie group.

Given fixed element  $h \in G$ , then

$$\begin{aligned} (35) \quad L_h : g &\mapsto hg \text{ i.e. } L_h(g) = hg \\ R_h : g &\mapsto gh \text{ i.e. } R_h(g) = gh \end{aligned}$$

where  $L_h \equiv$  left translation of the group

$R_h \equiv$  right translation of the group

Suppose  $g_t$  is a curve on  $G$  described in terms of parameter  $t$ .

Left translation of  $g_t$  by  $g_{\Delta t}$  for infinitesimal  $\Delta t$  given by  $g_{\Delta t} \circ g_t$ .

$$(36) \quad \dot{g}_t = \lim_{\Delta t \rightarrow 0} \frac{g_{t+\Delta t} - g_t}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{(g_{\Delta t} - 1)g_t}{\Delta t} = X \circ g_t \text{ with } X := \lim_{\Delta t \rightarrow 0} \frac{g_{\Delta t} - 1}{\Delta t} \in T_1 G$$

Thus, left translation  $L_{g_{\Delta t}}$  leads to **right**-invariant vector field.

Similarly, right translation  $R_{g_{\Delta t}}$  leads to **left**-invariant vector field.

In summary,

$\dot{g}_t$  is said to be tangent vector at  $g_t$ .

Vector field  $X^L, (X^R)$  on  $G$  is left invariant (right invariant) if it's invariant under all left-translations (right translations) respectively, i.e.  $\forall g, h \in G$ , if

$$\begin{aligned} (37) \quad (L_h)_* X_g^L &= X_{hg}^L \\ (R_h)_* X_g^R &= X_{hg}^R \end{aligned}$$

Given tangent vector  $X$  to  $G$ , at  $e$ , one may left-translate (right translate)  $X$  to every pt.  $g \in G$  as

$$\begin{aligned} (38) \quad X_g^L &= (L_g)_* X = g \circ X = gX \\ X_g^R &= (R_g)_* X = X \circ g = Xg \end{aligned}$$

Since

$$(R_h)_* X_g^R = (R_h)_*(R_g)_* X = X \cdot gh = X_{gh}^R$$

Then  $(R_h)_*$  operation does yield a right-invariant field.

Hence  $(R_h)_*$  transformation in Eq. 38 gives a right-invariant field generated by  $X$ .

In other words, in summary

$$(39) \quad \begin{aligned} & \text{Left translation of } g_t \text{ by } g_{\Delta t}: L_{g_{\Delta t}}: g_t \mapsto g_{\Delta t}g_t = g_{t+\Delta t} \\ & \dot{g}_t = Xg_t \text{ so } X \in T_1G \text{ is right-invariant vector field} \\ & \text{i.e. } X_g^R = (R_g)_*X = Xg \end{aligned}$$

Consider curves  $\xi_t: \mathbb{R} \rightarrow G$  with tangent  $\dot{\xi}_0 = X$  at  $t = 0$ .

$$\begin{aligned} & \xi: t \mapsto \xi(t), \quad t \in \mathbb{R} \\ & \text{Left-invariant field } X_s^L = \frac{d}{dt}(g \circ \xi_t) \quad \forall g_s \in G, (s \text{ a parameter}) \\ & \text{Right-invariant field } X_s^R = \frac{d}{dt}(\xi_t \circ g_s) \end{aligned}$$

Lie group  $G$  acts as a group of linear transformations on its own Lie algebra  $\mathfrak{g}$ , namely  $\forall g \in G, \exists$  operator  $\text{Ad}_g$  s.t.

$$(40) \quad \text{Ad}_g Y := (L_g)_* \circ (R_{g^{-1}})_* Y = gYg^{-1}, \quad \forall Y \in \mathfrak{g}$$

Operator  $\text{Ad}_g$  transforms  $Y \in \mathfrak{g}$  into  $\text{Ad}_g Y \in \mathfrak{g}$ , linearly.

Set of all such  $\text{Ad}_g$ , i.e.  $\text{Ad}(G)$  is called the adjoint representation of  $G$ , an adjoint group.

Setting  $g$  to be the inverse  $\xi_t^{-1} := (\xi_t)^{-1}$ , adjoint transformation  $\text{Ad}_{\xi_t^{-1}} Y$  is a function of  $t$ .

Its derivative with respect to  $t$  is a linear transformation from  $Y$  to  $\text{ad}_X Y$ :

$$(41) \quad \begin{aligned} \text{ad}_X Y &= \left. \frac{d}{dt} \xi_t^{-1} Y \xi_t \right|_{t=0} := [X, Y] \\ \text{ad}_X: \mathfrak{g} &\rightarrow \mathfrak{g} \end{aligned}$$

$\text{ad}_X$  is a linear transformation  $\mathfrak{g} \rightarrow \mathfrak{g}$ .

## 12. LIE GROUPS AND THEIR LIE ALGEBRAS

[Lie groups and their Lie algebras - Lec 13 - Frederic Schuller](#), Schuller (2015) [37]

### 12.1. Chapter 4. Lie Theory, Schuller (2015) [37].

#### 12.1.1. 4.1 Lie groups, Schuller (2015) [37].

**Definition 31** (Lie group, Schuller (2015) [37]). *A lie group  $(G, \cdot)$  is*

### Part 6. Cohomology; Stoke's Theorem

#### 13. STOKE'S THEOREM

**Theorem 17** (Stoke's Theorem). *Let  $M$  be oriented, smooth  $n$ -manifold with boundary, let  $\omega$  be a compactly supported smooth  $(n-1)$ -form on  $M$ , or if  $\omega \in A_c^{n-1}(M)$ , Then*

$$(42) \quad \int_M d\omega = \int_{\partial M} \omega$$

If  $\partial M = \emptyset$ , then  $\int_{\partial M} \omega = 0$   
 $\int_{\partial M} \omega$  interpreted as  $\int_{\partial M} i_{\partial M}^* \omega = \int_{\partial M} i^* \omega$  so

$$(43) \quad \int_M d\omega = \int_{\partial M} i^*(\omega)$$

where inclusion  $i: \partial M \hookrightarrow M$

*Proof.* Begin with very special case:

Suppose  $M = \mathbb{H}^n$  (upper half space),  $\partial M = \mathbb{R}^{n-1}$

$\omega$  has compact support, so  $\exists R > 0$  s.t.  $\text{supp} \omega \subseteq \text{rectangle } A = [-R, R] \times \cdots \times [-R, R] \times [0, R]$ .

$$(44) \quad \begin{aligned} & \forall \omega \in A_c^{n-1}(\mathbb{H}^n) \\ & \omega = \sum_{j=1}^n (-1)^{j-1} f_j dx^1 \wedge \cdots \wedge \widehat{dx}^j \wedge \cdots \wedge dx^n \equiv \sum_{i=1}^n \omega_i dx^1 \wedge \cdots \wedge \widehat{dx}^i \wedge \cdots \wedge dx^n \end{aligned}$$

with Conlon (2008) [19] and John Lee (2012) [5]'s notation, respectively, and where  $f_j$  has compact support.

$$i^* \omega = (f_1 \circ i) dx^2 \wedge \cdots \wedge dx^n \in A_c^{n-1}(\partial \mathbb{H}^n)$$

$$\begin{aligned} d\omega &= \sum_{i=1}^n d\omega_i \wedge dx^1 \wedge \cdots \wedge \widehat{dx}^i \wedge \cdots \wedge dx^n = \sum_{i,j=1}^n \frac{\partial \omega_i}{\partial x^j} dx^j \wedge dx^1 \wedge \cdots \wedge \widehat{dx}^i \wedge \cdots \wedge dx^n = \\ &= \sum_{i=1}^n (-1)^{i-1} \frac{\partial \omega_i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n \end{aligned}$$

i.e. (for another notation)

$$d\omega = \left( \sum_{j=1}^n \frac{\partial f_j}{\partial x^j} \right) dx^1 \wedge \cdots \wedge dx^n \in A_c^n(\mathbb{H}^n)$$

$$d\omega = \left( \sum_{j=1}^n \frac{\partial f_j}{\partial x^j} \right) dx^1 \wedge \cdots \wedge dx^n \in A_c^n(\mathbb{H}^n)$$

$$\int_{\mathbb{H}^n} d\omega = \sum_{i=1}^n (-1)^{i-1} \int_A \frac{\partial \omega_i}{\partial x^i} dx^1 \wedge \cdots \wedge dx^n = \sum_{i=1}^n (-1)^{i-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R dx^1 \cdots dx^n \frac{\partial \omega_i}{\partial x^i}(x)$$

We can change order of integration in each term so to do  $x^i$  integration first.

By fundamental thm. of calculus, terms for which  $i \neq n$  reduce to

$$\begin{aligned} \sum_{i=1}^{n-1} (-1)^{i-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \omega_i}{\partial x^i}(x) dx^1 \cdots dx^n &= \sum_{i=1}^{n-1} (-1)^{i-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R \frac{\partial \omega_i}{\partial x^i}(x) dx^i dx^1 \cdots \widehat{dx}^i \cdots dx^n = \\ &= \sum_{i=1}^{n-1} (-1)^{i-1} \int_0^R \int_{-R}^R \cdots \int_{-R}^R [\omega_i(x)]_{x^i=-R}^{x^i=R} dx^1 \cdots \widehat{dx}^i \cdots dx^n = 0 \end{aligned}$$

because we've chosen  $R$  large enough that  $\omega = 0$  when  $x^i = \pm R$ .

□

### Part 7. Prástaro

Prástaro (1996) [15]

13.0.1. *Affine Spaces.* cf. Sec. 1.2 - *Affine Spaces* of Prástaro (1996) [15]

**Definition 32** (affine space).

$$(45) \quad \begin{array}{ll} \text{affine space} & (M, \mathbf{M}, \alpha) \\ & \text{with} \\ M \equiv & \text{set (set of pts.)} \\ \mathbf{M} \equiv & \text{vector space (space of free vectors)} \\ \alpha \equiv & \mathbf{M} \times M \rightarrow M \equiv \text{translation operator} \\ \alpha : (v, p) \mapsto p' \equiv & p + v \end{array}$$

Note:  $\alpha$  is a **transitive** action and without fixed pts. (free).

i.e.  $\forall p \in M$ ,

$$\forall \text{ pt. } O \in M, \alpha : (v, O) \mapsto O' \equiv O + v, \alpha(\cdot, O) \equiv \alpha_O \equiv \alpha(O). \quad \alpha_O(v) = O' = O + \mathbf{v} \quad \forall O' \in M, \exists \mathbf{v} \in \mathbf{M} \text{ s.t. } O' = O + \mathbf{v} \\ \implies M \equiv \mathbf{M}.$$

$\forall (O, \{e_i\})_{1 \leq i \leq n}$ , where  $\{e_i\}$  basis of  $\mathbf{M}$ ,  $M \equiv \mathbf{M} = \mathbb{R}^n$  so isomorphism  $M \simeq \mathbb{R}^n$

i.e.  $\alpha$  is **without fixed pts.**, meaning,

Given pointed space  $(M, O)$ , where base pt.  $O \in M$ , we can associate  $\forall p \in M$ , vector  $\mathbf{x} \in \mathbf{M}$ , by 1-to-1 mapping  $M \rightarrow \mathbf{M}$ .

So for

$$\begin{array}{l} \alpha : \mathbf{M} \times M \rightarrow M \\ \alpha(\mathbf{x}, p) = p' = p + \mathbf{x} \end{array}$$

Consider

$$\alpha(\mathbf{x}, O) = p = \alpha_O(\mathbf{x}) = p \implies \exists \alpha_O^{-1}(p) = \mathbf{x} \in \mathbf{M}$$

- (1) tangent space of  $M$  in  $p \in M$  is vector space  $T_p M \equiv (\mathbf{M}, p) \cong M$
- (2) If  $\mathbf{M}$  Euclidean space, affine space  $(M, \mathbf{M}, \alpha)$  is Euclidean
- (3) Call dim. of affine space  $(M, \mathbf{M}, \alpha)$ , dim. of  $\mathbf{M} \equiv \dim \mathbf{M}$

$\{\mathbf{e}_i\}$  basis of  $\mathbf{M}$

**Definition 33.**  $(O, \{e_i\}) \equiv$  affine frame.

$\forall$  affine frame  $(O, \{e_i\})$ ,  $\exists$  coordinate system  $x^\alpha : M \rightarrow \mathbb{R}$ ,

where  $x^\alpha(p)$  is  $\alpha$ th component, in basis  $\{e_i\}$ , of vector  $p - O$

**Proposition 14** (1.6, Prástaro (1996) [15]).  $\forall O \in M$ , we have canonical identification  $M \equiv \mathbf{M}$ , since

$$\begin{array}{ll} \alpha_O^{-1} : M \rightarrow \mathbf{M} & \alpha_O : \mathbf{M} \rightarrow M \\ \alpha_O^{-1}(p) = \mathbf{x} & \alpha_O : \mathbf{x} = \alpha(\mathbf{x}, O) = p \end{array}$$

Furthermore,

$\forall$  **affine frame**  $(O, \{\mathbf{e}_i\})_{1 \leq i \leq d}$ , where  $\{\mathbf{e}_i\}$  basis of  $\mathbf{M}$ ,

$\exists$  isomorphism  $M \cong \mathbb{R}^d$ ,

Then,  $\forall (O, \{\mathbf{e}_i\})_{1 \leq i \leq d}$ ,

$\exists$  coordinate system  $x^\alpha : M \rightarrow \mathbb{R}$ ,

where  $x^\alpha(p) = \alpha$ th component, in basis  $\{\mathbf{e}_i\}$ , of vector  $p - O$ .

**Theorem 18** (1.4 Prástaro (1996) [15]). Let  $(x^\alpha), (\bar{a}^\alpha)$  2 coordinate systems correspond to affine frames  $(O, \{e_i\})$ ,  $(\bar{O}, \{\bar{e}_i\})$ , respectively.

$$(46) \quad \bar{x}^\alpha = A^\alpha_\beta x^\beta + y^\alpha$$

where

$$y^\alpha \in \mathbb{R}^n, \quad A^\alpha_\beta \in GL(n; \mathbb{R})$$

**Definition 34** (1.10 Prástaro (1996) [15]).

$$(47) \quad A(n) \equiv Gl(n, \mathbb{R}) \times \mathbb{R}^n$$

affine group of dim.  $n$

**Theorem 19** (1.5). symmetry group of  $n$ -dim. affine space, called affine group  $A(M)$  of  $M$ .  $\exists$  isomorphism,

$$(48) \quad A(M) \simeq A(n), \quad f \mapsto (f^\alpha_\beta, y^\alpha); \quad f^\alpha \equiv x^\alpha \circ f = f^\alpha_\beta x^\beta + y^\alpha$$

cf. Eq. 1.4 Prástaro (1996) [15]

**Definition 35** (metric). Let smooth manifold  $M$ ,  $\dim M = n$ ,  $\forall p \in M$ ,  $\exists$  vector space  $T_p M$ , and so for

$$(49) \quad \begin{array}{l} g_p(T_p M)^2 \rightarrow \mathbb{R} \\ g_p : (X_p, Y_p) \mapsto g_p(X_p, Y_p) \in \mathbb{R} \end{array}$$

with  $g_p$  being bilinear, symmetric (in  $X_p, Y_p$ ), nondegenerate (i.e. if  $g_p(X_p, Y_p) = 0$ , then  $X_p$  or  $Y_p = 0$ )

Note that

$$g \in \Gamma((TM \otimes TM)^*)$$

and that for  $X = X^i \frac{\partial}{\partial x^i}$  so

$$Y = Y^i \frac{\partial}{\partial x^i}$$

$$g(X, Y) = g_{ij} X^i Y^j$$

Now for

$$\begin{array}{ll} F : M \rightarrow N & DF \equiv F_* : T_p M \rightarrow T_{F(p)} N \\ F : x \mapsto y = y(x) & DF : X_p \mapsto (DF)(X^j \frac{\partial}{\partial x^j}) = X^j \frac{\partial y^i}{\partial x^j} \frac{\partial}{\partial y^i} \end{array}$$

$$(F^* g')(X, Y) = (F^* g')(X^i \frac{\partial}{\partial x^i}, Y^j \frac{\partial}{\partial x^j}) = (F^* g')_{ij} X^i Y^j = g'(F_* X, F_* Y) =$$

$$=$$

## Part 8. Connections

### 14. CONNECTIONS OF VECTOR BUNDLES

Morita (2001) [43]

**Definition 36** (Connection in a vector bundle). **connection in a vector bundle**  $\pi : E \rightarrow M$  over  $C^\infty$  manifold  $M$ , is a bilinear map

$$\nabla : \mathfrak{X}(M) \times \Gamma(E) \rightarrow \Gamma(E)$$

satisfying

- (i)  $\nabla_{fX} s = f \nabla_X s$
- (ii)  $\nabla_X (fs) = f \nabla_X s + (Xf)s$  where  $f \in C^\infty(M)$ ,  $X \in \mathfrak{X}(M)$ ,  $s \in \Gamma(E)$

$\nabla_X s$  is covariant derivative of  $s$  relative to  $X$  (for Morita (2001)[43])

Alternatively, Jost (2017) [44] defines in the following way:

Let vector bundle  $(E, \pi, M)$ , where differentiable manifold  $M$ , vector bundle  $E$  over  $M$ ; a covariant derivative or (linear) connection is map

$$(50) \quad D : \Gamma(E) \rightarrow \Gamma(E) \otimes \Gamma(T^* M) \equiv \Gamma(E) \otimes \mathfrak{X}(M)$$

(we may also consider  $D$  as map  $D : \mathfrak{X}(M) \otimes \Gamma(E) \rightarrow \Gamma(E)$ , with  $s \in \Gamma(E)$ ,  $X \in \mathfrak{X}(M)$ ,  $DX(s) \equiv D_X s$ ) s.t.



(1)  $D$  is tensorial in  $\mathfrak{X}(M)$ ,

$$(51) \quad \begin{aligned} D_{X+Y}s &= D_Xs + D_Ys \quad \forall X, Y \in T_xM, s \in \Gamma(E) \\ D_{fX}s &= fD_Xs \quad \forall f \in C^\infty(M, \mathbb{R}), X \in \Gamma(TM) \end{aligned}$$

(2)  $D$  is  $\mathbb{R}$ -linear in  $s$ ,

$$(52) \quad D_X(s+t) = D_Xs + D_Xt \quad \forall X \in T_xM, s, t \in \Gamma(E)$$

and product rule

$$(53) \quad D_X(fs) = X(f) \cdot s + fD_X(s) \quad \forall f \in C^\infty(M, \mathbb{R})$$

Claim: Any vector bundle admits a connection.

e.g. product bundle  $M \times \mathbb{R}^n$ . Let  $x_1, \dots, x_n$  be canonical coordinates in  $\mathbb{R}^n$ . Take frame field  $(s_1, \dots, s_n)$ , where  $s_i(p) = \frac{\partial}{\partial x^i}$ . Set  $\nabla_X s_i = 0$  ( $i = 1, \dots, n$ )  $\forall$  vector space  $X$ ,  $\forall s = \sum_i a_i s_i, \forall X \in \mathfrak{X}(M)$ , set

$$\nabla_X s = \sum_{i=1}^n (Xa_i) s_i$$

For this connection  $\nabla_X s$  is the partial derivative in direction of  $X$  if  $s$  is considered  $\mathbb{R}^n$ -valued function on  $M$ . Call it **trivial connection** in product bundle.

Indeed,

$$\begin{aligned} \nabla_X s &= \nabla_{X^i \frac{\partial}{\partial x^i}} \left( s^m \frac{\partial}{\partial x^m} \right) = X^i \nabla_{\frac{\partial}{\partial x^i}} \left( s^m \frac{\partial}{\partial x^m} \right) = X^i \left( \nabla_{\frac{\partial}{\partial x^i}} s^m \right) \frac{\partial}{\partial x^m} + X^i s^m \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^m} = \\ &= X^i \left( \frac{\partial}{\partial x^i} s^m \right) \frac{\partial}{\partial x^m} + X^i s^m \Gamma_{mi}^q \frac{\partial}{\partial x^q} = X^i \frac{\partial s^m}{\partial x^i} \frac{\partial}{\partial x^m} + 0 \end{aligned}$$

if  $\Gamma_{mi}^q = 0$  at a chosen point  $p$ .

For arbitrary vector bundle  $\pi : E \rightarrow M$ , take locally finite open covering  $\{U_\alpha\}_{\alpha \in A}$  s.t.  $\pi^{-1}(U_\alpha)$  trivial. Denote  $\nabla^\alpha$  trivial connection  $\forall \pi^{-1}(U_\alpha)$ . Let  $\{f_\alpha\}$  be a partition of unity for covering  $U_\alpha$ , define

$$\nabla_X s := \sum_\alpha f_\alpha \nabla_X^\alpha s$$

Verify this defines connection in  $E$ :

$$\begin{aligned} \nabla_X(gs) &= \sum_\alpha f_\alpha \nabla_X^\alpha(gs) = \sum_\alpha f_\alpha \left[ X^i \left( \frac{\partial g}{\partial x^i} \right) s + X^i g \frac{\partial s^m}{\partial x^i} \frac{\partial}{\partial x^m} \right] = \\ &= g \sum_\alpha f_\alpha \nabla_X^\alpha s + \sum_\alpha f_\alpha (Xg)s \end{aligned}$$

**Proposition 15** (5.18 Morita (2001)[43]). *Let  $\nabla_i$  ( $1 \leq i \leq k$ ) be  $k$  connections in a given vector bundle. Then  $\forall$  linear combination  $\sum_{i=1}^k t_i \nabla_i$ , where  $t_1 + \dots + t_k = 1$  is a connection.*

*Proof.* TODO: Ex. 5.5  $\square$

**14.1. Directional derivatives of tensor fields; Connections on a tensor product.** cf. Schuller (2015) [38], **Lecture 7: Connections (International Winter School on Gravity and Light 2015))**.

**Definition 37** (Connection on tensor fields). A **connection**  $\nabla$  on smooth manifolds  $(M, \mathcal{O}, \mathcal{A})$  is a map  $\nabla : \mathfrak{X}(M) \times \Gamma(TM^p \otimes T^*M^q) \rightarrow \Gamma(TM^p \otimes T^*M^q)$ ,

where  $\binom{p}{q}$  tensor-field section  $\equiv \Gamma(TM^p \otimes T^*M^q)$

- (1)  $\nabla_X f = Xf \quad \forall f \in C^\infty(M)$
- (2)  $\nabla_X(T+S) = \nabla_X T + \nabla_X S$

$$(3) \quad \nabla_X(T(\omega, Y)) = (\nabla_X T)(\omega, Y) + T(\nabla_X \omega, Y) + T(\omega, \nabla_X Y), \text{ where } T(\omega, Y) \in C^\infty(M)$$

*This is written for a (1,1) tensor field  $T$ , but is analogously true  $\forall (p, q)$  tensor field.*

$$(4) \quad \nabla_{fX+Z}T = f\nabla_X T + \nabla_Z T, \text{ where } f \in C^\infty(M)$$

*Consider also product rule  $\nabla_X(fT) = X(f)T + f\nabla_X T \quad \forall f \in C^\infty(M, \mathbb{R})$*

A manifold with connection is a quadruple  $(M, \mathcal{O}, \mathcal{A}, \nabla)$

Consider

$$\begin{aligned} \nabla_X Y &= \nabla_{X^i \frac{\partial}{\partial x^i}} \left( Y^m \frac{\partial}{\partial x^m} \right) = X^i \nabla_{\frac{\partial}{\partial x^i}} \left( Y^m \frac{\partial}{\partial x^m} \right) = X^i \left[ \left( \nabla_{\frac{\partial}{\partial x^i}} Y^m \right) \frac{\partial}{\partial x^m} + Y^m \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^m} \right] = \\ &= X^i \frac{\partial Y^m}{\partial x^i} \frac{\partial}{\partial x^m} + X^i Y^m \Gamma_{mi}^q \frac{\partial}{\partial x^q} \end{aligned}$$

so then

**Definition 38** (Connection coefficient functions, i.e. Christoffel Symbols).

$$(54) \quad \begin{aligned} \Gamma_{jk}^i : \mathcal{U} &\rightarrow \mathbb{R} \\ p &\mapsto \left( dx^i \left( \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} \right) \right)(p) \end{aligned}$$

Now since

$$\begin{aligned} \nabla_{\frac{\partial}{\partial x^m}} \left( dx^i \left( \frac{\partial}{\partial x^j} \right) \right) &= \nabla_{\frac{\partial}{\partial x^m}} (\delta_j^i) = 0 \text{ so} \\ \nabla_{\frac{\partial}{\partial x^m}} \left( dx^i \left( \frac{\partial}{\partial x^j} \right) \right) &= \left( \nabla_{\frac{\partial}{\partial x^m}} dx^i \right) \left( \frac{\partial}{\partial x^j} \right) + dx^i \left( \nabla_{\frac{\partial}{\partial x^m}} \frac{\partial}{\partial x^j} \right) = \left( \nabla_{\frac{\partial}{\partial x^m}} dx^i \right) \left( \frac{\partial}{\partial x^j} \right) + dx^i \Gamma_{jm}^q \frac{\partial}{\partial x^q} = 0 \\ &\implies \left( \nabla_{\frac{\partial}{\partial x^m}} dx^i \right) \left( \frac{\partial}{\partial x^j} \right) = -\Gamma_{jm}^i \text{ since} \\ dx^i \Gamma_{jm}^q \frac{\partial}{\partial x^q} &= \Gamma_{jm}^q dx^i \left( \frac{\partial}{\partial x^q} \right) = \Gamma_{jm}^q \delta_q^i = \Gamma_{jm}^i \end{aligned}$$

14.1.1. *Change of  $\Gamma$ 's under change of chart.* cf. **1:08:50 3. Change of  $\Gamma$ 's under change of chart**, Schuller (2015) [38]

$(U, x), (V, y) \in \mathcal{A}$  and  $U \cap V \neq \emptyset$

$$\Gamma_{jk}^i(y) := dy^i \left( \nabla_{\frac{\partial}{\partial y^k}} \frac{\partial}{\partial y^j} \right) = \frac{\partial y^i}{\partial x^q} dx^q \left( \nabla_{\frac{\partial x^p}{\partial y^k} \frac{\partial}{\partial x^p}} \frac{\partial x^s}{\partial y^j} \frac{\partial}{\partial x^s} \right)$$

because,

since  $y^i$  is a function, i.e.

$y^i = y^i(x)$ , and  $d$  acts on a function, then  $dy^i = dy^i(x) = \frac{\partial y^i}{\partial x^q} dx^q$ , and since

$$\frac{\partial f(y)}{\partial y^k} = \frac{\partial f(x(y))}{\partial y^k} = \frac{\partial x^p}{\partial y^k} \frac{\partial f(x)}{\partial x^p} \text{ so } \frac{\partial}{\partial y^k} = \frac{\partial x^p}{\partial y^k} \frac{\partial}{\partial x^p} \text{ by chain rule}$$

Note  $\nabla_{fX}$  is  $C^\infty$ -linear for  $fX$

covector  $dy^i$  is  $C^\infty$ -linear in its argument

$$\begin{aligned} \implies \Gamma_{jk}^i(y) &= \frac{\partial y^i}{\partial x^q} dx^q \left( \frac{\partial x^p}{\partial y^k} \left[ \left( \nabla_{\frac{\partial}{\partial x^p}} \frac{\partial x^s}{\partial y^j} \right) \frac{\partial}{\partial x^s} + \frac{\partial x^s}{\partial y^j} \left( \nabla_{\frac{\partial}{\partial x^p}} \frac{\partial}{\partial x^s} \right) \right] \right) = \\ &= \frac{\partial y^i}{\partial x^q} \underbrace{\frac{\partial x^p}{\partial y^k} \frac{\partial}{\partial x^p}}_{\frac{\partial}{\partial y^k}} \frac{\partial x^s}{\partial y^j} \delta_s^q + \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^k} \frac{\partial x^s}{\partial y^j} \Gamma_{sp}^q(x) \end{aligned}$$



In summary (Schuller says to commit this to memory):

$$(55) \quad \boxed{\Gamma_{jk}^i(y) = \frac{\partial y^i}{\partial x^q} \frac{\partial^2 x^q}{\partial y^j \partial y^k} + \frac{\partial y^i}{\partial x^q} \frac{\partial x^s}{\partial y^j} \frac{\partial x^p}{\partial y^k} \Gamma_{sp}^q(x)}$$

Eq. (55) is the change of connection coefficient function under the change of chart  $(U \cap V, x) \rightarrow (U \cap V, y)$  Notice especially the second-order partial derivative term: for any linear transformation, this term will "not be felt" (i.e. will be zero), but otherwise for a nonlinear transformation of coordinates, you'll "feel it."

14.1.2. *Normal Coordinates.* Let  $p \in M$  of  $(M, \mathcal{O}, \mathcal{A}, \nabla)$

Then one can construct a chart  $(U, x)$  with  $p \in U$  such that

$$(56) \quad \Gamma(x)^i_{(jk)}(p) = 0$$

at the point  $p$ . **Not** necessarily in any neighborhood. Note the  $(jk)$  subscript, meaning the symmetric part vanishes, i.e. equals zero. So Eq. 56 says we can choose a coordinate system such that at a point the symmetric part of the Christoffel Symbol is zero.

Recall the definition of symmetric and antisymmetric parts of some tensor  $T$ . Define the symmetric and antisymmetric parts, respectively:

$$T_{(ab)} = \frac{1}{2}(T_{ab} + T_{ba})$$

$$T_{[ab]} = \frac{1}{2}(T_{ab} - T_{ba})$$

*Proof.* Let  $(V, y)$  be any chart (with)  $p \in V$ .

Thus, in general:  $\Gamma(y)^i_{jk} \neq 0$

Then consider a new chart  $(U, x)$  to which one transits by virtue of

$$(x \circ y^{-1})^i(\alpha^1, \dots, \alpha^d) := \alpha^i - \frac{1}{2} \Gamma(y)^i_{(jk)}(p) \alpha^j \alpha^k$$

for  $p = x^{-1}(\alpha^1, \dots, \alpha^d)$

$$\left( \frac{\partial x^i}{\partial y^j} \right)_p = \partial_j(x^i \circ y^{-1}) = \delta_j^i - \Gamma(y)^i_{mj}(p) \alpha^m|_{\alpha=0} = \delta_j^i$$

$$\frac{\partial x^i}{\partial y^k \partial y^j}(p) = -\Gamma(y)^i_{(kj)}(p)$$

where the partial derivative acting on the second term with the Christoffel symbol is because of the following:

$$\begin{aligned} \frac{\partial}{\partial y^j} \frac{1}{2} \Gamma(y)^i_{(lk)}(p) \alpha^l \alpha^k &= \frac{\partial}{\partial \alpha^j} \frac{1}{2} \Gamma(y)^i_{(lk)}(p) \alpha^l \alpha^k = \frac{1}{2} \Gamma(y)^i_{(lk)}(p) \delta_j^l \alpha^k + \frac{1}{2} \Gamma(y)^i_{(lk)}(p) \alpha^l \delta_j^k = \\ &= \frac{1}{2} \Gamma(y)^i_{(jk)}(p) \alpha^k + \frac{1}{2} \Gamma(y)^i_{(lj)}(p) \alpha^l = \Gamma(y)^i_{(jm)}(p) \alpha^m \end{aligned}$$

where the last step is by the symmetry property of the symmetric part of  $\Gamma_{jm}^i, \Gamma_{(jm)}^i$ .

Then use Eq. (55), the change of Christoffel symbols equation we just derived:

$$\Gamma_{jk}^i(y) = \frac{\partial y^i}{\partial x^q} \frac{\partial^2 x^q}{\partial y^j \partial y^k} + \frac{\partial y^i}{\partial x^q} \frac{\partial x^s}{\partial y^j} \frac{\partial x^p}{\partial y^k} \Gamma_{sp}^q(x) = \frac{\partial y^i}{\partial x^q} \left( -\Gamma(y)^q_{(kj)}(p) \right) + \frac{\partial y^i}{\partial x^q} \Gamma_{jk}^q(x)$$

If  $\frac{\partial y^i}{\partial x^q} = -\delta^i_q$  then we obtain

$$\Gamma_{jk}^i(y) = +\Gamma(y)^i_{(jk)} - \Gamma_{jk}^i(x) \text{ or}$$

$$\Gamma_{jk}^i(x) = -\Gamma_{jk}^i(y) + \Gamma_{(jk)}^i(y) = \Gamma(y)^i_{[kj]}$$

Compare this against what Schuller (2015) [38] obtains:

$$\begin{aligned} \implies \Gamma(x)^i_{jk}(p) &= \Gamma(y)^i_{(jk)}(p) - \Gamma(y)^i_{kj}(p) = \\ &= \Gamma(y)^i_{[jk]}(p) = T(y)^i_{jk} \end{aligned}$$

Terminology:  $(U, x)$  is called a **normal coordinate chart** of  $\nabla$  at  $\mathbf{p} \in \mathbf{M}$ .

□

cf. pp. 164, 4.1 *Connections in Vector Bundles*, Jost (2017) [44]

Let  $x_0 \in M$ , let open neighborhood  $U$  of  $x_0$  s.t.  $\exists$  chart for  $M$ ,  $\exists$  bundle chart for  $E$  on  $U$ . Thus,  $\exists$  coordinate vector fields  $\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^d}$  and through identification

$$E|_U \simeq U \times \mathbb{R}^n$$

$n$  = fiber dim. of  $E$ .

For a basis of  $\mathbb{R}^n$  yields basis  $\mu_1 \dots \mu_n \equiv e_q \dots e_n$  of sections of  $E|_U \equiv \Gamma(E|_U)$ . Define Christoffel symbols  $\Gamma_{ij}^k$  ( $j, k = 1 \dots n, i = 1 \dots d$ ) (pp. 164, Jost (2017) [44])

$$(57) \quad \begin{aligned} D_{\frac{\partial}{\partial x^i}} \mu_j &:= \Gamma_{ij}^k e_k \quad (4.1.5, \text{Jost (2017)[44]}) \\ &\equiv D_{\frac{\partial}{\partial x^j}} e_i \Gamma_{ij}^k e_k \quad (\text{Schuller's notation}) \end{aligned}$$

14.2. **Parallel Transport.** Let  $\mu \equiv s \in \Gamma(E)$ , ( $\mu$  is the notation of Jost (2017) [44]) locally  $s(y) = a^k(y) e_k(y)$ , let  $c(t) =$  smooth curve in  $U$ .

Let  $s(t) := s(c(t))$ , to define a section of  $E$  along  $c$ .

Let  $X(t) = \dot{c}(t) = \dot{c}^i(t) \frac{\partial}{\partial x^i}$

Then

$$\begin{aligned} D_{X(t)} s(t) &= D_{\dot{c}^i \frac{\partial}{\partial x^i}} s(t) = \dot{c}^i D_{\frac{\partial}{\partial x^i}} a^k e_k = \dot{c}^i \left( \frac{\partial a^k}{\partial x^i} \right) e_k + \dot{c}^i a^k \Gamma_{ki}^q e_q = \\ &= (\dot{a}^k + \dot{c}^j a^i \Gamma_{ij}^k) e_k \end{aligned}$$

where we had used chain rule:

$$\dot{a}^k \equiv \frac{da^k}{dt} = \frac{\partial a^k}{\partial x^i}(x(t)) \frac{dc^i}{dt}$$

So if  $D_{x(t)} s(t) = 0$ , we have  $k = 1 \dots n$  first order linear equations, solving for  $a^1, \dots, a^n(t)$ .

So by a fundamental theorem for ordinary differential equations (ODEs), for given initial values  $s(0) \in E_{c(0)}$ ,  $\exists!$  solution to  $D_{x(t)} s(t) = 0$ .

**Definition 39** (Parallel transport of section of vector bundle, Def. 4.1.2, Jost (2017)[44]). *The solution set of  $D_{x(t)} s(t) = 0$  is called the parallel transport of  $s(0)$  along curve  $c$ .*

Given  $X \in T_x M$ , let curve  $c$  in  $M$  with  $c(0) = x$ ,  $\dot{c}(0) = X$ ,  $\forall s \in \Gamma(E)$ , then

$$D_X s := \lim_{t \rightarrow 0} \frac{P_{c,t}(s(c(t))) - s(c(0))}{t}$$

EY: This seems to be the covariant derivative.

pp. 165 Jost (2017) [44] claims that parallel transport on Riemannian manifold can obtain covariant derivative.

*Proof.* Let  $P_{c,t} : E_{c(t)} \rightarrow E_{c(0)}$  be identification by parallel transport along  $c$ .

Select a basis of parallel sections  $\mu_1(t), \dots, \mu_n(t) \equiv e_1(t) \dots e_n(t)$  of  $E$  along  $c$ , i.e.

$$D_{\dot{c}(t)} e_j(t) = 0 \quad \forall j = 1 \dots n$$

$\forall$  arbitrary section  $s$  of  $E$  along  $c$ , then

$$s(t) = a^k(t) e_k(t)$$

and for  $X = \dot{c}(0)$ ,

$$D_X s(t) = D_X(a^k(t)e_k(t)) = D_X(a^k(t))e_k(t) + a^k(t)D_X e_k(t) = \dot{a}^k(t)e_k(t)$$

where  $D_X e_k(t) = 0$  by parallel transport.

So

$$(D_X s)(c(0)) = \lim_{t \rightarrow 0} \frac{a^k(t) - a^k(0)}{t} e_k(0) = \lim_{t \rightarrow 0} \frac{P_{c,t}(s(t)) - s(0)}{t}$$

Consider tangent space at pt.  $\psi \equiv s$  to the total space  $E$  of a vector bundle  $T_s E$ .

Inside  $T_s E$ , there's a distinguished subspace, namely the tangent space to the fiber  $E_x$  containing  $s$  ( $x = \pi(s)$ ).

This space is called vertical space  $V_s$ .

However, there's no distinguished "horizontal space"  $H_s$  complementary to  $V_s$ , i.e.  $T_s E = V_s \oplus H_s$ .

If we have covariant derivative  $D$ , we can parallel transport  $s$ ,  $\forall X \in T_x M$  along curve  $c(t)$  with

$$c(0) = x, \dot{c}(0) = X$$

Thus,  $\forall X$ , we obtain a curve  $s(t)$  in  $E$ .

subspace of  $T_s E$  spanned by all tangent vectors to  $E$  at  $s$  of the form

$$\left. \frac{d}{dt} s(t) \right|_{t=0}$$

then is the horizontal space  $H_s$ .

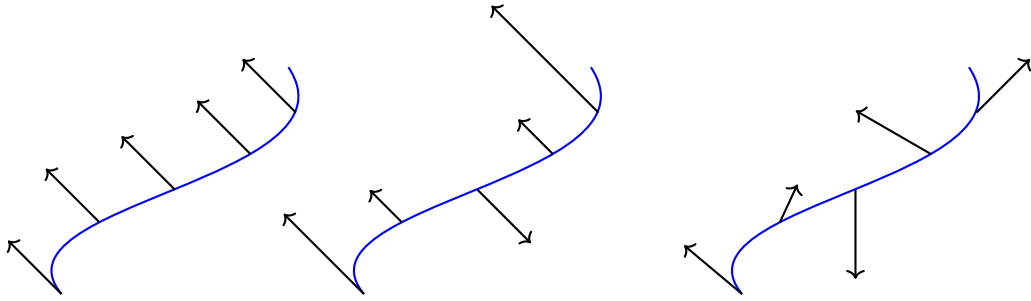
14.2.1. *Parallelity of vector fields.* cf. Lecture 8, Schuller (2015) [38].

Assume smooth affine manifold  $(M, \mathcal{O}, \mathcal{A}, \nabla)$  with connection.

**Definition 40** (Parallely Transported). *vector field  $X$  on  $M$  is **parallely transported** along smooth curve  $\gamma : \mathbb{R} \rightarrow M$  if*

$$(58) \quad \nabla_{v_\gamma} X = 0$$

**Definition 41** (Parallel). *vector field  $X$  in  $M$  is **parallel** along curve  $\gamma : \mathbb{R} \rightarrow M$  if  $\nabla_{v_\gamma} = \mu \cdot X$*



The left drawing is a parallely transported vector field, the middle drawing is a parallel vector field. For the middle drawing, it's important that the vector field vanishes in-between points when it points "up" vs. "down" because  $\mu : \mathbb{R} \rightarrow \mathbb{R}$  is smooth. The right drawing is not even parallel.

14.2.2. *Autoparallely Transported Curves.*

**Definition 42** (Autoparallely Transported). *smooth curve  $\gamma : \mathbb{R} \rightarrow M$  **autoparallely transported** if*

$$(59) \quad \nabla_{v_\gamma} v_\gamma = 0$$

**Definition 43.** *smooth curve  $\gamma : \mathbb{R} \rightarrow M$  is autoparallel*

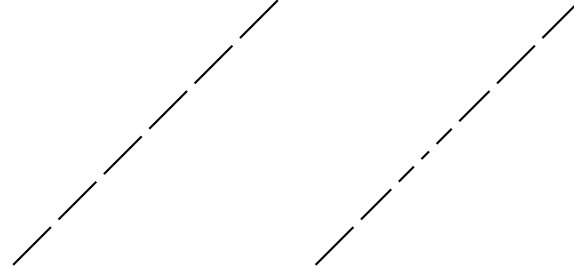
$$(60) \quad \nabla_{v_\gamma} v_\gamma = \mu v_\gamma$$

for some smooth  $\mu : \mathbb{R} \rightarrow \mathbb{R}$

For autoparallel, it means we move in the straightest line. We are not talking about the shortest line because we do not have a notion of distance.

Note that the straightest line might not actually look straight when "viewed from above." That is, if we embed the manifold into a higher dimensional one, and then just look at the curve, it might be curved. For example, on a sphere, a straight line traces out a portion of a circle around the sphere. This doesn't look straight in the (higher, Euclidean) embedding, but *on the surface* of the sphere, it's the straightest line.

□



Left is an autoparallely transported curve, and the right is just an autoparallel.

Consider

$$\dot{\gamma}(t) = v(\gamma(t)) = v^i(\gamma^1(t) \dots \gamma^n(t)) \frac{\partial}{\partial x^i} \in T_{\gamma(t)} U$$

Then

$$\frac{d}{dt} \dot{\gamma}(t) = \frac{d}{dt} v^i(\gamma^1(t) \dots \gamma^n(t)) \frac{\partial}{\partial x^i} = \frac{d}{dt} v^i(x^1 \dots x^n) \frac{\partial}{\partial x^i} = \frac{\partial v^i}{\partial x^j}(x) \dot{x}^j \frac{\partial}{\partial x^i} = \dot{\gamma}^j \frac{\partial \dot{\gamma}^i}{\partial x^j} \frac{\partial}{\partial x^i}$$

So

$$\nabla_{v_\gamma} v_\gamma = \nabla_{v^j \frac{\partial}{\partial x^j}} v^i \frac{\partial}{\partial x^i} = v^j \left[ \frac{\partial v^i}{\partial x^j} \frac{\partial}{\partial x^i} + v^i \Gamma_{ij}^k \frac{\partial}{\partial x^k} \right] = \ddot{\gamma}(t) + \dot{\gamma}^i \dot{\gamma}^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} \text{ i.e.}$$

(61)

$$\nabla_{v_\gamma} v_\gamma = (\ddot{\gamma}^k(t) + \dot{\gamma}^i \dot{\gamma}^j \Gamma_{ij}^k) \frac{\partial}{\partial x^k}$$

with  $\Gamma_{ij}^k = \Gamma_{ij}^k(\gamma(t))$ .

14.2.3. *Torsion.* cf. Lecture 8, Schuller (2015) [38]

**Definition 44** (torsion). *The **torsion** of a connection  $\nabla$  is the  $(1,2)$ -**tensor field***

$$(62) \quad T(\omega, X, Y) := \omega(\nabla_X Y - \nabla_Y X - [X, Y])$$

where  $[\cdot, \cdot] : \Gamma(TM) \times \Gamma(TM) \rightarrow \Gamma(TM)$  is the commutator on  $\Gamma(TM)$ , that yields a vector field,  $[X, Y]$  defined by

$$(63) \quad [X, Y]f := X(Yf) - Y(Xf)$$

**Theorem 20.**  *$T$  is  $C^\infty$ -linear in each entry*

*Proof.* check

$$T(f \cdot \omega, X, Y) = f \cdot \omega(\dots) = fT(\omega, X, Y)$$

$$T(\omega + \psi, X, Y) = \dots = T(\omega, X, Y) + T(\psi, X, Y)$$

$$T(\omega, fX, Y) = \omega(\nabla_{fX} Y - \nabla_Y(fX) - [fX, Y]) =$$

$$= \omega(f\nabla_X Y - (Yf)X - f\nabla_Y X - f[X, Y] + (Yf)X) = fT(\omega, X, Y)$$

since

$$[fX, Y]g = fX(Yg) - Y(fXg) = fX(Yg) - (Yf)(Xg) - fY(Xg) =$$

$$= f \cdot [X, Y]g - (Yf)Xg = (f \cdot [X, Y] - (Yf)X)g$$

$$T(\omega, X, Y) = -T(\omega, Y, X) \quad \checkmark$$

□

**Definition 45.** An affine manifold  $(M, \mathcal{O}, \mathcal{A}, \nabla)$  is called torsion-free if  $T = 0$ . This can also be written as

$$(64) \quad \nabla_X Y - \nabla_Y X = [X, Y] \quad \forall X, Y \in \Gamma(TM)$$

In a chart

$$\begin{aligned} T_{ab}^i &:= T\left(dx^i, \frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right) = dx^i(\dots) \\ &= \Gamma_{ba}^i - \Gamma_{ab}^i = -2\Gamma_{[ab]}^i \end{aligned}$$

since

$$\begin{aligned} \left[\frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right] &= 0 \text{ and} \\ \nabla_{\frac{\partial}{\partial x^a}} \frac{\partial}{\partial x^b} &= \Gamma_{ba}^k \\ \nabla_{\frac{\partial}{\partial x^b}} \frac{\partial}{\partial x^a} &= \Gamma_{ab}^k \end{aligned}$$

So what this means is that if the torsion is 0, i.e.  $T = T_{ab}^i = 0$ , then the *anti-symmetric* part of the connection coefficients, i.e. Christoffel symbols, is zero,  $\Gamma_{[ij]}^k = 0$ . So the Christoffel symbols only have a symmetric part and is thus symmetric in indices  $i, j$ .

From now on, in these lectures, we only use torsion-free connections.

Suppose  $X = X^i \frac{\partial}{\partial x^i} \in \mathfrak{X}(M)$ ,  $s = s^i e_i \in \Gamma(E)$ ,

$$D_X s = X^j D_{\frac{\partial}{\partial x^j}} s^i e_i = X^j \left( \frac{\partial s^i}{\partial x^j} e_i + s^i D_{\frac{\partial}{\partial x^j}} e_i \right) = X^j \left( \frac{\partial s^k}{\partial x^j} + s^i \Gamma_{ij}^k \right) e_k$$

Consider  $\Gamma_{ij}^k$  as an  $(n \times n)$  matrix valued 1-form on  $U \quad \forall j = 1 \dots d$ .

$$(65) \quad \Gamma_{ij}^k \in \Gamma(\mathfrak{gl}(n, \mathbb{R}) \otimes T^*M|_U)$$

cf. Eq. (4.1.12), pp. 166, Jost (2017) [44].

Abstract this to

$$D = d + A$$

where  $d$  is the exterior derivative and  $A \in \Gamma(\mathfrak{gl}(n, \mathbb{R}) \otimes T^*M|_U)$ .

Given  $a^j e_j$

$$D(a^j e_j) = e_k \otimes (da^k) + A_j a \otimes dx^j$$

where  $d(a^k)(X) = (da^k)(X) = X(a^k) = X^j \frac{\partial a^k}{\partial x^j}$ .

Let  $(U_\alpha)_{\alpha \in A}$  be covering of  $M$  by open sets which the bundle is trivial, with transition maps

$$\varphi_{\beta\alpha} : U_\alpha \cap U_\beta \rightarrow GL(n; \mathbb{R})$$

Let section  $s$  be represented by  $s_\alpha$  on  $U_\alpha$ .

Thus,

$$s_\beta = \varphi_{\beta\alpha} s_\alpha \text{ on } U_\alpha \cap U_\beta$$

But then

$$\begin{aligned} \varphi_{\beta\alpha}(d + A_\alpha)s_\alpha &= (d + A_\beta)s_\beta \text{ on } U_\alpha \cap U_\beta \\ \varphi_{\beta\alpha} ds_\alpha + \varphi_{\beta\alpha} A_\alpha s_\alpha &= d\varphi_{\beta\alpha} s_\alpha + A_\beta \varphi_{\beta\alpha} s_\alpha \\ A_\alpha s_\alpha &= -ds_\alpha + \varphi_{\beta\alpha}^{-1} d\varphi_{\beta\alpha} s_\alpha + \varphi_{\beta\alpha}^{-1} A_\beta \varphi_{\beta\alpha} s_\alpha \end{aligned}$$

14.3. **Curvature.** cf. Lecture 8, Schuller (2015) [38].

**Definition 46** (Riemann Curvature). The **Riemann curvature** of a connection  $\nabla$  is the  $(1,3)$ -tensor field

$$(66) \quad \text{Riem}(\omega, Z, X, Y) := \omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z)$$

**Definition 47** (Ricci Curvature). Let *Riem* be the Riemann curvature tensor of connection  $\nabla$ .

Define the Ricci curvature tensor as  $(0,2)$ -tensor field

$$(67) \quad \text{Ric}(X, Y) := \text{Riem}(f^k, Y, X, e_k)$$

where  $f^l(e_m) = \delta_m^l$

In terms of components, the Ricci curvature tensor is given by

$$(68) \quad \text{Ric}_{ab} = \text{Riem}_{amb}^m$$

**Theorem 21.** *Riem* is  $C^\infty$ -linear in each slot.

*Proof.*

since  $\text{Riem}(\omega, Z, X, Y) = \omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z)$

$$\text{Riem}(f\omega, Z, X, Y) = f\omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) = f\text{Riem}(\omega, Z, X, Y)$$

$$\text{Riem}(\omega + \sigma, Z, X, Y) = (\omega + \sigma)(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z) = \text{Riem}(\omega, Z, X, Y) + \text{Riem}(\sigma, Z, X, Y)$$

$$\begin{aligned} \text{Riem}(\omega, Z, fX, Y) &= \omega(f\nabla_X \nabla_Y Z - Y(f)\nabla_X Z - f\nabla_Y \nabla_X Z - \nabla_{f[X,Y] - (Yf)X} Z) = \\ &= \omega(f\nabla_X \nabla_Y Z - Y(f)\nabla_X Z - f\nabla_Y \nabla_X Z - f\nabla_{[X,Y]} Z + (Yf)\nabla_X Z) = f\text{Riem}(\omega, Z, X, Y) \end{aligned}$$

where the second and last term cancel each other out in the second to last expression. □

The Riemann curvature tensor *Riem* is antisymmetric in its last 2 entries. This is because

$$\begin{aligned} \text{Riem}(\omega, Z, Y, X) &= \omega(\nabla_Y \nabla_X Z - \nabla_X \nabla_Y Z - \nabla_{[Y,X]} Z) = -\omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z + \nabla_{-[X,Y]} Z) = \\ &= -\text{Riem}(\omega, Z, X, Y) \end{aligned}$$

since  $\nabla_{-[X,Y]} Z = -\nabla_{[X,Y]} Z$

Therefore, *Riem* has  $d^3(d-1)/2$  independent components. This is because  $d^4$  is the total number of components (there are 4 indices, each index running from 1 to d) if the indices are completely independent of each other.

Since two entries are antisymmetric, imagine a  $d \times d$  matrix with only the entries in the "top half triangle" being independent. It excludes the diagonal entries (they'd be all 0). There are  $\frac{d(d-1)}{2}$  entries. For each of those antisymmetric entries, there are still  $d^2$  ways to choose the other 2 entries of the *Riem*. Thus, the number of independent components are

$$d^4 = \left( \frac{d(d-1)}{2} \right) d^2 = d^4 - \frac{(d^2 - d)d^2}{2} = d^4 - \frac{d^4 - d^3}{2} = d^3(d-1)/2$$

## Part 9. Holonomy

**Definition 48** (Conlon, 10.1.2). If  $X, Y \in \mathfrak{X}(M)$ ,  $M \subset \mathbb{R}^m$ , **Levi-Civita connection** on  $M \subset \mathbb{R}^m$

$$\nabla : \mathfrak{X}(M) : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$$

(69)

$$\nabla_X Y := p(D_X Y)$$

with

$$\begin{aligned} D_X Y &:= \sum_{j=1}^m X(Y^j) \frac{\partial}{\partial x^j} = \sum_{i,j=1}^m X^i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial x^j} & \forall X &= \sum_{i=1}^m X^i \frac{\partial}{\partial x^i}, \\ & & \forall Y &= \sum_{i=1}^m Y^i \frac{\partial}{\partial x^i} \end{aligned}$$

$$\nabla_{fX}Y = f(D_{fX}Y) = p(fD_XY) = fpD_XY = f\nabla_XY$$

$$\nabla_X fY = p(D_X fY) = p\left(\sum_{i,j=1}^m \left(X^i f \frac{\partial Y^j}{\partial x^i} + X^i Y^j \frac{\partial f}{\partial x^i}\right) \frac{\partial}{\partial x^j}\right) = f\nabla_XY + p\sum_{j=1}^m X(f)Y^j \frac{\partial}{\partial x^j} = f\nabla_XY + X(f)p(Y)$$

**Definition 49** (Conlon, 10.1.4; Christoffel symbols).

$$(70) \quad \begin{aligned} \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} &= \Gamma_{ij}^k \frac{\partial}{\partial x^k} & (\text{Conlon's notation}) \\ \nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} &= \Gamma_{ij}^k \frac{\partial}{\partial x^k} & (F. Schuller's notation) \end{aligned}$$

**Definition 50** (torsion).

$$(71) \quad \begin{aligned} T : \mathfrak{X}(M) \times \mathfrak{X}(M) &\rightarrow \mathfrak{X}(M) \\ T(X, Y) &= \nabla_X Y - \nabla_Y X - [X, Y] \end{aligned}$$

If  $T = 0$ ,  $\nabla$  torsion-free or symmetric.

$$T(fX, Y) = f\nabla_X Y - (f\nabla_Y X + Y(f)X) - \{(fXY - (Y(f)X + fYX))\} = fT(X, Y)$$

$$T(X, fY) = f\nabla_X Y + X(f)Y - f\nabla_Y X - \{((X(f)Y + fXY) - fYX)\} = fT(X, Y)$$

Thus,  $T(X, Y)$   $C^\infty(M)$ -bilinear.

$$T \in \tau_1^2(M).$$

$$T(v, w) \in T_x M \text{ defined, } \forall v, w \in T_x M, \forall x \in M.$$

Thus, torsion is a **tensor**.

**Exercise 10.1.7 Conlon (2008)[19]** . .

If  $T(X, Y) = 0$ ,

$$T(e_i, e_j) = \Gamma_{ji}^k e_k - \Gamma_{ij}^k e_k - 0 = 0 \implies \Gamma_{ji}^k = \Gamma_{ij}^k$$

$$\text{If } \Gamma_{ij}^k = \Gamma_{ji}^k, \quad T(e_i, e_j) = 0.$$

**Exercise 10.1.8, Conlon (2008)[19].**

If  $M \subset \mathbb{R}^m$  smoothly embedded submanifold,  $\forall \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^i} \in T_x M$ , spanning  $T_x M$ , consider  $\frac{\partial}{\partial x^j} = X_j^k \frac{\partial}{\partial \tilde{x}^k}$ ,  $\frac{\partial}{\partial x^i} = X_i^k(\tilde{x}) \frac{\partial}{\partial \tilde{x}^k}$

$$\nabla_{\frac{\partial}{\partial x^j}} \frac{\partial}{\partial x^i} = pD_{X_j^k \frac{\partial}{\partial \tilde{x}^k}} X_i^l \frac{\partial}{\partial \tilde{x}^l} = p\left(X_j^k \frac{\partial X_i^l}{\partial \tilde{x}^k} \frac{\partial}{\partial \tilde{x}^l}\right) = X_j^k p\left(\frac{\partial X_i^l}{\partial \tilde{x}^k} \frac{\partial}{\partial \tilde{x}^l}\right)$$

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = X_i^k p\left(\frac{\partial X_j^l}{\partial \tilde{x}^k} \frac{\partial}{\partial \tilde{x}^l}\right)$$

If  $X \in \mathfrak{X}(M)$ , smooth  $s : [a, b] \rightarrow M$ ,  
then  $\forall s(t)$ ,

$$X'_{s(t)} = \nabla_{\dot{s}(t)} X \in T_{s(t)} M$$

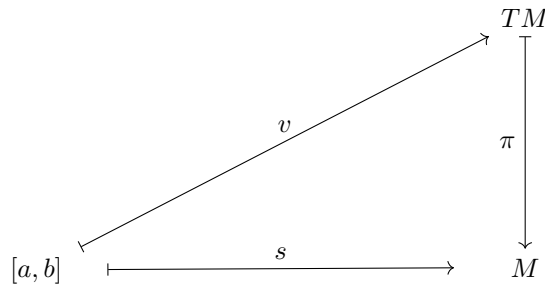
In fact, it's often natural to consider fields  $X_{s(t)}$  along  $s$ , parametrized by parameter  $t$ , allowing

$$X_{s(t_1)} \neq X_{s(t_2)}$$

each of  $s(t_1) = s(t_2)$ .

**Definition 51** (10.1.9). *Let smooth  $s : [a, b] \rightarrow M$ .*

*Vector field along  $s$  is smooth  $v : [a, b] \rightarrow TM$  s.t.*



*commutes.*

*Note that  $v \in \mathfrak{X}(s) \subset \mathfrak{X}(M)$*

e.g.  $(Y|s)(t) = Y_{s(t)}$ , restriction of  $Y \in \mathfrak{X}(M)$  to  $s$ .

e.g.  $\dot{s}(t) \in \mathfrak{X}(M)$ .

$\forall v, w \in \mathfrak{X}(s)$ ,  $v + w \in \mathfrak{X}(s)$ ,

$$(fv + gw)(t) := (f(s(t)) + g(s(t)))v(t) = f(s(t))v(t) + g(s(t))v(t) = (f + g)v(t)$$

Likewise,

$$f(v + w) = fv + fw$$

$\mathfrak{X}(s)$  is a real vector space and  $C^\infty[a, b]$ -module.

**Definition 52** (10.1.10). *Let conection  $\nabla$  on  $M$ .*

***Associated covariant derivative*** is operator

$$\frac{\nabla}{dt} \mathfrak{X}(s) \rightarrow \mathfrak{X}(s)$$

$\forall$  smooth  $s$  on  $M$ , s.t.

- (1)  $\frac{\nabla}{dt}$   $\mathbb{R}$ -linear
- (2)  $\left(\frac{\nabla}{dt}\right)(fv) = \frac{df}{dt}v + f\frac{\nabla}{dt}v$ ,  $\forall f \in C^\infty[a, b]$ ,  $\forall v \in \mathfrak{X}(s)$
- (3) If  $Y \in \mathfrak{X}(M)$ , then

$$\frac{\nabla}{dt}(Y|s)(t) = \nabla_{\dot{s}(t)} Y \in T_{s(t)} M, \quad a \leq t \leq b$$

**Theorem 22** (Conlon Thm. 10.1.11[19]).  $\forall$  connection  $\nabla$  on  $M$ ,  $\exists!$  associated covariant derivative  $\frac{\nabla}{dt}$

*Proof.* Consider arbitrary coordinate chart  $(U, x^1 \dots x^n)$ .

Consider smooth curve  $s : [a, b] \rightarrow U$ .

Let  $v \in \mathfrak{X}(s)$ ,  $v(t) = v^i(t) \frac{\partial}{\partial x^i}$ ;  $\dot{s}(t) = s^j \frac{\partial}{\partial x^j}$ .

$$\frac{\nabla v}{dt} = \frac{dv^i(t)}{dt} \frac{\partial}{\partial x^i} + v^i(t) \frac{\nabla}{dt} \frac{\partial}{\partial x^i} = \frac{dv^i}{dt} \frac{\partial}{\partial x^i} + v^i \nabla_{\dot{s}(t)} \frac{\partial}{\partial x^i} = \dot{v}^i \frac{\partial}{\partial x^i} + v^i \dot{s}^j \Gamma_{ij}^k \frac{\partial}{\partial x^k} = (\dot{v}^k + v^i \dot{s}^j \Gamma_{ij}^k) \frac{\partial}{\partial x^k}$$

This is an explicit, local formula in terms of connection, proving uniqueness.

Existence:  $\forall$  coordinate chart  $(U, x^1 \dots x^n)$ ,  $(\dot{v}^k + v^i \dot{s}^j \Gamma_{ij}^k) \frac{\partial}{\partial x^k} =: \frac{\nabla v}{dt}$ .

$$\frac{\nabla}{dt}(fv) = \dot{f}v^k + f\dot{v}^k + fv^i \dot{s}^j = \dot{f}v + f \frac{\nabla v}{dt}$$

If  $f$  constant, then  $\frac{\nabla}{dt}$  is  $\mathbb{R}$ -linear.

□

**Definition 53** (10.1.12 Conlon (2008)[19]). *Let  $(M, \nabla)$ . Let  $v \in \mathfrak{X}(s)$  for smooth  $s : [a, b] \rightarrow M$ .*

*If  $\frac{\nabla v}{dt} \equiv 0$  on  $s$ , then  $v$  is **parallel** along  $s$ .*

**Theorem 23** (10.1.13). *Let  $(M, \nabla)$ , smooth  $s : [a, b] \rightarrow M$ ,  $c \in [a, b]$ ,  $v_0 \in T_{s(c)}M$ .*

*Then  $\exists!$  parallel field  $v \in \mathfrak{X}(s)$  s.t.  $v(c) = v_0$ .*

*$v$  parallel transport along  $s$ .*

*Proof.*

$$\dot{s}(t) = \dot{s}^j(t)e_j$$

$$v(t) = v^i(t)e_i$$

$$v_0 = a^i e_i$$

$$0 = \left( \frac{dv^k}{dt}(t) + v^i(t)\dot{s}^j(t)\Gamma_{ij}^k(s(t)) \right) e_k$$

or equivalently

$$(72) \quad \frac{dv^k}{dt} = -v^i \dot{s}^j \Gamma_{ij}^k, \quad 1 \leq k \leq n \quad (10.1)$$

with initial conditions  $v^k(c) = a^k$ ,  $1 \leq k \leq n$ .

By existence and uniqueness of solutions of O.D.E.

$\exists \epsilon > 0$  s.t.  $\exists!$  solutions  $v^k(t)$ . For  $c - \epsilon < t < c + \epsilon$ .

In fact, these ODEs being linear in  $v^k$ , by ODE theory (Appendix C, Thm. C.4.1).

$\nexists$  restriction on  $\epsilon$ , so  $\exists! v^k(t) \quad \forall t \in [a, b]$ ,  $1 \leq k \leq n$

□

**14.4. Principal bundle, vector bundle case for parallel transport.** Recall the 2 different forms or viewpoints for Lie-algebra valued 1-forms, or vector-valued 1-forms, or sections of 1-form-valued endomorphisms:

$$\omega_{i\mu}^k dx^\mu \equiv \omega_i^k \in \Omega^1(M, \mathfrak{gl}(n, \mathbb{F})) = \Gamma(\mathfrak{gl}(n, \mathbb{R} \otimes T^*M|_U))$$

for  $i, k = 1 \dots n = \dim E$ .

$\mu = 1 \dots d = \dim E$

Now

$$D_X \mu = X^\mu D_{\frac{\partial}{\partial x^\mu}} \mu = X^\mu \left[ \left( \frac{\partial}{\partial x^\mu} \mu^k \right) e_k + \mu^i \omega_{i\mu}^k e_k \right] = (X(\mu^k) + \mu^i \omega_i^k(X)) e_k = (d\mu^k(X) + \mu^i \omega_i^k(X)) e_k$$

So then define

$$(73) \quad \begin{aligned} D : \Gamma(E) &\rightarrow \Gamma(E) \otimes \Gamma(T^*M) \\ D\mu &= D(\mu^i e_i) = e_k (d\mu^k + \mu^i \omega_i^k) \equiv (d + A)\mu \end{aligned}$$

Also,  $D$  can be defined for this case:

$$D : \Gamma(\text{End}(E)) \rightarrow \Gamma(\text{End}E) \otimes \Gamma(T^*M)$$

Let  $\sigma = \sigma_j^i e_i \otimes e^j \in \Gamma(\text{End}(E))$

$$(74) \quad \begin{aligned} D\sigma &= D(\sigma_j^i e_i) \otimes e^j + \sigma_j^i e_i \otimes D^* e^j = (d\sigma_j^k + \sigma^i A_i^k) e_k \otimes e^j + \sigma_j^i e_i \otimes (A^*)^j_k e^k = \\ &= (d\sigma_j^k + \sigma_j^i A_i^k) e_k \otimes e^j + \sigma_j^i e_j \otimes (-A^i_j) e^j = (d\sigma_j^k + [A, \sigma]_j^k) e_k \otimes e^j \end{aligned}$$

cf. Def. 4.1.4 of Jost (2017) [44], pp. 138.

For  $\mu \in \Gamma(E)$ , smooth  $s : [a, b] \rightarrow M$ ,  $X(t) = \dot{s}(t)$ ,

$$(75) \quad D_{\dot{s}(t)} \mu = \dot{s}^\mu D_{\frac{\partial}{\partial x^\mu}} \mu = \dot{s}^\mu \left[ \frac{\partial \mu^k}{\partial x^\mu} e_k + \mu^i \omega_{i\mu}^k e_k \right] = \left[ \dot{s}^\mu \frac{\partial \mu^k}{\partial x^\mu} + \dot{s}^\mu \mu^i \omega_{i\mu}^k \right] e_k = \frac{d}{dt} \mu(s(t)) + \mu^i \dot{s}^\mu \omega_{i\mu}^k e_k$$

Let  $D_{\dot{s}(t)} \mu = 0$ . Then,

$$(76) \quad \frac{d}{dt} \mu(s(t)) = -\mu^i \dot{s}^\mu \omega_{i\mu}^k e_k$$

Recall, given vector bundle  $E \xrightarrow{\pi} N$ , given  $\varphi : M \rightarrow N$ , then pullback

$$(77) \quad \varphi^* E \rightarrow M$$

i.e.

$$\begin{array}{ccc} \varphi^* E & \xleftarrow{\quad \varphi^* \quad} & E \\ \downarrow \psi & & \downarrow \pi \\ M & \xrightarrow{\quad \varphi \quad} & N \end{array} \quad \begin{array}{c} (\varphi^* E)_x = E_{\varphi(x)} \\ \uparrow \\ x \in M \end{array}$$

i.e. if  $s \in \Gamma(E)$ ,

$$\varphi^* s = s \circ \varphi \in \Gamma(\varphi^* E)$$

Thus,

$$\begin{array}{ccc} \gamma^* E & \xleftarrow{\quad \gamma^* \quad} & E \\ \downarrow & & \downarrow \pi \\ [a, b] & \xrightarrow{\quad \gamma \quad} & M \end{array} \quad \begin{array}{c} (\varphi^* E)_c = E_{\gamma(c)} \\ \uparrow \\ c \in [a, b] \end{array}$$

For

$$\dot{v}^k = -v^i \dot{s}^j \Gamma_{ij}^k$$

$$v^k(c) = v_0^k \quad 1 \leq k \leq m$$

$$\dot{v} = -v^i \dot{s}^j \Gamma_{ij}$$

$$(v + w) = -(v^i + w^i) \dot{s}^j \Gamma_{ij}(v + w)(c) = v(c) + w(c) = v_0 + w_0$$

so  $v + w \in \mathfrak{X}(s)$  is parallel transport of  $v_0 + w_0$ .

Likewise,  $\forall a \in \mathbb{F}$ ,  $av \in \mathfrak{X}(s)$  is the parallel transport of  $av_0$ .

$$\dot{\mu}^k = -\mu^i \dot{s}^\mu \omega_{i\mu}^k = -\mu^i \omega_i^k(\dot{s}^\mu)$$

Suppose  $\gamma^* E$  trivialized over  $[a, b]$ .

Closed interval is contractible, so this is always possible.

For chart  $(U, \varphi)$ ,

$$\begin{array}{ccc} \gamma^* E & \xleftarrow{\quad \gamma^* \quad} & E \\ \downarrow & & \downarrow \pi \\ [a, b] & \xrightarrow{\quad \gamma \quad} & M \end{array} \quad \begin{array}{ccc} E|_U & \xrightarrow{\quad \psi \quad} & U \times V \\ \pi^{-1} \uparrow & \nearrow & \\ U \subset M & & \end{array}$$

Consider

$$\varphi : [a, b] \times V \rightarrow \gamma^* E$$

$$\varphi(t, \cdot) = \gamma^* \circ \psi^{-1}(\gamma(t), \cdot)$$

$$\begin{aligned} \forall \mu \in \Gamma(E|_{x \in M}), \\ \mu = \mu^i e_i. \\ \varphi(t, e_i) = \epsilon_i \text{ is a basis for } \gamma^* E. \\ \forall \sigma \in \Gamma(\gamma^* E), \end{aligned}$$

$$\begin{aligned} \sigma &= \sigma^i \epsilon_i, \quad \sigma^i : [a, b] \rightarrow \mathbb{F} \\ \nabla_{\frac{\partial}{\partial x^\mu}} \sigma &= \frac{\partial \sigma^k}{\partial x^\mu} \epsilon_k + \omega^k_{j\mu} \sigma^j \epsilon_k = \left( \frac{\partial \sigma^k}{\partial x^\mu} + \omega^k_{j\mu} \sigma^j \right) \epsilon_k \\ \nabla \sigma &= \epsilon_k \otimes (d\sigma^k + \omega^k_{j\mu} dx^\mu \sigma^j) = \epsilon_k \otimes (d\sigma^k + \omega^k_j \sigma^j) \\ \nabla_{\frac{d}{dt}} \sigma &= \epsilon_k \otimes \left( \frac{d\sigma^k}{dt} + \omega^k_{j\mu} \dot{x}^\mu \sigma^j \right) \end{aligned}$$

Now

$$\frac{d}{dt} = \dot{x}^\nu \frac{\partial}{\partial x^\nu}$$

Then  $\sigma$  parallel along  $\gamma$  if

$$\frac{d\sigma^k}{dt} + \omega^k_{j\mu} \dot{x}^\mu \sigma^j = 0$$

**Definition 54** (3.1.4 [20]). *Parallel transport along  $\gamma$  is*

$$(78) \quad \begin{aligned} P_\gamma : E_{\gamma(a)} &\rightarrow E_{\gamma(b)} \\ P_\gamma(v) &\mapsto \sigma(b) \end{aligned}$$

where  $\sigma \in \Gamma(\gamma^* E)$ ,  $\sigma$  unique and s.t.  $\sigma(a) = v$ .

**Lemma 1** (10.1.16[19]). *holonomy*

$$h_s : T_x M \rightarrow T_{x_0} M$$

if  $\nabla$  around piecewise smooth loop  $s$  is a linear transformation.

**Lemma 2** (10.1.18 Conlon (2008)[19]). *Let piecewise smooth loop  $s : [a, b] \rightarrow M$  at  $x_0$ .*

*Let weak reparametrization  $\tilde{s} = s \circ r : [c, d] \rightarrow M$ .  
If reparametrization is orientation-preserving, then  $h_{\tilde{s}} = h_s$ ,  
If reparametrization is orientation-reversing, then  $h_{\tilde{s}} = h_s^{-1}$ ,*

*Proof.* Without loss of generality, assume smooth  $s, r$

$$\begin{aligned} \tilde{s}(\tau) &= s(r(\tau)) \\ \tilde{v}(\tau) &= v(r(\tau)) \\ \tilde{u}^j(\tau) &= \frac{dt}{d\tau}(\tau) u^j(r(\tau)) \\ \frac{d\tilde{v}^k}{d\tau}(\tau) &= \frac{dr}{d\tau}(\tau) \frac{dv^k}{dt}(r(\tau)) \\ \frac{d\tilde{v}^k}{d\tau} &= -\tilde{v}^i \tilde{u}^j \Gamma^k_{ij} \end{aligned}$$

since

$$\begin{aligned} \frac{dv^k}{dt} &= -v^i u^j \Gamma^k_{ij}; \quad 1 \leq k \leq n \\ v^k(c) &= a^k; \quad 1 \leq k \leq a \\ \frac{dr}{d\tau} \frac{dv^k}{dt} &= -v^i \frac{dr}{d\tau} u^j \Gamma^k_{ij} = \frac{d\tilde{v}^k}{d\tau} = -\tilde{v}^i \tilde{u}^j \Gamma^k_{ij} \end{aligned}$$

Thus, if  $r(c) = a, r(d) = b$

$$h_{\tilde{s}}(v_0) = \tilde{v}(d) = v(b) = h_s(v_0)$$

If  $r(c) = a, r(d) = b$ , then

$$\tilde{v}(c) = v(b) = h_s(v_0)$$

and

$$h_{\tilde{s}}(h_s(v_0)) = h_{\tilde{s}}(v(b)) = \tilde{v}(d) = v(a) = v_0$$

At this point, I will switch to my notation because it clarified to me, at least, what was going on, in that a holonomy  $h_s$  is *invariant* under orientation-preserving reparametrization, and its inverse is well-defined.

For  $\tilde{s} = s \circ t : [c, d] \rightarrow M$ ,  
piecewise smooth  $t$  is reparametrized, i.e.

$$(79) \quad t : [c, d] \rightarrow [a, b]$$

Now,

$$\begin{aligned} \frac{d}{d\tau} \tilde{s}(\tau) &= \frac{d}{d\tau} \tilde{s}(t(\tau)) = \dot{s}(t) \frac{dt}{d\tau}(\tau) \equiv \dot{s} \frac{dt}{d\tau} \\ v^k(t) &= v^k(t(\tau)) = v^k(\tau) \\ \frac{dv^k}{d\tau}(t(\tau)) &= \frac{dv^k}{dt} \frac{dt}{d\tau} = \frac{dt}{d\tau} (-v^i(\tau) \dot{s}^j(t) \Gamma^k_{ij}) = -v^i(\tau) \frac{d\tilde{s}^j}{d\tau} \Gamma^k_{ij} \end{aligned}$$

Consider

$$h_s(v_0) = v(b)$$

If  $t(c) = a,$   
 $t(d) = b$

$$h_{\tilde{s}}(v_0) = \tilde{v}(d) = v(t(d)) = v(b) = h_s(v_0)$$

If  $t(c) = b,$   
 $t(d) = a$

$$\begin{aligned} h_{\tilde{s}}(h_s(v_0)) &= h_{\tilde{s}}(v(b)) = h_{\tilde{s}}(v(t(c))) = h_{\tilde{s}}(\tilde{v}(c)) = \\ &= \tilde{v}(d) = v(t(d)) = v(a) = v_0 \end{aligned}$$

Thus,

$$h_{\tilde{s}} = h_s^{-1}$$

□

I am working through Conlon (2008) [19] , Clarke and Santoro (2012) [20], and Schreiber and Waldorf (2007)[21], concurrently, for holonomy.



## 15. PATH GROUPOID OF A SMOOTH MANIFOLD; GENERALIZATION OF PATHS

cf. Schreiber and Waldorf (2007)[21].

**Definition 55** (path). ***path** is a smooth map  $\gamma : [0, 1] \rightarrow M$ , between 2 pts.  $x, y \in M$ , which has a sitting instant; i.e. number  $0 < \epsilon < \frac{1}{2}$  s.t.*

$$(80) \quad \gamma(t) = \begin{cases} x & \text{for } 0 \leq t < \epsilon \\ y & \text{for } 1 - \epsilon < t \leq 1 \end{cases}$$

Denote the set of such paths by  $PM$ ,

$$(81) \quad PM \equiv \{\gamma \in \Gamma(M) \mid \text{smooth } \gamma : [0, 1] \rightarrow M \text{ s.t. } \exists 0 < \epsilon < \frac{1}{2} \text{ s.t. } \begin{cases} x & \text{for } 0 \leq t < \epsilon \\ y & \text{for } 1 - \epsilon < t \leq 1 \end{cases}\}$$

cf. Def. 2.1. of Schreiber and Waldorf (2007)[21]

Define *composition*:

Given paths  $\gamma_1, \gamma_2$ ;  $\gamma_1(0) = x$ ,  $\gamma_2(0) = y$ ,

$$\gamma_1(1) = y \quad \gamma_2(1) = z$$

define composition to be path

$$(82) \quad \begin{matrix} \gamma_2 \circ \gamma_1 \\ (\gamma_2 \circ \gamma_1)(t) := \begin{cases} \gamma_1(2t) & \text{for } 0 \leq t \leq \frac{1}{2} \\ \gamma_2(2t - 1) & \text{for } \frac{1}{2} \leq t \leq 1 \end{cases} \end{matrix}$$

$\gamma_2 \circ \gamma_1$  smooth since  $\gamma_1, \gamma_2$  both constant near gluing pt., due to sitting instants  $\epsilon_1, \epsilon_2$ , respectively.

Define *inverse*:

$$(83) \quad \begin{matrix} \gamma^{-1} : [0, 1] \rightarrow M \\ \gamma^{-1}(t) := \gamma(1 - t) \end{matrix}$$

$$(\text{so that } \gamma^1(t) = \begin{cases} y & \text{for } 1 - \epsilon < 1 - t \leq 1 \text{ or } 0 \leq t < \epsilon \\ x & \text{for } 0 \leq 1 - t < \epsilon \text{ or } 1 - \epsilon < t \leq 1 \end{cases})$$

**Definition 56** (thin homotopy equivalent). *2 paths  $\gamma_1, \gamma_2$  s.t.  $\gamma_1(0) = \gamma_2(0) = x$ ,  $\gamma_1, \gamma_2$  are thin homotopy equivalent,*

$$\gamma_1(1) = \gamma_2(1) = y$$

*if  $\exists$  smooth  $h : [0, 1] \times [0, 1] \rightarrow M$  s.t.*

$$(1) \quad \exists 0 < \epsilon < \frac{1}{2} \text{ with}$$

$$(a) \quad h(s, t) = x \text{ for } 0 \leq t < \epsilon$$

$$h(s, t) = y \text{ for } 1 - \epsilon < t \leq 1$$

(b)

$$(c) \quad h(s, t) = \gamma_1(t) \text{ for } 0 \leq s < \epsilon$$

$$h(s, t) = \gamma_2(t) \text{ for } 1 - \epsilon < s \leq 1$$

(2) *differential of  $h$  has at most rank 1 everywhere, i.e.*

$$(84) \quad \text{rank}(dh|_{(s,t)}) \leq 1 \quad \forall (s, t) \in [0, 1] \times [0, 1]$$

cf. Def. 2.2. of Schreiber and Waldorf (2007)[21]

$h(s, t) = \gamma_1(t)$  for  $0 \leq s < \epsilon$  is the homotopy from  $\gamma_1$  to  $\gamma_2$ , i.e.  $h(0, t) = \gamma_1(t)$

$h(s, t) = \gamma_2(t)$  for  $1 - \epsilon < s \leq 1$   $h(1, t) = \gamma_2(t)$

and define an equivalence relation on  $PM$ .

Note that for  $h : [0, 1] \times [0, 1] \rightarrow M$ ,

$$(Dh)|_{(s,t)} = \left[ \frac{\partial h^i}{\partial s}, \frac{\partial h^i}{\partial t} \right]$$

$P^1M \equiv$  set of thin homotopy classes of paths, i.e.

$$(85) \quad P^1M = \{[\gamma] \mid \gamma_1 \in PM, \text{ if } \exists \text{ smooth } h : [0, 1] \times [0, 1] \rightarrow M \text{ s.t. } h \text{ thin homotopy of } \gamma_1 \text{ and } \gamma_2, \gamma_1 \sim \gamma_2\}$$

pr :  $PM \rightarrow P^1M$  is projection to classes.

Denote thin homotopy class of path  $\gamma$ ,  $\gamma(0) = x$ , by  $\bar{\gamma}$ , or  $[\gamma]$ .

$$\gamma(1) = y$$

**15.1. Reparametrization of thin homotopies.** Let  $\beta : [0, 1] \rightarrow [0, 1]$ ,  $\beta(0) = 0$ .

$$\beta(1) = 1$$

Then  $\forall$  path  $\gamma$ ,  $\gamma(0) = x$ ,  $\gamma \circ \beta$  is also a path  $\gamma \circ \beta(0) = x$  and

$$\gamma(1) = y$$

$$\gamma \circ \beta(1) = y$$

(86)

$$h(s, t) := \gamma(t\beta(1 - s) + \beta(t)\beta(s))$$

defines a homotopy from  $\gamma$  to  $\gamma \circ \beta$ .

$$\gamma_1 \circ \gamma_2 \in PM \xrightarrow{\text{pr}} [\gamma_1 \circ \gamma_2] = [\gamma_1][\gamma_2] \in P^1M$$

Composition of thin homotopy classes of paths obeys following rules:

**Lemma 3.**  $\forall$  path  $\gamma$ ,  $\gamma(0) = x$

$$\gamma(1) = y$$

$$(1) \quad \bar{\gamma} \circ \overline{id_x} = \bar{\gamma} = \overline{id_y} \circ \bar{\gamma} \equiv [\gamma]1_x = [\gamma] = 1_y[\gamma]$$

$$(2) \quad \text{for paths } \gamma'; \gamma'(0) = y, \quad \gamma''(0) = z$$

$$\gamma'(1) = z \quad \gamma''(1) = w$$

(87)

$$(\bar{\gamma}'' \circ \bar{\gamma}') \circ \bar{\gamma} = \bar{\gamma}'' \circ (\bar{\gamma}' \circ \bar{\gamma}) \equiv ([\gamma''][\gamma'])[\gamma] = [\gamma'']([\gamma'][\gamma])$$

$$(3) \quad \bar{\gamma} \circ \bar{\gamma}^{-1} = \overline{id_y} \text{ and } \overline{\gamma^{-1}} \circ \bar{\gamma} = \overline{id_x} \equiv [\gamma][\gamma^{-1}] = 1_y \text{ and } [\gamma^{-1}][\gamma] = 1_x$$

cf. Lemma 2.3. of Schreiber and Waldorf (2007)[21]

**Definition 57** (path groupoid).  $\forall$  smooth manifold  $M$ ,

consider category whose set of objects is  $M$ ,

whose set of morphisms is  $P^1M$ , where class  $[\gamma]$ ,  $[\gamma](0) = x$  is a morphism from  $x$  to  $y$  and

$$[\gamma](1) = y$$

composition  $[\gamma_1][\gamma_2] = [\gamma_1 \circ \gamma_2] \in P^1M$  Lemma 3 are axioms of a category, 3rd. property says  $\forall$  morphism is invertible.

Hence, we've defined a groupoid, called **path groupoid** of  $M$ ,  $\mathcal{P}_1(M)$ .

So

$$\text{Obj}(\mathcal{P}_1(M)) = M$$

$$\text{Mor}(\mathcal{P}_1(M)) = P^1M$$

$\forall$  smooth  $f : M \rightarrow N$ , denote functor  $f_*$

$$(88) \quad f_* : \mathcal{P}_1(M) \rightarrow \mathcal{P}_1(N)$$

with

$$f_*(x) = f(x)$$

$$(f_*)([\gamma]) := [f \circ \gamma]$$

If  $\gamma \sim \gamma'$ , for  $f \circ \gamma$ ,  $f \circ \gamma'$ ,

$$f \circ h(s, t) \text{ with } f \circ h(0, t) = f \circ \gamma(t),$$

$$f \circ h(1, t) = f \circ \gamma'(t)$$

so  $f \circ h$  is a thin homotopy between  $f \circ \gamma$ ,  $f \circ \gamma'$  and so  $[f \circ \gamma]$  well-defined.

## Part 10. Complex Manifolds

EY : 20170123 I don't see many good books on Complex Manifolds for physicists other than Nakahara's. I will supplement this section on Complex Manifolds with external links to the notes of other courses that I found useful to myself.

[Complex Manifolds - Lecture Notes](#) Koppensteiner (2010) [16]

[Lectures on Riemannian Geometry, Part II: Complex Manifolds by Stefan Vandoren](#)

Vandoren (2008) [17]

## Part 11. Jets, Jet bundles, $h$ -principle, $h$ -Prinzipien

cf. Eliashberg and Misahchev (2002) [22]

cf. Ch. 1 Jets and Holonomy, Sec. 1.1 Maps and sections of Eliashberg and Misahchev (2002) [22].

Visualize  $f : \mathbb{R}^n \rightarrow \mathbb{R}^q$  as graph  $\Gamma_f \subset \mathbb{R}^n \times \mathbb{R}^q$ .

Consider this graph as image of  $\mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^q$ , i.e.

$$x \mapsto (x, f(x))$$

$\mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^q$  is called section (by mathematicians),

$$x \mapsto (x, f(x))$$

is called *field* or  $\mathbb{R}^q$ -valued field (by physicists).

cf. Ch. 1 Jets and Holonomy, Sec. 1.2 Coordinate definition of jets of Eliashberg and Misahchev (2002) [22].

**Definition 58** ( $r$ -jet). Given (smooth)  $f : \mathbb{R}^n \rightarrow \mathbb{R}^q$ , given  $x \in \mathbb{R}^n$ .

$r$ -jet of  $f$  at  $x$  - sequence of derivatives of  $f$ , up to order  $r$ ,  $\equiv$

$$(89) \quad J_f^r(x) = (f(x), f'(x) \dots f^{(r)}(x))$$

$f^{(q)}$  consists of all partial derivatives  $D^\alpha f$ ,  $\alpha = (\alpha_1 \dots \alpha_n)$ ,  $|\alpha| = \alpha_1 + \dots + \alpha_n = s$ , ordered lexicographically.

e.g.  $q = 1$ ,

$$f : \mathbb{R}^n \rightarrow \mathbb{R}.$$

$$\text{1-jet of } f \text{ at } x = J_f^1(x) = (f(x), f^{(1)}(x)).$$

$$f^{(1)}(x) = \{D^\alpha f | \alpha = (\alpha_1 \dots \alpha_n), |\alpha| = \alpha_1 + \dots + \alpha_n = 1\} = \left( \frac{\partial f}{\partial x^1}, \frac{\partial f}{\partial x^2}, \dots, \frac{\partial f}{\partial x^n} \right)$$

Let  $d_r = d(n, r) =$  number of all partial derivatives  $D^\alpha$  of order  $r$  of function  $\mathbb{R}^n \rightarrow \mathbb{R}$ .

Consider  $r$ -jet  $J_f^r(x)$  of map  $f : \mathbb{R}^n \rightarrow \mathbb{R}^q$  as pt. of space  $\mathbb{R}^q \times \mathbb{R}^{qd_1} \times \mathbb{R}^{qd_2} \times \dots \times \mathbb{R}^{qd_r} = \mathbb{R}^{qN_r}$ , where  $N_r = N(n, r) = 1 + d_1 + d_2 + \dots + d_r$ , i.e.

$$J_f^r(x) = (f(x), f^{(1)}(x), \dots, f^{(r)}(x)) \in \mathbb{R}^q \times \mathbb{R}^{qd_1} \times \dots \times \mathbb{R}^{qd_r} = \mathbb{R}^{qN_r}$$

### Exercise 1.

Given order  $r$ , consider  $n$ -tuple of (positive) integers  $(r_1, r_2 \dots r_n)$  s.t.  $r_1 + r_2 + \dots + r_n = r$ , and  $r_k \geq 0$ .

Imagine  $r_k =$  occupancy number, num ber of balls in  $k$ th cell.  $(r_1 \dots r_n)$  describes a positive onfiguration of occupancy numbers, with indistinguishable balls; 2 distributions are distinguishable only if corresponding  $n$ -tuples  $(r_1 \dots r_n)$  not identical.

Represent balls by stars, and indicate  $n$  cells by  $n$  spaces between  $n + 1$  bars.

With  $n + 1$  bars,  $r$  stars, 2 bars are fixed.  $n - 1$  bars and  $r$  stars to arrange linearly, so a total of  $n - 1 + r$  objects to arrange.  $r$  stars indistinguishable amongst themselves, so choose  $r$  out of  $n - 1 + r$  to be stars.

$$(90) \quad \implies d_r = d(n, r) = \binom{n-1+r}{r}$$

Use *induction* (cf. [Ch. 4 Binomial Coefficients](#)).

$$N_0 = N(n, 0) = \binom{n-1+0}{0} = 1$$

$$N_1 = N(n, 1) = 1 + \binom{n-1+1}{1} = 1 + n = \frac{(n+1)!}{n!1!}$$

Induction step:

$$N_{r-1} = N(n, r-1) = \sum_{k=1}^{r-1} d_k + 1 = \binom{n+r-1}{r-1}$$

and so

$$\begin{aligned} N_r &= N(n, r) = \sum_{k=1}^r d_k + 1 = \sum_{k=1}^r \binom{n-1+k}{k} + 1 = \sum_{k=1}^{r-1} \binom{n-1+k}{k} + \binom{n-1+r}{r} + 1 = \\ &= \binom{n+r-1}{r-1} + \binom{n-1+r}{r} = \frac{(n+r-1)!}{(r-1)!n!} + \frac{(n-1+r)!}{r!(n-1)!} = \frac{(n+r)!}{n!r!} = \binom{n+r}{r} \end{aligned}$$

$$\begin{array}{ccc} \mathbb{R}^{qN_r} & & J_f^r(x) \\ \uparrow J_f^r & & \uparrow J_f^r \\ \mathbb{R}^n & \xrightarrow{f} & \mathbb{R}^q \end{array} \quad \begin{array}{ccc} & & \\ & & \downarrow J_f^r \\ & & x \end{array} \quad \begin{array}{ccc} & & \\ & & \xrightarrow{f} f(x) \end{array}$$

**Definition 59** (space of  $r$ -jets). *space of  $r$ -jects of maps  $\mathbb{R}^n \rightarrow \mathbb{R}^q$  or space of  $r$ -jets of sections  $\mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^q \equiv$*

$$(91) \quad J^r(\mathbb{R}^n, \mathbb{R}^q) = \mathbb{R}^n \times \mathbb{R}^{qN_r} \equiv \mathbb{R}^n \times \mathbb{R}^q \times \mathbb{R}^{qd_1} \times \mathbb{R}^{qd_2} \times \dots \times \mathbb{R}^{qd_r}$$

e.g.  $J^1(\mathbb{R}^n, \mathbb{R}^q) = \mathbb{R}^n \times \mathbb{R}^q \times M_{q \times n}$ , where  $M_{q \times n} = \mathbb{R}^{qn}$  is the space of  $(q \times n)$ -matrices.

**Part 12. Morse Theory**

## 16. MORSE THEORY INTRODUCTION FROM A PHYSICIST

I needed some physical motivation to understand Morse theory, and so I looked at Hori, et. al. [18].  
cf. pp. 43, Sec. 3.4 Morse Theory, from Ch. 3. Differential and Algebraic Topology of Hori, et. al. [18].

Consider smooth  $f : M \rightarrow \mathbb{R}$ , with non-degenerate critical points.

If no critical values of  $f$  between  $a$  and  $b$  ( $a < b$ ), then subspace on which  $f$  takes values less than  $a$  is deformation retract of subspace where  $f$  less than  $b$ , i.e.

$$\{x \in M | f(x) < b\} \times [0, 1] \xrightarrow{F} \{x \in M | f(x) < b\}$$

$\forall x \in M$  s.t.  $f(x) < b$ ,

$$\begin{aligned} F(x, 0) &= x \\ F(x, 1) &\in \{x \in M | f(x) < a\} \end{aligned} \quad \text{and } F(a', 1) = a' \quad \forall a' \in M \text{ s.t. } f(a') < a$$

To show this, consider  $-\nabla f / |\nabla f|^2$

Morse lemma:  $\forall$  critical pt.  $p$  s.t.  $\exists$  choice of coordinates s.t.

$$(92) \quad f = -(x_1^2 + x_2^2 + \cdots + x_\mu^2) + x_{\mu+1}^2 + \cdots + x_n^2$$

where  $f(p) = 0$  and  $p$  is at origin of these coordinates.

- difference between

$$f^{-1}(\{x \leq -\epsilon\}), f^{-1}(\{x \leq +\epsilon\})$$

can be determined by local analysis and only depends on  $\mu$ ,  $\mu \equiv$  “Morse index” = number of negative eigenvalues of Hessian of  $f$  at critical pt.

Answer:

$f^{-1}(\{x \leq +\epsilon\})$  can be obtained from  $f^{-1}(\{x \leq -\epsilon\})$  by “attaching  $\mu$ -cell” along boundary  $f^{-1}(0)$

- “attaching  $\mu$ -cell to  $X$  mean, take  $\mu$ -ball  $B_\mu = \{|x| \leq 1\}$  in  $\mu$ -dim. space, identity pts. on boundary  $S^{\mu-1}$  with pts. in the space  $X$ , through cont.  $f : S^{\mu-1} \rightarrow X$ , i.e. take

$$X \coprod B_\mu$$

with  $x \sim f(x) \quad \forall x \in \partial B_\mu = S^{\mu-1}$ .

- find homology of  $M$ ,  
 $f$  defines chain complex  $C_f^*$ ,  $k$ th graded piece  $C^{\alpha_k}$ ,  $\alpha_k$  is number of critical pts. with index  $k$ .

$$(93) \quad \begin{aligned} \partial : C_p^k &\rightarrow C_p^{k-1} \\ \partial x_a &= \sum_b \Delta_{a,b} x_b \end{aligned}$$

where  $\Delta_{a,b} :=$  signed number of lines of gradient flow from  $x_a$  to  $x_b$ ,  $b$  labels pts. of index  $k-1$ .

Gradient flow line is path  $x(t)$  s.t.  $\dot{x} = \nabla(f)$ , with  $x(-\infty) = x_a$

$$x(+\infty) = x_b$$

- To define this number ( $\Delta_{a,b}?$ ), construct moduli space of such lines of flow (???)  
by intersecting outward and inward flowing path spaces from each critical point, and then show this moduli space is oriented, 0-dim. manifold (pts. with signs)

- $\partial^2 = 0$  proof

$\partial$ , boundary of space of paths connecting critical points, whose index differs by 2 = union over compositions of paths between critical pts. whose index differs by 1.

$\implies$  coefficients of  $\partial^2$  are sums of signs of pts. in 0-dim. space, which is boundary of 1-dim. space.

These signs must therefore add to 0, so  $\partial^2 = 0$ .

Hori, et. al. [18] is good for physics, but there isn’t much thorough, step-by-step explanations of the math. I will look at Hirsch (1997) [8] and Shastri (2011) [7] at the same time.

**16.1. Introduction, definitions of Morse Functions, for Morse Theory.** cf. Ch. 6, Morse Theory of Hirsch (1997) [8], Section 1. Morse Functions, pp. 143-

Recall for  $TM$ ,  $T_x M \xrightarrow{\varphi} \mathbb{R}^n$ .

Cotangent bundle  $T^*M$  defined likewise:

$$T_x^* M \xrightarrow{\varphi} \text{dual vector space } (\mathbb{R}^n)^* = L(\mathbb{R}^n, \mathbb{R})$$

i.e.

$$T^*M = \bigcup_{x \in M} (M_x^*) \quad M_x^* = L(M_x, \mathbb{R})$$

If chart  $(\varphi, U)$  on  $M$ , natural chart on  $T^*M$  is

$$T^*U \rightarrow \varphi(U) \times (\mathbb{R}^n)^*$$

$$\lambda \in M_x^* \mapsto (\varphi(x), \lambda \varphi_x^{-1})$$

Projection map

$$p : T^* \rightarrow M$$

$$M_x^* \mapsto x$$

Let  $C^{r+1}$  map,  $1 \leq r \leq \omega$ ,  $f : M \rightarrow \mathbb{R}$ ,  $\forall x \in M$ , linear map  $T_x f : M_x \rightarrow \mathbb{R}$  belongs to  $M_x^*$

$$T_x f = Df_x \in M_x^*$$

Then

$$Df : M \rightarrow T^*M$$

$$x \mapsto Df_x = Df(x)$$

is  $C^r$  section of  $T^*M$ .

**Definition 60. critical point**  $x$  of  $f$  is zero of  $Df$ , i.e.

$$(94) \quad Df(x) = 0$$

of vector space  $M_x^*$ .

Thus, set of critical pts. of  $f$  is counter-image of submanifold  $Z^* \subset T^*M$  of zeros.

Note  $Z^* \approx M$ , codim. of  $Z^*$  is  $n = \dim M$ .

**Definition 61. Morse function**  $f$  if  $\forall$  critical pts. of  $f$  are nondegenerate.

Note set of critical pts. closed discrete subset of  $M$ .

Let open  $U \subset \mathbb{R}^n$ , let  $C^2$  map  $g : U \rightarrow \mathbb{R}$ ,

critical pt.  $p \in U$  nondegenerate iff

- linear  $D(Dg)(p) : \mathbb{R}^n \rightarrow (\mathbb{R}^n)^*$  bijective
- identify  $L(\mathbb{R}^n, (\mathbb{R}^n)^*)$  with space of bilinear maps  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $\implies$  equivalent to condition that symmetric bilinear  $D^2g(p) : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  non-degenerate
- $n \times n$  Hessian matrix

$$\left[ \frac{\partial^2 g}{\partial x^i \partial x^j}(p) \right]$$

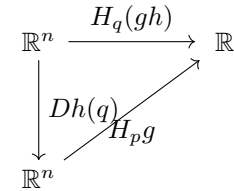
has rank  $n$

Hessian of  $g$  at critical pt.  $p$  is quadratic form  $H_p f$  associated to bilinear form  $D^2 g(p)$

$$\implies H_p f(y) = D^2 g(p)(y, y) = \sum_{i,j} \frac{\partial^2 g}{\partial x^i \partial x^j}(p) y^i y^j$$

Let open  $V \subset \mathbb{R}^n$ , suppose  $C^2$  diffeomorphism  $h : V \rightarrow U$ .

Let  $q = h^{-1}(p)$ , so  $q$  is critical pt. of  $gh : V \rightarrow \mathbb{R}$ .



(quadratic) form  $(H_p f)$  invariant under diffeomorphisms.

Let  $C^2 f : M \rightarrow \mathbb{R}$ .

$\forall$  critical pt.  $x$  of  $f$ , define

Hessian quadratic form

$$H_x f : M_x \rightarrow \mathbb{R}$$

$$H_x f : M_x \xrightarrow{D\varphi_x} \mathbb{R}^n \xrightarrow{H_{\varphi(x)}(f\varphi^{-1})} \mathbb{R}$$

where  $\varphi$  is any chart at  $x$ .

Thus, critical pt. of a  $C^2$  real-valued function nondegenerate iff associated Hessian quadratic form is nondegenerate.

Let  $Q$  nondegenerate quadratic form on vector space  $E$ .

$Q$  negative definite on subspace  $F \subset E$  if  $Q(x) < 0$  whenever  $x \in F$  nonzero.

Index of  $Q \equiv \text{Ind} Q$ , is largest possible dim. of subspace on which  $Q$  is negative definite.

cf. 1.1. Morse's Lemma of Ch. 6, pp. 145, Morse Theory of Hirsch (1997) [8]

**Lemma 4** (Morse's Lemma). *Let  $p \in M$  be nondegenerate critical pt. of index  $k$  of  $C^{r+2}$  map  $f : M \rightarrow \mathbb{R}$ ,  $1 \leq r \leq \omega$ . Then  $\exists C^r$  chart  $(\varphi, U)$  at  $p$  s.t.*

$$(95) \quad f\varphi^{-1}(u_1 \dots u_n) = f(p) - \sum_{i=1}^k u_i^2 + \sum_{i=k+1}^n u_i^2$$

Let  ${}^T Q \equiv Q^T$  denote tranpose of matrix  $Q$ .

**Lemma 5.** *Let  $A = \text{diag}\{a_1, \dots, a_n\}$  diagonal  $n \times n$  matrix, with diagonal entries  $\pm 1$ . Then  $\exists$  neighborhood  $N$  of  $A$  in vector space of symmetric  $n \times n$  matrices,  $C^\infty$  map*

$$(96) \quad P : N \rightarrow GL(n, \mathbb{R})$$

s.t.  $P(A) = I$ , and if  $P(B) = Q$ , then  $Q^T B Q = A$

*Proof.* Let  $B = [b_{ij}]$  be symmetri matrix near  $A$  s.t.  $b_{11} \neq 0$  and  $b_{11}$  has same sign as  $a_1$ .

Consider  $x = T y$  where

$$x_1 = \left[ y_1 - \frac{b_{12}}{b_{11}} y_2 - \dots - \frac{b_{1n}}{b_{11}} y_n \right] / \sqrt{|b_n|}$$
$$x_k = y_k \text{ for } k = 2, \dots, n$$

From *wikipedia:Lagrange multiplier*, [https://en.wikipedia.org/wiki/Lagrange\\_multiplier](https://en.wikipedia.org/wiki/Lagrange_multiplier), find local minima (maxima), pt.  $a \in N$ , s.t.  $\exists$  neighborhood  $U$  s.t.  $f(x) \geq f(a)$  ( $f(x) \leq f(a)$ )  $\forall x \in U$ .  
For  $f : U \rightarrow \mathbb{R}$ , open  $U \subset \mathbb{R}^n$ , find  $x \in U$  s.t.  $D_x f \equiv Df(x) = 0$ , check if Hessian  $H_x f < 0$ .  
Maxima may not exit since  $U$  open.

References:

[Relative Extrema and Lagrange Multipliers](#)

Other interesting links:

[The Lagrange Multiplier Rule on Manifolds and Optimal Control of nonlinear systems](#)

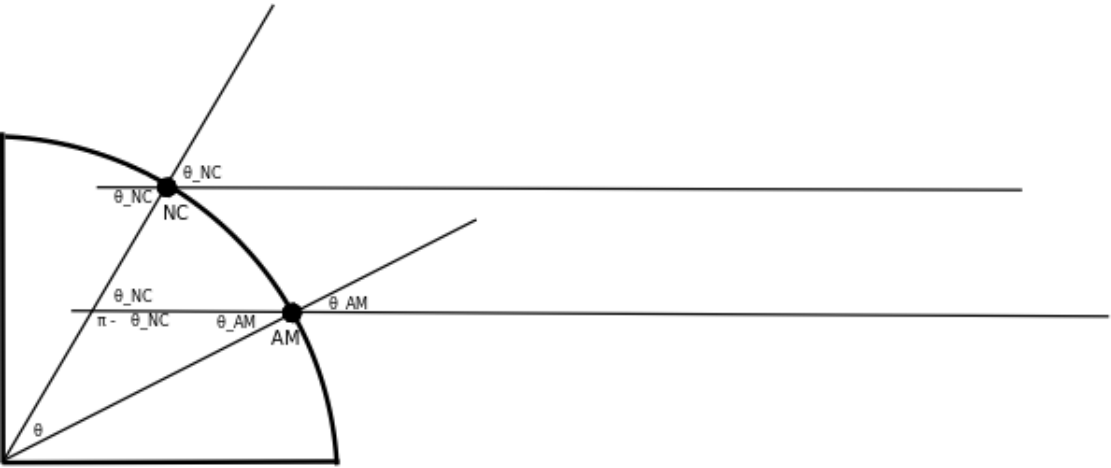
### Part 13. Classical Mechanics

**Problem 1.4 The altitude angle of the Sun, Japan Physics Olympiad (2014).** cf. Japan Physics Olympiad, (2014) [48].

**Solution 1.4 The altitude angle of the Sun, Japan Physics Olympiad (2014).**

$NC \equiv$  Niigata City,

$AM \equiv$  Amagi-san in Izu.



From all 3 internal angles of a triangle sum up to 180 degrees, or i.e.  $\pi$  radians,

$$\pi = \pi - \theta_{NC} + \theta_{AM} + \theta \implies \boxed{\theta = \theta_{NC} - \theta_{AM}}$$

So we've shown indeed that  $\theta$  is the **difference between altitude angles of the Sun** at NC and AM, respectively.

From

$$C = 2\pi R$$
$$l_M = \theta R$$
$$\frac{l_M}{l} = \frac{\theta R}{\frac{\pi}{2} R} \text{ or } \frac{l_M}{l} = \frac{2\theta}{\pi}$$
$$\implies \theta = \frac{\pi l_M}{2 l} = \frac{\pi}{2} \cdot \frac{334 \times 10^3 m}{10^7 m} \cdot \frac{180^\circ}{\pi} = 3^\circ$$

- Advice, remember to draw a *cross-section* to make diagram clear, rather than worry about "the other transverse dimension".
- Also, draw the lines literally parallel to the Sun because of how far away the Sun is and how large the Sun is.

## 18. STRUCTURE OF GALILEAN SPACE-TIME

cf. Sec. 3.1 - *Structure of Galilean Space-Time* of Prástaro (1996) [15].

Mechanics assumes a particular simple formulation if formulated with respect to some spacetime manifold.

In Galilean spacetime, it's possible to naturally recognize absolute objects, and others that depend on frames.

cf. Def. 3.1 of Prástaro (1996) [15]

**Definition 62** (Galilean spacetime structure). (1) *Galilean spacetime structure*  $:= (\mathcal{G}, g)$  where  $\mathcal{G}$  is (fiber bundle space-time)

$$(97) \quad \mathcal{G} \equiv \{\tau : M \rightarrow T\}$$

where

$M = 4\text{-dim. affine manifold } (\mathbf{space-time});$  corresponding structure is  $(M, \mathbf{M}, \alpha),$

(2)  $T = 1\text{-dim. affine space } (\mathbf{time}),$  corresponding affine structure is  $(T, \mathbf{T}, \beta)$

(3)  $\tau = \text{surjective affine mapping, of constant rank } 1,$  s.t.  $\forall p \in M$  associates its time  $\tau(p) \in T$

Put  $\mathbf{S} = \ker(\underline{\tau}) \equiv \ker(D\tau) \in M,$

where  $\underline{\tau} \equiv D\tau,$   $D$  is symbol of derivative.

Define

$$g : M \rightarrow vS_2^0(M) \equiv M \times S_2^0(\mathbf{S})$$

$$g(p) = (p, g) \equiv (p, Dg), \forall p \in M$$

where  $\underline{g} \equiv Dg$  is a Euclidean structure on  $\mathbf{S}.$   $g$  is called vertical metric field.

Thus, given  $(M, \mathbf{M}, \alpha), \forall (O, \{\mathbf{e}_i\}_{1 \leq i \leq d}, \{\mathbf{e}_i\}_{i=1 \dots d},$  is basis of  $\mathbf{M},$

$$M \cong \mathbb{R}^4, \text{ and } \exists \{x^\alpha : M = \mathbb{R}^4 \rightarrow \mathbb{R}\}_{\alpha=1 \dots 4}$$

cf. 1.1.3 Principle of Relativity, pp. 9, 1 Basic Principles of Classical Mechanics, Arnold, Kozlov, Neishtadt (2006) [32].

direct product  $E^3 \times \mathbb{R}\{t\}$  (spacetime) has natural structure of affine space

Galilean group is by def. group of all affine transformations of  $E^3 \times \mathbb{R}$  that preserve time intervals and are isometries of space  $E^3 \forall$  fixed  $t \in \mathbb{R}.$

Thus if

**Definition 63** (Galilean group). *Galilean group* is group of all affine transformations of  $E^3 \times \mathbb{R}$  that preserve time intervals and are isometries of space  $E^3 \forall$  fixed  $t \in \mathbb{R}.$

Thus, if Galilean transformation  $g,$

$$(98) \quad \begin{aligned} g : E^3 \times \mathbb{R}\{t\} &\rightarrow E^3 \times \mathbb{R}\{t\} \\ g : (s, t) &\mapsto (s', t') \end{aligned}$$

$$(1) \quad t_\alpha - t_\beta = t'_\alpha - t'_\beta$$

$$(2) \quad \text{if } t_\alpha = t_\beta, \text{ then } |s_\alpha - s_\beta| = |s'_\alpha - s'_\beta|$$

Consider  $(s_a, t_a), (s_b, t_b), t_a - t_b, |s_a - s_b|.$

$$g(s_a, t_a) = (s'_a, t'_a)$$

$$g(s_b, t_b) = (s'_b, t'_b)$$

$$t'_a - t'_b = \pi_t g(s_a, t_a) - \pi_t g(s_b, t_b) = t_a - t_b$$

$$|s'_a - s'_b| = |\pi_s g(s_a, t_a) - \pi_s g(s_b, t_b)| = |s_a - s_b|$$

Galilean group  $g$  acts on  $\mathbb{R}^3\{\setminus\} \times \mathbb{R}\{t\}.$

18.1. **3 examples of Galilean transformations.** uniform motion with constant velocity  $\mathbf{v}:$

$$(99) \quad g_1(\mathbf{r}, t) = (\mathbf{r} + \mathbf{v}t, t)$$

translation of origin in spacetime

$$(100) \quad g_2(\mathbf{r}, t) = (\mathbf{r} + \mathbf{x}, t + \alpha)$$

rotation of coordinate axes

$$(101) \quad g_3(\mathbf{r}, t) = (G\mathbf{r}, t)$$

where  $G : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  is orthonormal transformation.

**Proposition 16** (Galilean transformation decomposition).  $\forall$  galilean transformation  $g : \mathbb{R}^3 \times \mathbb{R} \rightarrow \mathbb{R}^3 \times \mathbb{R}$  can be uniquely represented as composition  $g_1 g_2 g_3$

cf. Prop. 1.1 of Arnold, Kozlov, Neishtadt (2006) [32]

*Proof.* Consider affine transformation of  $E^4 = E^3 \times \mathbb{R}$  of general form

$$\mathbf{x} = A\mathbf{x}' + \mathbf{v}t' + \mathbf{a}$$

$$t = \langle \mathbf{l}, \mathbf{x}' \rangle + kt' + s$$

$$A \in \text{Mat}_{\mathbb{R}}(3, 3), \mathbf{v}, \mathbf{a}, \mathbf{l} \in \mathbb{R}^3, k, s \in \mathbb{R}.$$

By galilean transformation (Def. 63),  $t_2 - t_1 = t'_2 - t'_1,$  so

$$t_2 - t_1 = \langle \mathbf{l}, \mathbf{x}'_2 \rangle + kt'_2 + s - \langle \mathbf{l}, \mathbf{x}'_1 \rangle - kt'_1 - s = \langle \mathbf{l}, \mathbf{x}'_2 - \mathbf{x}'_1 \rangle + k(t'_2 - t'_1) = t'_2 - t'_1$$

Since each of the 3 components of  $\mathbf{x}_2 - \mathbf{x}_1$  are arbitrary,  $\mathbf{l} = 0, k = 1$  clearly. Or, let  $t = 2 = t_1.$  Then clearly  $\mathbf{l} = 0.$

By galilean transformation (Def. 63),  $|\mathbf{x}_2 - \mathbf{x}_1| = |\mathbf{x}'_2 - \mathbf{x}'_1|$

Let  $t_2 = t_1.$

$$|\mathbf{x}_2 - \mathbf{x}_1| = |A\mathbf{x}'_2 + \mathbf{v}t'_2 + \mathbf{a} - A\mathbf{x}'_1 - \mathbf{v}t'_1 - \mathbf{a}| = |A(\mathbf{x}'_2 - \mathbf{x}'_1) + \mathbf{v}(t'_2 - t'_1)| = |A(\mathbf{x}'_2 - \mathbf{x}'_1)| = |\mathbf{x}'_2 - \mathbf{x}'_1|$$

Now let  $\mathbf{y} = \mathbf{x}'_2 - \mathbf{x}'_1.$

$$|A\mathbf{y}|^2 = y^T A^T A y = y^T y$$

only if  $A^T A = 1.$  So  $A$  orthonormal.

Thus we've found the *general form of Galilean transformations:*

$$(102) \quad \boxed{\begin{aligned} \mathbf{x} &= A\mathbf{x}' + \mathbf{v}t' + \mathbf{a}, & A &\in O(3), \mathbf{v}, \mathbf{a} \in \mathbb{R}^3, s \in \mathbb{R} \\ t &= t' + s \end{aligned}}$$

Since orthogonal matrices form a 3 parameter family,  $\dim A + \dim \mathbf{v} + \dim \mathbf{a} + \dim s = 3 + 3 + 3 + 1 = 10$  independent parameters.

Recall that  $g_1(\mathbf{r}, t) = (\mathbf{r} + \mathbf{v}t, t), g_2(\mathbf{r}, t) = (\mathbf{r} + \mathbf{a}, t + s), g_3(\mathbf{r}, t) = (A\mathbf{r}, t),$  then  $g = g_1 g_2 g_3.$

□

Introduce  $E^3$  a "fixed" reference frame by fixing pt.  $o \in E^3$  and choosing 3 mutually perpendicular axes.

*inertial frame* - frame moving uniformly and rectilinearly with respect to fixed pt.  $o \in E^3; \forall$  element of Galilean group transforms fixed frame into inertial frame.

action of galilean group on  $E^3 \times \mathbb{R}$  can be extended to action on  $E^3 \times \dots \times E^3 \times \mathbb{R}$  by the rule:

if  $g : (s, t) \rightarrow (s', t'),$  then

$$g(s_1, \dots s_n, t) = g(s'_1, \dots s'_n, t')$$



Galileo-Newton principle of relativity asserts that Newton’s equations invariant under Galilean transformation group in inertial reference frame.

$M$  = 4-dim. affine manifold (spacetime). Recall definition of an *affine manifold* from [wikipedia](#).  
Affine manifold  $M$  is real manifold with charts  $\psi_i : U_i \rightarrow \mathbb{R}^n$ , s.t.  $\psi_i \circ \psi_j^{-1} \in \text{Aff}(\mathbb{R}^n) \quad \forall i, j$  where  $\text{Aff}(\mathbb{R}^n)$  denotes Lie group of affine transformations;

Recall the definition of a *Lie group of affine transformations*,  $\text{Aff}(\mathbb{R}^n)$ , the ”affine structure”, <sup>5</sup>,

**Definition 64** (Lie group of affine transformations, affine structure).

(103) 
$$\text{Aff}(\mathbb{R}^n) = \left\{ \begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \mid A \in GL(n, \mathbb{R}), b \in \mathbb{R}^n \right\}$$

s.t.  $\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix}$  acts on real affine space  $\widetilde{\mathbb{R}^n}$  by

(104) 
$$\begin{pmatrix} A & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ 1 \end{pmatrix} = \begin{pmatrix} Av + b \\ 1 \end{pmatrix}$$

where  $A \in GL(n, \mathbb{R})$ ,  $b \in \mathbb{R}^n$ , where  $(v, 1)^T \in \widetilde{\mathbb{R}^n}$

If changing notation, and keeping in mind that  $n = 4$ ,  $\psi_j \equiv y(v_{(y)}) \rightarrow \mathbb{R}^n$ ,

$\psi_i \equiv x : U_{(x)} \rightarrow \mathbb{R}^n$   
 $x(p) \in \mathbb{R}^n, x_i(p) \in \mathbb{R}$

Rewrite the Eq. [102](#) in matrix form:

$$\begin{aligned} \mathbf{x} &= A\mathbf{x}' + \mathbf{v}t + \mathbf{b} \\ t &= t' + s, \quad A \in O(3), \mathbf{v}, \mathbf{b} \in \mathbb{R}^3, s \in \mathbb{R} \\ \begin{pmatrix} A & \mathbf{v} & \mathbf{b} \\ 0 & 1 & s \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}' \\ t' \\ 1 \end{pmatrix} &= \begin{pmatrix} \mathbf{x} \\ t \\ 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} A & \mathbf{v} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}' \\ t' \end{pmatrix} + \begin{pmatrix} \mathbf{b} \\ s \end{pmatrix} = \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix} \end{aligned}$$

Consider  $y$  being standard Euclidean coordinates in fixed frame  $o$ .  $x$  then is coordinates in any other inertial frame.

$T$  = 1-dim. affine space since affine structure  $(T, \mathbf{T}, \beta)$  is s.t.  $T \in \mathbb{R}, \mathbf{T} \in \mathbb{R}, \beta : \mathbf{T} \times T \rightarrow T$   
$$\beta(s, t') = t' + s = t$$

Holm (2011) [\[23\]](#)

19. KINEMATICS

19.1. Kinematics Problems from Irodov. cf. Irodov (1981) [\[47\]](#)

**Solution 1.1.** If we assume that the raft goes along downstream with no further ”rowing”,  
 $l \equiv$  distance raft travels for total time  $T = 6.0 \text{ km}$   
 $v_d \equiv$  downstream speed = raft speed  
 $\tau = 60 \text{ mins}$  = time before motorboat turning back  
 $v_m \equiv$  motorboat speed

<sup>5</sup>”Affine structures on Lie algebras” by Michel Goze and Elisabeth Remm

$$\begin{aligned} (v_m - v_d)l &= 2\tau v_m v_d - l v_d \\ v_m l &= 2\tau v_m v_d \\ \frac{l}{v_d} &= \tau + \frac{\tau(v_m + v_d) - l}{v_m - v_d} = \frac{\tau(2v_m) - l}{v_m - v_d} \quad \text{or} \quad \boxed{v_d = \frac{l}{2\tau}} = \frac{6.0 \text{ km}}{2 \cdot 60 \text{ mins}} = 0.05 \text{ km / min} \end{aligned}$$

**Solution 1.2.** The speed formula, recall, is  $L = \bar{v}T$ , or  $\bar{v} = \frac{L}{T}$ .

$$\begin{aligned} v_0 t_0 &= \frac{L}{2} \quad \text{or} \quad t_0 = \frac{L}{2v_0} \\ l_1 &= v_1 \frac{T_1}{2} \quad \text{and} \quad l_1 + l_2 = \frac{L}{2}, \quad \text{so} \quad L = v_1 T_1 + v_2 T_1 \\ l_2 &= v_2 \frac{T_1}{2} \\ \implies \bar{v} &= \frac{L}{\frac{L}{2v_0} + \frac{L}{v_1 + v_2}} = \frac{2v_0(v_1 + v_2)}{2v_0 + v_1 + v_2} \end{aligned}$$

**Solution 1.3.** Given  
 $w = 5.0 \text{ m/s}^2 =$  constant acceleration and deceleration  
 $\tau = 25 \text{ sec.}$  = total time of motion  
 $\bar{v} = 72 \text{ km / hr.}$  = average velocity during that time ( $\tau$ ?)

Recall that

$a = a(t) = \frac{dv}{dt}$  so if  $a$  constant, then  $v(t) - v_0 = at$  and since  $v(t) = \frac{dx}{dt}$  so

$$\begin{aligned} x(t) &= x_0 + v_0 t + \frac{1}{2} at^2 \\ d = x(t) - x_0 &= t(v_0 + \frac{a}{2}t) = \frac{v - v_0}{a}(v_0 + \frac{v - v_0}{2}) = \frac{v^2 - v_0^2}{2a} \quad \text{or} \quad v^2 - v_0^2 = 2ad \end{aligned}$$

Then

$$\frac{v^2}{2w} + \frac{v^2}{2w} + vT_2 = L \quad \text{or} \quad \frac{v^2}{w} + vT_2 = L$$

Also

$$\tau = \frac{v}{w} + T_2 + \frac{-v}{-w} = \frac{2v}{w} + T_2 \quad \text{or} \quad (\tau - T_2) \left( \frac{w}{2} \right) = v$$

Then

$$\begin{aligned} \bar{v} &= \frac{\frac{v^2}{w} + vT_2}{\tau} \quad \text{or} \quad \bar{v}\tau = \frac{v^2}{w} + wT_2 \\ \implies \bar{v}\tau &= v \left( \frac{v}{w} + T_2 \right) = v \left( \frac{\tau - T_2}{2} + T_2 \right) = \frac{w}{2} (\tau - T_2) \left( \frac{\tau + T_2}{2} \right) \\ 4\bar{v}\tau &= w(\tau^2 - T_2^2) \quad \text{or} \quad wT_2^2 = wT^2 - 4\bar{v}\tau \\ \implies \boxed{T_2} &= \sqrt{\tau^2 - \frac{4\bar{v}\tau}{w}} \end{aligned}$$

**Solution 1.5; particles collide.**

**Hint:** *change reference frames*, in particular, into frame of references of particles 1 or 2, respectively.  
Consider, by label symmetry, reference frame of particle 1:

$$\begin{aligned} r'_1 &= r_1 - r_1 = 0 & v'_1 &= v_1 - v_1 = 0 \\ r'_2 &= r_2 - r_1 & v'_2 &= v_2 - v_1 \end{aligned}$$



On order for a collision, in reference frame of particle 1, particle 2 will need to approach or travel towards the origin from

$r'_2 = r_2 - r_1$ ; consider

$v'_2 \parallel -r'_2$  i.e.  $v_2 - v_1 \parallel -(r_2 - r_1)$ , otherwise particle 2 would never reach origin where particle 1 is

$$\implies \boxed{\frac{\mathbf{v}_2 - \mathbf{v}_1}{\|\mathbf{v}_2 - \mathbf{v}_1\|} = \frac{-(\mathbf{r}_2 - \mathbf{r}_1)}{\|\mathbf{r}_2 - \mathbf{r}_1\|}}$$

**Solution 1.6.**  $v_0 = 30 \text{ km/hr}$  = ship moves along equator to the east with speed  $v_0$

$\varphi = 60^\circ$  = southeastern wind blows at angle  $\varphi$  to equator

$v_w = 15 \text{ km/h}$  = wind speed

$v_w = v_{w,x}\mathbf{e}_x + v_{w,y}\mathbf{e}_y$

If southeastern wind means *it comes **from** the south east*, then we have

$$v'_w = v_w - v_0 = (-v_w \cos \varphi - v_0)\mathbf{e}_x + v_w s \varphi \mathbf{e}_y$$

$$\|v'_w\|^2 = v_w^2 c^2 \varphi + 2v_w v_0 c \varphi + v_0^2 + v_w^2 s^2 \varphi = v_w^2 + v_0^2 + 2v_w v_0 c \varphi = v_w^2 + v_0^2 + 2v_w v_0 \frac{1}{2} \approx 39.7 \text{ km/h} \approx 40 \text{ km/h}$$

$$\varphi' = \arctan\left(\frac{-v_w s \varphi}{v_w c \varphi + v_0}\right) = \arctan\left(\frac{-v_w \sqrt{3}/2}{v_w/2 + v_0}\right) \approx -19.1^\circ$$

cf. <https://physics.stackexchange.com/questions/126268/relative-wind-velocity-explanation> for southeastern wind explanation.

**Solution 1.7.**

2 swimmers leave pt. A, pt. A lie right across point B.

stream velocity  $v_0 = 2.0 \text{ km/hr}$

velocity  $v'$  of each swimmer with respect to water =  $2.5 \text{ km/hr}$

Now  $(\mathbf{v}_A + \mathbf{v}_0)T_A = \mathbf{x}_A$  and so in components,  $v_{A,x}T_A = L$  and  $(v')^2 = (v_{A,x})^2 + v_0^2$  or  $v_{A,x} = \sqrt{(v')^2 - v_0^2}$

$T_{B,1}$  = time to swim across the river

$T_{B,2}$  = time to walk the distance he has been carried.

Since  $(\mathbf{v}_B + \mathbf{v}_0)T_{B,1} = \mathbf{x}_{B,1}$  with  $v'T_{B,1} = L$

$$v_0 T_{B,1} = y_B$$

$T_A = T_{B,1} + T_{B,2}$  since both swimmers reach the destination simultaneously.

$$\implies \frac{L}{v_{A,x}} = \frac{L}{v'} + \frac{v_0 L}{uv'} \text{ or } u = \frac{v_0/v'}{\frac{1}{v_{A,x}} - \frac{1}{v'}} = \boxed{\frac{v_0}{\frac{v'}{\sqrt{(v')^2 - v_0^2}} - 1}} = 3.0 \text{ km/hr}$$

**Solution 1.8.**

Buoy anchored at the middle of a river.

$v_s$  = stream velocity

$$(v_A + v_s)T_{A,1} = L = (v_A - v_s)T_{A,2}$$

$$\implies (\eta + 1)v_s T_{A,1} = L = (\eta - 1)v_s T_{A,2}$$

where  $T_{A,1} = \frac{L}{(\eta+1)v_s}$  is the time going away from the buoy to a distance L for person A, and  $T_{A,2} = \frac{L}{(\eta-1)v_s}$  is the time to go back.

B moves away along the mutually perpendicular straight lines across the river. Then

$$(\mathbf{v}_B + \mathbf{v}_s)T_{B,1} = \mathbf{x}_B$$

$$v_{B,x}T_{B,1} = L$$

$$v_B^2 = (\eta v_s)^2 = v_{B,x}^2 + v_{B,y}^2 = v_{B,x}^2 + v_s^2$$

Now  $T_{B,1} = T_{B,2}$ , i.e. the time to go away from the buoy is equal to the time to go back to the buoy, and so

$$2T_{B,1} = \frac{2L}{\sqrt{(\eta v_s)^2 - v_s^2}} = \frac{2L}{v_s \sqrt{\eta^2 - 1}}$$

Thus,

$$\frac{\tau_A}{\tau_B} = \frac{T_{A,1} + T_{A,2}}{T_{B,1} + T_{B,2}} = \frac{\frac{2\eta}{\eta^2-1}}{\frac{2}{\sqrt{\eta^2-1}}} = \boxed{\frac{\eta}{\sqrt{\eta^2-1}}} \approx 1.81 \approx 1.8$$

**Solution 1.10.**

For body 1 going straight up,  $y_1 = v_0 t - \frac{1}{2}gt^2$  and there's no horizontal,  $x_1$ , translation.

For body 2, the coordinates over time are the following:

$$y_2 = v_0 s \theta t - \frac{1}{2}gt^2$$

$$x_2 = v_0 c \theta t$$

$$\|\mathbf{x}_2 - \mathbf{x}_1\|^2 = (v_0 t(1 - s\theta))^2 + (v_0 c \theta t)^2 = (v_0 t)^2 [2 - 2s\theta] = 2(v_0 t)^2 (1 - s\theta)$$

$$\|\mathbf{x}_2 - \mathbf{x}_1\| = \boxed{v_0 t \sqrt{2} \sqrt{1 - s\theta}} = 25(1.7)\sqrt{2}\sqrt{1 - \sqrt{3}/2} \approx 22$$

**Solution 1.11.** 2 particles move in a uniform gravitational field with acceleration  $g$ .

At initial moment, the particles were located at 1 pt. and moved with velocities  $v_1 = 3.0 \text{ m/s}$ ,  $v_2 = 4.0 \text{ m/s}$  horizontally in opposite directions.

Typically, gravitational field is **vertical**:

$$v_{1y} = -gt \quad v_{2y} = -gt$$

$$v_{1x} = -v_1 \quad v_{2x} = v_2$$

$$y_1 = -\frac{1}{2}gt^2 \quad y_2 = -\frac{1}{2}gt^2$$

$$x_1 = -v_1 t \quad x_2 = v_2 t$$

$$\text{mutually perpendicular: } \mathbf{v}_1 \cdot \mathbf{v}_2 = 0 = -v_1 v_2 + g^2 t^2 \text{ or } t = \frac{\sqrt{v_1 v_2}}{g}$$

$$\|\mathbf{x}_2 - \mathbf{x}_1\|^2 = ((v_2 + v_1)t)^2 + 0 = (v_2 + v_1)^2 \frac{v_1 v_2}{g^2} \text{ or}$$

$$\boxed{\|\mathbf{x}_2 - \mathbf{x}_1\| = \frac{(v_1 + v_2)}{g} \sqrt{v_1 v_2}} = \frac{7}{9.8 \text{ m/s}^2} 2\sqrt{3} = 2.5 \text{ m}$$

**Problem 1.12.** Three points are located at the vertices of an equilateral triangle whose side equals  $a$ . They all start moving simultaneously with velocity  $v$  constant in modulus, with the first point heading continually for the second, the second for the third, and the third for the first. How soon will the points converge?

**Solution 1.12.** By symmetry at any time (point), consider component of  $\mathbf{v}$  towards triangle center (EY: this is the key insight; breakup components of  $\mathbf{v}$  to be along "radial" distance traveled, not just along Cartesian x and y).

By symmetry of the problem,  $v$  will always be along an equilateral triangle side, for each moving point.

Let  $\theta_0 = 30^\circ$ .

Then

$$\cos \theta_0 = \frac{v_r}{v} \text{ or } v \cos \theta_0 = v_r$$
$$\cos \theta_0 = \frac{a/2}{d} \text{ or } d = \frac{a/2}{\cos \theta_0} = \frac{a/2}{\sqrt{3}/2} = \frac{a}{\sqrt{3}}$$

where  $d$  is the distance from a point to the center of the triangle.

$$v_x = \frac{ds}{dt} \implies T = \int dt = \int \frac{ds}{v_x} = \int_{a/\sqrt{3}}^0 \frac{ds}{-v \cos \theta_0} = \frac{-a}{\sqrt{3}} \frac{1}{v\sqrt{3}/2} = \frac{2a}{3v}$$

**Problem 1.13.** Point  $A$  moves uniformly with velocity  $v$  so that the vector  $\mathbf{v}$  is continually "aimed" at point  $B$  which in its turn moves rectilinearly and uniformly with velocity  $u < v$ . At the initial moment of time  $\mathbf{v} \perp \mathbf{u}$  and the points are separated by a distance  $l$ . How soon will the points converge?

**Solution 1.13.** Choose  $\mathbf{u}$  to be in the  $x+$  direction:  $\mathbf{u} = u\mathbf{e}_x$ .

Condition: Point A reaches point B, so total x-distance traveled during time interval  $T$  is the same:

$$(105) \quad \int_0^T v \cos \alpha dt = uT$$

Now consider the "straight line distance" between point A and point B. We're given at the initial moment of time  $\mathbf{v} \perp \mathbf{u}$ , that the points are separated by distance  $l$ .

Now think in terms of displacement in very short time interval  $\Delta t$ . Point A is traveling distance  $v\Delta t$  towards point B in a straightline. But point B is moving away with component along the straight line by distance  $u \cos \alpha \Delta t$ .

The total straight line distance point A has to cover is distance  $l$ . The net distance traveled in  $\Delta t$ ,  $v\Delta t - u \cos \alpha \Delta t$  must sum up to  $l$  in order for point A and point B to converge. For a small time interval  $\Delta t$ , the  $-u \cos \alpha$  term already is accounting for how much point B is moving away and  $v$  term accounts for how fast point A is traveling the distance towards point B.

So

$$(106) \quad \int_0^T (v - u \cos \alpha) dt = l = vT - u \left( \frac{uT}{v} \right) \text{ or } \boxed{T = \frac{lv}{v^2 - u^2}}$$

<https://physics.stackexchange.com/questions/568531/why-is-the-following-equation-not-correct> <https://www.physicsforums.com/threads/a-question-of-kinematics-involving-a-changing-velocity-vector.975126/post-6210692>

20. FUNDAMENTAL THEOREMS OF (CLASSICAL) DYNAMICS

cf. Sec. 3.4 - *Fundamental Theorems of Dynamics* of Prástaro (1996) [15].

cf. Thm. 3.20 of Prástaro (1996) [15]

**Theorem 24** (Momentum Theorem). *Variation of the free part of momentum of the observed motion of 1 body, in time interval  $\Delta t \equiv [0, t]$  is equal to the corresponding impulse:*

$$(107) \quad I[0, t] \equiv \int_0^t F dt \equiv \left( \int_0^t F^j dt \right) \mathbf{e}_j$$

where  $\{\mathbf{e}_j\}_{1 \leq j \leq 3}$  is a fixed basis of  $\mathbf{S}$

$$(108) \quad \bar{p}_\psi(t) - \bar{p}_\psi(0) = I[0, t]$$

*Proof.*

$$\bar{p}_\psi = \mu \ddot{\mathbf{m}}_\psi = \bar{\mathbf{f}}_\psi \implies \dot{\bar{p}}_\psi = \dot{p}^j \mathbf{e}_j = F^j \mathbf{e}_j \implies \int_{[0, t]} \dot{p}^j dt = \int_{[0, t]} F^j dt$$

**20.1. Coordinate transformation.** For  $(U, \varphi)$ ,  $(V, \psi)$  charts on  $M$ , s.t.  $\varphi : U \rightarrow \mathbb{R}^n$ ,  $\varphi^i \equiv x^i$ , then consider  $\psi : V \rightarrow \mathbb{R}^n$   $\psi^j \equiv y^j$

$$\psi(p) = y = \psi \circ \varphi^{-1} \circ \varphi(p) = \psi \circ \varphi^{-1}(x) = y(x)$$

Consider  $f : M \rightarrow M$  that is a coordinate transformation (bijective). Under this coordinate transformation, transformation of the components of the same vector  $X = Y$  is the push-forward:

$$(109) \quad Y^j = X^i \frac{\partial y^j}{\partial x^i}$$

Correspondingly, basis transformation described by

$$(110) \quad \frac{\partial}{\partial x^j} = \frac{\partial y^k}{\partial x^j} \frac{\partial}{\partial y^k} \text{ or } \frac{\partial}{\partial y^k} = \left[ \left( \frac{\partial y}{\partial x} \right)^{-1} \right]_k^j \frac{\partial}{\partial x^j} = \left( \frac{\partial x^j}{\partial y^k} \right) \frac{\partial}{\partial x^j}$$

Then we can show that indeed  $Y = X$  under coordinate transformations:

$$(111) \quad Y = X = Y^k \frac{\partial}{\partial y^k} = X^j \frac{\partial}{\partial x^j}$$

Also, for  $a_i = b_j \frac{\partial y^j}{\partial x^i}$ , the pull back describing how 1-form components transform, consider the inner product:

$$(112) \quad b(Y) = b_i dy^i \left( Y^k \frac{\partial}{\partial y^k} \right) = b_i Y^i = b_i X^k \frac{\partial y^i}{\partial x^k} = b_i \frac{\partial y^i}{\partial x^k} X^k = a_k X^k = a(X)$$

$\implies$  inner product preserved under coordinate transformation, i.e. invariance of inner product.

**20.2. Transformation of reference frames.** cf. Sec. 4.2.1 "Transformation of reference frames", Kambe (2009) [24].

Let us take pt.  $\mathbf{y}_b$  fixed to body, located at  $\mathbf{x} = X^1 \mathbf{e}_1 + X^2 \mathbf{e}_2 + X^3 \mathbf{e}_3 = X^i \mathbf{e}_i$  at  $t = 0$ , where  $\mathbf{e}_i$  ( $i = 1, 2, 3$ ) are orthonormal basis fixed to space  $F \equiv$  nonrotating **fixed** space (inertial space)

By transformation matrix  $A = (A^i_j) \in \text{SO}(3)$ , initial pt.  $\mathbf{x} = (X^i)$  mapped to current pt.  $\mathbf{y}_b$  at time  $t$ .

$$(113) \quad \mathbf{y}_b(t) = A(t)\mathbf{x} = y^i(t)\mathbf{e}_i, \quad y^i(t) = A^i_j(t)X^j$$

cf. Eq. (4.14) of Kambe (2009) [24].  
i.e.

$$\mathbf{y}_b(t) = A(t)\mathbf{x} = A^i_k e_i \otimes e^k x^l e_l = A^i_k X^k e_i = y^i(t) e_i$$

EY: Observe that  $A(t)$ ,  $t \in [0, t]$  is a curve in  $\text{SO}(3)$ .

In other words, we considered a point in the fixed frame  $F$  given by frame  $\{\mathbf{e}_i\}$ :

$$\mathbf{x} = X^i \mathbf{e}_i \text{ at } t = 0$$

Then, we considered an "active" rotation  $A \in \text{SO}(3)$  of this point, so

$$(114) \quad \mathbf{y}(t) = A(t)\mathbf{x} = A^i_j(t)X^j \mathbf{e}_i = y^i(t)\mathbf{e}_i$$

Indeed, consider  $A = \begin{bmatrix} c & -s \\ s & c \\ & & 1 \end{bmatrix}$ . Then  $A$  would rotate points in the  $x - y$  plane, such as  $[1, 0, 0]$ ,  $[0, 1, 0]$  counter-clockwise in

□ the positive  $z$  axis direction.

On the other hand, relative to body frame  $F_B$ , which is the frame instantaneously fixed to the moving body. For the same pt.,  $\mathbf{y}_b(t)$ , fixed to body, is

$$\mathbf{Y}(=\mathbf{y}_b) = Y^1\mathbf{b}_1 + Y^2\mathbf{b}_2 + Y^3\mathbf{b}_3 = Y^i\mathbf{b}_i$$

where  $\mathbf{b}_i$  ( $i = 1, 2, 3$ ) are orthonormal basis fixed to body which coincided with  $\mathbf{e}_i$  ( $i = 1, 2, 3$ ) at  $t = 0$ .

From property of a rigid body, it's required that  $Y^i = X^i$  which don't change with  $t$ .

According to Sec. 1.5.1(b) (Kambe (2009) [24]), basis transformation written as (1.40) Kambe(2009) [24], where

$$(115) \quad \mathbf{b}_i = \mathbf{b}_i(t) = \mathbf{e}_j B_i^j(t) \\ \text{Then } \mathbf{Y} = Y^i \mathbf{b}_i = Y^i \mathbf{e}_j B_i^j = \mathbf{e}_j B_i^j Y^i$$

$$(4.10) \text{ Kambe (2009) [24], where } B = (B_k^j) \in \text{SO}(3), \text{ i.e. } BB^T = 1.$$

Note that

$$(116) \quad \mathbf{Y} = Y^i \mathbf{b}_i = \mathbf{e}_j B_i^j Y^i = A_k^j X^k \mathbf{e}_j$$

implies that  $B : F' \rightarrow F$  where  $F'$  rotating frame,  $F$  fixed frame, i.e.  $B$  is a coordinate transformation from coordinates in the  $F'$  rotating frame into coordinates in the  $F$  fixed frame.

Indeed, consider

$$B = \begin{bmatrix} c & -s & \\ s & c & \\ & & 1 \end{bmatrix}. \text{ Then } B \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = [c, s, 0] \text{ (in } \{\mathbf{e}_j\} \text{ frame)}$$

Now recall that  $\mathbf{x} = X^i \mathbf{e}_i$  at  $t = 0$ .  $\mathbf{Y} = Y^i \mathbf{b}_i$  and  $\{\mathbf{b}_i\}$  frame axes had coincided with  $\{\mathbf{e}_i\}$  frame axes at  $t = 0$ .  $Y^i$ 's are constant, so  $Y^i = X^i$

So for

$$(117) \quad \mathbf{Y} = Y^i \mathbf{b}_i = \mathbf{e}_j B_i^j Y^i = A_k^j(t) X^k \mathbf{e}_j \implies B_i^j Y^i = A_k^j(t) X^k = A_i^j Y^i \text{ so } \boxed{B = A} \text{ since } X^i \text{ arbitrary}$$

For the basis transformation derivation, recall the push forward transformation (cf. Kambe (2009) [24]),

$$\begin{aligned} \phi : M &\rightarrow N \\ \phi_* : T_x M &\rightarrow T_y N \\ y = \phi(x) &= F(x) \\ Y = \phi_* X = F_* X &= \left. \frac{d}{dt} (F(p(t))) \right|_{t=0} \\ Y^k &= (\phi_* X)^k = \left( \frac{\partial F^k}{\partial x^j} \right) X^j \\ Y &= \phi_* X = \phi_* \left[ X^j \frac{\partial}{\partial x^j} \right] = X^j \phi_* \left[ \frac{\partial}{\partial x^j} \right] = X^j \frac{\partial y^k}{\partial x^j} \frac{\partial}{\partial y^k} = Y^k \frac{\partial}{\partial y^k} \\ Y^k &= \frac{\partial y^k}{\partial x^j} X^j \end{aligned}$$

20.2.1. *Direction Cosine Matrix (DCM)*. cf. pp. 134, 4.1 "The Independent Coordinates of a Rigid Body", Goldstein, Poole, Safko (2001) [27]

Consider frame  $F$ ,  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ ,  
frame  $F'$ ,  $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$

If

$$\mathbf{e}'_i = C_{ij} \mathbf{e}_j$$

and so then

$$\begin{aligned} \mathbf{e}'_i \cdot \mathbf{e}_k &= C_{ij} \delta_{jk} = C_{ik} \\ \implies \mathbf{e}'_i &= (\mathbf{e}'_i \cdot \mathbf{e}_j) \mathbf{e}_j \end{aligned}$$

So for  $\mathbf{r}' = y^i \mathbf{e}'_i = x^j \mathbf{e}_j$ ,

$$(118) \quad \implies y^k = x^j (\mathbf{e}'_k \cdot \mathbf{e}_j) \text{ or } y^i = (\mathbf{e}'_i \cdot \mathbf{e}_j) x^j$$

So  $\cos \theta_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j$  completely specifies the orientation of  $F'$  relative to  $F$ .

Now use this fact, that  $\mathbf{e}_m \cdot \mathbf{e}'_{m'} = \delta_{mm'}$ , to show that

$$\mathbf{e}_m \cdot \mathbf{e}'_{m'} = \delta_{mm'} = (\mathbf{e}_m \cdot \mathbf{e}'_l) \mathbf{e}'_l \cdot (\mathbf{e}_{m'} \cdot \mathbf{e}'_{l'}) \mathbf{e}'_{l'} = (\mathbf{e}_m \cdot \mathbf{e}'_l) (\mathbf{e}_{m'} \cdot \mathbf{e}'_{l'}) \delta_{ll'} = (\mathbf{e}'_l \cdot \mathbf{e}_m) (\mathbf{e}'_l \cdot \mathbf{e}_{m'}) = \cos \theta_{lm} \cos \theta_{lm'}$$

So  $\cos \theta_{ij}$  are 9 coordinates that can't be generalized coordinates.

Be careful about the interpretation of the DCM. Note that we had derived the following from Eq. 118:

$$(119) \quad \boxed{y^i = C_{ij} x^j} \text{ with } \mathbf{e}'_i = C_{ij} \mathbf{e}_j \text{ and } \boxed{C_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j}$$

So this is a transformation of the same vector between two different frames, not an "active" rotation of the vector itself. Nevertheless,  $C_{ij} : F \rightarrow F'$ , a mapping from frame  $F$  to  $F'$ .

For example, consider

$$\begin{aligned} \mathbf{e}'_x &= \cos(\omega t) \mathbf{e}_x - \sin(\omega t) \mathbf{e}_y \\ \mathbf{e}'_y &= \sin(\omega t) \mathbf{e}_x + \cos(\omega t) \mathbf{e}_y \end{aligned}$$

This isn't an active rotation in the negative z, clockwise direction, but this is the same vector with coordinates in a frame that's rotating in the counterclockwise, positive z direction.

20.3. **Euler Angles.** cf. pp. 150, 4.4 "Euler Angles", Goldstein, Poole, Safko (2001) [27]

rotation 1: about vertical axis, gives **heading** or **yaw** angle

rotation 2: around perpendicular axis fixed in vehicle and normal to figure axis; it's measured by the *pitch* or *attitude* angle

rotation 3: about figure axis of vehicle, *roll* or *bank* angle.

$\implies$  xyz - convention, or *Tait-Bryan* angles.

**Solution Exercise 24, Ch. 4 Goldstein, Poole, Safko (2001) [27].**

Recall that

$$\mathbf{a}_r = \mathbf{a}_i - 2\omega \times \mathbf{v}_r - \omega \times (\omega \times \mathbf{r}) - \dot{\omega} \times \mathbf{r}$$

Let  $\dot{\omega} = 0$ .

Then consider, multiplying by  $m$ ,

$$\begin{aligned} m\mathbf{a}_r &= m\mathbf{a}_i - 2m\omega \times \mathbf{v}_r - m\omega \times (\omega \times \mathbf{r}) \text{ or} \\ \mathbf{F}_r &= \mathbf{F}_i - 2m\omega \times \mathbf{v}_r - m\omega \times (\omega \times \mathbf{r}) \end{aligned}$$

where  $\mathbf{F}_r$  is the apparent force experienced by the bug in the rotating frame, and  $\mathbf{F}_i$  are the forces in the inertial reference frame.

By the geometry of the wheel and the problem, and taking the instantaneous coordinate axis  $\mathbf{e}_r$  to be same for the rotating and inertial frames,

$$\begin{aligned} -\omega \times (\omega \times \mathbf{r}) &= \omega^2 R \mathbf{e}_r \\ -2\omega \times \mathbf{v}_r &= 2\omega v_r (- + \mathbf{e}_y) \end{aligned}$$

where the direction of the last expression for the Coriolis force is negative when the apparent velocity is in the positive radial direction and vice versa.

Suppose that the frictional force acts in the opposite direction of the force. In the rotating frame, consider the sum of the 2 fictitious forces, the Coriolis force and centripetal force, which happen to be orthogonal to each other:

$$\|\mathbf{F}_{\text{fic}}\| = \|(-2\omega \times \mathbf{v}_r + \omega \times (\omega \times \mathbf{r}))m\| = m\sqrt{(2\omega v_r)^2 + (\omega^2 R)^2} = m\omega\sqrt{(2v_r)^2 + (\omega R)^2}$$

When  $F_{\text{fric}} \geq F_{\text{fic}}$ , then

$$(\mu g)^2 \geq \omega^2(4v_r^2 + \omega^2 R^2) \text{ or } \left(\frac{\mu g}{\omega}\right)^2 \geq 4v_r^2 + \omega^2 R^2 \text{ or } \omega^2 R^2 \leq \left(\frac{\mu g}{\omega}\right)^2 - 4v_r^2$$

With  $v_r = 0.5 \text{ cm/s}^2$  while  $g = 9.8 \text{ m/s}^2$ , the Coriolis force is negligible compared to max. frictional force.

**20.4. Right Invariance, Left Invariance, Angular Velocity.** For an infinitesimal time increment  $\delta t$ , motion of the body from position  $g_t$  is described by the left translation  $g_{t+\delta t} = g_{\delta t} g_t$ .

This is interpreted as an infinitesimal rotation  $\delta t \bar{\Omega}$  at  $g_t$ , matrix  $\bar{\Omega}$  defined by

$$(120) \quad \begin{aligned} A(t + \delta t) - A(t) &= A(\delta t)A(t) - A(t) = (A(\delta t) - 1)A(t) = \delta t \bar{\Omega} A(t) \\ A(t + \delta t)^i{}_k &= (A(\delta t)A(t))^i{}_k = A(\delta t)^i{}_l A(t)^l{}_k \end{aligned}$$

cf. 4.16 Kambe (2009) [24]

$\bar{\Omega} = \frac{A(\delta t) - 1}{\delta t}$  is the tangent vector at 1.

Recall that, in general, mathematically,

$$\dot{g}_t = X \circ g_t, \text{ where } X := \lim_{\Delta t \rightarrow 0} \frac{(g_{\Delta t} - 1)}{\Delta t}$$

$$(121) \quad \dot{g}_t := \frac{dg_t}{dt} = \frac{dA}{dt} = \bar{\Omega} g_t$$

so

$$(122) \quad \bar{\Omega} = \dot{g}_t \circ (g_t)^{-1}$$

In my notation, for  $O \in SO(N)$ ,

$$(123) \quad \hat{\omega} = \dot{O} O^T$$

20.4.1. 2 *Physical Interpretations of rotations: frame transformation and "active" rotation.* Consider  $O \in SO(N)$ ,  $N = 3$ .  
 $O : \mathbb{R}^N \rightarrow \mathbb{R}^N$

$O$  can be interpreted as either  
 $O$  : rotating body frame  $\rightarrow$  lab frame, i.e.  
 $O : K \rightarrow K'$  (notation used in Hand and Finch (1998) [26] where  $K$  is the body frame,  $K'$  is the fixed frame)  
 $O : \searrow \mapsto \mathbf{r}'$

or  $O$  interpreted as "active" rotation of vector in  $\mathbb{R}^N$ .

Consider this example:

$$(124) \quad \begin{aligned} O &= \begin{bmatrix} \cos(\omega t) & -\sin(\omega t) & 0 \\ \sin(\omega t) & \cos(\omega t) & \\ & & 1 \end{bmatrix} & \dot{O} &= \omega \begin{bmatrix} -s & -c \\ c & -s \\ & & \end{bmatrix} \\ \dot{O} O^T &= \omega \begin{bmatrix} -s & -c \\ c & -s \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \\ & & 1 \end{bmatrix} = \omega \begin{bmatrix} & -1 \\ 1 & \end{bmatrix} = \hat{\omega} \end{aligned}$$

s.t.  $\omega = \omega \mathbf{e}_z$ .

$\dot{O} O^T$  is the angular velocity of the rotating frame, or angular velocity of how much the vector in  $\mathbb{R}^N$  is rotated by.

20.4.2. *Instantaneous Angular Velocity.* Given  $\mathbf{r} = (r^1, r^2, r^3) \in M$ , an infinitesimal transformation of  $\mathbf{r}$  by  $\xi(t) = 1 + t\mathbf{a} + O(t^2)$ .

$$\begin{aligned} \mathbf{r} &\mapsto \mathbf{r}' = \xi(t)\mathbf{r} = \mathbf{r} + t\mathbf{a}\mathbf{r} \\ \mathbf{a}\mathbf{r} &= \begin{bmatrix} & -a^3 & a^2 \\ a^3 & & -a^1 \\ -a^2 & a^1 & \end{bmatrix} \begin{bmatrix} r^1 \\ r^2 \\ r^3 \end{bmatrix} = \begin{bmatrix} -a^3 r^2 + a^2 r^3 \\ a^3 r^1 - a^1 r^3 \\ -a^2 r^1 + a^1 r^2 \end{bmatrix} = (\hat{\mathbf{a}} \times \mathbf{r}) = \epsilon_{ijk} a^j r^k \mathbf{e}_i \end{aligned}$$

So multiplication by a skew symmetric matrix is equivalent to taking the cross product with the axial vector representation.

Now  $\bar{\Omega}$  shown to be skew-symmetric (from a Taylor series expansion). See either Appendix C.3 of Kambe (2009) [24] or Eq. 34.

Make the following notation change from Kambe's (2009) [24]:

$$(125) \quad \hat{\Omega} \equiv \omega$$

where  $\omega$  is an axial vector of the angular velocity of body rotation relative to the *fixed* reference frame  $F$  (because in body frame,  $F_B$ , there's no rotation!).

**Solution Question 9:  $U$  Matrix, pp. 265, Ch. 7 "Rotating Coordinate Systems", Hand and Finch (1998) [26].**

Find the form of  $\mathbf{U}$  for a fixed rotation  $\theta$  about the  $X$  axis:

$$O = \begin{bmatrix} & c & -s \\ & s & c \end{bmatrix}$$

Then

$$\begin{aligned} \dot{O} &= \omega \begin{bmatrix} & -s & -c \\ & c & -s \end{bmatrix} \\ \dot{O} O^T &= \omega \begin{bmatrix} & -s & -c \\ & c & -s \end{bmatrix} \begin{bmatrix} & c & s \\ & -s & c \end{bmatrix} = \omega \begin{bmatrix} & & \\ & 1 & -1 \end{bmatrix} \end{aligned}$$

Thus

$$\dot{O} O^T = \hat{\omega} \mapsto \omega = \omega \mathbf{e}_x$$

So we get the form for  $\omega$ . Now consider

$$\hat{\omega}_b = O^T \hat{\omega} O = \begin{bmatrix} & c & s \\ & -s & c \end{bmatrix} \omega \begin{bmatrix} & & \\ & 1 & -1 \end{bmatrix} O = \omega \begin{bmatrix} & s & -c \\ & c & s \end{bmatrix} \begin{bmatrix} & c & s \\ & s & c \end{bmatrix} = \omega \begin{bmatrix} & & \\ & 1 & -1 \end{bmatrix}$$

$\hat{\omega}_b$  is  $\hat{\omega}$  expressed in body coordinates.  $\hat{\omega} = \hat{\omega}_b$  in this particular case because  $\hat{\omega}_b$  is the angular velocity measured in the inertial frame, but expressed in the body frame coordinates.

See <https://physics.stackexchange.com/questions/88398/why-is-body-frame-angular-velocity-nonzero/88401#88401>

For  $g_t$  operating on  $\mathbf{x}$ , let's follow Kambe's development in pp. 133-134, Eqns. (4.17)-(4.19) (Kambe (2009) [24]). So then

$$g_t \mathbf{x} = \mathbf{y}(t), \quad g_{t+\delta t} \mathbf{x} = \mathbf{y}(t + \delta t)$$

Then

$$\begin{aligned} \mathbf{y}(t + \delta t) &= \mathbf{y}(t) + \frac{d\mathbf{y}}{dt}(t) \delta t + O((\delta t)^2) = \mathbf{y}(t) + (\dot{g}_t \mathbf{x}) \delta t + \cdots = \mathbf{y}(t) + (\bar{\Omega} g_t \mathbf{x}) \delta t + \cdots = \\ &= g_t \mathbf{x} + (\omega \times (g_t \mathbf{x})) \delta t \end{aligned}$$

Let

$$\mathbf{v}_y = \frac{d\mathbf{y}}{dt} = \lim_{\delta t \rightarrow 0} \frac{\mathbf{y}(t + \delta t) - \mathbf{y}(t)}{\delta t} = \omega \times (g_t \mathbf{x}) = \omega \times \mathbf{y}$$

This means that point  $\mathbf{y}(t)$  is moving with velocity  $\mathbf{v}_y = \omega \times \mathbf{y}$ , i.e.  $\omega$  *angular velocity* in fixed space.

But typically,  $\mathbf{x} = \mathbf{x}(t)$ , i.e. the position on the body is time dependent, it just doesn't "stay on the body".

Recall that  $\mathbf{x} = \mathbf{x}(t)$  are coordinates in body frame  $F_B$ .  $\mathbf{y}(t)$  are coordinates in fixed reference frame  $F$ .

$$\mathbf{y}(t) = g_t \mathbf{x}(t)$$

Compare this notation to the one in Hand and Finch (1998) [26] in Eq. (7.18):

$$\mathbf{r}' = \tilde{U} \mathbf{r} \equiv U \mathbf{r}$$

where the rotated or body frame is  $K$ , and fixed, inertial reference frame  $K'$ , and I modify the notation for  $U$  to be an element in  $SO(3)$ .

So instead of how Kambe assumed above that the position is fixed in the body frame, assume that it's also time-dependent in the rotating frame. So use chain rule:

$$\frac{d\mathbf{y}}{dt}(t) = \frac{d}{dt}(g_t \mathbf{x}) = \dot{g}_t \mathbf{x} + g_t \dot{\mathbf{x}}(t) = \dot{g}_t \mathbf{x} + g_t \mathbf{v}_{\mathbf{x}} = \bar{\Omega} g_t \mathbf{x} + g_t \mathbf{v}_{\mathbf{x}}$$

$$(126) \quad \implies \mathbf{v}_{\mathbf{y}} = \omega \times \mathbf{y} + g_t \mathbf{v}_{\mathbf{x}}$$

Follow the development in Hand and Finch (1998) [26]:

$$(127) \quad \begin{aligned} \mathbf{v}|'_{\text{space}} &= \dot{U} \mathbf{r} + U \dot{\mathbf{r}}|_{\text{body}} = \dot{U} \mathbf{r} + U \mathbf{v}|_{\text{body}} = \dot{U} U^{-1} U \mathbf{r} + U \mathbf{v}|_{\text{body}} = \dot{U} U^{-1} \mathbf{r}' + U \mathbf{v}|_{\text{body}} = \\ &= \omega \times \mathbf{r} + \mathbf{v}|_{\text{body}} \end{aligned}$$

What we wrote is essentially a summary of Eqns. (7.28), (7.29) Hand and Finch (1998) [26].

Compare this, namely Eq. 126, with wikipedia ("**Rotating reference frame**"), where the velocities in the 2 reference frames are related by:

$$(128) \quad \mathbf{v}_i = \mathbf{v}_r + \boldsymbol{\Omega} \times \mathbf{r}$$

where  $i \equiv$  inertial frame of reference,  $r \equiv$  rotating reference frame.

Now

$$\begin{aligned} \mathbf{v}_{\mathbf{y}} &= \omega \times \mathbf{y} + g_t \mathbf{v}_{\mathbf{x}} \\ \mathbf{a}_{\mathbf{y}} &= \frac{d}{dt} \mathbf{v}_{\mathbf{y}} \\ \frac{d}{dt} \omega \times \mathbf{y} &= \frac{d}{dt} \epsilon_{ijk} \omega_j y_k = \epsilon_{ijk} \dot{\omega}_j y_k + \epsilon_{ijk} \omega_j \dot{y}_k \\ \text{For } \dot{y}_k, \text{ consider } \frac{d\mathbf{y}}{dt} &= \omega \times \mathbf{y} + g_t \mathbf{v}_{\mathbf{x}} \\ \text{so } \omega \times \frac{d\mathbf{y}}{dt} &= \omega \times (\omega \times \mathbf{y}) + \omega \times g_t \mathbf{v}_{\mathbf{x}} \\ \implies \frac{d}{dt} (\omega \times \mathbf{y}) &= \dot{\omega} \times \mathbf{y} + \omega \times (\omega \times \mathbf{y}) + \omega \times g_t \mathbf{v}_{\mathbf{x}} \\ \frac{d}{dt} (g_t \mathbf{v}_{\mathbf{x}}) &= \dot{g}_t \mathbf{v}_{\mathbf{x}} + g_t \mathbf{a}_{\mathbf{x}} = \omega \times g_t \mathbf{v}_{\mathbf{x}} + g_t \mathbf{a}_{\mathbf{x}} \\ \implies \mathbf{a}_{\mathbf{y}} &= \omega \times (\omega \times \mathbf{y}) + 2\omega \times g_t \mathbf{v}_{\mathbf{x}} + \dot{\omega} \times \mathbf{y} + g_t \mathbf{a}_{\mathbf{x}} \end{aligned}$$

Compare this to Hand and Finch (1998) [26], Eq. 7.32:

$$(130) \quad \mathbf{a}|_{\text{space}} = \omega \times (\omega \times \mathbf{r}) + 2\omega \times \mathbf{v}|_{\text{body}} + \dot{\omega} \times \mathbf{r} + \mathbf{a}|_{\text{body}}$$

cf. pp. 134 Kambe (2009) [24]

Now compare this, namely Eq. 129, with wikipedia ("**Rotating reference frame**"), where the accelerations in the 2 reference frames are related by:

$$(131) \quad \mathbf{a}_r = \mathbf{a}_i - 2\boldsymbol{\Omega} \times \mathbf{v}_r - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) - \frac{d\boldsymbol{\Omega}}{dt} \times \mathbf{r}$$

where  $\mathbf{a}_r := \left( \frac{d^2 \mathbf{r}}{dt^2} \right)_r$  is the apparent acceleration in the rotating reference frame.

**Solution Ch. 7, Angular Velocity, Problem 1, (Locomotive) Hand and Finch (1998) [26].**

(a)

(b)

(c)

$$\hat{\omega} = \dot{O} O^T$$

$$O = O_2 O_1 \text{ and } O^T = O_1^T O_2^T$$

$$\dot{O} = \dot{O}_2 O_1 + O_2 \dot{O}_1$$

$$\dot{O} O^T = \dot{O}_2 O_2^T + O_2 \dot{O}_1 O_1^T O_2^T = \hat{\omega}_2 + O_2 \hat{\omega}_1 O_2^T = \hat{\omega}$$

Corresponding to right-invariant  $\dot{g}_t = \omega g_t$  of the tangent vector, the angular momentum is written in the right-invariant way as

$$(132) \quad M_t = J \dot{g}_t = \bar{M} g_t, \quad \bar{M} = J \omega$$

where  $J$  is an inertia operator and time dependent in fixed space. cf. Eqn. (4.20), pp. 134, Kambe (2009) [24].

Make the following notation change:

$$(133) \quad L_t = I \dot{g}_t = \bar{L} g_t, \quad \bar{L} = I \hat{\omega}$$

where  $I$  is an inertia operator.

In mechanics, angular momentum is defined by

$$M_t = \int (\mathbf{y} \times \mathbf{v}_{\mathbf{y}}) \rho d^3 \mathbf{y} = \int (g_t \mathbf{x} \times (\hat{\omega} g_t \mathbf{x})) g_t (\rho d^3 \mathbf{x})$$

and so  $M_t := \bar{M} g_t$ .

Relative to body frame  $F_B$ , the same velocity (relative to fixed space) is represented as  $\frac{d\mathbf{y}_b}{dt} = v_b^i \mathbf{b}_i$ , where  $\mathbf{v}_b = \hat{\Omega}_b \times Y$ .  $\hat{\Omega}_b$  is angular velocity relative to body frame. Now  $\bar{\Omega}_b$  derived by *left translation* of tangent vector.

$$(134) \quad \bar{\Omega}_b = g_t^{-1} \circ \dot{g}_t = L_{g_t^{-1}} \dot{g}_t = g_t^{-1} \bar{\Omega} g_t$$

as a change of notation, I will also write

$$\omega_b = g_t^{-1} \omega g_t$$

An important point is that tangent vector  $\dot{g}_t$  is represented by **left-invariant** form

$$\hat{\omega}_b = g_t^{-1} \dot{g}_t \xrightarrow{L_{g_t}} L_{g_t} \hat{\omega}_b = g_t \hat{\omega}_b = \dot{g}_t \text{ or } \dot{g}_t = g_t \hat{\omega}_b$$

by  $\omega_b$  at 1, whereas it's also represented by right-invariant form  $\dot{g}_t = \hat{\omega} g_t$  with  $\hat{\omega}$  at 1 in Eqn. (4.17) of Kambe (2009) [24], where, recall  $\dot{g}_t = \frac{dg}{dt} = \frac{dA}{dt}$ .

Indeed, note that for

$$\mathbf{b}_i(t) = \mathbf{e}_j B^j{}_{\phantom{j}i}(t) \text{ then}$$

$$\mathbf{b}_i(t + \delta t) = \mathbf{e}_j B^j{}_{\phantom{j}i}(t + \delta t) = \mathbf{e}_j B^j{}_{\phantom{j}k}(t) B^k{}_{\phantom{k}i}(\delta t)$$

where the last step is because since  $\mathbf{e}_j$  on the left forces an infinitesimal change to be from the right.

Then

$$B(t + \delta t) - B(t) = B(t) B(\delta t) - B(t) = B(t) (B(\delta t) - 1) \text{ so } \dot{B}(t) = B(t) Y_1$$

where  $Y_1 \in \mathfrak{so}(N)$  and  $Y_1 := \lim_{\delta t \rightarrow 0} \frac{B(\delta t) - 1}{\delta t}$ .



Also, recall that

$$\mathbf{Y} = Y^i \mathbf{b}_i = \mathbf{e}_j B^j{}_i Y^i = X^j \mathbf{e}_j \text{ implies that } B : F' \rightarrow F$$

where  $F'$  is a rotating frame,  $F$  is a fixed frame, i.e.  $B$  is a coordinate frame transformation.

Due to right-translation property of  $B(t)$ , then  $\frac{dB}{dt} = B(t)\widehat{\Omega}_b$ . Thus, we obtain a left-invariant vector field.

From Eq. 117,  $A = B$  but

(135)

$$\begin{aligned} \frac{dA}{dt} &= \widehat{\Omega}A \\ \frac{dA}{dt} &= \frac{dB}{dt} = B(t)\widehat{\Omega}_b = \widehat{\Omega}A(t) \text{ or } \widehat{\Omega}_b = A^{-1}\widehat{\Omega}A \end{aligned}$$

21. RIGID BODY DYNAMICS, RIGID BODY MOTION

**21.1. Moment of Inertia. Solution Hand and Finch (1998) [26], Problem 3, pp. 327.** Hand and Finch (1998) [26], Problem 3: (Moment of inertia tensor) from "Moment of Inertia Tensor", pp. 327. Which of the symmetric 3x3 matrices below could represent a physical moment of inertia tensor?

(136)

$$I_1 = \begin{pmatrix} 1 & 2 & 1 \\ \dots & 0 & 2 \\ \dots & \dots & 1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1.94791 & .0347273 & -.394509 \\ \dots & 2.42924 & -.823746 \\ \dots & \dots & 1.62285 \end{pmatrix}$$

See Hand and Finch (1998) [26], (8.131)

Recall that

$$\begin{aligned} I_{ij} &= \sum m(r_i^2 \delta_{ij} - x_i x_j) \\ I_{ii} &= \sum m(r_j^2 + r_i^2) \end{aligned}$$

$I_1$  is not physical since

$$\begin{aligned} \sum m(x^2 + z^2) &= 0 \\ \implies \sum mx^2 &= -\sum mz^2 \end{aligned}$$

but both sums on the left and right-hand sides are positive.

21.1.1. Moments of Inertia of Some Simple Objects.

(a) **Uniform thin ring of mass  $M$  and radius  $R$ , around the axis of symmetry of the ring**

See Kleppner and Kolenkow (2014) [25], pp. 247, "Angular Momentum and Fixed Axis Rotation, Example 7.3 Moments of Inertia of Some Simple Objects; Hand and Finch (1998) [26] for Problem 4: ( $I$  for a circular hoop), pp. 327, Ch. 8 "The Dynamics of Rigid Bodies".

Get the mass density  $\lambda$  of a thin ring or i.e. circular hoop.

$$M = \int \rho(\mathbf{x}) d\mathbf{x} \equiv \int \rho(\mathbf{s}) ds = \int_0^{2\pi} \rho(\theta) R d\theta = R\lambda 2\pi \implies \boxed{\lambda = \frac{M}{2\pi R} = \rho(s)}$$

where  $\lambda$  = mass density of a thin ring of total mass  $M$ , radius  $R$  and where we can say  $\lambda$  is constant because it's a *uniform* thin ring.

(137)

$$I = \int r^2 dm = \int_0^{2\pi} R^2 \lambda R d\theta = MR^2$$

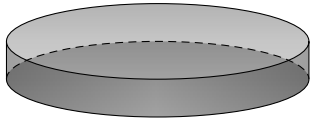
Now from Hand and Finch (1998) [26], Problem 4, pp. 327, for the principal axes and moments, recall that

$$\begin{aligned} I_3 &= \int (x^2 + y^2) dm \\ I_2 &= \int (x^2 + z^2) dm = \int x^2 dm \\ I_1 &= \int (y^2 + z^2) dm = \int y^2 dm \end{aligned}$$

and that

$$I_3 = \int R^2 dm = M^2 \implies I_3 = I_1 + I_2 \text{ so that by symmetry } I_1 = I_2 = \frac{MR^2}{2}$$

(b) **Uniform thin disk of mass  $M$  and radius  $R$ , around the axis of symmetry of the disk**



Using the fact that the disk is uniform, so mass density  $\sigma$  is uniform

$$M = \int \rho(\mathbf{x}) d\mathbf{x} = \int \rho(\mathbf{s}) ds = \sigma \int r d\theta dr = \sigma \frac{1}{2} R^2 2\pi = \sigma \pi R^2$$

$$\sigma = \frac{M}{\pi R^2} = \text{mass per unit area}$$

(138)

$$I = \int r^2 dm = \int r^2 \sigma r d\theta dr = \sigma \frac{1}{4} r^4 2\pi = \boxed{\frac{MR^2}{2}}$$

**21.2. Euler's equations of motion, Euler's rotation equations.** In the body frame, inertia tensor  $J \equiv I$  is time-independent.

$$\frac{d}{dt} \widehat{\Omega} - \text{ad}^*_{\widehat{\Omega}} \widehat{\Omega} = 0$$

Use (Eq. (4.36) Kambe (2009) [24]

22. WAVES

**22.1. Forced Oscillation and Resonance, and Damped Oscillations.** See Georgi (1992) [34].

22.1.1. *Damped Oscillators.* See Sec. 2.1 "Damped Oscillations", Georgi (1992) [34].

For small velocities, frictional force takes the form of

(139)

$$-m\Gamma v$$

See (2.1.1) of Georgi (1992) [34].  $1/\Gamma$  has the dimensions of time.

Now

$$\sum F = m\ddot{x} = -m\Gamma v - Kx$$

or

(140)

$$\ddot{x} + \Gamma \dot{x} + \frac{K}{m} x = \ddot{x} + \Gamma \dot{x} + \omega_0^2 x = 0$$

where  $\omega_0 = \sqrt{\frac{K}{m}}$ . See (2.1.2) of Georgi (1992) [34].



Since  $\ddot{x} + \Gamma\dot{x} + \omega_0^2 x = 0$  is 1. time translation invariant, i.e. if  $x(t+a)$  is a solution, then  $x(t)$  is a solution, and 2. linearity, then find irreducible solutions of form  $z(t) = e^{\alpha t}$ .

$$\implies \alpha^2 + \Gamma\alpha + \omega_0^2 = 0 \text{ or } \alpha = \frac{-\Gamma \pm \sqrt{\Gamma^2 - 4\omega_0^2}}{2}$$

22.1.2. *Overdamped oscillations.* See Sec. 2.1.1 "Overdamped Oscillations" of Georgi (1992) [34].

If  $\frac{\Gamma}{2} > \omega_0$ , let

$$(141) \quad \Gamma_{\pm} = \frac{\Gamma}{2} \pm \sqrt{\frac{\Gamma^2}{4} - \omega_0^2}$$

See (2.1.8) of Georgi (1992) [34]

$$(142) \quad x(t) = A_+ e^{-\Gamma_+ t} + A_- e^{-\Gamma_- t}$$

See (2.1.7) of Georgi (1992) [34].

22.1.3. *Underdamped Oscillations.* See Sec. 2.1.2 "Underdamped Oscillations" of Georgi (1992) [34].

If  $\frac{\Gamma}{2} < \omega_0$ , so

$$\alpha = \frac{-\Gamma}{2} \pm i\sqrt{\omega_0^2 - \frac{\Gamma^2}{2}}$$

so

$$(143) \quad \omega^2 = \omega_0^2 - \frac{\Gamma^2}{4}$$

$$(144) \quad x(t) = Ae^{\frac{-\Gamma}{2}t} \cos(\omega t - \theta)$$

See (2.1.10) of Georgi (1992) [34].

$$x(t) = Ae^{\frac{-\Gamma}{2}t} c(\omega t - \theta) = Ae^{\frac{-\Gamma}{2}t} (c(\omega t)c(\theta) + s(\omega t)s(\theta))$$

$$(145) \quad x(t) = e^{-\frac{\Gamma}{2}t} (c \cos(\omega t) + d \sin(\omega t))$$

See (2.1.11) of Georgi (1992) [34], where  $c = A \cos \theta$ ,  $d = A \sin \theta$ .

22.1.4. *Critically Damped Oscillations.* See Sec. 2.1.3 "Critically Damped Oscillations" of Georgi (1992) [34].

If  $\frac{\Gamma}{2} = \omega_0$ , we only get  $e^{\frac{-\Gamma}{2}t}$ . We know that there will be 2 solutions to the 2nd-order differential equation.

One way to find the other solution is to approach this situation from underdamped case as a limit, i.e.

$$\omega \rightarrow 0 \text{ from } \omega > 0$$

$$\lim_{\omega \rightarrow 0} e^{\frac{-\Gamma}{2}t} c \cos(\omega t) = e^{\frac{-\Gamma}{2}t}$$

$$\lim_{\omega \rightarrow 0} e^{\frac{-\Gamma}{2}t} d \sin \omega t = 0$$

Divide the second term by  $\omega$ , because  $\omega$  doesn't depend on  $\omega$ ,  $\frac{1}{\omega} e^{\frac{-\Gamma}{2}t} d \sin \omega t$  is still a solution.

$$\lim_{\omega \rightarrow 0} e^{\frac{-\Gamma}{2}t} \frac{\sin \omega t}{\omega} = t e^{\frac{-\Gamma}{2}t} \text{ (by L'Hopital's rule)}$$

So

$$x(t) = ce^{-\Gamma t/2} + dt e^{-\Gamma t/2}$$

22.1.5. *Forced oscillations.* See Sec. 2.2 "Forced Oscillations" of Georgi (1992) [34]. Equation of motion is

$$(146) \quad \ddot{x}(t) + \Gamma\dot{x} + \omega_0^2 x = F(t)/m$$

where

$$(147) \quad F(t) = F_0 \cos(\omega_d t)$$

See (2.1.5), (2.1.4) of Georgi (1992) [34], respectively.

$\frac{\omega_d}{2\pi}$  = driving frequency, driving frequency can be tuned completely independently of the other parameters of the system.

If we let

$$\ddot{z}(t) + \Gamma\dot{z}(t) + \omega_0^2 z(t) = \mathcal{F}(t)/m$$

where  $\mathcal{F}(t) = F_0 e^{-i\omega_0 t}$ .

For in particular numerical computation applications, we can transform Eq. 146 into a system of first order equations:

Let

$$x_1(t) = x(t)$$

$$x_2(t) = \dot{x}(t)$$

so that

$$\dot{x}_1 = x_2$$

$$\dot{x}_2 = -\Gamma x_2 - \omega_0^2 x_1 + f(t)$$

Then if we look for steady state solution of the form

$$z(t) = Ae^{-i\omega_d t}$$

One can think like the steady state solution is such that the system is driven with driving frequency  $\omega_d$ .

$$(148) \quad (-\omega_d^2 + (-i\omega_d)\Gamma + \omega_0^2)A = \frac{F_0}{m}$$

See (2.20) of Georgi (1992) [34]

or

$$(149) \quad A = \frac{F_0/m}{\omega_0^2 - i\omega_d\Gamma - \omega_d^2}$$

See (2.21) of Georgi (1992) [34]

The amplitude  $A$  of the displacement is proportional to amplitude of driving force; this is as expected from linearity.

$$A = \frac{\frac{F_0}{m}(\omega_0^2 + i\omega_d\Gamma - \omega_d^2)}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d\Gamma)^2} \text{ with}$$

$$\Re A = \frac{(\omega_0^2 - \omega_d^2)F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d\Gamma)^2} \quad \Im A \equiv B = \frac{\Gamma\omega_d F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma\omega_d)^2}$$

$$\mathcal{A}e^{-i\omega_d t} = \mathcal{A}(c(\omega_d t) - is(\omega_d t)) \xrightarrow{\Re} \Re A c(\omega_d t) + \Im A s(\omega_d t)$$

where I use  $\mathcal{A} \equiv A$  notation interchangeably, relying on context on whether it's real or complex, and  $\cos \omega_d t \equiv c(\omega_d t)$ ,  $\sin \omega_d t \equiv s(\omega_d t)$  as shorthand notation for sin and cos.

For

$$x(t) = \Re z(t) = (\Re A)c(\omega_d t) + Bs(\omega_d t)$$

Term proportional to  $B$  is 90° out of phase.

Real part of  $\mathcal{A}$ ,  $A = \Re A$  is the *elastic amplitude*.

Imaginary part of  $\mathcal{A}$ ,  $B = \Im A$  is called the *absorptive amplitude*.

In the case where

$$F(t) = F_0 \sin(\omega_d t)$$

we'll see that the fundamental physical characteristics don't change, but the initial conditions for the solution do change, which is important for actual modeling (but gets looked over a bit in physics that concern about high-level concepts).

If  $F(t) = F_0 \sin(\omega_d t)$ , while the following doesn't change:

$$z(t) = \mathcal{A}e^{-i\omega_d t}$$

$$\dot{z}(t) = (-i\omega_d)z$$

$$\ddot{z}(t) = (-\omega_d^2)z$$

but now consider

$$\mathcal{F}(t) = +iF_0 e^{-i\omega_d t} = +iF_0(c(\omega_d t) - is(\omega_d t))$$

$$\Re(\mathcal{F}(t)) = F_0 s(\omega_d t)$$

so then

$$(-\omega_d^2 + (-i\omega_d)\Gamma + \omega_0^2)\mathcal{A} = \frac{iF_0}{m} \text{ or}$$

$$\mathcal{A} = \frac{iF_0/m}{\omega_0^2 - \omega_d^2 - i\omega_d\Gamma} = \frac{\frac{iF_0}{m}(\omega_0^2 - \omega_d^2 + i\omega_d\Gamma)}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d\Gamma)^2}$$

So then in this case,

$$(150) \quad \Re A = \frac{-\Gamma\omega_d F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma\omega_d)^2} \equiv A$$

$$\Im A = \frac{(\omega_0^2 - \omega_d^2)F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d\Gamma)^2} \equiv B$$

and so the general form of the steady-state solution doesn't change, but the values for the amplitudes get switched:

$$x(t) = \Re z(t) = Ac(\omega_d t) + Bs(\omega_d t)$$

$$\dot{x}(t) = \omega_d [-As(\omega_d t) + Bc(\omega_d t)]$$

22.1.6. *Resonances.* See Sec. 2.3 "Resonances" of Georgi (1992) [34].

$$\omega_d = \omega_0$$

so  $\Re A = 0$ ,  $B = \frac{F_0/m}{\Gamma\omega_d}$ .

22.1.7. *Work, Work done by forced oscillation on the system.* Work is a nonlinear function of force, so we must use real force and real displacement.

$$P(t) = F(t)\frac{\partial}{\partial t}X(t) = F_0 c(\omega_d t)(\omega_d)(-\Re As(\omega_d t) + Bc(\omega_d t))$$

The first term is proportional to  $s(2\omega_d t)$  and averages to 0 over a half period.

$$\int_{t_0}^{t_0+\pi/\omega_d} s(2\omega_d t) dt = \left. \frac{-c(2\omega_d t)}{2\omega_d} \right|_{t_0}^{t_0+\pi/\omega_d} = 0$$

This is why  $\Re A$  is called elastic amplitude. If  $A$  dominates, then energy fed into the system at 1 time is returned at a later time, as in elastic collision in mechanics.

$$P_{\text{average of 2nd. term}} = P_{\text{average}}$$

Consider

$$F_0\omega_d B \int_{t_0}^{t_0+\pi/\omega_d} c^2(\omega_d t) dt = F_0\omega_d B \int_{t_0}^{t_0+\pi/\omega_d} \frac{c(2\omega_d t) + 1}{2} dt =$$

$$F_0\omega_d B \left( \frac{s(2\omega_d t)}{4\omega_d} + \frac{t}{2} \right) \Big|_{t_0}^{t_0+\frac{\pi}{\omega_d}} = F_0\omega_d B \frac{\pi}{2\omega_d} = \frac{F_0 B \pi}{2}$$

$$P_{\text{average}} = \frac{\int_{t_0}^{t_0+\pi/\omega_d} P(t) dt}{\Delta t} \text{ where } \Delta t = \pi/\omega_d$$

$$\implies P_{\text{average}} = \frac{F_0 B \pi / 2}{\pi / \omega_d} = \boxed{\frac{F_0 \omega_d B}{2}}$$

Recall that

$$B = \frac{\Gamma\omega_d F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma\omega_d)^2}$$

$$P_{\text{average}} = \frac{F_0\omega_d B}{2} = \frac{(F_0\omega_d)^2 \Gamma / 2m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma\omega_d)^2} \xrightarrow{\omega_d \rightarrow \omega_0} \frac{F_0^2}{2m\Gamma}$$

The maximum average power is inversely proportional to  $\Gamma$ .

This analysis doesn't change when  $F(t) = F_0 \sin(\omega_d t)$ .

$$P(t) = F(t)\dot{x}(t) = F_0 s(\omega_d t)\omega_d [-As(\omega_d t) + Bc(\omega_d t)]$$

$$\int_{t_0}^{t_0+\frac{\pi}{\omega_d}} dt c(\omega_d t) s(\omega_d t) = \left. \frac{1}{2\omega_d} s^2(\omega_d t) \right|_{t_0}^{t_0+\frac{\pi}{\omega_d}} = 0$$

$$\int_{t_0}^{t_0+\frac{\pi}{\omega_d}} dt s^2(\omega_d t) = \int_{t_0}^{t_0+\frac{\pi}{\omega_d}} dt \frac{1 - c(2\omega_d t)}{2} = \frac{\pi}{2\omega_d} - \left. \frac{s(2\omega_d t)}{4\omega_d} \right|_{t_0}^{t_0+\frac{\pi}{\omega_d}} = \frac{\pi}{2\omega_d}$$

$$P_{\text{average}} = \frac{\int_{t_0}^{t_0+\frac{\pi}{\omega_d}} dt P(t)}{\pi/\omega_d} = \frac{-F_0 A \pi}{2/\pi/\omega_d} = \frac{-F_0 A \omega_d}{2}$$

So  $B$  is now the elastic amplitude and  $A$  is the absorptive amplitude, but the same values are obtained, since  $A$  and  $B$  are effectively switched for  $F(t) = F_0 \sin \omega_d t$ .

To find the width (for fixed height) of the resonance curve (curve of  $P_{\text{average}}$ , depending on  $\omega_d$ ), find the value  $\omega_d = \omega_{1/2}$  s.t.

$$P_{\text{average}} = \frac{P_{\text{average, max}}}{2} = \frac{F_0^2}{4m\Gamma}$$

Then

$$\frac{P_{\text{average, max}}}{2} = \frac{F_0^2}{4m\Gamma} = \frac{(F_0^2 \Gamma / 2m) \omega_{1/2}^2}{(\omega_{1/2}^2 - \omega_0^2)^2 + (\Gamma \omega_{1/2})^2} \text{ or } \omega_{1/2}^4 - 2\omega_{1/2}^2 \omega_0^2 + \omega_0^2 + \Gamma^2 \omega_{1/2}^2 = 2\Gamma^2 \omega_{1/2}^2 \text{ or}$$

$$\omega_{1/2}^4 - (2\omega_0^2 + \Gamma^2) \omega_{1/2}^2 + \omega_0^4 = 0$$

$$\omega_{1/2}^2 = \frac{(2\omega_0^2 + \Gamma^2) \pm \sqrt{(\Gamma^2 + 2\omega_0^2)^2 - 4\omega_0^4}}{2} = \omega_0^2 + \frac{\Gamma^2}{2} \pm \sqrt{\frac{\Gamma^4}{4} + \Gamma^2 \omega_0^2} = \omega_0^2 + \frac{\Gamma^2}{2} \pm \Gamma \sqrt{\frac{\Gamma^2}{4} + \omega_0^2}$$

Now (EY: it seems like you would have needed to have known the answer beforehand to have tried this factorization)

$$\omega_{1/2}^2 = \left( \sqrt{\omega_0^2 + \frac{\Gamma^2}{4}} \right)^2 \pm 2\frac{\Gamma}{2}\sqrt{\frac{\Gamma^2}{4} + \omega_0^2} + \left( \frac{\Gamma}{2} \right)^2 = \left( \sqrt{\omega_0^2 + \frac{\Gamma^2}{4}} \pm \frac{\Gamma}{2} \right)^2$$

Then

$$(151) \quad \boxed{\omega_{1/2} = \sqrt{\omega_0^2 + \frac{\Gamma^2}{4}} \pm \frac{\Gamma}{2}}$$

The "free oscillation" of the system would be the case of underdamped oscillations; recall

$$x(t) = A \exp\left(\frac{-\Gamma t}{2}\right) \cos(\omega t - \theta)$$

Notice the exponential decay behavior of the amplitude or i.e. magnitude of  $x(t)$  due to the  $\exp\left(\frac{-\Gamma t}{2}\right)$  factor. Let's just say the "lifetime" of free oscillation is the time for the amplitude to drop to half its value, i.e.

$$\frac{1}{2} = \exp\left(\frac{-\Gamma t_{1/2}}{2}\right) \text{ or } -\ln 2 = \frac{-\Gamma t_{1/2}}{2}, \quad \text{or } t_{1/2} = \frac{2 \ln 2}{\Gamma}$$

For the time dependence of the free oscillation, i.e.  $x(t) = A \exp\left(\frac{-\Gamma t}{2}\right) \cos(\omega t - \theta)$ , then we say that the lifetime of the state in free oscillation is of order  $1/\Gamma$ , i.e.

$$t_{1/2} = \frac{2 \ln 2}{\Gamma}$$

The width of the resonance peak in forced oscillation is

$$\Delta\omega_{1/2} = \Gamma$$

so width of the resonance peak in forced oscillation is inversely proportional to the lifetime of the corresponding normal mode of free oscillation.

**22.2. Apply Fourier Transform to Forced Oscillations.** So for the equation of motion, Eq. [146](#),

$$\ddot{x} + \Gamma \dot{x} + \omega_0^2 x = F(t)$$

Assume we drive the system at frequency  $\omega_d$  and that system's response's frequency is also  $\omega_d$ . Then

$$(i\omega_d)^2 \hat{x} + (i\omega_d)\Gamma \hat{x} + \omega_0^2 \hat{x} = \hat{F}(\omega_d) \text{ or } \frac{\hat{x}(\omega_d)}{\hat{F}(\omega_d)} = \frac{1}{-(\omega_d^2 - i\omega_d\Gamma - \omega_0^2)}$$

So in general,  $\frac{\hat{x}(\omega_d)}{\hat{F}(\omega_d)}$  has a real part and imaginary part.

**22.3. Apply Laplace Transform to Forced Oscillations.** Recall Eq. [146](#),

$$\ddot{x} + \Gamma \dot{x} + kx = mF(t)$$

so that if  $\frac{k}{m} = \omega_0^2$ , then

$$\xrightarrow{\mathcal{L}} s^2 \hat{x}(s) - sx(0) - x'(0) + \Gamma(s\hat{x}(s) - x(0)) + \omega_0^2 \hat{x}(s) = \hat{F}(s) \text{ or}$$

$$(s^2 + \Gamma s + \omega_0^2) \hat{x}(s) - (\Gamma + s)x(0) - x'(0) = \hat{F}(s)$$

If  $x'(0) = 0$ , and we start system at  $x(0) = 0$ , so that we rely upon  $F(t)$  to drive system, then we get a ratio of output or "response" to input, the driving, forced oscillation, per mass,

$$\frac{\hat{x}(s)}{\hat{F}(s)} = \frac{1}{s^2 + \Gamma s + \omega_0^2} = \frac{1}{\left(s + \frac{\Gamma + \sqrt{\Gamma^2 - 4\omega_0^2}}{2}\right) \left(s + \frac{\Gamma - \sqrt{\Gamma^2 - 4\omega_0^2}}{2}\right)}$$

Now look at where the denominator "blows up", the singularities. We had the case that if  $\Gamma^2 - 4\omega^2 \geq 0$  or  $\frac{\Gamma}{2} \geq \omega_0$ , we had overdamped or critically damped exponential decay. Then, the value of  $s$  where a singularity occurs would be completely real. For the underdamped case with free oscillation, the value of  $s$  where a singularity occurs would have an imaginary part.

If initial conditions can't be ignored, then if  $x'(0) = 0$ ,

$$(s^2 + \Gamma s + \omega_0^2) \hat{x}(s) - x'(0) = \hat{F}(s) \text{ or } \frac{\hat{x}(s)}{\hat{F}(s)} = \frac{1 + x'(0)/\hat{F}(s)}{s^2 + \Gamma s + \omega_0^2}$$

But if  $x'(0) = 0$ ,

$$\frac{\hat{x}(s)}{\hat{F}(s)} = \frac{1 + (\Gamma + s)x(0)/\hat{F}(s)}{s^2 + \Gamma s + \omega_0^2}$$

**22.3.1. General solution to forced oscillation.** For subscript  $s$  meaning steady-state driven solution, see Eq. [149](#)

$$x_s(t) = Ac(\omega_d t) + Bs(\omega_d t)$$

$$A = \frac{(\omega_0^2 - \omega_d^2)F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d\Gamma)^2}$$

$$B = \frac{\Gamma\omega_d F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma\omega_d)^2}$$

$$x_s(0) = A$$

$$\dot{x}_s(0) = \omega_d B$$

Let  $x_{\text{free}}$  denote the response, or solution, or the resulting motion for the "free oscillation", the underdamped, exponential decaying solution, for  $\omega_0 > \frac{\Gamma}{2}$ . See Eq. [145](#),

$$x_{\text{free}}(t) = \exp\left(\frac{-\Gamma t}{2}\right) (c \cos(\omega t) + d \sin(\omega t)), \quad \omega^2 = \omega_0^2 - \frac{\Gamma^2}{4}$$

$$x_{\text{free}}(0) = c$$

$$\dot{x}_{\text{free}}(t) = \frac{-\Gamma}{2} x_{\text{free}}(t) + e^{-\frac{\Gamma}{2}t} \omega (-cs(\omega t) + dc(\omega t))$$

$$\dot{x}_{\text{free}}(0) = \frac{-\Gamma}{2} c + \omega d$$

The most general solution  $x(t)$  is of the form

$$x(t) = x_{\text{free}}(t) + x_s(t)$$

with initial conditions

$$x(0) = c + A \text{ so } c = x(0) - A$$

$$\dot{x}(0) = \frac{-\Gamma}{2} c + \omega d + \omega_d B \text{ so } d = \frac{1}{\omega} \left[ \dot{x}(0) + \frac{\Gamma}{2} (x(0) - A) - \omega_d B \right]$$

And so the full, general solution is

$$x(t) = e^{\frac{-\Gamma t}{2}}(c \cos(\omega t) + d \sin(\omega t)) + A \cos(\omega_d t) + B \sin(\omega_d t)$$

where

$$A = \frac{(\omega_0^2 - \omega_d^2)F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d \Gamma)^2}$$
$$B = \frac{\Gamma \omega_d F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma \omega_d)^2} \qquad \omega^2 = \omega_0^2 - \frac{\Gamma^2}{4},$$
$$c = x(0) - A,$$
$$d = \frac{1}{\omega} \left[ \dot{x}(0) + \frac{\Gamma}{2}(x(0) - A) - \omega_d B \right]$$

(152)

For the case of  $F(t) = F_0 \sin(\omega_d t)$ , i.e. the force is proportional to sin, then the exact same analysis leads to the same result as above, but for  $A, B$  values switched to be as Eq. [150](#).

$$\Re A = \frac{-\Gamma \omega_d F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\Gamma \omega_d)^2} \equiv A$$
$$\Im A = \frac{(\omega_0^2 - \omega_d^2)F_0/m}{(\omega_0^2 - \omega_d^2)^2 + (\omega_d \Gamma)^2} \equiv B$$

23. FLUID MECHANICS, FLUID FLOW

23.0.1. *Courses to consider.* [MIT 16.01](#), [16.02](#)

23.0.2. *Videos to watch.* [Engineering MAE 130A. Intro to Fluid Mechanics. Lecture 01.](#) UCI. Follow the rest of the videos.

23.1. **Buoyancy.** From Problem 1.3 of the Japan Physics Olympiad (2014) [\[48\]](#), Ch. 1, General Physics, Elementary Problems.  
**Problem 1.3.** The part of the iceberg above the sea.

$$\rho_{sw} = 1024 \text{ kg/m}^3$$
$$\rho_{ice} = 917 \text{ kg/m}^3$$

with  $sw \equiv$  sea water.  
Find the ratio of the volume of the part of the iceberg above the sea to whole volume of the iceberg.

**Solution 1.3.**  
**The buoyancy force exerted on the iceberg is equal to the weight of the seawater displaced by the iceberg.**

$$\rho_{sw} g V_{is} = \rho_{ice} g V_{ice} \implies \frac{\rho_{sw}}{\rho_{ice}} = \frac{V_{ice}}{V_{is}}$$

with iceberg in sea  $\equiv$  is.  
iceberg above sea  $\equiv$  as

$$V_{as}/V_{ice} = \frac{V_{ice} - V_{is}}{V_{ice}} = 1 - \frac{V_{is}}{V_{ice}} = 1 - \frac{\rho_{ice}}{\rho_{sw}} = 1 - \frac{917 \text{ kg/m}^3}{1024 \text{ kg/m}^3} = \frac{1024 - 917}{1024} = \frac{107}{1024}$$

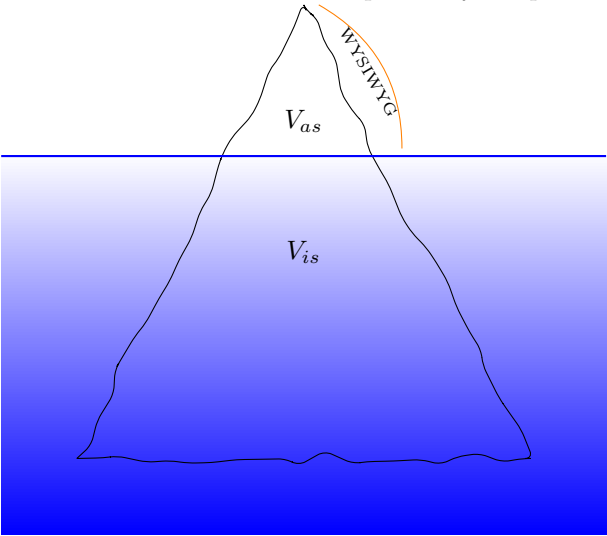
**The buoyancy on a body equals the resultant force due to the pressure exerted by the surrounding fluid**, i.e. **The buoyancy force exerted on a body (iceberg) is equal to the resultant force due to pressure exerted by the surrounding fluid (seawater).**  
Pressure on body of volume  $V$  due to its surrounding fluid (whose density is  $\rho$ ) acts perpendicularly to the boundary surface between the body and the fluid.  
Since fluid pressure at a deep location is greater than that at shallow location, resultant force due to pressure on boundary surface points upward. This resultant force is the buoyancy,  $F$ , acting on body.

Consider region of fluid with same volume  $V$  as body.

buoyancy  $F \equiv$  force exerted vertically on body by its surrounding fluid

$$\underbrace{\hspace{10em}}_{\text{balance gravitational force acting on } V} \rho g V$$

For a body floating in a fluid, the magnitude of the buoyancy force acting on the body is equal to the magnitude of the gravitational force on the fluid displaced by the part of the body submerged in the fluid.



23.2. **Fluids as a State of Matter.** cf. Pert (2013) [\[28\]](#), Ch. 1 Introduction, 1.1 Fluids as a State of Matter.  
fluid - isotropic, (no directional preference; identical values of a property in all directions),  
locally homogeneous, macroscopic material,  
whose particles are free to move within constraints established by the dynamical laws of continuum physics.  
*requirement that fluid be a continuum* - if volume of fluid successively subdivided into smaller elements, each element will remain structurally similar to its parent, and this subdivision can be carried out down to infinitesimal volumes.

**fluid particle** - fictitious particle fixed within fluid continuum and moving with flow velocity;  
represents average over large number of microscopic particles.

**fluid point** - fixed in fluid moving with flow velocity;  
- fluid particle always situated at same fluid point.

**infinitesimal volume** - within continuum of fluid, large compared with microscopic scales, but small compared with macroscopic ones.

fluid is made up of discrete microscopic particles, namely molecules, distributed randomly, with distribution of velocities characteristic of fluid in thermal equilibrium, typically Maxwell-Boltzmann distribution in a gas.

At typical experimental densities,  
intermolecular separation extremely small and very much less than lab scale, i.e.  $l_s \ll L_0$   
possible to average over small volumes which contain very large number of particles, yet very small on lab scale,  $N - V$  in  $V_0$ ,  $V_0^{1/3} \ll L_0$   
 $\implies$  allow us to recover continuum approximation

obtain terms which characterize fluid as bulk material

Typical average quantities:

Density,  $\rho$ : number of mass particles per unit volume.

temperature,  $\tau$ : average energy of random motion per particle in thermal equilibrium

pressure  $p$  average momentum flow associated with random motion per unit area.

flow velocity  $u$  - mean velocity of molecules averaging out random motion.

role of collisions amongst particles important in defining irreversibility through loss of correlation between particles.

*mean free path*  $l_0$  - particles collide on average after distance equal to  $l_0$  and time after collision interval

since fluid mechanics assumes fluid particles in thermal equilibrium and randomly distributed, this condition requires that spatial and temporal averages be taken to include large number of collisions; i.e. lab scale length  $L_0$  large compared with  $l_0$ , and time to collision interval

effects of collisions on fluid transport (momentum and energy) are averaged over thermal distribution to yield bulk properties of material, namely viscosity and thermal conduction, respectively.

theory: flow of basic fluid may be calculated using Newtonian mechanics, classical thermodynamics, viscosity and thermal conductivity values.

Conditions under which this theory may be applied:

- lab-scale lengths  $L_0$  must be large compared with intermolecule separation and mean free path  $l_0$
- characteristic lab times  $T_0$  must be large compared with collision interval
- fluid locally in thermal equilibrium

cf. 1.2 The Fundamental Equations for Flows of a Dissipationless Fluid. Pert (2013) [28]

**inviscid limit** - viscosity and thermal conduction are neglected

**adiabatic flow** - specific entropy of fluid particle constant in time

**isentropic flow** - specific entropy of each fluid particle has same initial value

many flows are both isentropic and adiabatic, e.g. ideal steady fluid flow, whose specific entropy an entry is everywhere constant.

*Eulerian frame* - lab frame, coordinates fixed in space and time, derivatives are usual partial derivatives

$$\left. \frac{\partial}{\partial t} \right|_{\mathbf{r}}, \nabla|_t$$

Consider coordinate system  $(t, x^1, x^2, x^3)$  and so each vector field  $\mathbf{e}_j$  forming basis of tangent space at each pt. expressed as linear combination of coordinate vector fields

$$\left. \frac{\partial}{\partial t} \right|_{\mathbf{r}}, \left. \frac{\partial}{\partial x^1} \right|_t, \left. \frac{\partial}{\partial x^2} \right|_t, \left. \frac{\partial}{\partial x^3} \right|_t,$$

in Lagrangian frame, spatial variation seen by particle due to its motion is absorbed into time derivative:

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla$$

Rather, consider total time derivative of integral of  $n$ -form:

$$\frac{d}{dt} \int_{V_0} \omega = \int_{V_0} \mathcal{L}_{\frac{\partial}{\partial t} + \mathbf{v}} \omega$$

Volume form  $\Omega^n(M) = \bigwedge^n(T^*M)$

$$\text{vol}^n \in \Gamma(\Omega^n(M)) = \Gamma(\bigwedge^n T^*M)$$

i.e.  $\text{vol}^n$  is  $n$ -form, a section of the line bundle,  $\Omega^n(M) = \bigwedge^n T^*M$

Now

$$\text{vol}^n = \frac{\sqrt{g}}{n!} \epsilon_{i_1, \dots, i_n} dx^{i_1} \wedge \dots \wedge dx^{i_n}$$

Consider

$$\text{vol}^{n+1} = \frac{\sqrt{g}}{(n+1)!} \epsilon_{i_0 i_1, \dots, i_n} dx^{i_0} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_n}$$

23.3. **Lagrangian frame.** cf. Pert (2013) [28] 1.3 Lagrangian Frame.

**Lagrangian frame** - frame of the moving particle;

Lagrangian reference frame considers fluid from point of view of an observer on a fluid particle.

fluid particle may be conveniently identified by coordinate set, which is fixed on the particle,  $\mathbf{\Lambda} = (\lambda, \mu, \nu)$  i.e. triad of numbers.

e.g.  $\mathbf{\Lambda}$  maybe initial position of particle  $\mathbf{r}_0 = (x_0, y_0, z_0)$ , i.e. fluid particle labeled by some (time-independent) vector field  $\mathbf{x}_0 \equiv \mathbf{\Lambda}$ .

$\implies$  position, velocity, thermodynamic state of particles are therefore functions of time alone.

$\implies$  leads to simple set of kinematic and dynamic relations governing motion of particle

$$\mathbf{r} = \mathbf{r}(t) = (x(t), y(t), z(t)) = (x^1(t), x^2(t), x^3(t))$$

$$\dot{\mathbf{r}} \equiv \dot{\mathbf{r}}(t) \equiv \frac{d\mathbf{r}}{dt} \equiv \frac{d\mathbf{r}}{dt}(t) = \mathbf{v} \equiv \mathbf{v}(t)$$

or

$$(153) \quad \begin{aligned} \frac{d\mathbf{r}}{dt} &= \mathbf{v} \\ \frac{d\mathbf{v}}{dt} &= \frac{\mathbf{F}}{m} \end{aligned}$$

cf. Eq. (1.2) of Pert (2013) [28], i.e.

flow described by function  $\mathbf{X}(\mathbf{x}_0, t)$  giving position of particle labeled at  $\mathbf{x}_0$  at time  $t$ .

23.4. **Momentum conservation in Lagrangian Frame-Euler's Equation.** Recall that a connection  $\nabla$  on smooth manifold  $(M, \mathcal{O}, \mathcal{A})$  is a map that takes a pair consisting of a vector (field)  $X$  and  $(p, q)$ -tensor field  $T$  and sends them to a  $(p, q)$ -tensor (field)  $\nabla_X T$ , i.e.

$$\nabla : \mathfrak{X}(M) \times (\otimes^p TM \otimes^q T^*M) \rightarrow \otimes^p TM \otimes^q T^*M$$

s.t.

$$\nabla_X f = Xf \quad \forall f \in C^\infty(M)$$

$$\nabla_X(T + S) = \nabla_X T + \nabla_X S$$

$$\nabla_X(T(\omega, Y)) = (\nabla_X T)(\omega, Y) + T(\nabla_X \omega, Y) + T(\omega, \nabla_X Y)$$

$$\nabla_{fX+Z} T = f \nabla_X T + \nabla_Z T, \quad f \in C^\infty(M) \quad \text{''Leibnitz'' rule}$$

cf. Schuller, International Winter School on Gravity and Light held central lectures.

Remember that in the Lagrangian frame, we are an observer following an individual fluid parcel or i.e. "fluid particle" as it moves through space and time. This fluid particle will obey Newton's 2nd. law:  $\nabla_u u = \frac{F}{m} \iff m \cdot \mathbf{a} = F$ .

$$\begin{aligned}\nabla_u u &= \nabla_{u^\mu \frac{\partial}{\partial x^\mu}} (u^\nu \frac{\partial}{\partial x^\nu}) = u^\mu \nabla_{\frac{\partial}{\partial x^\mu}} (u^\nu \frac{\partial}{\partial x^\nu}) = u^\mu \left( \nabla_{\frac{\partial}{\partial x^\mu}} u^\nu \right) \frac{\partial}{\partial x^\nu} + u^\mu u^\nu \nabla_{\frac{\partial}{\partial x^\mu}} \frac{\partial}{\partial x^\nu} = \\ &= u^\mu \frac{\partial u^\nu}{\partial x^\mu} \frac{\partial}{\partial x^\nu} + u^\mu u^\nu \Gamma_{\mu\nu}^\gamma \frac{\partial}{\partial x^\gamma}\end{aligned}$$

Let  $u^0 = 1$ .

$$\nabla_u u = \frac{\partial u^a}{\partial t} \frac{\partial}{\partial x^a} + u^b \frac{\partial u^a}{\partial x^b} \frac{\partial}{\partial x^a} + u^\mu u^\nu \Gamma_{\mu\nu}^a \frac{\partial}{\partial x^a}$$

where  $\Gamma_{ba}^0 = 0$  since  $\nabla dt = 0$  and we choose a stratified atlas  $\mathcal{A}_{\text{stratified}}$ .

Then let

$$\int \rho \text{vol}^n \nabla_u u = F_{\text{tot}} = \int f^i \text{vol}^n \otimes e_i$$

From Sec. 2.4 Equations of Motion Neglecting Viscosity, pp. 38, Sabersky, et. al. (1998) [29], the principal forces are those due to pressures acting on surfaces and to forces acting directly on a mass of the particle:

$$\begin{aligned}\int_{\partial V} p dS^i \otimes e_i &= \int_V \mathbf{d}(p dS^i \otimes e_i) = \int_V \frac{\partial p}{\partial x^j} dx^j \wedge \left( \frac{\sqrt{g}}{(n-1)!} \epsilon_{i_2, \dots, i_n}^i dx^{i_2} \wedge \dots \wedge dx^{i_n} \right) \otimes e_i = \\ &\int_V \frac{\partial p}{\partial x^j} g^{ij} \text{vol}^n \otimes e_i = \int_V (\text{grad} p)^i \text{vol}^n \otimes e_i \\ \implies \int \rho \text{vol}^n \nabla_u u &= - \int_V (\text{grad} p)^i \text{vol}^n \otimes e_i + \int_V f^i \text{vol}^n \otimes e_i \\ \implies \rho (\nabla_u u)^i &= -(\text{grad} p)^i + f^i \\ (\nabla_u u)^i &= -\frac{1}{\rho} (\text{grad} p)^i + f^i\end{aligned}\tag{154}$$

describes the motion of a perfect fluid.

Assuming a non-rotating, inertial frame (so that we can set the Christoffel symbols to 0),

$$\frac{\partial}{\partial t} u^i + u^b \frac{\partial u^i}{\partial x^b} = -\frac{1}{\rho} (\text{grad} p)^i + f^i\tag{155}$$

which describes the motion of the perfect fluid in an inertial ref. frame. Compare this to Eqns. (2.10a- 2.10c) of Sabersky (1998) [29]. Sabersky says that no assumptions were made about density so these results apply to compressible fluid as well as incompressible ones. Also Eqns. 154, 155 are called either Euler’s Equation by Pert (2013) [28] or Eulerian equation by Sabersky, et. al. (1998) [29]

**23.5. Mass Conservation for Fluid Flow, Continuum media.** The mass of fluid in some volume  $V_0 \subset N$  is  $\int_{V_0} \rho \text{vol}^n$ , where  $\rho$  is fluid density,  $\rho \in C^\infty(N)$ .

The total mass of fluid flowing out of volume  $V_0$  is

$$\begin{aligned}\frac{d}{dt} \int_{V_0} \rho \text{vol}^n &= \int_{V_0} \mathcal{L}_{\frac{\partial}{\partial t} + \mathbf{u}} (\rho \text{vol}^n) = \int_{V_0} \frac{\partial}{\partial t} \rho \text{vol}^n + \int_{V_0} \mathcal{L}_u \rho \text{vol}^n \\ \int_{V_0} \mathcal{L}_u \rho \text{vol}^n &= \int_{V_0} di_{\mathbf{u}} \rho \text{vol}^n + i_{\mathbf{u}} d\rho \text{vol}^n = \int_{V_0} di_{\mathbf{u}} \rho \text{vol}^n + 0 = \int_{V_0} di_{\mathbf{u}} \rho \text{vol}^n = \int_{\partial V_0} i_{\mathbf{u}} \rho \text{vol}^n\end{aligned}$$

Now

$$\begin{aligned}i_u \text{vol}^n &= i_u \frac{\sqrt{g}}{n!} \epsilon_{i_1 \dots i_n} dx^{i_1} \wedge \dots \wedge dx^{i_n} \\ i_u dx^{i_1} \wedge \dots \wedge dx^{i_n} &= u^{i_1} dx^{i_2} \wedge \dots \wedge dx^{i_n} - dx^{i_1} \wedge u^{i_2} dx^{i_3} \wedge \dots \wedge dx^{i_n} + \dots + (-1)^{n+1} dx^{i_1} \wedge \dots \wedge dx^{i_{n-1}} u^{i_n} = \epsilon_{j_1 \dots j_n}^{i_1 \dots i_n} u^{j_1} dx^{j_2} \wedge \dots \wedge dx^{j_n} \\ \implies i_u \text{vol}^n &= \frac{\sqrt{g}}{(n-1)!} \epsilon_{j_1 \dots j_n} u^{j_1} dx^{j_2} \wedge \dots \wedge dx^{j_n}\end{aligned}\tag{156}$$

We can also rewrite Eq. 156 to be a ”surface differential”:

$$\begin{aligned}\int_{V_0} \mathbf{d} i_{\mathbf{u}} \rho \text{vol}^n &= \int_{\partial V_0} i_{\mathbf{u}} \rho \text{vol}^n = \\ \int_{\partial V_0} \rho \frac{\sqrt{g}}{(n-1)!} u^{j_1} \epsilon_{j_1 j_2 \dots j_n} dx^{j_2} \wedge \dots \wedge dx^{j_n} &\equiv \int_{\partial V_0} \rho \mathbf{u} \cdot d\mathbf{S} \equiv \int_{\partial V_0} \rho \langle \mathbf{u}, d\mathbf{S} \rangle\end{aligned}\tag{157}$$

If  $\sqrt{g} = 1$ ,  $n = 2$ ,

$$i_u \text{vol}^2 = (u^1 dx^2 - u^2 dx^1) = u \cdot n_1 dx^2 + u \cdot n_2 dx^1 = u \cdot ndS$$

with  $n_1 = e_1$  and  $n_2 = -e_2$ .

Now

$$\begin{aligned}di_u \rho \text{vol}^n &= \\ = \frac{\partial(\sqrt{g} \rho u^{j_1})}{\partial x^k} \frac{\epsilon_{j_1 \dots j_n}}{(n-1)!} dx^k \wedge dx^{j_2} \wedge \dots \wedge dx^{j_n} &= \frac{\partial(\sqrt{g} \rho u^k)}{\partial x^k} \frac{\epsilon_{j_1 \dots j_n}}{n!} dx^{j_1} \wedge \dots \wedge dx^{j_n} = \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g} \rho u^k)}{\partial x^k} \text{vol}^n = \\ = \frac{\partial(\rho u^k)}{\partial x^k} \text{vol}^n + \rho u^k \frac{\partial \ln \sqrt{g}}{\partial x^k} \text{vol}^n &= \text{div}(\rho u) \text{vol}^n + \rho u^k \frac{\partial \ln \sqrt{g}}{\partial x^k} \text{vol}^n\end{aligned}$$

Now if  $\sqrt{g} = 1$ , then

$$\frac{d}{dt} \int_{V_0} \rho \text{vol}^n = \int_{V_0} \frac{\partial \rho}{\partial t} \text{vol}^n + \int_{V_0} di_u \rho \text{vol}^n = \int_{V_0} \frac{\partial \rho}{\partial t} \text{vol}^n + \int_{V_0} \text{div}(\rho u) \text{vol}^n \implies \frac{\partial \rho}{\partial t} + \text{div}(\rho u) = 0$$

which is the so-called mass continuity equation.  $j = \rho u$  is the mass flux density.

Thus,

$$\begin{aligned}&(\text{mass conservation}) \\ m = m(t) &:= \int_{V_0} \rho \text{vol}^n, V_0 \subset N \\ \dot{m} \equiv \frac{d}{dt} m(t) &= \int_{V_0} \left( \frac{\partial \rho}{\partial t} \text{vol}^n + \mathbf{d} i_{\mathbf{u}} \rho \text{vol}^n \right) = \boxed{\int_{V_0} \frac{\partial \rho}{\partial t} \text{vol}^n + \int_{\partial V_0} \rho \mathbf{u} \cdot d\mathbf{S}} \\ \text{if } \sqrt{g} = 1, \text{ and } \dot{m} = 0, \text{ then} & \\ \frac{\partial \rho}{\partial t} + \text{div}(\rho \mathbf{u}) &= 0\end{aligned}\tag{158}$$

Sec. 1.4 Eulerian Frame, 1.4.1 Conservation of Mass-Equation of Continuity, pp. 8 of Pert (2013) [28] says that Eq. 158 is for the Eulerian Frame.

From [wikipedia](#), recall that the *Eulerian specification* focuses on specific locations in space through which fluid flow as time passes. Visualize by sitting on bank of river and watch water pass fixed location. In the Eulerian specification of field, field represented as a function of position  $\mathbf{x}$ , time  $t$ , e.g.  $\mathbf{u}(\mathbf{x}, t) \equiv$  flow velocity.

TODO: 20190804 Frankel (2012) [42] in pp. 138 and onwards, for Sec. 4.3. Differentiation of Integrals posed the rightful question, ”How does one compute the rate of change of an integral when the domain of integration is also changing?” Revisit the derivation from a Lie derivative and 1-parameter flow point of view.

Force should not be represented by a vector but rather by a 1-form. Then,

$$f \in \Omega^1(N), f = f(t, \mathbf{x}) = f_j dx^j \quad j = 1, \dots, n, \mathbf{x} = (x^1, \dots, x^n), \dim N = n$$

Indeed, the reason for  $f$  to be a 1-form is that we integrate differential forms, we don’t integrate vectors.

$$W = \int_C f \quad (\text{line integral})\tag{159}$$

If  $f$  conservative,  $\exists$  scalar  $U = U(x)$ ;  $x \in N$  s.t.  $f = -dU$ .



Let  $\mathbf{u} = u^j(t, \mathbf{x}) \frac{\partial}{\partial x^j} \in \mathfrak{X}(N)$  generate flow  $\phi_t$ .

The Lie derivative along vector field  $\mathbf{u}$ ,  $\mathcal{L}_{\mathbf{u}}$ , can be calculated in at least 2 ways: from the definition:

$$\mathcal{L}_{\mathbf{u}}\omega = \frac{d}{dt} \Big|_{t=0} \phi_t^*\omega$$

or Cartan's magic formula:

$$\begin{aligned} \mathcal{L}_{\mathbf{u}}\text{vol}^n &= (i_{\mathbf{u}}\mathbf{d} + \mathbf{d}i_{\mathbf{u}})\text{vol}^n = 0 + \mathbf{d}i_{\mathbf{u}}\text{vol}^n = \mathbf{d}i_{\mathbf{u}}\left(\frac{\sqrt{g}}{n!}\epsilon_{i_1\dots i_n}dx^{i_1}\wedge\dots\wedge dx^{i_n}\right) = \\ &= \mathbf{d}\left(\frac{\sqrt{g}}{(n-1)!}\epsilon_{j_1\dots j_n}u^{j_1}dx^{j_2}\wedge\dots\wedge dx^{j_n}\right) = \frac{\partial(\sqrt{g}u^{j_1})}{\partial x^k}\frac{\epsilon_{j_1\dots j_n}}{(n-1)!}dx^k\wedge dx^{j_2}\wedge\dots\wedge dx^{j_n} = \\ (160) \quad &= \frac{\partial(\sqrt{g}u^{j_1})}{\partial x^k}\frac{\epsilon_{j_1\dots j_n}}{(n-1)!}\underbrace{\epsilon_{k_1,\dots,k_n}^{kj_2\dots j_n}}_{=1}dx^{k_1}\wedge\dots\wedge dx^{k_n}\left(\frac{\sqrt{g}}{\sqrt{g}}\right) = \frac{1}{\sqrt{g}}\frac{\partial(\sqrt{g}u^k)}{\partial x^k}\text{vol}^n = \\ &= (\text{div}\mathbf{u})\text{vol}^n \end{aligned}$$

cf. <https://math.stackexchange.com/questions/2566381/lie-derivative-of-volume-form?rq=1>

Note that

$$\text{div} : \mathfrak{X}(N) \rightarrow C^\infty(N)$$

$$\mathcal{L}_{\mathbf{u}} : \Omega^n(N) \rightarrow \Omega^n(N)$$

Frankel (2012) [42] on pp. lvii of the "Elasticity and Stresses" section offered this analogy: "While work in particle mechanics pairs a force covector ( $f_i$ ) with a contravariant tangent vector ( $dx^i/dt$ ) to a curve, work done by traction in elasticity pairs the contravariant stress force 2-form  $\mathcal{S}$  with the covector valued deformation 1-form  $\mathcal{E}$ , to yield a scalar valued 3-form.

Instead of pushing along a curve (line) particle to do work, what does it mean to do work on a moving fluid element? One could possibly try to consider the deformation of a volume element under fluid flow. Frankel (2012) [42] in Sec. 4.2. The Lie Derivative of a Form, on pp. 132, asks "If a flow deforms some attribute, say volume, how does one measure the deformation?  $\mathcal{L}_{\mathbf{u}}\text{vol}^n$  "is the  $n$ -form that reads off the rate of change of volume of a parallelopiped spanned by  $n$  vectors that are pushed forward by the flow  $\phi_t$ ." So, "in other words,  $\mathcal{L}_{\mathbf{u}}\text{vol}^n$  measures how volumes are changing under the flow  $\phi_t$  generated by  $\mathbf{X}$ " on pp. 133 on Frankel (2012) [42].

Then rate of work done on a fluid element could be the following:

$$W = \int_V \mathcal{L}_{\mathbf{u}}(p\text{vol}^n) = \int_V \mathbf{d}i_{\mathbf{u}}p\text{vol}^n = \int_{\partial V} i_{\mathbf{u}}p\text{vol}^n = \int_{\partial V} p\mathbf{u} \cdot d\mathbf{S}$$

TODO: work on a fluid element?

### 23.6. Bernoulli's equation.

23.6.1. *Simple cases.* cf. Ch. 3 The Bernoulli Equation, Sec. 3.2 "A Simple Form of the Bernoulli Equation", pp. 90, Sabersky, et. al. (1998) [29]

Start from Eq. 154. So

$$(\nabla_u u)^i = \frac{-1}{\rho}(\text{grad}p)^i + f^i$$

Suppose  $f^i = 0$  (no external forces),  $\frac{\partial u^i}{\partial t} = 0$  (steady flow). Suppose  $\Gamma_{\mu\nu}^a = 0$  (Cartesian coordinates? TODO: find out the conditions when this is true).

Suppose the flow is irrotational (TODO: find out the condition for irrotational flow; is it when having streamlines or ?). For the curl, recall that an external derivative  $\mathbf{d}$  takes a 1-form into a 2-form. So consider

$$\begin{aligned} \mathbf{u} \in \mathfrak{X}(V), \mathbf{u} \xrightarrow{b} u_idx^i = g_{ij}u^jdx^i \\ d\mathbf{u}^b \in \Omega^2(V), d\mathbf{u}^b = \frac{\partial u_j}{\partial x^i}dx^i \wedge dx^j = \frac{\partial(g_{ij}u^j)}{\partial x^a}dx^a \wedge dx^i \end{aligned}$$

If we have the condition that the flow is irrotational, i.e.

$$(161) \quad d\mathbf{u}^b = 0$$

then

$$d\mathbf{u}^b(\mathbf{u}) = \frac{\partial(u_i)}{\partial x^a}u^a dx^i - \frac{\partial(u_i)}{\partial x^a}dx^a u^i = 0 \implies u^a \frac{\partial(u_i)}{\partial x^a}dx^i = u^i \frac{\partial u_i}{\partial x^a}dx^a$$

Take the exterior derivative  $\mathbf{d}$  of the inner product of a vector with itself in the most general case on a metric manifold:

$$\begin{aligned} d(u_i u^i) &= d(g_{ij}u^j u^i) = \frac{\partial g_{ij}}{\partial x^k}u^j u^i dx^k + g_{ij} \frac{\partial u^j}{\partial x^k}u^i dx^k + g_{ij}u^j \frac{\partial u^i}{\partial x^k}dx^k = \\ &= 2u_j \frac{\partial u^j}{\partial x^k}dx^k + u^i u^j \frac{\partial g_{ij}}{\partial x^k}dx^k \end{aligned}$$

Since the Christoffel symbols are the connection coefficients of a *metric-compatible* affine connection (i.e. they can be used to construct a connection  $\nabla$  s.t.  $\nabla g = 0$ ), then <sup>6</sup>

$$(162) \quad \nabla_i g_{jk} = \frac{\partial g_{jk}}{\partial x^i} - \Gamma_{ji}^s g_{sk} - \Gamma_{ki}^s g_{js} = 0$$

and so

$$(163) \quad u^i u^j \frac{\partial g_{ij}}{\partial x^k} = u^i u^j [\Gamma_{ik}^s g_{sj} + \Gamma_{jk}^s g_{is}] = u^i u_s \Gamma_{ik}^s + u_s u^j \Gamma_{jk}^s = 2u_s u^i \Gamma_{ik}^s$$

Then

$$(164) \quad \frac{1}{2}d(u^2) = u_j \frac{\partial u^j}{\partial x^k}dx^k + u_s u^i \Gamma_{ik}^s dx^k$$

Eq. 164 only depended upon the fact that the connection  $\nabla$  is *metric-compatible*.

So if we suppose that the Christoffel symbols  $\Gamma_{ik}^s$  are 0, and have irrotational flow, as well,

$$d(u^2/2) = u_j \frac{\partial u^j}{\partial x^k}dx^k + 0 = u^a \frac{\partial u_i}{\partial x^a}dx^i$$

Now consider the RHS of Eq. 154, with  $f_i = 0$ ,

$$\frac{-1}{\rho}(\text{grad}p)^i \xrightarrow{b} \frac{-1}{\rho}(\text{grad}p)^i g_{ij}dx^j = \frac{-1}{\rho} \frac{\partial p}{\partial x^k} g^{ki} g_{ij}dx^j = \frac{-1}{\rho} \frac{\partial p}{\partial x^j} dx^j = \frac{-1}{\rho} d\mathbf{p}$$

So if the fluid is incompressible, then  $\rho$  is constant and so we can write

$$(165) \quad d\left(\frac{u^2}{2} + \frac{p}{\rho}\right) = 0$$

for any irrotational, incompressible, fluid flow with no (external) body forces.

<sup>6</sup>"Prove Christoffel Symbol Identity", Physics StackExchange

23.6.2. *Venturi tube example.* cf. Example 3.1 in Sabersky, et. al. (1998) [29]

Consider flow of liquid through the tube shown. Determine pressure difference between points 1 and 2 as a function of flow rate  $Q$ , density  $\rho$ . Velocity assumed constant over cross sections, and pressure will also be assumed uniform over each of these 2 areas  $A_1, A_2$ . Also assume there’s no viscosity.

**Solution Example 3.1 in Sabersky, et. al. (1998) [29].**

Since density is constant, continuity equation requires that for  $Q$  the volume flow rate,

$$\text{flow rate} Q = u_1 A_1 = u_2 A_2$$

$$\begin{aligned} \frac{u_1^2}{2} + \frac{p_1}{\rho} &= \frac{u_2^2}{2} + \frac{p_2}{\rho} \implies \frac{Q^2}{2A_1^2} + -\frac{Q^2}{2A_2^2} = \frac{-(p_1 - p_2)}{\rho} \quad \text{or } p_1 - p_2 = \frac{+\rho Q^2}{2A_2^2} \left(1 - \frac{A_2^2}{A_1^2}\right) \\ &\frac{Q^2}{2} \left(\frac{1}{A_1^2} - \frac{1}{A_2^2}\right) = \frac{-(p_1 - p_2)}{\rho} \end{aligned}$$

23.6.3. *Conservative body forces with Bernoulli equation.* cf. Sec. 3.3 in Sabersky, et. al. (1998) [29]

Consider  $U = U(x)$ , force potential.

$$\begin{aligned} -\mathbf{d}U &= \frac{-\partial U}{\partial x^i} dx^i = f_i dx^i \\ \mathbf{d}\left(\frac{u^2}{2} + \frac{p}{\rho} + U\right) &= 0 \end{aligned}$$

23.6.4. *Barotropic Fluids ( $\rho = \rho(p)$ ) with Bernoulli equation.* cf. Sec. 3.4 in Sabersky, et. al. (1998) [29]

*Barotropic fluid:* density is single-valued function of  $p$ . In the absence of heat transfer and absence of entropy change, by thermodynamics this can be shown to be barotropic.

$$\mathbf{d}\left(\frac{u^2}{2}\right) = -\frac{1}{\rho}\mathbf{d}p - \mathbf{d}U \text{ or } \mathbf{d}\left(\frac{u^2}{2} + U\right) + \frac{1}{\rho}\mathbf{d}p$$

Suppose  $i = i(p)$  s.t.  $di = \frac{\partial i}{\partial p} dp = \frac{1}{\rho(p)} = \frac{di}{dp} dp$ . So

$$\begin{aligned} \frac{di}{dp} &= \frac{1}{\rho(p)} \text{ or } i = \int \frac{1}{\rho} dp + \text{const} \\ \implies \mathbf{d}\left(\frac{u^2}{2} + U + i\right) &= 0 \end{aligned}$$

In special case of isentropic flow of perfect gas,  $\frac{P}{\rho^\nu} = \text{const.}$  and  $i = \int \frac{dp}{\rho^{1/\nu}} = \frac{kp^{1-1/\nu}}{1-1/\nu} = \text{enthalpy.}$

23.6.5. *Nonsteady flow  $\frac{\partial u}{\partial t}$  with Bernoulli equation.* First, use the musical isomorphism to transform the (1,0) tensor:

$$\frac{\partial u^i}{\partial t} \frac{\partial}{\partial x^i} \xrightarrow{\flat} g_{ij} \frac{\partial u^i}{\partial t} dx^j$$

Along a curve  $\gamma = \gamma^i(t) = x^i(t)$  s.t. . Then

$$\begin{aligned} \dot{\gamma} = \dot{\gamma}^i(t) \frac{\partial}{\partial x^i} &= u^i(t, \gamma(t)) \frac{\partial}{\partial x^i} \\ \int_{\gamma} g_{ij} \frac{\partial u^i}{\partial t} dx^j &= \int g_{ij} \frac{\partial u^i}{\partial t} \dot{\gamma}^j(t) dt = \int g_{ij} \frac{\partial u^i}{\partial t} u^j dt = \int \frac{\partial \|u\|}{\partial t} dt \end{aligned}$$

The last step can be made only if the metric  $g$  is time-independent.

cf. Example 3.2 in Sabersky, et. al. (1998) [29].

Consider the *efflux* of liquid of density  $\rho$  through an orifice of area  $A_0$  at the bottom of a large tank of cross-sectional area  $A_t$ .

The depth of liquid at the given instant is  $y$ .

pressure exerted on the liquid in the tank is maintained at  $p_t$ ,

pressure outside the orifice is the atmospheric pressure  $p_a$ .

The action of gravity is to be considered, but frictional effects may be neglected. Velocity of efflux  $u_j$  is to be determined for these conditions.

efflux - something given off in or as if in a stream (Webster dictionary)

**Solution Example 3.2 in Sabersky, et. al. (1998) [29].**

In Sec. 3.8(b) flows through orifices discussed in great detail. For this example, assume the orifice is carefully designed so streamlines at orifice exit are parallel and thus orifice area  $A_0$  and cross-sectional area of jet  $A_j$  are equal there.

cf. Example 3.3 in Sabersky, et. al. (1998) [29]. Example of a problem in which nonsteady terms cannot be neglected.

## 24. VISCOUS FLUIDS, NAVIER-STOKES EQUATION APPLICATIONS

cf. Ch. 2, Viscous Fluids, Sec. 15. The equations of motions of a viscous fluid, Landau and Lifshitz (1987) [30]

Recall Navier-Stokes equation for *compressible, viscous* fluid flow:

$$(166) \quad \boxed{\rho \left( \frac{\partial u^i}{\partial t} + u^j \frac{\partial u^i}{\partial x^j} \right) = -\frac{\partial p}{\partial x^i} + (\lambda + \mu) \frac{\partial}{\partial x^i} \text{div} u + \mu \Delta u^i}$$

Compare this with the expression in Eq. 15.6, Landau and Lifshitz (1987) [30]

$$(167) \quad \rho \left[ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \mathbf{grad}) \mathbf{v} \right] = -\mathbf{grad} p + \eta \Delta \mathbf{v} + \left( \zeta + \frac{1}{3} \eta \right) \mathbf{grad} \text{div} \mathbf{v}$$

$\mu \equiv \eta$ , fluid isotropic, so properties described by scalar quantities.

$\mu \equiv \eta$  viscosity coefficients,

$\lambda + \mu \equiv \zeta + \frac{1}{3} \eta$ ,  $\zeta \equiv$  second viscosity.

### 24.1. Flow in a pipe.

## 25. THERMODYNAMICS

Let  $\Sigma$  be a (topological) manifold.

Suppose  $U$  is a global coordinate on  $\Sigma$ :

$$(168) \quad \begin{aligned} &\text{(First Law of Thermodynamics (energy conservation))} \\ &\boxed{dU = Q - W \text{ or } Q = dU + W} \end{aligned}$$

where  $dU, Q, W \in \Omega^1(\Sigma)$  (i.e.  $dU, Q, W$  are 1-forms over manifold  $\Sigma$ ).

Consider a path in  $\Sigma$ ,  $\gamma, \gamma : \mathbb{R} \rightarrow \Sigma$ ,

and using a chart  $(U, S^1, \dots, S^n)$  (e.g.  $n = 1$ ,  $S^1 = v$  for volume)

$$\gamma(t) = (U(t), S^1(t), \dots, S^n(t))$$

$$\dot{\gamma} \in \mathfrak{X}(\Sigma), \dot{\gamma} = \dot{U} \frac{\partial}{\partial U} + \dot{S}^i \frac{\partial}{\partial S^i}$$

Now

$$\begin{aligned} dU(\dot{\gamma}) &= \dot{\gamma}(U) = \dot{U} \frac{\partial}{\partial U} U + 0 = \dot{U} \\ Q(\dot{\gamma}) &= Q(t)dt \left( \dot{\gamma} \frac{\partial}{\partial t} \right) = Q(t)\dot{\gamma} \\ \implies \dot{U} &= Q(t)\dot{\gamma} - W(t)\dot{\gamma} \end{aligned}$$

Recall that for enthalpy  $H$ ,  $H := U + pV$ ,  $H = H(\sigma, p)$

TODO 20190804 Derive and check convection form of enthalpy against both Kittle and Kroemer plus thermodynamics and Sonntag, et. al.

Incomplete:

$$\begin{aligned} dH &= Q + Vdp \\ W = pdV &= pdV + Vdp - Vdp = d(pV) - Vdp = \\ dE &= Q - W + \mu dH \\ dE &= Q - W + \hat{h}dN \end{aligned}$$

$U \rightarrow E$  notation is to promote the internal energy  $U$  to include kinetic and potential energies, so that possibly,  $E = U + \text{K.E.} + \text{P.E.}$ , or, i.e.,  $E$  includes internal energy and mechanical energy.

$$\begin{aligned} G &= \mu N \\ G = U - \tau\sigma + pV &\xrightarrow{U \rightarrow E} G = E - \tau\sigma + pV \implies \tau\check{h} \end{aligned}$$

$$dE = Q + W + \mu dN = Q + W + (\check{h} - \tau\check{\sigma})dN$$

$$Q = \tau d\sigma = \tau d(N\check{\sigma}) = \tau Nd\check{\sigma} + \tau\check{\sigma}dN$$

$\tau Nd\check{\sigma}$  is the entropy change due to change in entropy per particle; i.e. **conduction term**

$\tau\check{\sigma}dN$  is entropy change due to change in number of particles, i.e. **convection term**

$$dE = Q + W + \check{h}dN - \tau\check{\sigma}dN = \tau Nd\check{\sigma} + \tau\check{\sigma}dN + W + \check{h}dN - \tau\check{\sigma}dN = \tau Nd\check{\sigma} + W + \check{h}dN$$

$$m(t) = MN$$

Assume only 1 chemical species:

$$\begin{aligned} \check{Q} &:= \tau Nd\check{\sigma} \quad \hat{h} := \frac{H}{m} = \frac{H}{MN} \\ \implies dE &= \check{Q} + W + \check{h}dN \frac{M}{M} = \check{Q} + W + \hat{h}dm \end{aligned}$$

The Gibbs free energy equilibrium is given by,

$$dG = \mu dN - \sigma d\tau + Vdp$$

For a throttling valve (don't all valves throttle?) the pressure drop is accounted for by the Gibbs free energy, *not* by an isenthalpy condition.

#### Part 14. Classical Mechanics applications

cf. Arnold, Kozlov, Neishtadt (2006) [32].

If known forces  $\mathbf{F}_1 \dots \mathbf{F}_n$  acts on points, then

$$(169) \quad \sum_{i=1}^n \langle m_i \ddot{\mathbf{r}}_i - \mathbf{F}_i, \xi_i \rangle = 0$$

cf. Eq. (1.26) of Arnold, Kozlov, Neishtadt (2006) [32], where  $\xi_1, \dots, \xi_n$  are arbitrary tangent vectors to  $M$ ,  $\xi_i, \dots, \xi_n \in TM$ .  
 $\sum_{i=1}^n \langle m_i \ddot{\mathbf{r}}_i - \mathbf{F}_i, \xi_i \rangle$  called "general equation of dynamics" or d'Alembert-Lagrange principle.

#### 26. ORBITAL MECHANICS

References: Arnold, Kozlov, Neishtadt (2006) [32] Choquet-Bruhat (2009) [39], pp. 39, 3. General relativity Bate, Mueller, White (1971) [46], Sec. 1.2, pp. 5, [https://en.wikipedia.org/wiki/N-body\\_problem](https://en.wikipedia.org/wiki/N-body_problem) Abraham and Marsden (2008) [33], Ch. 9 Prástaro (1996) [15], Remark 3.20 many-bodies dynamics, pp. 346

26.1. **Conic sections.** The following is an amalgamation or summary of the development in Ch. 13, Sections 13.18-22, from pp. 497, Apostol (1991) [45]

**Definition 65** (General definition of conic sections (vector-formulated definition)). *Given line  $L$ , called the directrix, pt.  $F$ , the focus, such that  $F \notin L$ ,*

*eccentric  $e > 0$ ,*

*$d =$  distance of  $L$  from  $F$ , then*

*the **conic section**  $C = \{X\}$  (set of all points) such that*

$$(170) \quad \|X - F\| = ed(X, L)$$

*Now,*

*$e < 1$  ellipse,*

*$e = 1$  parabola,*

*$e > 1$  hyperbola.*

*Further, if  $N$  unit normal to  $L$ , then  $\forall P \in L$ ,*

$$(171) \quad d(X, L) = |(X - P) \cdot N|$$

*If  $(X - P) \cdot N \geq 0$ ,  $X \in$  positive half plane (negative half-plane).*

**Theorem 25** (Further conic section development). *Building upon Def. 65,*

*If  $F \in$  positive half plane determined by  $N$ , then  $F - dN \in L$ .*

*If  $F \in$  negative half plane determined by  $N$ , then  $F + dN \in L$ .*

*Choose this convenient, specific point:  $F + dN = P \in L \implies \|X - F\| = e|(X - (F + dN)) \cdot N|$*

*If  $F = 0$ ,  $\|X\| = e|X \cdot N - d| = r = e(d - r \cos \theta)$ . Then*

$$(172) \quad \frac{ed}{r} = 1 + e \cos \theta \text{ or } \boxed{r = \frac{ed}{1 + e \cos \theta}} \text{ for } F = 0$$

*Consider symmetry about the origin (i.e. of  $X, \exists -X$ ).*

*As a caveat, let  $A = e(d + F \cdot N)$ . Then*

$$\|X - F\| = e|(X - (F + dN)) \cdot N| = ed(X, L) = e|X \cdot N - \frac{A}{e}|$$

*Then*

$$(173) \quad \|X - F\|^2 = \|X\|^2 - 2X \cdot F + \|F\|^2 = e^2((X \cdot N)^2 - 2\frac{A}{e}X \cdot N + \left(\frac{A}{e}\right)^2)$$

*Using the symmetry about the origin,  $X \rightarrow -X$ ,*

$$\implies X \cdot F = AeX \cdot N \text{ or } X \cdot (F - AeN) = 0$$

*Then  $F = AeN$ . Note that  $A = e(F \cdot N \pm d)$  (for  $F$  in negative (positive) half-plane).*

Then

$$Ae = e^2(Ae + d) \text{ or } \frac{A}{e} = Ae + d \text{ or } A\left(\frac{1}{e} - e\right) = d \text{ or } A = \frac{ed}{1 - e^2}$$

$$\implies \|X - AeN\| = e|(X - (AeN + dN)) \cdot N| = e|X \cdot N - (Ae + d)|$$

Let  $X = aN$ . Plug this into Eq. [173](#).

$$(174) \quad a^2 - 2aAe + A^2e^2 = e^2\left(a^2 - \frac{2A}{e}a + \left(\frac{A}{e}\right)^2\right) = e^2a^2 - 2Aae + A^2 \implies a = A$$

$$\implies a = \frac{ed}{1 - e^2}$$

Thus, we've shown that  $A$  is equal to the semi-major axis  $a$  and that

$$(175) \quad \boxed{F = aeN}$$

Thus, the semi-major axis is obtained when  $X = \pm aN$  (vertices), so that  $\|X\|^2 + (ae)^2 = e^2(X \cdot N)^2 + a^2$  is satisfied. Likewise, the semi-minor axis is obtained if  $X = \pm bN_\perp$ , so that

$$(176) \quad b^2 + (ae)^2 = 0 + a^2 \implies \boxed{b = a\sqrt{1 - e^2}}$$

We can then rewrite Eq. [172](#), with Eq. [174](#) as

$$\boxed{r = \frac{a(1 - e^2)}{1 + e \cos \theta}}$$

With Eq. [173](#), and using the symmetry of  $X$  about the origin, then

$$\|X\|^2 + \|F\|^2 = e^2(X \cdot N)^2 + a^2$$

Choose  $\mathbf{N} = \mathbf{e}_x$ . With  $\|X\|^2 = x^2 + y^2$ , and  $\|F\| = Ae = ae$ , then

$$(177) \quad x^2 + y^2 + a^2e^2 = e^2x^2 + a^2 \text{ or } x^2(1 - e^2) + y^2 = a^2(1 - e^2)$$

cf. 13.23 "Cartesian equations for the conic sections", Apostol (1991) [[45](#)].

26.1.1. *Ellipse*. If  $e < 1$ , then from Eq. [177](#),

$$\xrightarrow{\text{if } e < 1} \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

Observe that what Greiner (2004) [[31](#)] in pp. 249, Ch. "Planetary Motions", (a) Ellipse, denotes  $c$  is what we and Apostol denote as  $\|F\| = ae$ , the distance from the origin to the focus. Observe that

$$c^2 + b^2 = a^2e^2 + a^2(1 - e^2) = a^2$$

26.1.2. *Hyperbola*. If  $e > 1$ , then from Eq. [177](#),

$$\xrightarrow{\text{if } e > 1} \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Let  $b = |a|\sqrt{e^2 - 1}$ . The foci of a hyperbola are at  $(c, 0)$ ,  $(-c, 0)$  where  $c = |a|e = \sqrt{a^2 + b^2}$ .

26.1.3. *Parabola*. cf. pp. 507 of 13.23 "Cartesian equations for the conic sections", Apostol (1991) [[45](#)]

Go back to Eq. [170](#), where

$$\|X - F\| = d(X, L)$$

Let  $L$  be a vertical line at  $x = \mp c$ , place focus  $F = (\pm c, 0)$ .

If  $X = (x, y)$ ,  $X - F = (x \mp c, y)$ , so

$$\|X - F\|^2 = (x \mp c)^2 + y^2 = |x \pm c|^2 = x^2 \pm 2xc + c^2 = x^2 \mp 2xc + c^2 + y^2 \implies y^2 = \pm 4xc$$

So midway between the focus and directrix, at the origin, is the vertex of the parabola.

## 26.2. Central Force Problem via Energy Methods.

26.2.1. *Reduction to the Equivalent One-Body Problem*. cf. pp. 70 of Goldstein, Poole, Safko (2001) [[27](#)]

Let  $\mathbf{R}$  = radius vector to the center of mass.

Let  $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$  for the relative position of two masses. Then the Lagrangian

$$(178) \quad \mathcal{L} = T(\dot{\mathbf{R}}, \dot{\mathbf{r}}) - U(\mathbf{r}, \dot{\mathbf{r}}, \dots)$$

cf. Eq. (3.1) of Goldstein, Poole, Safko (2001) [[27](#)]

The kinetic energy  $T$  is the sum of the kinetic energy of the center of mass motion plus the kinetic energy of motion, about the center of mass  $T'$ .

$$T = \frac{1}{2}(m_1 + m_2)\dot{\mathbf{R}}^2 + T' \text{ with } T' = \frac{1}{2}m_1\dot{\mathbf{r}}_1'^2 + \frac{1}{2}m_2\dot{\mathbf{r}}_2'^2$$

$\mathbf{r}'_1, \mathbf{r}'_2$  are radii vectors relative to the center of mass, since

$$\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{M}$$

so

$$(179) \quad \begin{aligned} \mathbf{r}_1 - \mathbf{R} &= \frac{m_2(\mathbf{r}_1 - \mathbf{r}_2)}{M} = -\frac{m_2}{M}\mathbf{r} \equiv \mathbf{r}'_1 \\ \mathbf{r}'_2 \equiv \mathbf{r}_2 - \mathbf{R} &= \frac{m_1(\mathbf{r}_2 - \mathbf{r}_1)}{M} = \frac{m_1}{M}\mathbf{r} \end{aligned}$$

cf. Eqn. (3.2) of Goldstein, Poole, Safko (2001) [[27](#)].

Thus  $T' = \frac{1}{2}\frac{m_1m_2}{m_1+m_2}\dot{\mathbf{r}}^2$

So

$$(180) \quad \mathcal{L} = \frac{M}{2}\dot{\mathbf{R}}^2 + \frac{1}{2}\frac{m_1m_2}{M}\dot{\mathbf{r}}^2 - U(\mathbf{r}, \dot{\mathbf{r}}, \dots)$$

cf. Eqn. (3.3) of Goldstein, Poole, Safko (2001) [[27](#)].

3 coordinates  $\mathbf{R}$  are cyclic, so CM is at rest or moving uniformly. Drop the first term. The rest of the Lagrangian is exactly what would be expected if we had a fixed center of force with a single particle at distance  $\mathbf{r}$  from it.

Let

$$\mu = \frac{m_1m_2}{m_1 + m_2} \equiv \text{reduced mass}$$

26.2.2. *Equations of Motion from Lagrangian and First Integrals.* cf. Sec. 3.2 "The Equations of Motions and First Integrals, cf. Eqn. (3.15) of Goldstein, Poole, Safko (2001) [27], with Chapter 3 The Central Force Problem, pp. 72 of Goldstein, Poole, Safko (2001) [27].

Goldstein, Poole, Safko argues: assume potential is of form  $V(r)$  and so  $V(r)$  has only radial distance, so the problem has spherical symmetry. Then, Goldstein, et. al. argues  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$  conserved. Thus  $\mathbf{r}$  and  $\mathbf{p} \perp \mathbf{L}$  always.

So

$$(181) \quad \mathcal{L} = T - V = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2) - V(r)$$

cf. Eqn. (3.6) Goldstein, Poole, Safko (2001) [27].

Now

$$p_{\theta} = \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \mu r^2 \dot{\theta}$$

Since  $\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = \frac{\partial \mathcal{L}}{\partial \theta}$ , then

$$(182) \quad \dot{p}_{\theta} = \frac{d}{dt}(\mu r^2 \dot{\theta}) = 0 = \frac{\partial \mathcal{L}}{\partial \theta}$$

cf. Eqn. (3.7) Goldstein, Poole, Safko (2001) [27].

So

$$(183) \quad \mu r^2 \dot{\theta} = l$$

cf. Eqn. (3.8) Goldstein, Poole, Safko (2001) [27], where  $l$  is the constant magnitude of the angular momentum.

So then  $\frac{d}{dt} \left( \frac{1}{2} r^2 \dot{\theta} \right) = 0$  where  $\frac{1}{2} r^2 \dot{\theta}$  is the areal velocity, the area swept out by radius vector per unit time.

Goldstein argues that  $\frac{1}{2} r^2 \dot{\theta}$  follows from  $dA = \frac{1}{2} r(r d\theta)$ , so  $\frac{dA}{dt} = \frac{1}{2} r^2 \frac{d\theta}{dt}$ . Thus, angular momentum conservation is equal to areal velocity constant. Do keep in mind that this argument is based upon the diagram in Goldstein, Poole, Safko (2001) [27] and presumably the small angle approximation (perhaps using  $\sin \theta \approx \theta$  for small  $\theta$ ). Thus, Kepler's 2nd. law: the radius vector sweeps out equal areas in equal times.

For the  $r$  coordinate:

$$\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{r}} \right) = \frac{\partial \mathcal{L}}{\partial r} \implies \frac{d}{dt}(\mu \dot{r}) = \mu r \dot{\theta}^2 - \frac{\partial V}{\partial r} \text{ or } \mu \ddot{r} - \mu r \dot{\theta}^2 = -\frac{\partial V}{\partial r}$$

Denote

$$(184) \quad f(r) = \mu \ddot{r} - \mu r \dot{\theta}^2$$

cf. Eqn. (3.11) of Goldstein, Poole, Safko (2001) [27].

Since  $\mu r^2 \dot{\theta} = l$ , then

$$(185) \quad f(r) = \mu \ddot{r} - \frac{l^2}{\mu r^3}$$

cf. Eqn. (3.12) of Goldstein, Poole, Safko (2001) [27].

Since

$$(186) \quad \mu \ddot{r} = \frac{-\partial}{\partial r} \left( V + \frac{1}{2} \frac{l^2}{\mu r^2} \right)$$

cf. Eqn. (3.14) of Goldstein, Poole, Safko (2001) [27] and since  $\mu \ddot{r} \dot{r} = \frac{d}{dt} \left( \frac{1}{2} \mu \dot{r}^2 \right)$  and  $\frac{d}{dt} g(r) = \frac{\partial g}{\partial r} \dot{r}$ , then

$$\frac{d}{dt} \left( \frac{1}{2} \mu \dot{r}^2 \right) = \frac{-d}{dt} \left( V + \frac{1}{2} \frac{l^2}{\mu r^2} \right) \text{ or } \frac{d}{dt} \left( \frac{1}{2} \mu \dot{r}^2 + \frac{1}{2} \frac{l^2}{\mu r^2} + V \right) = 0$$

and so

$$(187) \quad \frac{1}{2} \mu \dot{r}^2 + \frac{1}{2} \frac{l^2}{\mu r^2} + V = \text{constant}$$

$$\frac{1}{2} \frac{l^2}{\mu r^2} = \frac{1}{2\mu r^2} \mu^3 r^4 \dot{\theta}^2 = \frac{\mu r^2 \dot{\theta}^2}{2}$$

and so

$$(188) \quad \frac{1}{2} \mu \dot{r}^2 + \frac{1}{2} \mu r^2 \dot{\theta}^2 + V = E = T + V$$

cf. Eqn. (3.13) of Goldstein, Poole, Safko (2001) [27].

For  $E = \frac{1}{2} \mu \dot{r}^2 + \frac{l^2}{2\mu r^2} + V$ , then

$$(189) \quad \dot{r} = \sqrt{\frac{2}{\mu} \left( E - V - \frac{l^2}{2\mu r^2} \right)}$$

cf. Eqn. (3.16) of Goldstein, Poole, Safko (2001) [27], or

$$(190) \quad dt = \frac{dr}{\sqrt{\frac{2}{\mu} \left( E - V - \frac{l^2}{2\mu r^2} \right)}}$$

cf. Eqn. (3.17) of Goldstein, Poole, Safko (2001) [27].

Let

$$t = \int_{r_0}^r \frac{dr}{\sqrt{\frac{2}{\mu} \left( E - V - \frac{l^2}{2\mu r^2} \right)}}$$

with initial value  $r_0$ .

From  $l = \mu r^2 \dot{\theta}$ ,

$$(191) \quad d\theta = \frac{l dt}{\mu r^2}$$

cf. Eqn. (3.19) of Goldstein, Poole, Safko (2001) [27] so for  $\theta_0 \equiv$  initial value of  $\theta$ .

$$(192) \quad \theta = l \int_0^t \frac{dt'}{\mu r^2(t')} + \theta_0$$

cf. Eqn. (3.20) of Goldstein, Poole, Safko (2001) [27].

In Sec. 3.3, "The Equivalent 1-Dimensional Problem and Classification of Orbits" of Goldstein, Poole, Safko (2001) [27], consider  $V' = V + \frac{1}{2} \frac{l^2}{\mu r^2}$ , so that  $F' = \frac{-\partial V'}{\partial r} = f(r) + \frac{l^2}{\mu r^3}$ .

Let  $V = \frac{-k}{r}$ , e.g.  $V = \frac{-GMm}{r}$  with  $k = GMm$ .

Then consider  $V' = \frac{-k}{r} + \frac{l^2}{2\mu r^2}$  and its plot.

Compare this, Thm. 25, against (26.6) of Greiner (2004) [31]:

$$(193) \quad r = \frac{ed}{1 + e \cos \theta}, \quad e = \frac{C \widehat{L}_0^2}{GM}, \quad d = 1/C$$

cf. Eq. (26.32) of Greiner (2004) [31].

Investigate which physical quantities (e.g.  $E, L$ ) eccentricity  $e$  depends on.



$$\begin{aligned}
u &= \frac{GM}{\widehat{L}_0^2} + C \cos(\theta - \theta_0) \implies \frac{du}{d\theta} = -C \sin(\theta - \theta_0) \\
\frac{1}{2}mL_0^2 \left[ C^2 \sin^2(\theta - \theta_0) + \left( \frac{GM}{L_0^2} \right)^2 + 2C \frac{GM}{L_0^2} \cos(\theta - \theta_0) + C^2 \cos^2(\theta - \theta_0) \right] &= \\
\implies &= \frac{1}{2}m\widehat{L}_0^2 \left[ C^2 + \left( \frac{GM}{\widehat{L}_0^2} \right)^2 + 2C \frac{GM}{\widehat{L}_0^2} \cos(\theta - \theta_0) \right] = \\
&= E - V\left(\frac{1}{u}\right) = E + GMmu = E + \frac{(GM)^2}{\widehat{L}_0^2}m + CGMm \cos(\theta - \theta_0) \\
&\implies \frac{1}{2}m\widehat{L}_0^2 C^2 = E + \frac{1}{2} \frac{(GM)^2}{\widehat{L}_0^2}m
\end{aligned}$$

Solve for the dummy constant  $C$ :

$$(194) \quad C = \sqrt{\frac{2E}{m\widehat{L}_0^2} + \frac{(GM)^2}{\widehat{L}_0^4}}$$

cf. Eq. (26.36) of Greiner (2004) [31].

From Eq. (26.32) of Greiner (2004) [31] or Eq. 193,

$$e = \frac{C\widehat{L}_0^2}{GM} = \sqrt{\frac{2E}{m\widehat{L}_0^2} + \frac{(GM)^2}{\widehat{L}_0^4}} \left( \frac{\widehat{L}_0^2}{GM} \right)$$

So

$$(195) \quad e = \sqrt{\frac{2E\widehat{L}_0^2}{m(GM)^2} + 1}$$

cf. Eq. (26.37) of Greiner (2004) [31].

Hence, shape of the path depends on total energy  $E$ , angular momentum  $L_0 = m\widehat{L}_0$  of moving body (recall that  $\widehat{L} := \mathbf{r} \times \mathbf{v}$ ).

It holds that for a

$$\begin{aligned}
&\text{parabola } e = 1, & E &= 0 \\
&\text{ellipse } 0 < e < 1, & E &< 0, & \frac{(GM)^2 m}{-2\widehat{L}_0^2} < E < 0 \\
&\text{circle } e = 0 & E &= \frac{-(GM)^2 m}{2\widehat{L}_0^2} \\
&\text{hyperbola } e = 1, & E &> 0
\end{aligned}$$

The general form of initial energy  $E_0$  is, recall, give by

$$E_0 = \frac{1}{2}mv^2 + V(r) = \frac{1}{2}mv^2 - \frac{GMm}{r_0} = m \left( \frac{v^2}{2} - \frac{GM}{r_0} \right)$$

Notice that for the case of the parabola that  $E_0 = 0$  when

$$(196) \quad v = \sqrt{\frac{2GM}{r_0}}$$

Eq. 196 is the expression for the **escape velocity**.

Knowing what we know now, let's rewrite some of the expressions from before:

$$\begin{aligned}
r(\theta) &= \frac{\widehat{L}_0^2/GM}{1 + \frac{C\widehat{L}_0^2}{GM} \cos(\theta - \theta_0)} \\
\widehat{L}_0^2 &:= r_0^2 \dot{\theta}_0 \\
C &= \sqrt{\frac{2E}{m\widehat{L}_0^2} + \frac{(GM)^2}{\widehat{L}_0^4}}
\end{aligned}$$

26.3. **Path parameters (i.e. orbital parameters).** cf. pp. 262, Greiner (2004) [31].

$$a = \frac{1}{2}(r(\theta = 0) + r(\theta = \pi)) = \frac{1}{2} \left[ \frac{\widehat{L}_0^2/GM}{1+e} + \frac{\widehat{L}_0^2/GM}{1-e} \right] = \frac{\widehat{L}_0^2/GM}{1-e^2}$$

The addition above for  $a$  is needed because we center origin at a focus, for this expression.

$$(197) \quad a = \frac{\widehat{L}_0^2/GM}{\frac{-2E\widehat{L}_0^2}{m(GM)^2}} = \frac{GMm}{-2E}$$

cf. Eq. (26.39) of Greiner (2004) [31].

$$(198) \quad b = \sqrt{a^2 - f^2} = \sqrt{a^2 - e^2 a^2} = a\sqrt{1 - e^2} = a\sqrt{\frac{-2E\widehat{L}_0^2}{m(GM)^2}} = \widehat{L}_0 \sqrt{\frac{-m}{2E}}$$

cf. Eq. (26.40) of Greiner (2004) [31].

We can also calculate the perigee (point of closest approach to mass  $M$ ):

$$(199) \quad r(\theta = 0) = \frac{\widehat{L}_0^2/GM}{1+e} = \frac{\widehat{L}_0^2/GM}{1 + \sqrt{\frac{2E\widehat{L}_0^2}{m(GM)^2} + 1}}$$

Given Eq. 197 for  $a$ , plug in for  $E$ :

$$(200) \quad a = \frac{-GMm}{2E} = -\frac{GMm}{2\left(\frac{1}{2}mv_0^2 - \frac{GMm}{r}\right)} = \frac{-GM}{v_0^2 - \frac{2GM}{r}} = \frac{-GM}{v_0^2 \left(1 - \frac{2GM}{rv^2}\right)}$$

Notice when the denominator is 0. It is when  $v_0$  is the the escape velocity (cf. Eq. 196).

26.3.1. *Eccentricity from  $r_0$ ,  $\mathbf{v}_0$ .* Consider being given  $\mathbf{r}$ ,  $\mathbf{v}$ , with  $\mathbf{r}$  in a coordinate frame with its origin at the center of mass (or, if  $M$  large enough, from the focus where  $M$  is centered at).

The first step is to use a coordinate *frame* (set of orthonormal unit vectors) with  $\mathbf{r}$ ,  $\mathbf{v}$  forming a *plane*. Note this important fact from geometry: a plane is spanned by 2 non-collinear vectors. So  $\mathbf{r}$ ,  $\mathbf{v}$  better not be parallel to each other.

Let

$$\begin{aligned}
\mathbf{v} &= v^x \mathbf{e}_x + v^y \mathbf{e}_y + v^z \mathbf{e}_z = v^r \mathbf{e}_r + v^\varphi \mathbf{e}_\varphi \\
\mathbf{r} &= x \mathbf{e}_x + y \mathbf{e}_y + z \mathbf{e}_z = r \mathbf{e}_r
\end{aligned}$$

Now

$$(\mathbf{r} \times \mathbf{v})_i = \epsilon_{jki} r_j v_k$$

To get the vector *orthogonal* to the plane that is formed by  $\mathbf{r}$ ,  $\mathbf{v}$ ,

$$\begin{aligned}
((\mathbf{r} \times \mathbf{v}) \times \mathbf{r})_m &= \epsilon_{ijm} \epsilon_{jki} r_j v_k r_l = \epsilon_{ilm} \epsilon_{ijk} r_j v_k r_l = \\
&= r^2 v_m - v_k r_k r_m = r^2 v_m - v_k r_k r_m
\end{aligned}$$



Then

$$\widehat{L}_0 = r^2 \dot{\theta} = r \left( \mathbf{v} \cdot \frac{((\mathbf{r} \times \mathbf{v}) \times \mathbf{r})}{|(\mathbf{r} \times \mathbf{v}) \times \mathbf{r}|} \right) = r \left( \frac{r^2 v^2 - (\mathbf{v} \cdot \mathbf{r})^2}{|r^2 \mathbf{v} - (\mathbf{v} \cdot \mathbf{r}) \mathbf{r}|} \right)$$

Thus, we can find the following:

$$e^2 = 1 + \frac{2E\widehat{L}_0^2}{m(GM)^2} = 1 + \frac{2\left(\frac{1}{2}mv_0^2 - \frac{GMm}{r}\right)\widehat{L}_0^2}{m(GM)^2} = 1 + \frac{(v_0^2 - \frac{2GMm}{r})\widehat{L}_0^2}{(GM)^2}$$

$$a = \frac{-GMm}{2E} = -\frac{GMm}{2\left(\frac{1}{2}mv_0^2 - \frac{GMm}{r}\right)} = \frac{-GM}{v_0^2 - \frac{2GM}{r}} = \frac{-GM}{v_0^2\left(1 - \frac{2GM}{rv_0^2}\right)}$$

TODO: Calculate period of revolution.

Here are some calculations (forms) that can help with further calculations:

$$(201) \quad \widehat{L}_0^2 = b^2 \left( \frac{2E}{-m} \right) \text{ or } \frac{\widehat{L}_0^2}{GM} = b^2 \left( \frac{2E}{-GMm} \right) = \frac{b^2}{a} = a(1 - e^2)$$

using Eq. [176](#), for  $e^2 = 1 - \left(\frac{b}{a}\right)^2$ .

**26.4. Path in coordinates.** For  $\mathbf{r}$  centered at focus  $F$  and  $\mathbf{r}, \mathbf{v}$  expressed in coordinates in the plane of the conic section path, i.e. in the plane formed by  $\mathbf{r}, \mathbf{v}$ , with the axes expressed as  $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ ,

$$(202) \quad \mathbf{r} = r \cos \theta \frac{\partial}{\partial x} + r \sin \theta \frac{\partial}{\partial y}$$

Compare this to Eq. (2.5-1) of Sec. 2.5.1 "Expressing  $r$  and  $v$  in the Perifocal System" of Bate, Mueller, White (1971) [\[46\]](#).

Recall the equation for the conic section in polar coordinates, with the coordinates centered at focus  $F$ :

$$(203) \quad r = \frac{a(1 - e^2)}{1 + e \cos \theta}$$

Compare this to Eq. (1.5-4) of Bate, Mueller, White (1971) [\[46\]](#).

Now

$$(204) \quad \mathbf{v} = \dot{\mathbf{r}} = (\dot{r} \cos \theta - r \dot{\theta} \sin \theta) \frac{\partial}{\partial x} + (\dot{r} \sin \theta + r \dot{\theta} \cos \theta) \frac{\partial}{\partial y}$$

Compare this to the equation on pp. 72, of Sec. 2.5.1 "Expressing  $r$  and  $v$  in the Perifocal System" of Bate, Mueller, White (1971) [\[46\]](#).

Now

$$(205) \quad \dot{r} = \frac{a(1 - e^2)}{(1 + e \cos \theta)^2} e \sin \theta \dot{\theta} = \frac{r^2}{a(1 - e^2)} e \sin \theta \frac{\widehat{L}_0}{r^2} = \frac{\widehat{L}_0 e \sin \theta}{a(1 - e^2)} = \frac{\widehat{L}_0 e \sin \theta}{\frac{-GMm}{2E} \left( \frac{L_0^2}{\left( \frac{(GM)^2 m}{-2E} \right)} \right)} =$$

$$\implies \dot{r} = \frac{GM}{\widehat{L}_0} e \sin \theta$$

Compare this to Eq. (2.5-2) of Sec. 2.5.1 "Expressing  $r$  and  $v$  in the Perifocal System" of Bate, Mueller, White (1971) [\[46\]](#).

Then

$$(206) \quad \implies \frac{a(1 - e^2)}{(1 + e \cos \theta)^2} \dot{\theta} = \frac{GM}{\widehat{L}_0} \text{ or } r \dot{\theta} = \frac{GM}{\widehat{L}_0} (1 + e \cos \theta)$$

Compare this to Eq. (2.5-3) on pp. 73 of Sec. 2.5.1 "Expressing  $r$  and  $v$  in the Perifocal System" of Bate, Mueller, White (1971) [\[46\]](#).

Then

$$\mathbf{v} = \frac{GM}{\widehat{L}_0} \left[ (e \sin \theta \cos \theta - (1 + e \cos \theta) \sin \theta) \frac{\partial}{\partial x} + (e \sin^2 \theta + (1 + e \cos \theta) \cos \theta) \frac{\partial}{\partial y} \right] =$$

$$(207) \quad \mathbf{v} = \frac{GM}{\widehat{L}_0} \left[ \sin \theta \frac{\partial}{\partial x} + (e + \cos \theta) \frac{\partial}{\partial y} \right] = \frac{GM}{\widehat{L}_0} \left[ \frac{y}{r} \frac{\partial}{\partial x} + \left( e + \frac{x}{r} \right) \frac{\partial}{\partial y} \right]$$

Compare this to Eq. (2.5-4) on pp. 73 of Sec. 2.5.1 "Expressing  $r$  and  $v$  in the Perifocal System" of Bate, Mueller, White (1971) [\[46\]](#).

Consider this ansatz from subsection 4.4.3 "The  $f$  and  $g$  expressions from Ch. 4 section on "Position and velocity - A function of time" in Bate, Mueller, White (1971) [\[46\]](#).

Because  $\mathbf{r}, \mathbf{v}$  lie in the plane (i.e. coplanar) containing the entire conic section trajectory, then make the following ansatz:

$$(208) \quad \begin{aligned} \mathbf{r} &= f \mathbf{r}_0 + g \mathbf{v}_0 \\ \dot{\mathbf{r}} &=: \mathbf{v} = \dot{f} \mathbf{r}_0 + \dot{g} \mathbf{v}_0 \end{aligned}$$

Solve for  $f$  and  $g$  by taking the appropriate cross products:

$$(209) \quad \begin{aligned} \mathbf{r} \times \mathbf{v}_0 &= f \mathbf{r}_0 \times \mathbf{v}_0 \\ \mathbf{r} \times \mathbf{r}_0 &= g \mathbf{v}_0 \times \mathbf{r} \end{aligned}$$

**26.5. Eccentricity anomaly; Eccentricity anomaly coordinates.** Let the point of interest be P. Choose the origin for the coordinates  $(x, y)$  to have *its origin* at the *center of the ellipse*.

Let point C be the point at the center of the ellipse (origin), let point A be the projection of point P onto the  $x$ -axis, point F be the point at the focus.

Recall that the length from the center C to the focus,  $\overline{CF}$ , is  $ae$ , where  $a$  is the semi-major axis, and eccentricity is  $e$ .

The eccentricity anomaly  $E$  (it's just an angle) is given by

$$(210) \quad \cos E = \frac{x}{a}$$

Remember  $E$  is *not* the angle formed by triangle CAP, but rather, first take the line through P and A. This line is perpendicular to the  $x$ -axis. Then intersect that line with a circle, centered at C, of radius  $a$ . Then, the line from C to that intersection forms an angle  $E$  with the  $x$ -axis.

By using the ellipse equation and trigonometry,

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \implies \cos^2 E + \frac{y^2}{b^2} = 1 \text{ or } \frac{y^2}{b^2} = \sin^2 E \text{ or}$$

$$(211) \quad \sin E = \frac{y}{b}$$

Now

$$\overline{AF} = \overline{CF} - \overline{CA} = ae - a \cos E$$

Use the right triangle PAF and the hypotenuse being  $r$  (distance from the focus F to point P) to find

$$\overline{AF} = r \cos(\pi - \theta) = -r \cos \theta$$

So

$$\cos \theta = \frac{a(\cos E - e)}{r}$$

Then, plug into the conic section equation, Eq. 172,

$$\begin{aligned} r &= \frac{ed}{1+e\cos\theta} = \frac{a(1-e^2)}{1+e\cos\theta} = \frac{b^2/a}{1+e\cos\theta} = \frac{b^2/a}{1+e\left(\frac{a(\cos E - e)}{r}\right)} \\ &\implies 1 = \frac{b^2/a}{r+e(a(\cos E - e))} \\ \implies r &= \frac{b^2}{a} - e(a(\cos E - e)) = \frac{b^2}{a} + ae^2 - ea\cos E = a(1 - e\cos E) \end{aligned}$$

(212)  $r = a(1 - e\cos E)$

Compare this to Eq. (4.6-9) on pp. 215 of Sec. 4.6 in "The Kepler Problem" of Bate, Mueller, White (1971) [46].

For the coordinates with the origin at the *focus*  $F$ , it is given by

$$\begin{aligned} x &= a(\cos E - e) \\ y &= b\sin E = a\sqrt{1-e^2}\sin E \end{aligned}$$

(213)

Compare this to Eqns. (4.6-7), (4.6-8) on pp. 215 of Sec. 4.6 in "The Kepler Problem" of Bate, Mueller, White (1971) [46].

Indeed, for  $E = 0$ ,  $x = a - ae$  (the distance from the focus  $F$  to the vertex is  $a - ae$ ),  $E = \pi/2$ ,  $x = -ae$  (the distance from the focus  $F$  and back to the center  $C$ ) and for  $E = \pi$ ,  $x = -a - ae$  (the distance to the other vertex).

26.5.1. *Path in terms of eccentric anomaly.* Take the time derivative of Eq. 213:

$$\begin{aligned} \dot{x} &= -a\sin E\dot{E} \\ \dot{y} &= a\sqrt{1-e^2}\cos E\dot{E} \end{aligned}$$

(214)

To determine  $\dot{E}$ , consider both expressions for the radial distance  $r$  for Eq. 212 and Eq. 172:

$$\begin{aligned} r &= \frac{a(1-e^2)}{1+e\cos\theta} = \frac{b}{1+e\cos\theta} \\ \implies \dot{r} &= \frac{-b}{(1+e\cos\theta)^2}(-e\sin\theta)\dot{\theta} = \frac{be\sin\theta\dot{\theta}}{(1+e\cos\theta)^2} = \\ &\dot{r} = \frac{r^2e}{b}\sin\theta\dot{\theta} \\ r &= a(1-e\cos E) \\ \implies \dot{r} &= +ae\sin E\dot{E} \\ \implies ae\sin E\dot{E} &= \frac{r^2e}{b}\sin\theta\dot{\theta} \text{ or } a\sin E\dot{E} = \frac{r^2}{b}\sin\theta\dot{\theta} = \frac{\hat{L}_0\sin\theta}{b} = \frac{\hat{L}_0}{b}\frac{y}{r} \\ \implies \dot{E} &= \frac{\hat{L}_0}{ab}\frac{y}{r\sin E} = \frac{\hat{L}_0}{ar} = \frac{b}{a^{3/2}r} \end{aligned}$$

Compare the expression for  $\dot{E}$  with Eq. (4.6-11) on pp. 216 of Bate, Mueller, White (1971) [46].

Armed with an expression for  $\dot{E}$ , plug this back into the expression for  $\dot{r}$ :

$$\dot{r} = ae\sin E \left( \frac{\hat{L}_0}{ar} \right) = \frac{e\sin E\hat{L}_0}{r}$$

(215)

Compare this to the equation on pp. 216, "Position and Velocity - a function of time" in Ch. 4 of Bate, Mueller, White (1971) [46].

Use the expressions above in Eqns. 209. For  $\mathbf{r}, \mathbf{r}_0$  centered at focus  $F$ :

$$\begin{aligned} xv_{0,y} - yv_{0,x} &= a(\cos E - e)b\cos E_0\dot{E}_0 - b\sin E(-a\sin E_0\dot{E}_0) = \\ &= ab\dot{E}_0[(\cos E - e)\cos E_0 + \sin E\sin E_0] = \frac{b^2}{\sqrt{ar}}[\cos(E - E_0) - e\cos E_0] = \frac{\hat{L}^2\left(\frac{-m}{2E}\right)}{\sqrt{GM}\sqrt{\frac{m}{-2E}}r}[\cos(E - E_0) - e\cos E_0] = \\ &= \frac{\hat{L}_0b}{r\sqrt{GM}}[\cos(E - E_0) - e\cos E_0] \end{aligned}$$

Comparing this against the "right-hand side (RHS)" with  $f(\mathbf{r}_0 \times \mathbf{v}_0)$ , the RHS has a magnitude of  $f\hat{L}_0$  in the same direction as  $(\mathbf{r} \times \mathbf{v}_0)$ . Thus

$$f = \frac{b}{r\sqrt{GM}}[\cos(E - E_0) - e\cos E_0]$$

(216)

For the expression involving  $g$  in Eqns. 209,

$$\mathbf{r} \times \mathbf{r}_0 = xy_0 - yx_0 = a(\cos E - e)b\sin E_0 - b\sin Ea(\cos E_0 - e) = -ab[(\cos E_0)\sin E - (\cos E - e)\sin E_0]$$

with

$$ab = a\hat{L}_0\sqrt{\frac{-m}{2E}} = a\hat{L}_0\left(\frac{a}{GM}\right)^{1/2} = \left(\frac{a^3}{GM}\right)^{1/2}\hat{L}_0$$

For the RHS with  $g\mathbf{v}_0 \times \mathbf{r}_0$ ,  $\mathbf{v}_0 \times \mathbf{r}_0$ , by the right-hand rule or the right-hand orientation, it is in the *opposite* direction as  $\mathbf{r} \times \mathbf{r}_0$ . Thus, apply a negative side on the RHS.

$$\implies g = \left(\frac{a^3}{GM}\right)^{1/2}[(\cos E_0 - e)\sin E - (\cos E - e)\sin E_0]$$

(217)

## Part 15. Quantum Field Theory

### Part 16. The Standard Model

gauged symmetry group

$$\begin{array}{ccc} SU_c(3) \times & & SU_i(2) \times U_1(1) \\ \downarrow & & \downarrow \downarrow \\ 8G_\mu^\alpha & & 3W_\mu^a B_\mu \\ \alpha = 1 \dots 8 & a = 1, 2, 3 & \end{array}$$

8 spin-1 particles  $G_\mu^\alpha(x) \equiv$  gluons,  $c$  is meant to denote "color"

Gluons are strong to be massless,

strong interaction

3 spin-1 particles  $W_\mu^a(x)$  associated with  $SU_L(2)$ ,  $L$  indicates that only left-handed fermions turn out to carry this quantum number.

1  $B_\mu(x)$  with factor  $U_Y(1)$ , subscript  $Y$  to distinguish group associated with the quantum number (defined below) of weak hypercharge, from group associated with ordinary electric charge,  $Q \equiv U_{em}(1)$ .

4 spin-1 bosons associated with factors  $SU_L(2) = U_Y(1)$  are related to physical bosons that mediate weak interactions,  $W^\pm$  and  $Z^0$ , and familiar photon  $\gamma$  from QED.

## Part 17. General Relativity

$$\hbar = c = 1$$

Classically, entropy is a measure of the phase space available to the system in question. Then for  $S/E$ , it should have an upper bound, since,

let its energy be no more than  $E$ : this amounts to a limitation on momentum space available to system’s components, provided potential energy bounded from below. Bekenstein (1981) [40]

From [wikipedia](#), ”**Bekenstein bound**”

Suppose we have a black hole of mass  $M$ ,  
box of energy  $E$  outside of it of size  $R$   
radius of the black hole  $R_{\text{bh}}$  goes as  $GM$   
entropy of black hole goes as  $\frac{R_{\text{bh}}^2}{G} \equiv \frac{kc^3 R_{\text{bh}}^2}{\hbar G} \sim GM^2$

If we dump the box into the black hole, mass of black hole goes up to  $M + E$

entropy goes up by  $\frac{kGME}{\hbar c^3} \equiv GME$   
since entropy can’t decrease  $S \lesssim GME$   
In order for the box to fit inside black hole  $R \lesssim GM$   
If sizes are comparable, then  $S \lesssim \frac{kRE}{\hbar c} \equiv RE$

26.5.2. *The case of a weakly gravitating system falling into a black hole.* cf. Bekenstein (2018) [41]

Consider a weakly self-gravitating system  $\mathcal{U}$  of proper mass energy  $E$ , entropy  $\mathcal{S}$ .  
Assume spherical box of radius  $R$  enclosing  $\mathcal{U}$ , concentric to its center of mass of radius  $R$ .  
Box dropped into Schwarzschild black hold of mass  $M$  satisfying  $M \gg E$ ,  $2M \gg R$ .  
Assume black hole absorbs  $\mathcal{U}$ , and increases its mass by  $E$ .

Black hole’s entropy, originally  $S_{BH} = \frac{4\pi M^2}{l_p^2}$  increases by  $\Delta S_{BH} = \frac{8\pi ME}{l_p^2} + O(E^2)$

By GSL (generalized Second Law),  $\Delta S_{BH}$ + radiation entropies are at least as large as  $\mathcal{S}$

$$\implies S \leq \frac{8\pi ME}{l_p^2}$$

Choose  $M = \frac{1}{8}R^2/E$

$$S \leq \frac{\pi R^2}{l_p^2}$$

Claim: justify neglect of energy losses to Hawking’s radiance (small)  
approximate power emitted by Schwarzschild black hole by sphere with radius  $2M$  radiating according to Stefan-Boltzmann law at Hawking temperature  $T_H = \frac{l_p^2}{8\pi M}$

$$(218) \quad \dot{M} \equiv \frac{dM}{dt} \approx \frac{-Nl_p^3}{15360\pi M^2}$$

where  $N \equiv$  effective number of radiated species (photons contribute unity to  $N$ ).  
If  $\mathcal{U}$  dropped at Schwarzschild time  $t = 0$ , when its CM (center of mass) was at Schwarzschild radial coordinate  $r = \alpha 2M$  ( $\alpha \gg 1$ ), then bottom of sphere reaches horizon  $r = 2M$ , and infall practically finished, at time

$$(219) \quad t \approx 4M \left[ \frac{1}{3}\alpha^{3/2} - \frac{1}{4} \ln \left( \frac{R}{4M} \right) + O(\alpha^{1/2}) \right]$$

## Part 18. WE Heraeus International Winter School on Gravity and Light

INTRODUCTION (FROM EY)

The International Winter School on Gravity and Light held *central lectures* given by Dr. Frederic P. Schuller. These lectures on General Relativity and Gravity are unequivocally and undeniably, the best and most lucid and well-constructed lecture series on General Relativity and Gravity. The mathematical foundation from topology and differential geometry from which General Relativity arises from is solid, well-selected in rigor. The lectures themselves are well-thought out and clearly explained.

Even more so, the International Winter School provided accompanying Tutorial Sessions for each of the lectures. I had given up hopes in seeing this component of the learning process ever be put online so that anyone and everyone in the world could learn through the Tutorial process as well. I was afraid that nobody would understand how the Tutorial or “Office Hours” session was important for students to digest and comprehend and work out-doing exercises-the material presented in the lectures. This International Winter School gets it and shows how online education has to be done, to do it in an excellent manner, moving forward.

For anyone who is serious about learning General Relativity and Gravity, I would simply point to these video lectures and tutorials.

What I want to do is to build upon the material presented in this International Winter School. Why it’s important to me, and to the students and practicing researchers out there, is that the material presented takes the student from an introduction to the research frontier. That is the stated goal of the International Winter School. I want to dig into and help contribute to the cutting edge in research and this entire program with lectures and tutorials appears to be the most direct and sensible route directly to being able to do research in General Relativity and Gravity. -EY 20150323

20200321 update: Someone collected all the lecture videos and tutorials:

<http://mathswithphysics.blogspot.com/2016/07/the-we-heraeus-international-winter.html>

## 27. LECTURE 1: TOPOLOGY

### 27.1. Lecture 1: Topological Spaces.

**Definition 66.** *Let  $M$  be a set.*

A **topology**  $\mathcal{O}$  is a subset  $\mathcal{O} \subseteq \mathcal{P}(M)$ ,  $\mathcal{P}(M)$  power set of  $M$ : set of all subsets of  $M$ . satisfying

- (i)  $\emptyset \in \mathcal{O}$ ,  $M \in \mathcal{O}$
- (ii)  $U \in \mathcal{O}$ ,  $V \in \mathcal{O} \implies U \cap V \in \mathcal{O}$
- (iii)  $U_\alpha \in \mathcal{O}$ ,  $\alpha \in \mathcal{A} \implies (\bigcup_{\alpha \in \mathcal{A}} U_\alpha) \in \mathcal{O}$

$\mathcal{O}$  } utterly useless

**Definition 67.**  $\mathcal{O}_{\text{standard}} \subseteq \mathcal{P}(\mathbb{R}^d)$

EY : 20150524

I’ll fill in the proof that  $\mathcal{O}_{\text{standard}}$  is a topology.

*Proof.*  $\emptyset \in \mathcal{O}_{\text{standard}}$   
since  $\forall p \in \emptyset$ ,  $\exists r \in \mathbb{R}^+$ :  $\mathcal{B}_r(p) \subseteq \emptyset$  (i.e. satisfied “vacuously”)

Suppose  $U, V \in \mathcal{O}_{\text{standard}}$ .

Let  $p \in U \cap V$ . Then  $\exists r_1, r_2 \in \mathbb{R}^+$  s.t.  $\mathcal{B}_{r_1}(p) \subseteq U$   
 $\mathcal{B}_{r_2}(p) \subseteq V$

Let  $r = \min \{r_1, r_2\}$ .  
Clearly  $\mathcal{B}_r(p) \subseteq U$  and  $\mathcal{B}_r(p) \subseteq V$ . Then  $\mathcal{B}_r(p) \subseteq U \cap V$ . So  $U \cap V \in \mathcal{O}_{\text{standard}}$ .

Suppose,  $U_\alpha \in \mathcal{O}_{\text{standard}}, \forall \alpha \in \mathcal{A}$ .  
Let  $p \in \bigcup_{\alpha \in \mathcal{A}} U_\alpha$ . Then  $p \in U_\alpha$  for at least 1  $\alpha \in \mathcal{A}$ .  
 $\exists r_\alpha \in \mathbb{R}^+$  s.t.  $\mathcal{B}_{r_\alpha}(p) \subseteq U_\alpha \subseteq \bigcup_{\alpha \in \mathcal{A}} U_\alpha$ . So  $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{O}_{\text{standard}}$

27.2. **2. Continuous maps.**

27.3. **3. Composition of continuous maps.**

27.4. **4. Inheriting a topology.** EY : 20150524

I'll fill in the proof that given  $f$  continuous (cont.), then the restriction of  $f$  onto a subspace  $S$  is cont. If you want a reference, check out Klaus Jänich [?, pp. 13, Ch. 1 Fundamental Concepts, Sec. Continuous Maps]

If cont.  $f : M \rightarrow N, S \subseteq M$ , then  $f|_S$  cont.

*Proof.* Let open  $V \subseteq N$ , i.e.  $V \in \mathcal{O}_N$  i.e.  $V$  in the topology  $\mathcal{O}_N$  of  $N$ .

$$f|_S^{-1}(V) = \{m \in M \mid f|_S(m) \in V\}$$

Now  $f^{-1}(V) = \{m \in M \mid f(m) \in V\}$ .

So  $f^{-1}(V) \cap S = f|_S^{-1}(V)$

Now  $f$  cont. So  $f^{-1}(V) \in \mathcal{O}_N$ .

and recall  $\mathcal{O}_S := \{U \cap S \mid U \in \mathcal{O}_M\}$ .

so  $f^{-1}(V) \cap S = f|_S^{-1}(V) \in \mathcal{O}_S$  i.e.  $f|_S^{-1}(V)$  open.

$\implies f|_S$  cont.

□

TOPOLOGY TUTORIAL SHEET

filename : `main.pdf`

The WE-Heraeus International Winter School on Gravity and Light: Topology

EY : 20150524

What I won't do here is retype up the solutions presented in the Tutorial (cf. [https://youtu.be/\\_XkhZQ-hNLs](https://youtu.be/_XkhZQ-hNLs)): the presenter did a very good job. If someone wants to type up the solutions and copy and paste it onto this LaTeX file, in the spirit of open-source collaboration, I would encourage this effort.

Instead, what I want to encourage is the use of as much CAS (Computer Algebra System) and symbolic and numerical computation because, first, we're in the 21st century, second, to set the stage for further applications in research. I use Python and Sage Math alot, mostly because they are open-source software (OSS) and fun to use. Also note that the structure of Sage Math modules matches closely to Category Theory.

In checking whether a set is a topology, I found it strange that there wasn't already a function in Sage Math to check each of the axioms. So I wrote my own; see my code snippet, which you can copy, paste, edit freely in the spirit of OSS here, titled `topology.sage`:

[gist github ernestyalumni topology.sage](#)

[Download topology.sage](#)

Loading `topology.sage`, after changing into (with the usual Linux terminal commands, `cd, ls`) by

```
sage: load('topology.sage')
```

**Exercise 2: Topologies on a simple set.**

**Question** Does  $\mathcal{O}_1 := \dots$  constitute a topology ...?

**Solution:** Yes, since we check by typing in the following commands in Sage Math:

```
emptyset in O_1
Axiom2check(O_1) # True
Axiom3check(O_1) # True
```

**Question** What about  $\mathcal{O}_2 \dots$ ?

**Solution:** No since the 3rd. axiom fails, as can be checked by typing in the following commands in Sage Math:

```
emptyset in O_2
Axiom2check(O_2) # True
Axiom3check(O_2) # False
```

28. LECTURE 2: TOPOLOGICAL MANIFOLDS

**Lecture 2: Manifolds.** Topological spaces:  $\exists$  so man that mathematicians cannot even classify them.

For spacetime physics, we may focus on topological spaces  $(M, \mathcal{O})$  that can be charted, analogously to how the surface of the earth is charted in an atlas.

28.1. **Topological manifolds.**

**Definition 68.** A topological space  $(M, \mathcal{O})$  is called a *d-dimensional topological method* if

$$\forall p \in M : \exists U \in \mathcal{O}, U \ni p : \exists x : U \subseteq M \rightarrow x(U) \subseteq \mathbb{R}^d \quad (M, \mathcal{O}), (\mathbb{R}^d, \mathcal{O}_{std})$$

(i)  $x$  *invertible* :

$$x^{-1} : x(U) \rightarrow U$$

(ii)  $x$  *continuous*

(iii)  $x^{-1}$  *continuous*

28.2. **Terminology.**

28.3. **3. Chart transition maps.** Imagine 2 charts  $(U, x)$  and  $(V, y)$  with overlapping regions.

28.4. **4. Manifold philosophy.** Often it is desirable (or indeed the way) to define properties (“continuity”) of real-world object (“ $\mathbb{R} \xrightarrow{\gamma} M$ ”) by judging suitable coordinates not on the “real-world” object itself, but on a chart-representation of that real world object.

**EY's add-ons.** This lecture gives me a good excuse to review Topology and Topological Manifolds from a mathematician's point of view. I find John M. Lee's **Introduction to Topological Manifolds** book good because it's elementary and thorough and it's fairly recent (2010) so it's up to date [?]. See my notes and solutions for the book; it's a file titled `LeeJM_IntroTopManifolds_sol.pdf` of which I'll try to keep the pdf and LaTeX file available for download on my ernestyalumni Google Drive (so try to search for it on Google).

filename: Sheet\_1.2.pdf

Exercise 4: Before the invention of the wheel.

Another one-dimensional topological manifold. Another one?

Consider set  $F^1 := \{(m, n) \in \mathbb{R}^2 | m^4 + n^4 = 1\}$ , equipped with subset topology  $\mathcal{O}_{\text{std}}|_{F^1}$

Question  $x : F^1 \rightarrow \mathbb{R}$  is what?.

Solution . EY : 20150525 The tutorial video [https://youtu.be/ghfEQ3u\\_B6g](https://youtu.be/ghfEQ3u_B6g) is really good and this solution is how I'd write it, but it's really the same (I needed the practice).

$$\begin{array}{l} x : F^1 \rightarrow \mathbb{R} \\ (m, n) \mapsto m \end{array}$$

If  $m = 0$ ,  $n^4 = 1$  so  $n = \pm 1$  so it's not injective.

Let the closed  $n$ -dim. upper half-space  $\mathbb{H}^n \subseteq \mathbb{R}^1$ . Then

$$\begin{aligned} \mathbb{H}^n &= \{(x_1 \dots x_n) \in \mathbb{R}^n | x_n \geq 0\} \\ \text{int}\mathbb{H}^n &= \{(x_1 \dots x_n) \in \mathbb{R}^n | x_n > 0\} \\ -\mathbb{H}^n &= \{(x_1 \dots x_n) \in \mathbb{R}^n | x_n \leq 0\} \\ -\text{int}\mathbb{H}^n &= \{(x_1 \dots x_n) \in \mathbb{R}^n | x_n < 0\} \end{aligned}$$

Question This map  $x$  may be made injective by restricting its domain to either of 2 maximal open subsets of  $F^1$ . Which ones?.

Solution .

Let

$$\begin{aligned} U_+ &= F^1 \cap \text{int}\mathbb{H}^2 \\ U_- &= F^1 \cap -\text{int}\mathbb{H}^2 \end{aligned}$$

Look at

$$\begin{aligned} x^4 &= 1 - n^4 \\ \implies x &= \pm(1 - n^4)^{1/4} \end{aligned}$$

Then for

$$\begin{aligned} x_+^{-1} : (-1, 1) &\subseteq \mathbb{R} \rightarrow U_+ \\ m &\mapsto (m, (1 - m^4)^{1/4}) \\ x_-^{-1} : (-1, 1) &\subseteq \mathbb{R} \rightarrow U_- \\ m &\mapsto (m, -(1 - m^4)^{1/4}) \end{aligned}$$

$x_+, x_-$  injective (since left inverse exists).

Question Construct injective  $y$ .

Solution .

Let

$$\begin{aligned} V_+ &= F^1 \cap \text{int}\mathbb{H}^1 \\ V_- &= F^1 \cap -\text{int}\mathbb{H}^1 \end{aligned}$$

Then

$$\begin{aligned} y_+ : V_+ &\rightarrow (-1, 1) \subseteq \mathbb{R} \\ (m, n) &\mapsto n \\ y_- : V_- &\rightarrow (-1, 1) \subseteq \mathbb{R} \\ (m, n) &\mapsto n \end{aligned}$$

Question Construct inverse  $y^{-1}$ . Solution .

For

$$\begin{aligned} y_+^{-1} : (-1, 1) &\subseteq \mathbb{R} \rightarrow V_+ \\ n &\mapsto ((1 - n^4)^{1/4}, n) \\ y_-^{-1} : (-1, 1) &\subseteq \mathbb{R} \rightarrow V_- \\ n &\mapsto (-(1 - n^4)^{1/4}, n) \end{aligned}$$

$y_+, y_-$  injective (since left inverse exists).

Note  $(-1, 0) \notin U_+, U_-$

$$(1, 0) \notin U_+, U_-$$

and

$$(0, 1) \notin V_+, V_-$$

$$(0, -1) \notin V_+, V_-$$

Question construct *transition map*  $x \circ y^{-1}$ .

Solution .

$$\begin{aligned} x_+ y_+^{-1} : (0, 1) &\subseteq \mathbb{R} \rightarrow (0, 1) \subseteq \mathbb{R} \\ n &\mapsto (1 - n^4)^{1/4} \\ x_- y_+^{-1} : (-1, 0) &\subseteq \mathbb{R} \rightarrow (0, 1) \subseteq \mathbb{R} \\ n &\xrightarrow{y_+^{-1}} ((1 - n^4)^{1/4}, n) \xrightarrow{x_-} (1 - n^4)^{1/4} \\ x_+ y_-^{-1} : (0, 1) &\subseteq \mathbb{R} \rightarrow (-1, 0) \subseteq \mathbb{R} \\ n &\mapsto -(1 - n^4)^{1/4} \\ x_- y_-^{-1} : (-1, 0) &\subseteq \mathbb{R} \rightarrow (-1, 0) \subseteq \mathbb{R} \\ n &\mapsto -(1 - n^4)^{1/4} \end{aligned}$$

Question ... Does the collection of these domains and maps form an atlas of  $F^1$ ?.

Yes, with atlas

$$\mathcal{A} = \left\{ \begin{array}{cc} (U_+, x_+) & (V_+, y_+) \\ (U_-, x_-) & (V_-, y_-) \end{array} \right\}$$

Clearly

$$\begin{aligned} U_+ \cup U_- \cup V_+ \cup V_- &= (F^1 \cap \text{int}\mathbb{H}^2) \cup (F^1 \cap -\text{int}\mathbb{H}^2) \cup (F^1 \cap \text{int}\mathbb{H}^1) \cup (F^1 \cap -\text{int}\mathbb{H}^1) = \\ &= F^1 \cap \mathbb{R}^2 \setminus \{(0,0)\} = F^1 \end{aligned}$$

and (the point is that)  $x_{\pm}, y_{\pm}$  are homeomorphisms of open sets of  $F^1$  onto open sets of 1 dim.  $\mathbb{R}^1$  (namely  $(-1,1) \subseteq \mathbb{R}^1$ ), and so  $\mathcal{A}$  is an atlas of  $F^1$ .

29. LECTURE 3: MULTILINEAR ALGEBRA

Lecture 3: Multilinear Algebra (International Winter School on Gravity and Light 2015)

We will **not** equip space(time) with a vector space structure. Do you know where

$$\begin{array}{ccc} 5 \cdot \text{Paris} & =? & \\ \text{lie ?} & & \text{Paris} + \text{Vienna} =? \end{array}$$

Moreover, the tangent spaces  $T_pM$  (lecture 5) smooth manifolds (Lecture 4)

Beneficial to first study vector spaces abstractly for two reason

- (i) for construction of  $T_pM$  one needs an intermediate vector space  $C^\infty(M)$
- (ii) tensor technique are most easily understood in an abstract setting.

29.1. Vector spaces.

**Definition 69.** A vector space  $(V, +, -)$  is

- (i) a set  $V$
- (ii)  $+: V \times V \rightarrow V$  “addition”
- (iii)  $\cdot: \mathbb{R} \times V \rightarrow V$  “s-multiplication” EY : 20160317 s for “scalar”

satisfying:

CANIADDU

$$\begin{array}{ll} C^+ : & v + w = w + v \\ A^+ : & (u + v) + w = u + (v + w) \\ N^+ : & \exists 0 \in V : \forall v \in V : v + 0 = v \\ I^+ : & \forall v \in V : \exists (-v) \in V : v + (-v) = 0 \\ A : & \lambda \cdot (\mu + v) = (\lambda \cdot \mu) \cdot v \quad (\forall \lambda, \mu \in \mathbb{R}) \\ D : & (\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v \\ D : & \lambda \cdot v + \lambda \cdot w = \lambda \cdot (v + w) \\ U : & 1 \cdot v = v \end{array}$$

Terminology. An element of a vector space is often referred to, informally as a vector.

Example. def. **set** of polynomials (fixed) degree  $\mathcal{P} := \{p: (-1,+1) \rightarrow \mathbb{R} | p(x) = \sum_{n=0}^N p_n \cdot x^n\}$

Thought bubble: is  $\square$  a vector?

$$\square(x) = x^2$$

No  $\square \in \mathcal{P}$ .

$$+: \mathcal{P} \times \mathcal{P} \rightarrow \mathcal{P}$$

$$(p,q) \mapsto p+q$$

where  $(p+q)(x) = p(x) +_{\mathbb{R}} q(x)$

$$\cdot: \mathbb{R} \times \mathcal{P} \rightarrow \mathcal{P}$$

$$(\lambda,p) \mapsto \lambda \cdot p$$

where  $(\lambda \cdot p)(x) := \lambda \cdot_{\mathbb{R}} p(x)$

Thought bubble:  $\square$  a vector?

Yes, but who cares?

$(\mathcal{P}, +, \cdot)$  is a vector space.

$$\square \in \mathcal{P}$$

29.2. Linear maps. These are the structure-respecting maps between vector spaces.

EY : 20160316 out of tradition, they’re called “linear” maps

**Definition 70.**  $(V, +_V, \cdot_V)$  and  $(W, +_W, \cdot_W)$  vector spaces

Then a map

$$\varphi: V \rightarrow W$$

is called **linear** if

- (i)  $\varphi(v+_V \tilde{v}) = \varphi(v) +_W \varphi(\tilde{v})$
- (ii)  $\varphi(\lambda \cdot_V v) = \lambda \cdot_W \varphi(v)$

$$\delta: \mathcal{P} \rightarrow \mathcal{P}$$

Example. :

$$p \mapsto \delta(p) := p'$$

linear:

- (i)  $\delta(p+q) = (p+_P q)'$   $\underbrace{\quad}_{\text{sum rule}} p' +_P q' = \delta(p) +_P \delta(q)$
- (ii)  $\delta(\lambda p) = (\lambda p)' = \lambda \cdot p' = \lambda \cdot \delta(p)$

Notation:  $\varphi: V \rightarrow W$  linear  $\iff \varphi: V \xrightarrow{\sim} W$

$$\begin{array}{ccccc} V & \xrightarrow[\sim]{\psi} & W & \xrightarrow[\sim]{\varphi} & U \\ & & \searrow \scriptstyle \sim & \nearrow \scriptstyle \sim & \\ & & \varphi \circ \psi & & \end{array}$$

29.2.1. Example\*.  $\delta \circ \delta: \mathcal{P} \xrightarrow{\sim} \mathcal{P}$

29.3. Vector space of Homomorphisms. fun fact:  $(V, +, \cdot)$   $(W, +, \cdot)$  vector spaces

def.  $\text{Hom}(V, W) := \{\varphi: V \xrightarrow{\sim} W\}$  set.

We can make this into a vector spaces.

$$\begin{array}{l} \oplus: \text{Hom}(V, W) \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W) \\ (\varphi, \psi) \mapsto \varphi \oplus \psi \end{array}$$

where  $(\varphi \otimes \psi)(v) := \varphi(v) +_W \psi(v)$

$\otimes: \dots$  similarly.

$(\text{Hom}(V, W), \oplus, \otimes)$  **is** a vector space.

29.3.1. Example\*.  $\text{Hom}(\mathcal{P}, \mathcal{P})$  is a vector space.

$$\delta \in \text{Hom}(\mathcal{P}, \mathcal{P})$$

$$\delta \circ \delta \in \text{Hom}(\mathcal{P}, \mathcal{P})$$

$\vdots$

$$\underbrace{\delta \circ \dots \circ \delta}_M \in \text{Hom}(\mathcal{P}, \mathcal{P})$$

$$\implies 5 \circ \delta \oplus_{\text{Hom}(\mathcal{P}, \mathcal{P})} \delta \circ \delta \in \text{Hom}(\mathcal{P}, \mathcal{P})$$



29.4. **Dual vector space.** heavily used special case:  
( $V, +, \cdot$ ) vector space:

**Definition 71.**

$$V^* := \{\varphi : V \xrightarrow{\sim} \mathbb{R}\} = Hom(V, \mathbb{R})$$

$\underbrace{(V^*, \oplus, \otimes)}_{\text{dual vector space (to } V)}$  is a vector space

Terminology:  $\varphi \in V^*$  is called, informally, a covector.

*Example.*  $I : \mathcal{P} \xrightarrow{\sim} \mathbb{R}$

i.e.  $I \in \mathcal{P}^*$

def.  $I(p) := \int_0^1 dx p(x)$

linear: 
$$I(p+q) = \int_0^1 dx \underbrace{(p+q)(x)}_{p(x)+q(x)} \\ = \dots = I(q) + I(p)$$

$$I(\lambda p) = \lambda \cdot I(p)$$

i.e.  $I = \int_0^1 dx$

29.5. **Tensors.**

**Definition 72.** Let  $(V, +, \cdot)$  be a vector space.

An  $\underline{(r, s)\text{-tensor } T \text{ over } V}$   $r, s \in \mathbb{N}_0$

is a multi-linear map

$$T : \underbrace{V^* \times \dots \times V^*}_r \times \underbrace{V \times \dots \times V}_s \xrightarrow[\sim]{\sim} \mathbb{R}$$

29.5.1. *Example.* :  $T$  (1,1)-tensor

$$T(\varphi + \psi, v) = T(\varphi, v) + T(\psi, v)$$

$$T(\varphi, v + w) = T(\varphi, v) + T(\varphi, w)$$

$$T(\lambda \varphi, v) = \lambda \cdot T(\varphi, v)$$

$$T(\varphi, \lambda \cdot v) = \lambda T(\varphi, v)$$

$$T(\varphi + \psi, v + w) = \\ = t(\varphi, v) + T(\varphi, w) + T(\psi, v) + t(\psi, w)$$

Excursion: Given  $T : V^* \times V \xrightarrow{\sim} \mathbb{R}$

$$\phi_T : V \xrightarrow{\sim} (V^*)^* \underset{\dim V < \infty}{=} V$$

Define  $v \mapsto \underbrace{T(\cdot, v)}_{V^* \xrightarrow{\sim} \mathbb{R}}$

Given  $\phi : V \xrightarrow{\sim} V$

Construct  $T_\phi : V^* \times V \xrightarrow{\sim} \mathbb{R}$

$$(\varphi, v) \mapsto \varphi(\phi(v))$$

$\implies$  given  $T : T = T_{\varphi_T}$

given  $\phi : \phi = \phi_{T_\phi}$

*Example.*  $g : P \times P \xrightarrow{\sim} \mathbb{R}$

$$(p, q) \mapsto \int_{-1}^1 dx p(x) q(x)$$

is a (0, 2)-tensor over  $P$ .

Info: If  $T \in \text{Hom}(V, W)$

29.6. **Vectors and covectors as tensors.**

**Theorem 26.** (including proof)

“covector”  $\varphi \in V^* \iff \varphi : V \xrightarrow{\sim} \mathbb{R} \iff \varphi(0, 1)\text{-tensor}.$

**Theorem 27.**  $v \in V \underset{\dim V < \infty}{=} (V^*)^* \iff v : V^* \xrightarrow{\sim} \mathbb{R} \iff v \text{ is } (1, 0)\text{-tensor}.$

29.7. **Bases.**

**Definition 73.**  $(V, +, \cdot)$  vector space.

A subset  $B \subset V$  is called

a basis if

*Thought bubble: Hamel (L.A.) EY : 20160316 Hamel basis, Linear Algebra*

$$\forall v \in V \quad \exists \text{ finite } \underbrace{F}_{\{f_1, \dots, f_n\}} \subset B : \exists ! \underbrace{v^1, v^2, \dots, v^n}_{\in \mathbb{R}}, \quad v = v^1 f_1 + \dots + v^n f_n$$

**Definition 74.** If  $\exists$  basis  $\mathcal{B}$  with finitely many elements, say  $d$  many, then we call  $d =: \dim V$

This is well-defined.

Remark:  $(V, +, \cdot)$  be a finite-dim. vector space.

Having chosen a basis  $e_1, \dots, e_n$  of  $(V, +, \cdot)$  we may uniquely associate

(Thought bubble: this requires a chosen basis)

$v \mapsto (v^1, \dots, v^n)$  called the components of  $v$  w.r.t. chosen basis

where:  $v^1 e_1 + \dots + v^n e_n = v$

29.8. **Basis for the dual space.** choose Basis  $e_1, \dots, e_n$  for  $V$

can choose Basis  $\epsilon^1, \dots, \epsilon^n$  for  $V^*$

However, more economical to require

once  $e_1, \dots, e_n$  on  $V$  has been chosen, that

$$\epsilon^a(e_b) = \delta_b^a$$

This uniquely determines choice of  $\epsilon^1, \dots, \epsilon^n$  from choice of  $e_1, \dots, e_n$

**Definition 75.** If a basis  $\epsilon^1, \dots, \epsilon^n$  of  $V^*$  satisfies this, it is called the **dual basis** (of the dual space)

Example:  $P$  ( $N = 3$ )

$$e_0(x) = 1$$

$$e_1(x) = x$$

$e_0, e_1, e_2, e_3$  basis if  $e_2(x) = x^2 \{e_a(x) := x^a$

$$e_3(x) = x^3$$

$$\epsilon^0, \epsilon^1, \epsilon^2, \epsilon^3 \text{ dual basis } \epsilon^a := \frac{1}{a!} \partial^a|_{x=0}$$

*Proof.*  $\epsilon^a(e_b) = \delta_b^a$

□

**29.9. Components of tensors.** Let  $T$  be an  $(r, s)$ -tensor on a finite-dim. vs.  $V$ . Then define the  $(r + s)^{\dim V}$  many real numbers.

$$\underbrace{T^{i_1 \dots i_r}_{j_1 \dots j_s}}_{\in \mathbb{R}} := T(\epsilon^{i_1}, \dots, \epsilon^{i_r}, e_{j_1}, e_{j_2}, \dots, e_{j_s})$$

$i_1 \dots i_r, j_1 \dots j_s \in \{1, \dots, \dim V\}$   
Thought bubble:  $\underbrace{T^{i_1 \dots i_r}_{j_1 \dots j_s}}_{\in \mathbb{R}}$  are the components of the tensor w.r.t. chosen basis

Useful: Knowing components (and basis) one can reconstruct the entire tensor.

*Example.*  $T(1, 1)$ - tensor

$$T^i_j := T(\epsilon^i, e_j)$$

reconstruct

$$\begin{aligned} T(\varphi, v) &= T\left(\sum_{i=1}^{\dim V} \varphi_i \epsilon^i, \sum_{j=1}^{\dim V} v^j e_j\right) \quad \varphi_i \in \mathbb{R}, v^j \in \mathbb{R} \\ &= \sum_{i=1}^{\dim V} \sum_{j=1}^{\dim V} \varphi_i v^j \underbrace{T(\epsilon^i, e_j)}_{T^i_j} \\ &=: \varphi_i v^j T^i_j \end{aligned}$$

30. LECTURE 4: DIFFERENTIABLE MANIFOLDS

so far: top. mfd.  $(M, \mathcal{O})$

$$\dim M = d$$

we wish to define a notion of differentiable

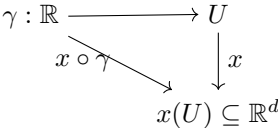
curves  $\mathbb{R} \rightarrow M$

function  $M \rightarrow \mathbb{R}$

maps  $M \rightarrow N$

**30.1. 1. Strategy.** choose a chart  $(U, x)$

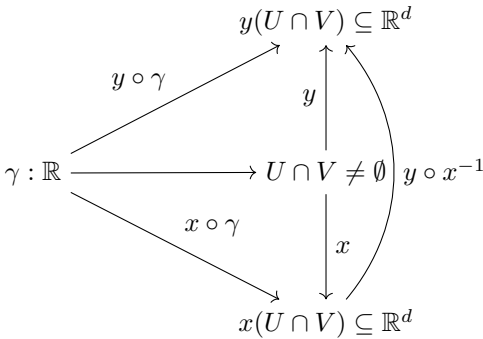
$\gamma : \mathbb{R} \rightarrow M$  portion of curve in chart domain



idea. try to “lift” the undergraduate notion of differentiability of a curve on  $\mathbb{R}^d$  to a notion of

differentiability of a curve on  $M$

Problem Can this be well-defined under change of chart?



$x \circ \gamma$  undergraduate differentiable (“as a map  $\mathbb{R} \rightarrow \mathbb{R}^d$ ”)

$$\underbrace{y \circ \gamma}_{\text{maybe only continuous, but not undergraduate differentiable}} = \underbrace{(y \circ x^{-1})}_{\text{continuous}} \circ \underbrace{(x \circ \gamma)}_{\text{undergrad differentiable}} = y \circ (x^{-1} \circ x) \circ \gamma$$

At first sight, strategy does not work out.

**30.2. 2. Compatible charts.** In section 1, we used any imaginable charts on the top. mfd.  $(M, \mathcal{O})$ . To emphasize this, we may say that we took  $U$  and  $V$  from the *maximal atlas*  $\mathcal{A}$  of  $(M, \mathcal{O})$ .

**Definition 76.** Two charts  $(U, x)$  and  $(V, y)$  of a top. mfd. are called  $\ast$ -compatible if either

- (a)  $U \cap V = \emptyset$  or
- (b)  $U \cap V \neq \emptyset$

chart transition maps have undergraduate  $\ast$  property.

*EY : 20151109 e.g. since  $\mathbb{R}^d \rightarrow \mathbb{R}^d$ , can use undergraduate  $\ast$  property such as continuity or differentiability.*

$$\begin{aligned} y \circ x^{-1} : x(U \cap V) \subseteq \mathbb{R}^d &\rightarrow y(U \cap V) \subseteq \mathbb{R}^d \\ x \circ y^{-1} : y(U \cap V) \subseteq \mathbb{R}^d &\rightarrow x(U \cap V) \subseteq \mathbb{R}^d \end{aligned}$$

Philosophy:

**Definition 77.** An atlas  $\mathcal{A}_\ast$  is a  $\ast$ -compatible atlas if any two charts in  $\mathcal{A}_\ast$  are  $\ast$ -compatible.

**Definition 78.** A  $\ast$ -manifold is a triple  $(\underbrace{M, \mathcal{O}}_{\text{top. mfd.}}, \mathcal{A}_\ast)$   $\mathcal{A}_\ast \subseteq \mathcal{A}_{\text{maximal}}$

$\ast$	undergraduate $\ast$	
$C^0$	$C^0(\mathbb{R}^d \rightarrow \mathbb{R}^d) =$	continuous maps w.r.t. $\mathcal{O}$
$C^1$	$C^1(\mathbb{R}^d \rightarrow \mathbb{R}^d) =$	differentiable (once) and is continuous
$C^k$		$k$ -times continuously differentiable
$D^k$		$k$ -times differentiable
$\vdots$		
$C^\infty$	$C^\infty(\mathbb{R}^d \rightarrow \mathbb{R}^d)$	
$\bigcup$		
$C^\omega$	$\exists$ multi-dim. Taylor exp.	
$\mathbb{C}^\infty$	satisfy Cauchy-Riemann equations, pair-wise	

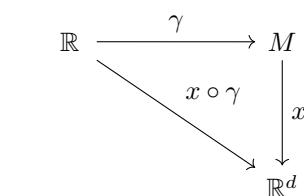
<sup>7</sup><http://mathoverflow.net/questions/8789/can-every-manifold-be-given-an-analytic-structure>

EY : 20151109 Schuller says:  $C^k$  is easy to work with because you can judge  $k$ -times cont. differentiability from existence of all partial derivatives **and** their continuity. There are examples of maps that partial derivatives exist but are not  $D^k$ ,  $k$ -times differentiable.

**Theorem 28** (Whitney). *Any  $C^{k \geq 1}$ -atlas,  $\mathcal{A}_{C^{k \geq 1}}$  of a topological manifold contains a  $C^\infty$ -atlas.*  
Thus we may w.l.o.g. always consider  $C^\infty$ -manifolds, “smooth manifolds”, unless we wish to define Taylor expandibility/-complex differentiability ...

EY : 20151109 Hassler Whitney <sup>7</sup>

**Definition 79.** A smooth manifold  $(\underbrace{M, \mathcal{O}}_{\text{top. mfd.}}, \underbrace{\mathcal{A}}_{C^\infty\text{-atlas}})$



EY: 20151109 Schuller was explaining that the trajectory is real in  $M$ ; the coordinate maps to obtain coordinates is  $x \circ \gamma$

30.3. 4. Diffeomorphisms.  $M \xrightarrow{\phi} N$

If  $M, N$  are naked sets, the structure preserving maps are the bijections (invertible maps).  
e.g.  $\{1, 2, 3\} \rightarrow \{a, b\}$

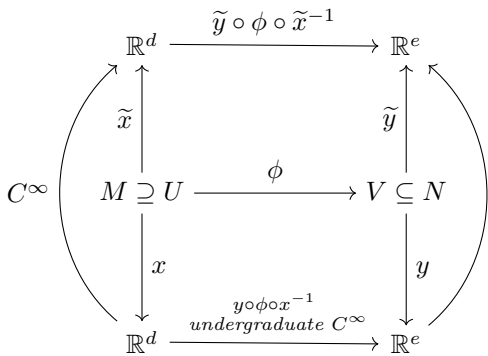
**Definition 80.**  $M \cong_{\text{set}} N$  (set-theoretically) isomorphic if  $\exists$  bijection  $\phi : M \rightarrow N$

Examples.  $\mathbb{N} \cong_{\text{set}} \mathbb{Z}$   
 $\mathbb{N} \cong_{\text{set}} \mathbb{Q}$  (EY: 20151109 Schuller says from diagonal counting)  
 $\mathbb{N} \not\cong_{\text{set}} \mathbb{R}$   
Now  $(M, \mathcal{O}_M) \cong_{\text{top}} (N, \mathcal{O}_N)$  (topl.) isomorphic = “homeomorphic”  $\exists$  bijection  $\phi : M \rightarrow N$   
 $\phi, \phi^{-1}$  are continuous.  
 $(V, +, \cdot) \cong_{\text{vec}} (W, +_w, \cdot_w)$  (EY: 20151109 vector space isomorphism) if  
 $\exists$  bijection  $\phi : V \rightarrow W$  linearly  
finally

**Definition 81.** Two  $C^\infty$ -manifolds  $(M, \mathcal{O}_M, \mathcal{A}_M)$  and  $(N, \mathcal{O}_N, \mathcal{A}_N)$  are said to be **diffeomorphic** if  $\exists$  bijection  $\phi : M \rightarrow N$  s.t.

$$\begin{aligned}\phi &: M \rightarrow N \\ \phi^{-1} &: N \rightarrow M\end{aligned}$$

are both  $C^\infty$ -maps



**Theorem 29.**  $\#$  = number of  $C^\infty$ -manifolds one can make out of a given  $C^0$ -manifolds (if any) - up to diffeomorphisms.

$\dim M$	$\#$	
1	1	Morse-Radon theorems
2	1	Morse-Radon theorems
3	1	Morse-Radon theorems
4	uncountably infinitely many	
5	finite	surgery theory
6	finite	surgery theory
$\vdots$	finite	surgery theory

EY : 20151109 cf. <http://math.stackexchange.com/questions/833766/closed-4-manifolds-with-uncountably-many-differen>  
The wild world of 4-manifolds

TUTORIAL 4 DIFFERENTIABLE MANIFOLDS

EY : 20151109 The [gravity-and-light.org](http://gravity-and-light.org) website, where you can download the tutorial sheets *and* the full length videos for the tutorials and lectures, are no longer there. = (  
Hopefully, the YouTube video will remain: [https://youtu.be/FXPdKxOq1KA?list=PLFeEvEPtX\\_ORQ1ys-7VIsKlBWz7RX-FaL](https://youtu.be/FXPdKxOq1KA?list=PLFeEvEPtX_ORQ1ys-7VIsKlBWz7RX-FaL)

**Exercise 1: True or false?.** These basic questions are designed to spark discussion and as a self-test.

Tick the correct statements, but not the incorrect ones!

- (a) The function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , ...
- - 
  - ... , defined by  $f(x) = |x^3|$ , lies in  $C^3(\mathbb{R} \rightarrow \mathbb{R})$ .

EY : 20151109 **Solution 1a3.** For  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = |x^3| = \begin{cases} x^3 & \text{if } x \geq 0 \\ -x^3 & \text{if } x < 0 \end{cases}$

$$\begin{aligned}f'(x) &= \begin{cases} 3x^2 & \text{if } x \geq 0 \\ -3x^2 & \text{if } x < 0 \end{cases} \\ f''(x) &= \begin{cases} 6x & \text{if } x \geq 0 \\ -6x & \text{if } x < 0 \end{cases}\end{aligned}$$

Thus,

$f(x) = |x^3| \in C^1(\mathbb{R})$  but  $f(x) \notin C^2(\mathbb{R}) \subseteq C^3(\mathbb{R})$

- 
-

- (b)
- (c)

**Short Exercise 4: Undergraduate multi-dimensional analysis .**

*A good notation and basic results for partial differentiation.*

For a map  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  we denote by the map  $\partial_i f : \mathbb{R}^d \rightarrow \mathbb{R}$  the partial derivative with respect to the  $i$ -th entry.

**Question ::** Given a function

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}; (\alpha, \beta, \delta) \mapsto f(\alpha, \beta, \delta) := \alpha^3 \beta^2 + \beta^2 \delta + \delta$$

calculate the values of the following derivatives:

**Solution ::**

- $(\partial_2 f)(x, y, z) =$
- $(\partial_1 f)(\square, \circ, *) =$
- $(\partial_1 \partial_2 f)(a, b, c) =$
- $(\partial_3^2 f)(299, 1222, 0) =$

EY: 20151110

For  $f(\alpha, \beta, \delta) := \alpha^3 \beta^2 + \beta^2 \delta + \delta$ , or  $f(x, y, z) = x^3 y^2 + y^2 z + z$ ,

$$(\partial_2 f) = 2(x^3 y + yz)$$

$$(\partial_1 f) = 3x^2 y^2$$

$$(\partial_1 \partial_2 f) = 6x^2 y$$

$$(\partial_3^2 f) = 0$$

and so

- $(\partial_2 f)(x, y, z) = 2(x^3 y + yz)$
- $(\partial_1 f)(\square, \circ, *) = 3\square^2 \circ^2$
- $(\partial_1 \partial_2 f)(a, b, c) = 6a^2 b$
- $(\partial_3^2 f)(299, 1222, 0) = 0$

**Exercise 5: Differentiability on a manifold.**

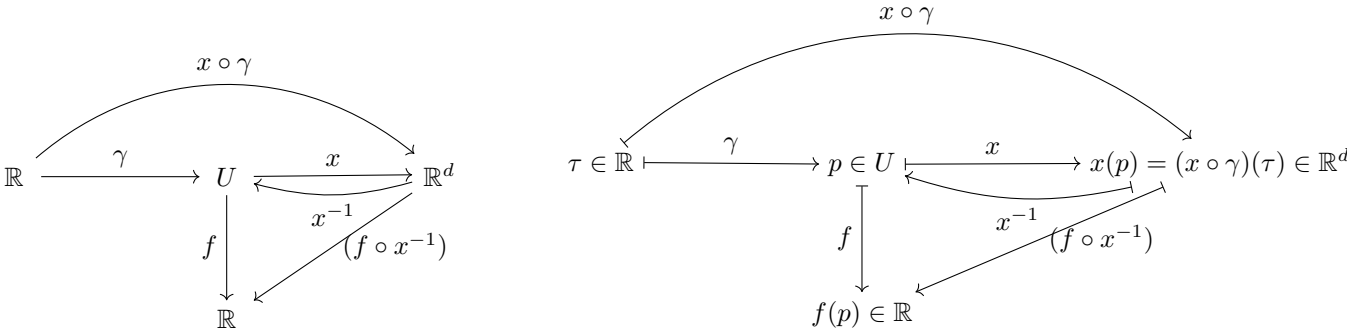
*How to deal with functions and curves in a chart*

Let  $(M, \mathcal{O}, \mathcal{A})$  be a smooth  $d$ -dimensional manifold. Consider a chart  $(U, x)$  of the atlas  $\mathcal{A}$  together with a smooth curve  $\gamma : \mathbb{R} \rightarrow U$  and a smooth function  $f : U \rightarrow \mathbb{R}$  on the domain  $U$  of the chart.

**Question ::** Draw a commutative diagram containing the chart domain, chart map, function, curve and the respective represen-

tatives of the function and the curve in the chart.

**Solution ::**



**Question ::** Consider, for  $d = 2$ ,

$$(x \circ \gamma)(\lambda) := (\cos(\lambda), \sin(\lambda)) \text{ and } (f \circ x^{-1})((x, y)) := x^2 + y^2$$

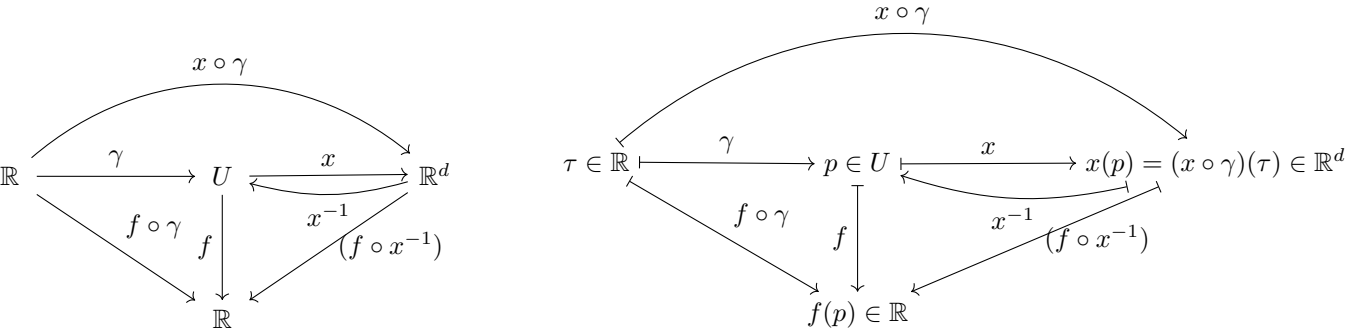
Using the chain rule, calculate

$$(f \circ \gamma)'(\lambda)$$

explicitly.

**Solution ::**

EY : 20151109 Indeed, the domains and codomains of this  $f \circ \gamma$  mapping makes sense, from  $\mathbb{R} \rightarrow \mathbb{R}$  for



$$(f \circ \gamma)'(\lambda) = (Df) \cdot \dot{\gamma}(\lambda) = \frac{\partial f}{\partial x^j} \dot{\gamma}^j(\lambda) = 2x(-\sin \lambda) + 2y \cos \lambda = 2(-\cos \lambda \sin \lambda + \sin \lambda \cos \lambda) = 0$$

31. LECTURE 5: TANGENT SPACES

lead question: “what is the velocity of a curve  $\gamma$  point  $p$ ?

31.1. Velocities.

**Definition 82.**  $(M, \mathcal{O}, \mathcal{A})$  smooth mfd.

curve  $\gamma : \mathbb{R} \rightarrow M$  at least  $C^1$ .

Suppose  $\gamma(\lambda_0) = p$

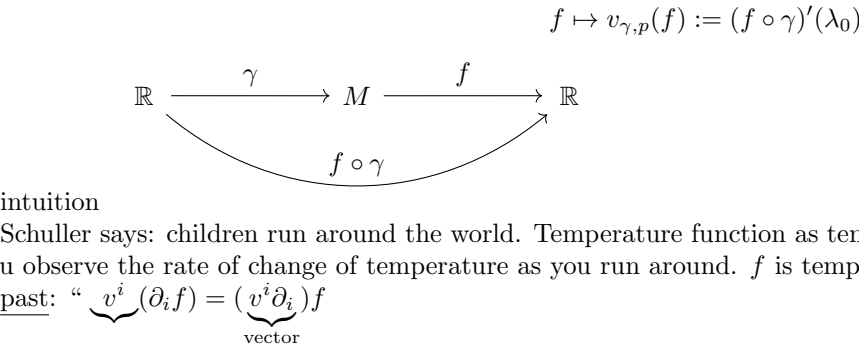
The **velocity** of  $\gamma$   $p$  is the linear map

$$v_{\gamma,p} : C^\infty(M) \xrightarrow{\sim} \mathbb{R}$$

$$C^\infty(M) := \{f : M \rightarrow \mathbb{R} | f \text{ smooth function} \} \text{ equipped with } (f \oplus g)(p) := f(p) + g(p)$$

$$(\lambda \otimes g)(p) := \lambda \cdot g(p)$$

$\sim$  denotes linear map on top of  $\rightarrow$ .



intuition

Schuller says: children run around the world. Temperature function as temperature contour lines. You feel the temperature.

You observe the rate of change of temperature as you run around.  $f$  is temperature.

past: “  $\underbrace{v^i}_{\text{vector}}(\partial_i f) = (\underbrace{v^i \partial_i}_{\text{vector}})f$

31.2. Tangent vector space.

**Definition 83.** For each point  $p \in M$   
def the **set** “tangent space  $\neq_0 M$   $p$  “

$$v_{\gamma,p} \oplus v_{\delta,p} = v_{\sigma,p}$$

$$T_pM := \{v_{\gamma,p} | \gamma \text{ smooth curves} \}$$

□

picture:  
rather  $\overline{M}$  than (embedded)  $p \in T_pM$  EY : 20151109 see [https://youtu.be/pepU\\_7NJSGM?t=12m38s](https://youtu.be/pepU_7NJSGM?t=12m38s) for the picture  
Observation:  $T_pM$  can be made into a vector space.

$$\begin{aligned} \oplus : T_pM \times T_pM &\rightarrow \\ (v_{\gamma,p} \oplus v_{\delta,p}) \big( \underbrace{f}_{\in C^\infty(M)} \big) &:= v_{\gamma,p}(f) +_{\mathbb{R}} v_{\delta,p}(f) \\ \odot : \mathbb{R} \times T_pM &\rightarrow \text{Hom}(C^\infty(\mathbb{R}), \mathbb{R}) \\ (\alpha \odot v_{\gamma,p})(f) &:= \alpha \cdot_{\mathbb{R}} v_{\gamma,p}(f) \end{aligned}$$

Remains to be shown that

- (i)  $\exists \sigma$  curve :  $v_{\gamma,p} \oplus v_{\delta,p} = v_{\sigma,p}$
- (ii)  $\exists \tau$  curve :  $\alpha \odot v_{\gamma,p} = v_{\tau,p}$

Claim:  $\tau : \mathbb{R} \rightarrow M$  where  $\mu_\alpha : \mathbb{R} \rightarrow \mathbb{R}$ , does the trick.  
 $\mapsto \tau(\lambda) := \gamma(\alpha\lambda + \lambda_0) = (\gamma \circ \mu_\alpha)(\lambda)$   $r \mapsto \alpha \cdot r + \lambda_0$   
 $\tau(0) = \gamma(\lambda_0) = p$

$$\begin{aligned} v_{\tau,p} &:= (f \circ \tau)'(0) = (f \circ \gamma \circ \mu_\alpha)'(0) \\ &= (f \circ \gamma)'(\lambda_0) \cdot \alpha = \\ &= \alpha \cdot v_{\gamma,p} \end{aligned}$$

Now for the sum:  
 $v_{\gamma,p} \oplus v_{\delta,p} \stackrel{?}{=} v_{\sigma,p}$   
make a choice of chart  $(\underbrace{U}_{\ni p}, x)$  In cloud: ill definition alarm bells.

and define:  
Claim:

$$\begin{aligned} \sigma : \mathbb{R} &\rightarrow M \\ \sigma(\lambda) &:= x^{-1} \big( \underbrace{(x \circ \gamma)(\lambda_0 + \lambda) + (x \circ \delta)(\lambda_1 + \lambda) - (x \circ \gamma)(\lambda_0)}_{\mathbb{R} \rightarrow \mathbb{R}^d} \big) \end{aligned}$$

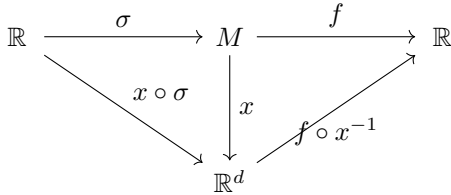
does the trick.

*Proof.* Since:

$$\begin{aligned} \sigma_x(0) &= x^{-1}((x \circ \gamma)(\lambda_0) + (x \circ \delta)(\lambda_1) - (x \circ \gamma)(\lambda_0)) \\ &= \delta(\lambda_1) = p \end{aligned}$$

Now:

$$\begin{aligned} v_{\sigma_x,p}(f) &:= (f \circ \sigma_x)'(0) = \\ &= \underbrace{((f \circ x^{-1}) \circ (x \circ \sigma_x))'(\gamma)}_{\mathbb{R}^d \rightarrow \mathbb{R}} = \underbrace{(x \circ \sigma_x)'(0)}_{(x \circ \gamma)'(\lambda_0) + (x \circ \delta)'(\lambda_1)} \cdot (\partial_i(f \circ x^{-1}))(x(\underbrace{\sigma(0))}_p) = \\ &= (x \circ \gamma)'(\lambda_0)(\partial_i(f \circ x^{-1}))(x(p)) + (x \circ \delta)'(\lambda_1)(\partial_i(f \circ x^{-1}))(x(p)) \\ &= (f \circ \gamma)'(\lambda_0) + (f \circ \delta)'(\lambda_1) = \\ &= v_{\gamma,p}(f) + v_{\delta,p}(f) \quad \forall f \in C^\infty(M) \end{aligned}$$



picture: (cf. [https://youtu.be/pepU\\_7NJSGM?t=39m5s](https://youtu.be/pepU_7NJSGM?t=39m5s))

$$\begin{aligned} \gamma : \mathbb{R} &\rightarrow M \\ \delta : \mathbb{R} &\rightarrow M \end{aligned}$$

$(\gamma \oplus)(\lambda) := \gamma(\lambda) + \delta(\lambda)$   
EY : 20151109 Schuller says adding trajectories is chart dependent, bad. Adding velocities is good.

31.3. Components of a vector wrt a chart.

**Definition 84.** Let  $(U, x) \in \mathcal{A}_{smooth}$ .  
 $\gamma : \mathbb{R} \rightarrow U$   
Let  $\gamma(0) = p$ .  
Calculate

$$\begin{aligned} v_{\gamma,p}(f) &:= (f \circ \gamma)'(0) = \underbrace{((f \circ x^{-1}) \circ (x \circ \gamma))'(\gamma)}_{\mathbb{R}^d \rightarrow \mathbb{R}} = \underbrace{(x \circ \gamma)^{i'}(0)}_{\dot{\gamma}_x^i(0)} \cdot \underbrace{(\partial_i(f \circ x^{-1}))(x(p))}_{=:(\frac{\partial f}{\partial x^i})_p} \end{aligned}$$

think cloud  $f : M \rightarrow \mathbb{R}$

$$= \dot{\gamma}_x^i(0) \cdot \left( \frac{\partial}{\partial x^i} \right)_p f \quad \forall f \in C^\infty(M)$$

$\therefore$  as a map.

$$v_{\gamma,p} \underbrace{=}_{\text{use of chart}} \underbrace{\gamma_x^i(0)}_{\text{“components of the velocity } v_{\gamma,p} \text{”}} \underbrace{\left( \frac{\partial}{\partial x^i} \right)}_{\text{basis elements of the } T_pM \text{ wrt which the components need to be understood. “chart induced basis of } T_pM \text{”}}$$

Picture: [https://youtu.be/pepU\\_7NJSGM?t=1h16s](https://youtu.be/pepU_7NJSGM?t=1h16s)

31.4. 4. Chart-induced basis.

**Definition 85.**  $(U, x) \in \mathcal{A}_{smooth}$   
the  $\left(\frac{\partial}{\partial x^1}\right)_p, \dots, \left(\frac{\partial}{\partial x^d}\right)_p \in T_p U \subseteq T_p M$   
constitute a **basis** of  $T_p U$

*Proof.* remains: linearly independent

$$\begin{aligned} \lambda^i \left(\frac{\partial}{\partial x^i}\right)_p &\stackrel{!}{=} 0 \\ \implies \lambda^i \left(\frac{\partial}{\partial x^i}\right)_p (x^j) &= \lambda^i \partial_i (\underbrace{x^j \circ x^{-1}})(x(p)) = & x^j \circ x^{-1} : \mathbb{R}^d \rightarrow \mathbb{R} \\ & & (\alpha^1, \dots, \alpha^d) \mapsto \alpha^j \\ &= \lambda^i \delta_i^j = \lambda^j & j = 1, \dots, d \end{aligned}$$

in cloud:  $x^j : U \rightarrow \mathbb{R}$  differentiable

**Corollary 1.**  $\dim T_p M = d = \dim M$

Terminology:  $X \in T_p M \rightarrow \exists \gamma : \mathbb{R} \rightarrow M : X = v_{\gamma, p}$  and  
 $\exists \underbrace{X_1^1, \dots, X^d}_\in \mathbb{R} : X = X^i \left(\frac{\partial}{\partial x^i}\right)_p$

31.5. 5. Change of vector components under a change of chart. ✖ vector does **not** change under change of chart.

Let  $(U, x)$  and  $(V, y)$  be overlapping charts and  $p \in U \cap V$ .  
Let  $X \in T_p M$

$$X^i_{(y)} \cdot \left(\frac{\partial}{\partial y^i}\right)_p \underbrace{=}_p \underbrace{X}_{(V, y)} \underbrace{=}_{(U, x)} X^i_x \left(\frac{\partial}{\partial x^i}\right)_p$$

to study change of components formula:

$$\begin{aligned} \left(\frac{\partial}{\partial x^i}\right)_p f &= \partial_i (f \circ x^{-1})(x(p)) = \\ &= \partial_i (\underbrace{(f \circ y^{-1})}_{\mathbb{R}^d \rightarrow \mathbb{R}} \circ \underbrace{(y \circ x^{-1})}_{\mathbb{R}^d \rightarrow \mathbb{R}^d})(x(p)) \\ &= (\partial_i (y^i \circ x^{-1}))(x(p)) \cdot (\partial_j (f \circ y^{-1}))(y(p)) = \\ &= \boxed{\left(\frac{\partial y^p}{\partial x^i}\right)_p \cdot \left(\frac{\partial f}{\partial y^j}\right)_p} f \\ \implies X^i_{(x)} \left(\frac{\partial y^j}{\partial x^i}\right)_p \left(\frac{\partial}{\partial y^j}\right)_p &= X^j_{(y)} \left(\frac{\partial}{\partial y^j}\right)_p \\ \implies \boxed{X^j_{(y)} = \left(\frac{\partial y^j}{\partial x^i}\right)_p X^i_{(x)}} \end{aligned}$$

31.6. 6. Cotangent spaces.  $T_p M = V$

trivial  $(T_p M)^* := \{\varphi : T_p M \xrightarrow{\sim} \mathbb{R}\}$   
Example:  $f \in C^\infty(M)$

$$\begin{aligned} (df)_p : T_p M &\xrightarrow{\sim} \mathbb{R} \\ X &\mapsto (df)_p(X) \end{aligned}$$

i.e.  $\boxed{(df)_p \in T_p M^*}$

$(df)_p$  called the gradient of  $f \quad p \in M$ .  
Calculate components of gradient w.r.t. chart-induced basis  $(U, x)$

$$\begin{aligned} ((df)_p)_j &:= (df)_p \left(\left(\frac{\partial}{\partial x^j}\right)_p\right) \\ &= \left(\frac{\partial f}{\partial x^j}\right)_p = \partial_j (f \circ x^{-1})(x(p)) \end{aligned}$$

**Theorem 30.** Consider chart  $(U, x) \implies x^i : U \rightarrow \mathbb{R}$

□ Claim:  $(dx^1)_p, (dx^2)_p, \dots, (dx^d)_p$  basis of  $T_p^* M$   
 $\implies$  In fact: dual basis:

$$(dx^a)_p \left(\left(\frac{\partial}{\partial x^b}\right)_p\right) = \left(\frac{\partial x^a}{\partial x^b}\right)_p = \dots = \delta_b^a$$

31.7. 7. Change of components of a covector under a change of chart:

$$\underbrace{T_p^* M}_{\ni \omega} \text{ with } \omega_{(y)}(dy^j)_p = \omega = \omega_{(x)i}(dx^i)_p$$

$$\implies \boxed{\omega_{(y)i} = \frac{\partial x^j}{\partial y^i} \omega_{(x)j}}$$

LECTURE 6: FIELDS

cf. **Lecture 6: Fields (International Winter School on Gravity and Light 2015)**

So far:

$$\begin{array}{c} T_p M \\ \vdots \downarrow \\ T_p^* M \\ \vdots \downarrow \\ \vdots \end{array},$$

now  
in Thought Cloud: theory of bundles

31.8. Bundles.

**Definition 86.** A **bundle** is a triple

$$E \xrightarrow{\pi} M$$

$E$  smooth manifold “**total space**”  
 $\pi$  smooth map (surjective) “**projection map**”  
 $M$  smooth manifold “**base space**”

Example  $E = \text{cylinder}$   $M = \text{circle}$



**Definition 87.**

$$E \overset{T}{\longrightarrow} M$$

bundle.

$p \in M$   
define ***fibre over***  $p$   
 $:= \text{preim}_\pi(\{p\})$

**Definition 88.** A *section*  $\sigma$  of a bundle

$$\begin{array}{ccc} E & & \\ \pi \downarrow \nearrow \sigma & & \\ M & & \end{array}$$

require  $\pi \circ \sigma = id_M$

Schuller says: in quantum mechanics, Aside:  $\psi : M \rightarrow \mathbb{C}$

31.9. **Tangent bundle of smooth manifold.**  $(M, \mathcal{O}, \mathcal{A})$  smooth manifold

(a) as a **set**  $TM := \dot{\bigcup}_{p \in M} T_p M$

(b) surjective  $\pi : TM \rightarrow M$  the *unique* point  $p \in M, X \in T_p M$

$$\begin{array}{ccccc} X \mapsto p & & & & \\ \text{situation: } \underbrace{TM}_{\text{set}} & \xrightarrow{\pi} & \underbrace{M}_{\text{smooth manifold}} & & \end{array}$$

surjective map

(c) Construct topology on  $TM$  that is the coarsest topology such that  $\pi$  (just) continuous. (“initial topology with respect to  $\pi$ ”).

$$\mathcal{O}_{TM} := \{\text{preim}_\pi(U) | U \in \mathcal{O}\}$$

Show: Tutorial  $\mathcal{O}_{TM}$  Schuller says this is shown in the tutorial  $(TM, \mathcal{O}_{TM})$

Construction of a  $C^\infty$ -atlas on  $TM$  from the  $C^\infty$ -atlas  $\mathcal{A}$  on  $M$ .

$$\mathcal{A}_{TM} := \{(T\mathcal{U}, \xi_x) | (U, x) \in \mathcal{A}\}$$

where

$$\begin{array}{l} \xi_x : T\mathcal{U} \rightarrow \mathbb{R}^{2 \cdot \dim M} \\ X \mapsto \underbrace{((x^1 \circ \pi)(X), \dots, (x^d \circ \pi)(X), (dx^1)_{\pi(X)}(X), \dots, (dx^d)_{\pi(X)}(X))}_{(U, x) - \text{ coords of } \pi(X) \text{ (d many) }} \end{array}$$

where  $X \in T_{\pi(X)} M$   
 $X = X^i_{(x)} \left( \frac{\partial}{\partial x^i} \right)_{\pi(X)}$

$$\begin{aligned} (dx^j)_{\pi(X)}(X) &= (dx^j)_{\pi(X)} \left( X^i_{(x)} \left( \frac{\partial}{\partial x^i} \right)_{\pi(X)} \right) = \\ &= X^i_{(x)} \delta_i^j = X^j_{(x)} \end{aligned}$$

Write  $\xi_x^{-1} : \xi_x(TU) \subseteq \mathbb{R}^{2 \dim M} \rightarrow TU$

$$(\alpha^1, \dots, \alpha^d, \beta^1, \dots, \beta^d) := \beta^i \left( \frac{\partial}{\partial x^i} \right) \underbrace{x^{-1}(\alpha^1, \dots, \alpha^d)}_{\pi(X)}$$

Check:

$$\begin{aligned} (\xi_y \circ \xi_x^{-1})(\alpha^1, \dots, \alpha^d, \beta^1, \dots, \beta^d) &= \\ &= \xi_y \left( \beta^i \left( \frac{\partial}{\partial x^i} \right)_{x^{-1}(\alpha^1, \dots, \alpha^d)} \right) \\ &= \left( \dots, (y^i \circ \pi)(\beta^m \cdot \left( \frac{\partial}{\partial x^m} \right)_{x^{-1}(\alpha^1 \dots \alpha^d)}), \dots, \dots (dy^i)_{x^{-1}(\alpha^1, \dots, \alpha^d)} \left( \beta^m \left( \frac{\partial}{\partial x^m} \right)_{x^{-1}(\alpha^1 \dots \alpha^d)} \right), \dots \right) = \\ &= \underbrace{(\dots, (y^i \circ x^{-1})(\alpha^1, \dots, \alpha^d), \dots, \dots, \beta^m (dy^i)_{x^{-1}(\alpha^1, \dots, \alpha^d)} \left( \left( \frac{\partial}{\partial x^m} \right)_{x^{-1}(\alpha^1 \dots \alpha^d)} \right))}_{\beta^m \left( \frac{\partial y}{\partial x^m} \right)_{x^{-1}(\alpha^1 \dots \alpha^d)}} \end{aligned}$$

Check transition map:  $(U, x), (V, y), \quad U \cap V \neq \emptyset$   
 $\left( \frac{\partial y}{\partial x^m} \right)_{x^{-1}(\alpha^1 \dots \alpha^d)} = \partial_m (y^i \circ x^{-1})(x \circ (x^{-1}(\alpha^1 \dots \alpha^d))) = \partial_m (y^i \circ x^{-1})(\alpha^1 \dots \alpha^d)$  smooth.  
upshot

$$\underbrace{TM}_{\text{smooth manifold}} \xrightarrow{\pi} \underbrace{M}_{\text{smooth manifold}}$$

smooth map

bundle, called the tangent bundle.

3. Vector fields.

**Definition 89.** A *smooth vector field*  $\chi$  is a smooth map, (where)

$$\begin{array}{ccc} TM & & \\ \pi \downarrow \nearrow \chi & & \\ M & & \end{array}$$

Example:

$$\begin{array}{ccc} TS^1 & & \\ \downarrow \pi & & \\ S^1 & & \end{array}$$

4. The  $C^\infty(M)$ -module  $\Gamma(TM)$ .  
 $C^\infty(M)$ -module  $\leftarrow (C^\infty(M), +, \cdot)$  (satisfies)  $C^+, A^+, N^+, I+, C^+, A^+, N^+, D^+$ . Not a field. A ring.

$$\begin{array}{l} \text{set } \Gamma(TM) = \{ \chi \mid M \rightarrow TM \mid \text{smooth section} \} \\ (\chi \oplus \tilde{\chi})(f) := (\chi f) \underbrace{+}_{C^\infty(M)} \tilde{\chi}(f) \end{array}$$

$$\left( \underbrace{g}_{C^\infty(M)} \odot \xi \right)(f) := \underbrace{g}_{C^\infty(M)} \cdot \chi(f)$$

$$\begin{aligned}\chi:M&\rightarrow TM \\ p&\mapsto \chi(p) \\ \chi f:M&\rightarrow \mathbb{R} \\ p&\mapsto \chi(p)f\end{aligned}$$

$(\Gamma(TM), \oplus, \odot)$   $C^\infty(M)$  - **module**

upshot: set of all smooth vector fields can be made into a  $C^\infty(M)$ -module.

Fact:

- (1)  $\text{ZFC} \implies$  every vector space has a basis. (You have to have C - axiom of choice in set theory)
- (2) no such result exists for modules.

This is a shame, because otherwise, we could have chosen (for any manifolds) vector fields,

$$\chi_{(1)}, \dots, \chi_{(d)} \in \Gamma(TM)$$

and would be able to write every vector field  $\Xi$

$$\chi = \underbrace{\qquad\qquad\qquad}_{\text{component functions}} \cdot \chi_{(i)}$$

Simple counterexample

Schuller says: Take a sphere, Morse Theorem, every smooth vector field must vanish at 2 pts. “mustn’t choose a global basis”

However:  $\frac{\partial}{\partial x^i} : U \xrightarrow{\text{smooth}} TU$

$$p \mapsto \left( \frac{\partial}{\partial x^i} \right)_p$$

- 31.10. **Tensor fields.** so far
- $\Gamma(M)$  =”set of vector fields”  $C^\infty(M)$ -module
- $\Gamma(T^*M)$  = “covector fields”  $C^\infty(M)$ -module

**Definition 90.** An  $(r, s)$ -tensor field  $T$  is a multi-linear map

$$T : \underbrace{\Gamma(T^*M) \times \dots \times \Gamma(T^*M)}_r \times \Gamma(TM) \times \dots \times \Gamma(TM) \xrightarrow{\sim} C^\infty(M)$$

Example:  $f \in C^\infty(M)$

$$\begin{aligned}df : \Gamma(TM) &\xrightarrow{\sim} C^\infty(M) \\ \chi &\mapsto df(\chi) := \chi[f]\end{aligned}$$

$df$  (0, 1)-T.F. (tensor field)  
where  $(\chi f) \underbrace{\binom{p}{\in M}} := \underbrace{\chi(p)}_{\in T_p M} f$   
can check:  $df$  is  $C^\infty$ –linear

LECTURE 7: CONNECTIONS

cf. [Lecture 7: Connections \(International Winter School on Gravity and Light 2015\)](#)  
So far: saw that a vector field  $X$  can be used to provide a directional derivative

$$\nabla_X f := Xf$$

of a function  $f \in C^\infty(M)$ .

Remark: from now on: consider mostly vector fields.  
Notational overkill?

$$\nabla_X f = Xf = (df)(X)$$

In Thought Bubble:  $\nabla_X(f \cdot g) = X(fg) = (Xf) \cdot g + fX(g)$  Product rule, because it’s a derivative.  
not quite:

$$\begin{aligned}X:C^\infty(M)&\rightarrow C^\infty(M) \\ df:\Gamma(TM)&\rightarrow C^\infty(M) \\ \nabla_X:C^\infty(M)&\rightarrow C^\infty(M)\end{aligned}$$

$$\begin{array}{ccc} \nabla_X : C^\infty(M) & \longrightarrow & C^\infty(M) \\ \vdots \downarrow & & \vdots \downarrow \\ \nabla_X : \begin{matrix} TM^p \otimes T^*M^q \text{ i.e.} \\ \binom{p}{q} \text{ tensor field} \end{matrix} & \longrightarrow & \begin{matrix} TM^p \otimes T^*M^q \text{ i.e.} \\ \binom{p}{q} \text{ tensor field} \end{matrix} \end{array}$$

**1. Directional derivatives of tensor fields.** We formulate a wish list of properties which the  $\nabla_X$  acting on a tensor field should have.

In Thought Bubble: Any remaining freedom in choosing  $\nabla$ , will need to be provided as additional structure beyond  $(M, \mathcal{O}, \mathcal{A})$

**Definition 91** (connection). In Thought Bubble: linear connection, *covariant derivative*, affine connection  
A **connection**  $\nabla$  on a smooth manifold  $(M, \mathcal{O}, \mathcal{A})$  is a map that takes a pair consisting of a vector (field)  $X$  and a  $(p, q)$ -tensor field  $T$  and sends them to a  $(p, q)$ -tensor (field)  $\nabla_X T$  satisfying

- (i)  $\nabla_X f = Xf \quad \forall f \in C^\infty(M)$
- (ii)  $\nabla_X(T + S) = \nabla_X T + \nabla_X S$
- (iii)  $\nabla_X(\underbrace{T(\omega, Y)}_{\in C^\infty(M)}) = (\nabla_X T)(\omega, Y) + T(\nabla_X \omega, Y) + T(\omega, \nabla_X Y)$

In Thought Bubble: for  $(1, 1)$ -TF  $T$ , but analogously for any  $(p, q)$  - TF

- ”Leibnitz” rule.
- (iv)  $\nabla_{fX+Z} T = f\nabla_X T + \nabla_Z T$   
 $f \in C^\infty(M)$

A manifold with connection is quadruple  $(M, \mathcal{O}, \mathcal{A}, \nabla)$   
Remark:  $\nabla_X$  is the extension of  $X$ .  
 $\nabla$  — ” — of  $d$

## 2. New structure on $(M, \mathcal{O}, \mathcal{A})$ required to fix $\nabla$ . Q: How much freedom do we have in choosing such a structure.

Consider  $X, Y$  vector fields

$$\begin{aligned} \nabla_X Y &\stackrel{\text{In Thought Bubble: } (U, x)}{=} \nabla_{X^i \frac{\partial}{\partial x^i}} \left( Y^m \frac{\partial}{\partial x^m} \right) \\ &\stackrel{\text{(iii)}}{=} \underbrace{X^i \left( \nabla_{\frac{\partial}{\partial x^i}} Y^m \right)}_{\text{(iii)}} \frac{\partial}{\partial x^m} + X^i Y^m \underbrace{\nabla_{\frac{\partial}{\partial x^i}} \left( \frac{\partial}{\partial x^m} \right)}_{\text{(iii)}} \\ &\stackrel{\text{(i)}}{=} \underbrace{X^i \left( \frac{\partial}{\partial x^i} Y^m \right)}_{\text{(i)}} \frac{\partial}{\partial x^m} + X^i Y^m \underbrace{\Gamma_{mi}^q(x)}_{\text{connection coefficient functions (on } M) \text{ of } \nabla \text{ wrt } (U, x)} \frac{\partial}{\partial x^q} \end{aligned}$$

$$\begin{aligned} (\underbrace{T}_{(p,q)} \otimes \underbrace{S}_{(r,s)})(\omega, \dots, X, \dots) &:= T(\omega, \dots, X, \dots) \underbrace{\cdot}_{C^\infty(M)} S(\dots, \dots) \\ \nabla_X(T \otimes S) &= (\nabla_X T) \otimes S + T \otimes (\nabla_X S) \end{aligned}$$

**Definition 92** (Connection coefficient functions).  $(M, \mathcal{O}, \mathcal{A}, \nabla)$ ,  $(\mathcal{U}, x) \in \mathcal{A}$ .

Then the **connection coefficient functions** ("Γ"s) with respect to (wrt)  $(U, x)$  on the  $(\dim(M))^3$  many functions

$$\begin{aligned} \Gamma_{jk}^i : \mathcal{U} &\rightarrow \mathbb{R} \\ p &\mapsto \left( dx^i \left( \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} \right) \right) (p) \end{aligned}$$

Thus:

$$(\nabla_X Y)^i = X^m \left( \frac{\partial}{\partial x^m} Y^i \right) + \Gamma_{nm}^i \underbrace{\cdot}_{C^\infty(M)} Y^n X^m$$

*Remark:* On a chart domain  $U$ , choice of the  $(\dim M)^3$  functions  $\Gamma_{jk}^i$  suffices to fix the action of  $\nabla$  on a *vector* field.

Fortunately, the same  $(\dim M)^3$  functions fix the action of  $\nabla$  on any tensor field.

*key observation:*

$$\nabla_{\frac{\partial}{\partial x^m}} (dx^i) = \sum_{jm}^i dx^j$$

but now:

$$\begin{aligned} \underbrace{\nabla_{\frac{\partial}{\partial x^m}} \left( dx^i \left( \frac{\partial}{\partial x^j} \right) \right)}_{\delta_j^i} &= \frac{\partial}{\partial x^m} (\delta_j^i) = 0 \\ \underbrace{\parallel}_{\text{(iii)}} & \\ &= \left( \nabla_{\frac{\partial}{\partial x^m}} dx^i \right) \left( \frac{\partial}{\partial x^j} \right) + dx^i \underbrace{\left( \nabla_{\frac{\partial}{\partial x^m}} \frac{\partial}{\partial x^j} \right)}_{\Gamma_{jm}^q \frac{\partial}{\partial x^q}} = 0 \\ \implies \left( \nabla_{\frac{\partial}{\partial x^m}} dx^i \right) \left( \frac{\partial}{\partial x^j} \right) &= -\Gamma_{jm}^i \end{aligned}$$

Summary so far:

$$\begin{aligned} (\nabla_X Y)^i &= X(Y^i) \underbrace{+}_{\text{act on } \mathbf{vector} \text{ field}} \Gamma_{jm}^i Y^j X^m \\ (\nabla_X \omega)_i &= X(\omega_i) + -\Gamma_{im}^j \omega_j X^m \end{aligned}$$

Note that for the immediately above expression for  $(\nabla_X Y)^i$ , in the second term on the right hand side,  $\Gamma_{jm}^i$  has the last entry at the bottom,  $m$  going in the direction of  $X$ , so that it matches up with  $X^m$ . This is a good mnemonic to memorize the index positions of  $\Gamma$ .

similarly, by further application of Leibnitz

$T$  a  $(1, 2)$ -TF (tensor field)

$$(\nabla_X T)_{jk}^i = X(T_{jk}^i) + \Gamma_{sm}^i T_{jk}^s X^m - \Gamma_{jm}^s T_{sk}^i X^m - \Gamma_{km}^s T_{js}^i X^m$$

Question: If in a Euclidean space, the  $\Gamma$ s all vanish in a (then existing) global chart.

Answer: Yes, but: What is a Euclidean space:

$(M = \mathbb{R}^n, \mathcal{O}_{\text{st}}, \mathcal{A})$  smooth manifold.

Assume  $(\mathbb{R}^n, \text{id}_{\mathbb{R}^n}) \in \mathcal{A}$  and

$$(\Gamma_{(x)}^i)_{jk} = dx^i \left( (\nabla_{\underline{E}})_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} \right) \stackrel{!}{=} 0$$

*Intuition:*

$\mathbb{R}^2$ :  $\nabla_{\text{Euclidean}}$

$\mathbb{R}^2$ :  $\nabla_{\text{Hyperbolic}}$

**Definition 93.**  $X$  vector field on  $(M, \mathcal{O}, \mathcal{A}, \nabla)$ .

Then divergence of  $X$  is the function:

$$\text{div}(X) := \left( \nabla_{\frac{\partial}{\partial x^i}} X \right)^i$$

*Claim:* chart-independent.

### 3. Change of $\Gamma$ 's under change of chart.

$(U, x), (V, y) \in \mathcal{A}$  and  $U \cap V \neq \emptyset$

$$\Gamma_{jk}^i(y) := dy^i \left( \nabla_{\frac{\partial}{\partial y^k}} \frac{\partial}{\partial y^j} \right) = \frac{\partial y^i}{\partial x^q} dx^q \left( \nabla_{\frac{\partial x^p}{\partial y^k} \frac{\partial}{\partial x^p}} \frac{\partial x^s}{\partial y^j} \frac{\partial}{\partial x^s} \right)$$

Note  $\nabla_{fX}$  is  $C^\infty$ -linear for  $fX$

covector  $dy^i$  is  $C^\infty$ -linear in its argument

$$\begin{aligned} \implies \Gamma_{jk}^i(y) &= \frac{\partial y^i}{\partial x^q} dx^q \left( \frac{\partial x^p}{\partial y^k} \left[ \left( \nabla_{\frac{\partial}{\partial x^p}} \frac{\partial x^s}{\partial y^j} \right) \frac{\partial}{\partial x^s} + \frac{\partial x^s}{\partial y^j} \left( \nabla_{\frac{\partial}{\partial x^p}} \frac{\partial}{\partial x^s} \right) \right] \right) = \\ &= \frac{\partial y^i}{\partial x^q} \underbrace{\frac{\partial x^p}{\partial y^k} \frac{\partial}{\partial x^p}}_{\frac{\partial}{\partial y^k}} \frac{\partial x^s}{\partial y^j} \delta_s^q + \frac{\partial y^i}{\partial x^q} \frac{\partial x^p}{\partial y^k} \frac{\partial x^s}{\partial y^j} \Gamma_{sp}^q(x) \end{aligned}$$

in summary:

$$(220) \quad \Gamma_{jk}^i(y) = \frac{\partial y^i}{\partial x^q} \frac{\partial^2 x^q}{\partial y^j \partial y^k} + \frac{\partial y^i}{\partial x^q} \frac{\partial x^s}{\partial y^j} \frac{\partial x^p}{\partial y^k} \Gamma_{sp}^q(x)$$

Eq. (220) is the change of connection coefficient function under the change of chart  $(U \cap V, x) \rightarrow (U \cap V, y)$

4. **Normal Coordinates.** Let  $p \in M$  of  $(M, \mathcal{O}, \mathcal{A}, \nabla)$

Then one can construct a chart  $(U, x)$  with  $p \in U$  such that

$$\Gamma(x)^i_{(jk)}(p) = 0$$

at the point  $p$ . **Not** necessarily in any neighborhood.

*Proof.* Let  $(V, y)$  be any chart (with)  $p \in V$ .

Thus, in general:  $\Gamma(y)^i_{jk} \neq 0$

Then consider a new chart  $(U, x)$  to which one transits by virtue of

$$(x \circ y^{-1})^i(\alpha^1, \dots, \alpha^d) := \alpha^i - \frac{1}{2}\Gamma(y)^i_{(jk)}(p)\alpha^j\alpha^k$$

$$p = x^{-1}(\alpha^1, \dots, \alpha^d)$$

$$\left(\frac{\partial x^i}{\partial y^j}\right)_p = \partial_j(x^i \circ y^{-1}) = \delta_j^i - \Gamma(y)^i_{mj}(p) \alpha^m|_{\alpha=0} = \delta_j^i$$

$$\frac{\partial x^i}{\partial y^k \partial y^j}(p) = -\Gamma(y)^i_{kj}(p)$$

$$\begin{aligned} \implies \Gamma(x)^i_{jk}(p) &= \Gamma(y)^i_{jk}(p) - \Gamma(y)^i_{kj}(p) = 0 \\ &= \Gamma(y)^i_{[jk]}(p) = T(y)^i_{jk} \end{aligned}$$

Terminology:  $(U, x)$  is called a **normal coordinate chart** of  $\nabla$  at  $\mathbf{p} \in \mathbf{M}$ .

**Tutorial 7 Connections. Exercise 1. : True or false?**

- (a)
- $\nabla_{fX}Y = f\nabla_XY$  by definition so  $\nabla_{fX} = f\nabla_X$  i.e.  $\nabla_X$  is  $C^\infty(M)$ -linear in  $X$
  - $f \in C^\infty(M)$  is a  $(0,0)$ -tensor field.  $\nabla_Xf = Xf \equiv X(f)$  by definition.
  - If the manifold is flat, I'm assuming that means that the manifold is globally a Euclidean space, and by definition,  $\Gamma = 0$ .

$$\nabla_XY = X^j \frac{\partial}{\partial x^j}(Y^i) \frac{\partial}{\partial x^i} + \Gamma_{jk}^i Y^k X^j \frac{\partial}{\partial x^i} = X^j \frac{\partial Y^i}{\partial x^j} \frac{\partial}{\partial x^i} + 0$$

and similarly for any  $(p,q)$ -tensor field, i.e.

$$\nabla_XT = X^j \frac{\partial T^{i_1 \dots i_p}_{j_1 \dots j_q}}{\partial x^j}$$

- 

$$\nabla_Xf = X^j \frac{\partial f}{\partial x^j} = X \cdot \text{grad}(f)$$

- $\forall (U, x) \in \mathcal{A}$ , locally (after working out the first few cases, and doing induction, one can look up the expression for the local form; I found it in Nakahara's **Geometry, Topology and Physics**, Eq. 7.26, and it needs to be modified for the convention of order of bottom indices for  $\Gamma$ :

$$\nabla_\nu t^{\lambda_1 \dots \lambda_p}_{\mu_1 \dots \mu_q} = \partial_\nu t^{\lambda_1 \dots \lambda_p}_{\mu_1 \dots \mu_q} + \Gamma_{\kappa \nu}^{\lambda_1} t^{\kappa \lambda_2 \dots \lambda_p}_{\mu_1 \dots \mu_q} + \dots + \Gamma_{\kappa \nu}^{\lambda_p} t^{\lambda_1 \dots \lambda_{p-1} \kappa}_{\mu_1 \dots \mu_q} - \Gamma_{\mu_1 \nu}^{\kappa} t^{\lambda_1 \dots \lambda_p}_{\kappa \mu_2 \dots \mu_q} - \dots - \Gamma_{\mu_q \nu}^{\kappa} t^{\lambda_1 \dots \lambda_p}_{\mu_1 \dots \mu_{q-1} \kappa}$$

Clearly,  $\nabla_X$  is uniquely fixed  $\forall p \in M$  by choosing each of the  $(\dim M)^3$  many connection coefficient functions  $\Gamma$ .

- (b)
- $\nabla : \mathfrak{X}(M) \rightarrow \mathfrak{X}(M)$   
 $\nabla : (p,q)$ -tensor field  $\mapsto (p,q)$ -tensor field
  - By definition,  $\nabla$  satisfies the Leibniz rule.
  -

- 
- 

**Exercise 2. : Practical rules for how  $\nabla$  acts** Torsion-free covariant derivative boils down to a connection coefficient function  $\Gamma$  that is symmetric in the bottom indices.

- 

$$\nabla_Xf = X(f) = X^i \frac{\partial f}{\partial x^i}$$

- 

$$(\nabla_XY)^a = X^i \frac{\partial Y^a}{\partial x^i} + \Gamma_{jk}^a Y^j X^k$$

- 

$$(\nabla_X\omega)_a = X^i \frac{\partial \omega_a}{\partial x^j} - \Gamma_{ak}^i \omega_i X^k$$

- 

$$(\nabla_mT)^a_{bc} = \frac{\partial}{\partial x^m}(T^a_{bc}) + \Gamma_{im}^a T^i_{bc} - \Gamma_{bm}^i T^a_{ic} - \Gamma_{cm}^j T^a_{bj}$$

- 

$$(\nabla_{[m}A)_{n]} = (\nabla_mA)_n - (\nabla_nA)_m = \frac{\partial A_n}{\partial x^m} - \Gamma_{nm}^i A_i - \left(\frac{\partial A_m}{\partial x^n} - \Gamma_{mn}^i A_i\right) = \frac{\partial A_m}{\partial x^m} - \frac{\partial A_m}{\partial x^n}$$

- 

$$(\nabla_m\omega)_{nr} = \frac{\partial \omega_{nr}}{\partial x^m} - \Gamma_{nm}^i \omega_{ir} - \Gamma_{rm}^i \omega_{ni}$$

□

**Exercise 3. : Connection coefficients**

**Question .**

The connection coefficient functions  $\Gamma$  in chart  $(U \cap V, y)$  is given, in terms of chart  $(U \cap V, x)$  as follows:  
Recall Eq. (220)

$$\Gamma_{jk}^i(y) = \frac{\partial y^i}{\partial x^q} \frac{\partial^2 x^q}{\partial y^j \partial y^k} + \frac{\partial y^i}{\partial x^q} \frac{\partial x^s}{\partial y^j} \frac{\partial x^p}{\partial y^k} \Gamma_{sp}^q(x)$$

LECTURE 8: PARALLEL TRANSPORT & CURVATURE (INTERNATIONAL WINTER SCHOOL ON GRAVITY AND LIGHT 2015)

**Experiment.**  $\nabla_{v_\gamma}Y \stackrel{!}{=} 0$   
(please move in this fashion)

round sphere  
 $(S^2, \mathcal{O}, \mathcal{A}, \nabla)$

ellipsoid  
 $(S^2, \mathcal{O}, \mathcal{A}, \tilde{\nabla})$

ellisoid embedded in  $\mathbb{R}^3$  with  $\mathbb{R}^3$  Euclidean connection  
 $(\mathbb{R}^3, \nabla_E)$

**1. Parallellity of vector fields.**  $(M,\mathcal{O},\mathcal{A},\nabla)$  manifold with connection.

**Definition 94** (parallelly transported, parallel). (1) *A vector field  $X$  on  $M$  is said to be **parallelly transported** along a smooth curve  $\gamma : \mathbb{R} \rightarrow M$  if*

$$\nabla_{v_\gamma} X = 0$$

*i.e.*

$$\left(\nabla_{v_{\gamma,\gamma(\lambda)}} X\right)_{\gamma(\lambda)} = 0$$

(2) *A slightly weaker condition than "parallelly transported" is "parallel":*

$$(\nabla_{v_{\gamma,\gamma(\lambda)}} X)_{\gamma(\lambda)} = \mu(\lambda) X_{\gamma(\lambda)}$$

*for  $\mu : \mathbb{R} \rightarrow \mathbb{R}$*

Example: Euclidean plane  $(\mathbb{R}^2,\mathcal{O},\mathcal{A},\nabla_E)$ ,  
parallelly transported  $X$  along  $\gamma$ , parallel  $X$  along  $\gamma$ , not parallel, not parallelly transported.

**2. Autoparallely transported curves.** Thought bubble: self- is "Auto"

Child: "How do I get to the toy store." Adult: "Just follow your nose."

(Move in direction of nose means move along curve whose tangent vector is your nose. Means move in the straightest line. We are not talking about shortest line because we do not have a notion of distance.)

**Definition 95.** *A (smooth) curve  $\gamma : \mathbb{R} \rightarrow M$  is called **autoparallely transported** if*

(221)

$$\nabla_{v_\gamma} v_\gamma = 0$$

$$\Longleftrightarrow \left(\nabla_{v_{\gamma,\gamma(\lambda)}} v_\gamma\right)_{\gamma(\lambda)} = 0$$

Remark: Define weaker notion of autoparallel curve  
Thought bubble: !

$$\nabla_{v_\gamma} v_\gamma = \mu v_\gamma$$

Example. Euclidean plane  $(\mathbb{R}^2,\mathcal{O},\mathcal{A},\nabla_E)$   
autoparallely transported (uniform straight motion, Newton's First law), autoparallel (straight motion) but not autoparallely transported,

**3. Autoparallel equation.**

Autoparallely transported curve  $\gamma$

Consider that portion of the curve that lies in  $U$ , where  $(U,x) \in \mathcal{A}$ .

$$\nabla_{v_\gamma} v_\gamma = 0$$

Express  $\nabla_{v_\gamma} v_\gamma = 0$  in terms of chart representation.

$$(\nabla_{v_\gamma} v_\gamma) = \left(\nabla_{\dot{\gamma}^m(x)\left(\frac{\partial}{\partial x^m}\right)_{\gamma(\lambda)}} \dot{\gamma}^n(x) \frac{\partial}{\partial x^n}\right)$$

where

$$v_{\gamma,\gamma(\lambda)} = \dot{\gamma}^m(x) \cdot \left(\frac{\partial}{\partial x^m}\right)$$
$$\gamma^m(x) := x^m \circ \gamma$$

$$\begin{aligned} &\implies (\nabla_{v_\gamma} v_\gamma) = \\ &= \underbrace{\dot{\gamma}^m \frac{\partial \dot{\gamma}^q}{\partial x^m}}_{\ddot{\gamma}^q(x)^m} \cdot \frac{\partial}{\partial x^q} + \dot{\gamma}^m \dot{\gamma}^n \Gamma^q_{nm} \frac{\partial}{\partial x^q} \end{aligned}$$

20200306 EY: For further clarification,

$$\begin{aligned} \dot{\gamma}(t) &= v(\gamma(t)) \in T_{\gamma(t)}U \\ \dot{\gamma}(t) &= v^i(\gamma^1(t), \dots, \gamma^n(t)) \\ \frac{d}{dt} \dot{\gamma}^i(t) &= \frac{d}{dt} v^i(\gamma^1(t), \dots, \gamma^n(t)) = \frac{d}{dt} v^i(x^1, \dots, x^n) \\ &= \frac{\partial v^i}{\partial x^j}(x) \dot{x}^j = \dot{\gamma}^j \frac{\partial \dot{\gamma}^i}{\partial x^j} \end{aligned}$$

in summary:

(222)

$$\ddot{\gamma}^m_{(x)}(\lambda) + (\Gamma^m_{(x)})_{ab}(\gamma(\lambda)) \dot{\gamma}^a_{(x)}(\lambda) \dot{\gamma}^b_{(x)}(\lambda) = 0$$

chart expression of the condition that  $\gamma$  be autoparallely transported.

*Examples.*

(a) Euclidean plane,  $U = \mathbb{R}^2$ ,  $x = \text{id}_{\mathbb{R}}$ ,  $\Gamma(x)^i_{jk} = 0$

$$\implies \ddot{\gamma}(x)^m = 0 \implies \gamma(x)^m(\lambda) = a^m \lambda + b^m a, b \in \mathbb{R}^d$$

(b) round sphere  $(S^2,\mathcal{O},\mathcal{A},\nabla_{\text{round}})$

Consider a chart:

$$\begin{aligned} x(p) &= (\theta, \varphi) \\ \theta &\in (0, \pi) \\ \varphi &\in (0, 2\pi) \end{aligned}$$

all other  $(\Gamma^1_{11} = 0)$  vanish.

$$\Gamma(x)^1_{22}(x^{-1}(\theta,p)) := -\sin \theta \cos \theta \quad \Gamma(x)^2_{21} = \Gamma(x)^2_{12} := \cot \theta$$

Sloppy notation (classical mechanics)

$$\begin{aligned} x^1(p) &= \theta(p) \\ x^2(p) &= \varphi(p) \end{aligned}$$

autoparallel equation

$$\begin{aligned} \ddot{\theta} + \Gamma^1_{22} \dot{\varphi} \dot{\varphi} &= 0 \\ \ddot{\varphi} + 2\Gamma^2_{12} \dot{\theta} \dot{\varphi} &= 0 \end{aligned}$$

$$\Longleftrightarrow \begin{aligned} \ddot{\theta} - \sin(\theta) \cos(\theta) \dot{\varphi} \dot{\varphi} &= 0 \\ \ddot{\varphi} + 2 \cot(\theta) \dot{\theta} \dot{\varphi} &= 0 \end{aligned}$$

Solve: e.g.  $\theta(\lambda) = \frac{\pi}{2} \quad \varphi(\lambda) = \underline{\omega} \cdot \lambda + \varphi_0$   
Similarly,  
 $\nabla_{\text{potato}}, (S^2,\mathcal{O},\mathcal{A})$

**3. Torsion.** Q: Can one use  $\nabla$  to define tensors on  $(M, \mathcal{O}, \mathcal{A}, \nabla)$ ?

**Definition 96.** The *torsion* of a connection  $\nabla$  is the  $(1, 2)$ -**tensor field**

(223) 
$$T(\omega, X, Y) := \omega(\nabla_X Y - \nabla_Y X - [X, Y])$$

(Inside a cloud)  
[X, Y] vector field defined by

$$[X, Y]f := X(Yf) - Y(Xf)$$

*Proof.* check  $T$  is  $C^\infty$ -linear in each entry

$$\begin{aligned} T(f \cdot \omega, X, Y) &= f \cdot \omega(\dots) = fT(\omega, X, Y) \\ T(\omega + \psi, X, Y) &= \dots = T(\omega, X, Y) + T(\psi, X, Y) \\ T(\omega, fX, Y) &= \omega(\nabla_{fX} Y - \nabla_Y(fX) - [fX, Y]) = \\ &= \omega(f\nabla_X Y - (Yf)X - f\nabla_Y X - f[X, Y] + (Yf)X) = fT(\omega, X, Y) \end{aligned}$$

since

$$\begin{aligned} [fX, Y]g &= fX(Yg) - Y(fXg) = \\ &= fX(Yg) - (Yf)(Xg) - fY(Xg) = \\ &= f \cdot [X, Y]g - (Yf)Xg = \\ &= (f \cdot [X, Y] - (Yf)X)g \\ T(\omega, X, Y) &= -T(\omega, Y, X) \quad \checkmark \end{aligned}$$

**Definition 97.** A  $(M, \mathcal{O}, \mathcal{A}, \nabla)$  is called *torsion-free* if  $T = 0$

In a chart

$$\begin{aligned} T^i_{ab} &:= T\left(dx^i, \frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b}\right) = dx^i(\dots) \\ &= \Gamma^i_{ab} - \Gamma^i_{ba} = 2\Gamma^i_{[ab]} \end{aligned}$$

From now on, in these lectures, we only use torsion-free connections.

**4. Curvature.**

**Definition 98.** The *Riemann curvature* of a connection  $\nabla$  is the  $(1, 3)$ -tensor field

(224) 
$$Riem(\omega, Z, X, Y) := \omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z)$$

*Proof.* do it:  $C^\infty$ -linear in each slot.

Tutorials  $Riem^i_{jab} = \dots$

*Algebraic relevance of Riem.*

$$(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z) = R(\cdot, Z, X, Y) + \nabla_{[X, Y]} Z$$

In one chart  $(U, x)$

$$\underbrace{\left(\nabla_a\right)}_{=\nabla_{\frac{\partial}{\partial x^a}}} \nabla_b Z - \nabla_b \nabla_a Z)^m = Riem^m_{\phantom{m}nab} Z^n + \nabla \left[ \begin{array}{c} \frac{\partial}{\partial x^a}, \frac{\partial}{\partial x^b} \\ \underbrace{\hspace{1cm}}_{=0} \end{array} \right] Z$$

$$(\nabla_a \nabla_b Z)^m - (\nabla_b \nabla_a Z)^m = Riem^m_{\phantom{m}nab} Z^n$$

*Geometric significance.*

Idea:  
[X, Y] = 0

If [X, Y]  $\neq$  0

$$(\delta Z)^m = \dots = R^m_{\phantom{m}nab} X^a Y^b Z^n \delta s \delta t + \mathcal{O}(\delta s^2 \delta t, \delta t^2 \delta s)$$

Holonomy

TUTORIAL 8 PARALLEL TRANSPORT & CURVATURE

**Exercise 1.**

**Exercise 2. : Where connection coefficients appear**

It was suggested in the tutorial sheets and hinted in the lecture that the following should be committed to memory.

**Question : Recall the autoparallel equation for a curve  $\gamma$ .**

(a)

$$\nabla_{v_\gamma} v_\gamma = 0$$

(b)

□

$$\begin{aligned} \nabla_{v_\gamma} v_\gamma &= \nabla_{\dot{\gamma} \frac{\partial}{\partial x^\mu}} v_\gamma = \dot{\gamma}^\nu \nabla_{\partial_\nu} v_\gamma = \dot{\gamma}^\nu \left[ \frac{\partial v^\mu_\gamma}{\partial x^\nu} + \Gamma^\rho_{\mu\nu} v^\mu_\gamma \right] \frac{\partial}{\partial x^\rho} = \dot{\gamma}^\nu \left[ \frac{\partial \dot{\gamma}^\rho}{\partial x^\nu} + \Gamma^\rho_{\mu\nu} \dot{\gamma}^\mu \right] \frac{\partial}{\partial x^\rho} = 0 \\ &\implies \boxed{\ddot{\gamma}^\rho + \Gamma^\rho_{\mu\nu} \dot{\gamma}^\mu \dot{\gamma}^\nu} \end{aligned}$$

as, for example, for  $F(x(t))$ ,

$$\frac{dF(x(t))}{dt} = \dot{x} \frac{\partial F}{\partial x} = \frac{d}{dt} F$$

so that

$$\dot{\gamma}^\nu \frac{\partial v^\mu_\gamma}{\partial x^\nu} = \frac{d}{d\lambda} v^\mu_\gamma = \frac{d^2}{d\lambda^2} \gamma^\mu$$

**Question : Determine the coefficients of the Riemann tensor with respect to a chart  $(U, x)$ .**

Recall this manifestly covariant definition

□

$$Riem(\omega, Z, X, Y) = \omega(\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X, Y]} Z)$$

We want  $R^i_{\phantom{i}jab}$ .  
now

$$\nabla_X \nabla_Y Z = \nabla_X \left( (Y^\mu \frac{\partial}{\partial x^\mu} Z^\rho + \Gamma^\rho_{\mu\nu} Z^\mu Y^\nu) \frac{\partial}{\partial x^\rho} \right) = (X^\alpha \frac{\partial}{\partial x^\alpha} (Y^\mu \frac{\partial}{\partial x^\mu} Z^\rho + \Gamma^\rho_{\mu\nu} Z^\mu Y^\nu) + \Gamma^\rho_{\alpha\beta} (Y^\mu \frac{\partial}{\partial x^\mu} Z^\alpha + \Gamma^\alpha_{\mu\nu} Z^\mu Y^\nu) X^\beta) \frac{\partial}{\partial x^\rho}$$

For  $X = \partial_a, Y = \partial_b, Z = \partial_j$ , then the partial derivatives of the coefficients of the input vectors become zero.

$$\implies \nabla_{\partial_a} \nabla_{\partial_b} \partial_j = \frac{\partial}{\partial x^a} (\Gamma^i_{\phantom{i}jb}) + \Gamma^i_{\phantom{i}a\alpha} \Gamma^\alpha_{\phantom{\alpha}jb}$$

Now

$$[X, Y]^i = X^j \frac{\partial}{\partial x^j} Y^i - Y^j \frac{\partial X^i}{\partial x^j}$$

For coordinate vectors,  $[\partial_i, \partial_j] = 0 \ \forall i, j = 0, 1 \dots d$ .



Thus

$$R^i{}_{jab} = \frac{\partial}{\partial x^a} \Gamma^i_{jb} - \frac{\partial}{\partial x^b} \Gamma^i_{ja} + \Gamma^i_{\alpha a} \Gamma^\alpha_{jb} - \Gamma^i_{\alpha b} \Gamma^\alpha_{ja}$$

**Question :** $\text{Ric}(X, Y) := \text{Riem}^m_{amb} X^a Y^b$  **define**  $(0, 2)$ -**tensor?**.

Yes, transforms as such:

**EY developments.** I roughly follow the spirit in Theodore Frankel’s **The Geometry of Physics: An Introduction** Second Ed. 2003, Chapter 9 Covariant Differentiation and Curvature, Section 9.3b. The Covariant Differential of a Vector Field. P.S. EY : 20150320 I would like a copy of the Third Edition but I don’t have the funds right now to purchase the third edition: go to my tilt crowdfunding campaign, <http://ernestyalumni.tilt.com>, and help with your financial support if you can or send me a message on my various channels and ernestyalumni gmail email address if you could help me get a hold of a digital or hard copy as a pro bono gift from the publisher or author. EY : 20221220 Update: I’m gainfully employed now and I’ve obtained a digital copy of the 3rd Ed. of Frankel! I don’t have the tilt campaign, but if you’d like to donate still, PayPal me [@ernestyalumni](#) or better yet Venmo me @ernestyalumni

The spirit of the development is the following:

“How can we express connections and curvatures in terms of forms?” -Theodore Frankel.

From Lecture 7, connection  $\nabla$  on vector field  $Y$ , in the “direction”  $X$ ,

$$\nabla_{\frac{\partial}{\partial x^k}} Y = \left( \frac{\partial Y^i}{\partial x^k} + \Gamma^i_{jk} Y^j \right) \frac{\partial}{\partial x^i}$$

Make the ansatz (approche, impostazione) that the connection  $\nabla$  acts on  $Y$ , the vector field, first:

$$\nabla Y(X) = \left( X^k \frac{\partial Y^i}{\partial x^k} + \Gamma^i_{jk} Y^j X^k \right) \frac{\partial}{\partial x^i} = X^k \left( \nabla_{\frac{\partial}{\partial x^k}} Y \right)^i \frac{\partial}{\partial x^i} = (\nabla_X Y)^i \frac{\partial}{\partial x^i} = \nabla_X Y$$

Now from Lecture 7, Definition for  $\Gamma$ ,

$$dx^i \left( \nabla_{\frac{\partial}{\partial x^k}} \frac{\partial}{\partial x^j} \right) = \Gamma^i_{jk}$$

Make this ansatz (approche, impostazine)

$$\nabla \frac{\partial}{\partial x^j} = (\Gamma^i_{jk} dx^k) \otimes \frac{\partial}{\partial x^i} \in \Omega^1(M, TM) = T^*M \otimes TM$$

where  $\Omega^1(M, TM) = T^*M \otimes TM$  is the set of all  $TM$  or vector-valued 1-forms on  $M$ , with the 1-form being the following:

$$\Gamma^i_{jk} dx^k = \Gamma^i_j \in \Omega^1(M) \quad \begin{array}{l} i = 1 \dots \dim(M) \\ j = 1 \dots \dim(M) \end{array}$$

So  $\Gamma^i_j$  is a  $\dim M \times \dim M$  matrix of 1-forms (EY !!!).

Thus

$$\nabla Y = (d(Y^i) + \Gamma^i_j Y^j) \otimes \frac{\partial}{\partial x^i}$$

So the connection is a (smooth) map from  $TM$  to the set of all vector-valued 1-forms on  $M$ ,  $\Omega^1(M, TM)$ , and then, after “eating” a vector  $Y$ , yields the “covariant derivative”:

$$\begin{aligned} \nabla : TM &\rightarrow \Omega^1(M, TM) = T^*M \otimes TM \\ \nabla : Y &\mapsto \nabla Y \\ \nabla Y : TM &\rightarrow TM \\ \nabla Y(X) &\mapsto \nabla Y(X) = \nabla_X(Y) \end{aligned}$$

Now

$$\left[ \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right] f = \frac{\partial}{\partial x^i} \left( \frac{\partial}{\partial x^j} \right) - \frac{\partial}{\partial x^j} \left( \frac{\partial}{\partial x^i} \right) = 0$$

(this is okay as on  $p \in (U, x)$ ;  $x$ -coordinates on same chart  $(U, x)$ )

EY : 20150320 My question is when is this nontrivial or nonvanishing (i.e. not equal to 0).

$$[e_a, e_b] = ?$$

for a frame  $(e_c)$  and would this be the difference between a tangent bundle  $TM$  vs. a (general) vector bundle?

Wikipedia helps here. cf. wikipedia, “Connection (vector bundle)”

$$\nabla : \Gamma(E) \rightarrow \Gamma(T^*M \otimes E) = \Omega^1(M, E)$$

$$\nabla e_a = \omega^c_{ab} f^b \otimes e_c$$

$f^b \in T^*M$  (this is the dual basis for  $TM$  and, note, this is for the manifold,  $M$

$$\nabla_{f_b} e_a = \omega^c_{ab} e_c \in E$$

$$\omega^c_a = \omega^c_{ab} f^b \in \Omega^1(M)$$

is the connection 1-form, with  $a, c = 1 \dots \dim V$ . EY : 20150320 This  $V$  is a vector space living on each of the fibers of  $E$ . I know that  $\Gamma(T^*M \otimes E)$  looks like it should take values in  $E$ , but it’s meaning that it takes vector values of  $V$ . Correct me if I’m wrong: ernestyalumni at gmail and various social media.

Let  $\sigma \in \Gamma(E)$ ,  $\sigma = \sigma^a e_a$

$$\nabla \sigma = (d\sigma^c + \omega^c_{ab} \sigma^a f^b) \otimes e_c \text{ with}$$

$$d\sigma^c = \frac{\partial \sigma^c}{\partial x^b} f^b$$

$$\implies \nabla_X \sigma = \left( X^b \frac{\partial \sigma^c}{\partial x^b} + \omega^c_{ab} \sigma^a X^b \right) e_c = X^b \left( \frac{\partial \sigma^c}{\partial x^b} + \omega^c_{ab} \sigma^a \right) e_c$$

LECTURE 9: NEWTONIAN SPACETIME IS CURVED!

**Axiom 1** (Newton I:). *A body on which no force acts moves uniformly along a straight line*

**Axiom 2** (Newton II:). *Deviation of a body’s motion from such uniform straight motion is effected by a force, reduced by a factor of the body’s reciprocal mass.*

Remark:

- (1) 1st axiom - in order to be relevant - must be read as a measurement prescription for the geometry of space ...
- (2) Since gravity universally acts on every particle, in a universe with at least two particles, gravity must not be considered a force if Newton I is supposed to remain applicable.

31.11. **Laplace’s questions.** Laplace \*1749

†1827

Q: “Can gravity be encoded in a curvature of space, such that its effects show if particles under the influence of (no other) force we postulated to move along straight lines in this curved space?”

Answer: No!

*Proof.* gravity is a force point of view

$$\begin{aligned} m\ddot{x}^\alpha(t) &= F^\alpha(x(t)) \\ m\dot{x}^\alpha(t) &= \underbrace{mf^\alpha}_{F^\alpha}(x(t)) \end{aligned}$$

$-\partial_\alpha f^\alpha = 4\pi G\rho$  (Poisson)  
 $\rho$  mass density of matter

$$m\ddot{x}^\alpha(t) = \underbrace{mf^\alpha}_{F^\alpha}(x(t))$$

True?  
(EY : 20150330) You know this,  $F = Gm_1m_2/r^2$   
Yes  
weak equivalence principle

$$\ddot{x}^\alpha(t) - f^\alpha(x(t)) = 0$$

Laplace asks: Is this ( $\ddot{x}(t)$ ) of the form

$$\ddot{x}^\alpha(t) + \Gamma^\alpha_{\beta\gamma}(x(t))\dot{x}^\beta(t)\dot{x}^\gamma(t) = 0$$

Conclusion: One cannot find  $\Gamma$ ’s such that Newton’s equation takes the form of an autoparallel.

□

Question (from audience) We can evaluate the autoparallel equation pointwise?! But at each point, we can set the Gammas to zero?!

Then, one should be able to write Newton’s second law in the usual form?

Prof. Schuller: you (observer) fall with the mass (i.e. accelerated reference frame) and so you transform  $\Gamma$ ’s to be 0. The problem with this is if you do the same experiment in the North pole and fall with the body. If someone else at the South Pole does the same experiment at the same time, with that same transformation (of reference frames), the effect of gravity cannot be transformed out.

In a homogeneous gravitational field, you can possibly transform away gravity,  $\Gamma = 0$ . But in an inhomogeneous gravitational field, no.

**2. The full wisdom of Newton I.** use also the information from Newton’s first law that particles (no force) move uniformly introduce the appropriate setting to talk about the difference easily  
insight: in spacetime uniform & straight motion in space is simply **straight** motion  
So let’s try in spacetime:  
let  $x : \mathbb{R} \rightarrow \mathbb{R}^3$

$$\begin{aligned} X : \mathbb{R} &\rightarrow \mathbb{R}^4 \\ t &\mapsto (t, x^1(t), x^2(t), x^3(t)) := \\ &:= (X^0(t), X^1(t), X^2(t), X^3(t)) \end{aligned}$$

be a particle’s trajectory in space  $\longleftrightarrow$  worldline (history) of the particle

That’s all it takes:

Trivial rewritings:

$$\dot{X}^0 = 1$$

$$\begin{aligned} \implies \begin{cases} \ddot{X}^0 &= 0 \\ \ddot{X}^\alpha - f^\alpha(X(t)) \cdot \dot{X}^0 \cdot \dot{X}^0 &= 0 \end{cases} & (\alpha = 1, 2, 3) \implies \begin{cases} a = 0, 1, 2, 3 \\ \ddot{X}^a + \Gamma^a_{bc} \dot{X}^b \dot{X}^c = 0 \end{cases} \\ & \text{antoparallel eqn in } \underline{\text{spacetime}} \end{aligned}$$

Yes, choosing  $\Gamma^0_{ab} = 0$   
 $\Gamma^\alpha_{\beta\gamma} = 0 = \Gamma^\alpha_{0\beta} = \Gamma^\alpha_{\beta 0}$

only:  $\Gamma^\alpha_{00} \stackrel{!}{=} -f^\alpha$

Question: Is this a coordinate-choice artifact?

No, since  $R^\alpha_{0\beta 0} = -\frac{\partial}{\partial x^\beta} f^\alpha$  (only non-vanishing components) (tidal force tensor, – the Hessian of the force component)

Ricci tensor  $\implies R_{00} = R^m_{0m0} = -\partial_\alpha f^\alpha = 4\pi G\rho$

Poisson:  $-\partial_\alpha f^\alpha = 4\pi G \cdot \rho$

writing:  $T_{00} = \frac{1}{2}\rho$

$$\implies \boxed{R_{00} = 8\pi GT_{00}}$$

Einstein in 1912  $R_{ab} \longrightarrow 8\pi GT_{ab}$

Conclusion: Laplace’s idea works in spacetime

Remark

$$\begin{aligned} \Gamma^\alpha_{00} &= -f^\alpha \\ R^\alpha_{\beta\gamma\delta} &= 0 \quad \alpha, \beta, \gamma, \delta = 1, 2, 3 \\ \boxed{R_{00} = 4\pi G\rho} \end{aligned}$$

Q: What about transformation behavior of LHS of

$$\underbrace{\ddot{x}^a + \Gamma^a_{bc} \dot{X}^b \dot{X}^c}_{\underbrace{(\nabla_{v_X} v_X)^a}_{:= a^a \text{ “acceleration vector”}}} = 0$$

**3. The foundations of the geometric formulation of Newton’s axiom.** new start

**Definition 99.** A *Newtonian spacetime* is a quintuple

$$(M, \mathcal{O}, \mathcal{A}, \nabla, t)$$

where  $(M, \mathcal{O}, \mathcal{A})$  4-dim. smooth manifold

$$t : M \rightarrow \mathbb{R} \text{ smooth function}$$

(i) “There is an absolute space”

$$(dt)_p \neq 0 \quad \forall p \in M$$

(ii) “absolute time flows uniformly”

$$\nabla dt \underbrace{=}_\text{space of (0,2)-tensor fields} 0 \quad \text{everywhere}$$

$\nabla dt$  is a (0,2)-tensor field

(iii) add to axioms of Newtonian spacetime  $\nabla = 0$  torsion free

**Definition 100.** *absolute space at time  $\tau$*

$$S_\tau := \{p \in M | t(p) = \tau\}$$
$$\xrightarrow{dt \neq 0} M = \coprod S_\tau$$

**Definition 101.** *A vector  $X \in T_pM$  is called*

(a) *future-directed if*

$$dt(X) > 0$$

(b) *spatial if*

$$dt(X) = 0$$

(c) *past-directed if*

$$dt(X) < 0$$

picture  
Newton I: The worldline of a particle under the influence of no force (gravity isn't one, anyway) is a future-directed autoparallel i.e.

$$\nabla_{v_X} v_X = 0$$
$$dt(v_X) > 0$$

and (iii)  $\nabla$  is torsion-free.

Newton II:

$$\nabla_{v_X} v_X = \frac{F}{m} \iff m \cdot \mathfrak{a} = F$$

where  $F$  is a spatial vector field:

$$dt(F) = 0$$

**Convention:** restrict attention to atlases  $\mathcal{A}_{\text{stratified}}$  whose charts  $(\mathcal{U}, x)$  have the property

$$\begin{array}{l} x^0 : \mathcal{U} \rightarrow \mathbb{R} \\ x^1 : \mathcal{U} \rightarrow \mathbb{R} \\ \vdots \\ x^3 \end{array} \quad x^0 = t|_{\mathcal{U}} \quad \implies \quad \begin{array}{l} 0 \\ 0 = \nabla_{\frac{\partial}{\partial x^a}} dx^0 = -\Gamma_{ba}^0 \end{array} \quad \begin{array}{l} \text{“absolute time flows uniformly”} \\ a = 0, 1, 2, 3 \end{array} \quad \nabla dt$$

Let's evaluate in a chart  $(\mathcal{U}, x)$  of a stratified atlas  $\mathcal{A}_{\text{sheet}}$ : Newton II:

$$\nabla_{v_X} v_X = \frac{F}{m}$$

in a chart.

$$(X^0)'' + \Gamma_{cd}^0 (X^a)' (X^b)' \Big|_{\text{stratified atlas}} = 0$$
$$(X^\alpha)'' + \Gamma_{\gamma\delta}^\alpha X^{\gamma'} X^{\delta'} + \Gamma_{00}^\alpha X^{0'} X^{0'} + 2\Gamma_{\gamma 0}^\alpha X^{\gamma'} X^{0'} = \frac{F^\alpha}{m} \quad \alpha = 1, 2, 3$$

EY: 20160623: where the factor of 2 comes from torsion free  $\nabla$

$$\implies (X^0)''(\lambda) = 0 \implies X^0(\lambda) = a\lambda + b \quad \text{constants } a, b \quad \text{with}$$

$$X^0(\lambda) = (x^0 \circ X)(\lambda) \stackrel{\text{stratified}}{=} (t \circ X)(\lambda)$$

convention parametrize worldline by absolute time

$$\frac{d}{d\lambda} = a \frac{d}{dt}$$

$$a^2 \ddot{X}^\alpha + a^2 \Gamma_{\gamma\delta}^\alpha \dot{X}^\gamma \dot{X}^\delta + a^2 \Gamma_{00}^\alpha \dot{X}^0 \dot{X}^0 + 2a^2 \Gamma_{\gamma 0}^\alpha \dot{X}^\gamma \dot{X}^0 = \frac{F^\alpha}{m}$$
$$\implies \underbrace{\ddot{X}^\alpha + \Gamma_{\gamma\delta}^\alpha \dot{X}^\gamma \dot{X}^\delta + \Gamma_{00}^\alpha \dot{X}^0 \dot{X}^0 + 2\Gamma_{\gamma 0}^\alpha \dot{X}^\gamma \dot{X}^0}_{\mathfrak{a}^\alpha} = \frac{1}{a^2} \frac{F^\alpha}{m}$$

Student question: How can you get the Lagrange equations in this formalism?

Schuller says: In *all* systems they take the form like this:

$$L = \underbrace{T}_{\frac{m}{2} g(v_x, v_x)} + V$$

where  $v_x$  is velocity to the spatial curves  $x$ .

Schuller says:  $\frac{m}{2} \sqrt{g(v_x, v_x)}$ , parameter invariance, you get straight curves; if you take  $\sqrt{\phantom{x}}$  away,  $\frac{m}{2} g(v_x, v_x)$  you get straight, *uniform* curves.

Schuller says: Only in inertial systems, which is defined as:

$$\nabla_{\frac{\partial}{\partial x^a}} \frac{\partial}{\partial x^b} = 0$$

inertial frame, non-accelerating, non-rotating.

Schuller says: Apart from  $g$ , a 2-tensor, you need to know  $\omega$ .

Schuller references 2nd. year Mechanik in German: <https://www.video.uni-erlangen.de/course/id/272>

32. LECTURE 10: METRIC MANIFOLDS

cf. [Lecture 10: Metric Manifolds \(International Winter School on Gravity and Light 2015\)](#)

We establish a structure on a smooth manifold that allows one to assign vectors in each tangent space a length (and an angle between vectors in the same tangent space).

From this structure, one can then define a notion of length of a curve.

Then we can look at shortest curves.

Requiring then that the shortest curves coincide with the straightest curves (wrt  $\nabla$ ) will result in  $\nabla$  being determined by the metric structure.

$$g \stackrel{\text{straight=short}}{\underset{T=0}{\rightsquigarrow}} \nabla \rightsquigarrow \text{Riem}$$

32.1. Metrics.

**Definition 102.** *A metric  $g$  on a smooth manifold  $(M, \mathcal{O}, \mathcal{A})$  is a  $(0, 2)$ -tensor field satisfying*

- (i) *symmetry*  $g(X, Y) = g(Y, X) \quad \forall X, Y \text{ vector fields}$
- (ii) *non-degeneracy: the musical map*

$$\text{“flat” } \flat : \Gamma(TM) \rightarrow \Gamma(T^*M)$$
$$X \mapsto \flat(X)$$

$$\text{where } \flat(X)(Y) := g(X, Y)$$

$$\flat(X) \in \Gamma(T^*M)$$

$$\text{In thought bubble: } \flat(X) = g(X, \cdot)$$

... is a  $C^\infty$ -isomorphism in other words, it is invertible.

Remark:  $(\flat(X))_a$  or

$X_a$

$$(\flat(X))_a := g_{am} X^m$$

Thought bubble:  $\flat^{-1} = \sharp$

$\flat^{-1}(\omega)^a := g^{am}\omega_m$   
 $\flat^{-1}(\omega)^a := (g^{\text{“}^{-1}\text{”}})^{am}\omega_m \implies \text{not needed. (all of this is not needed)}$

**Definition 103.** The  $(2,0)$ -tensor field  $g^{\text{“}^{-1}\text{”}}$  with respect to a metric  $g$  is the symmetric

$$g^{\text{“}^{-1}\text{”}} : \Gamma(T^*M) \times \Gamma(T^*M) \rightarrow C^\infty(M)$$
$$(\omega, \sigma) \mapsto \omega(\flat^{-1}(\sigma)) \qquad \flat^{-1}(\sigma) \in \Gamma(TM))$$

chart:  $g_{ab} = g_{ba}$   
 $(g^{-1})^{am}g_{mb} = \delta_b^a$

Example:  $(S^2, \mathcal{O}, \mathcal{A})$   
chart  $(\mathcal{U}, x)$   
 $\varphi \in (0, 2\pi)$   
 $\theta \in (0, \pi)$   
define the metric

$$g_{ij}(x^{-1}(\theta, \varphi)) = \begin{bmatrix} R^2 & 0 \\ 0 & R^2 \sin^2 \theta \end{bmatrix}_{ij}$$

$R \in \mathbb{R}^+$   
“the metric of the round sphere of radius  $R$ ”

32.2. **Signature.** Linear algebra:

$A^a{}_m v^m = \lambda v^a$

$g_{am}v^m = \lambda \cdot v^a? \rightsquigarrow$

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & & & & & & 0 \\ & & & & & & & \ddots \\ & & & & & & & & 0 \end{pmatrix}$$

$(1,1)$  tensor has eigenvalues  
 $(0,2)$  has signature  $(p,q)$  (well-defined)

$\left. \begin{matrix} (+ + +) \\ (+ + -) \\ (+ - -) \\ (- - -) \end{matrix} \right\} d + 1 \text{ if } p + q = \dim V$

**Definition 104.** A metric is called **Riemannian** if its signature is  $(+ + \cdots +)$   
**Lorentzian** if  $(+ - \cdots -)$

**32.3. Length of a curve.** Let  $\gamma$  be a smooth curve.  
Then we know its veloci<sup>ti</sup>y  $v_{\gamma, \gamma(\lambda)}$  at each  $\gamma(\lambda) \in M$ .

**Definition 105.** On a Riemannian metric manifold  $(M, \mathcal{O}, \mathcal{A}, g)$ , the **speed** of a curve at  $\gamma(\lambda)$  is the number

$$(\sqrt{g(v_\gamma, v_\gamma)})_{\gamma(\lambda)} = s(\lambda)$$

F. Schuller: “I feel the need for speed.” -Top Gun.  
(I feel the need for speed, then I feel the need for a metric)

Aside:  $[v^a] = \frac{1}{T}$   
 $[g_{ab}] = L^2$   
 $[\sqrt{g_{ab}v^av^b}] = \sqrt{\frac{L^2}{T^2}} = \frac{L}{T}$

**Definition 106.** Let  $\gamma : (0,1) \rightarrow M$  a smooth curve.  
Then the **length of**  $\gamma$  is the number

$$\mathbb{R} \ni L[\gamma] := \int_0^1 d\lambda s(\lambda) = \int_0^1 d\lambda \sqrt{(g(v_\gamma, v_\gamma))_{\gamma(\lambda)}}$$

F. Schuller: “velocity is more fundamental than speed, speed is more fundamental than length”

Example: reconsider the round sphere of radius  $R$   
Consider its equator:

$\theta(\lambda) := (x^1 \circ \gamma)(\lambda) = \frac{\pi}{2}$   
 $\varphi(\lambda) := (x^2 \circ \gamma)(\lambda) = 2\pi\lambda^3$   
 $\theta'(\lambda) = 0$   
 $\varphi'(\lambda) = 6\pi\lambda^2$

on the same chart  $g_{ij} = \begin{bmatrix} R^2 & \\ & R^2 \sin^2 \theta \end{bmatrix}$

F.Schuller: do everything in this chart

$$L[\gamma] = \int_0^1 d\lambda \sqrt{g_{ij}(x^{-1}(\theta(\lambda), \varphi(\lambda)))(x^i \circ \gamma)'(\lambda)(x^j \circ \gamma)'(\lambda)} = \int_0^1 d\lambda \sqrt{R^2 \cdot 0 + R^2 \sin^2(\theta(\lambda))36\pi^2\lambda^4} =$$
$$= 6\pi R \int_0^1 d\lambda \lambda^2 = 6\pi R [\frac{1}{3}\lambda^3]_0^1 = 2\pi R$$

**Theorem 31.**  $\gamma : (0,1) \rightarrow M$  and  
 $\sigma : (0,1) \rightarrow (0,1)$  smooth bijective and increasing “reparametrization”

$$L[\gamma] = L[\gamma \circ \sigma]$$

*Proof.*  $\implies$  Tutorials

□

32.4. **Geodesics.**

**Definition 107.** A curve  $\gamma : (0,1) \rightarrow M$  is called a **geodesic** on a Riemannian manifold  $(M, \mathcal{O}, \mathcal{A}, g)$  if its a stationary curve with respect to a length functional  $L$ .

Thought bubble: in classical mechanics, deform the curve a little,  $\epsilon$  times this deformation, to first order, it agrees with  $L[\gamma]$

**Theorem 32.**  $\gamma$  geodesic iff it satisfies the Euler-Lagrange equations for the Lagrangian

$$\mathcal{L} : TM \rightarrow \mathbb{R}$$
$$X \mapsto \sqrt{g(X, X)}$$

In a chart, the Euler Lagrange equations take the form:

$$\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^m}\right)^{\cdot} - \frac{\partial \mathcal{L}}{\partial x^m} = 0$$

F.Schuller: this is a chart dependent formulation  
here:

$$\mathcal{L}(\gamma^i, \dot{\gamma}^i) = \sqrt{g_{ij}(\gamma(\lambda)) \dot{\gamma}^i(\lambda) \dot{\gamma}^j(\lambda)}$$

Euler-Lagrange equations:

$$\frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m} = \frac{1}{\sqrt{\dots}} g_{mj}(\gamma(\lambda)) \dot{\gamma}^j(\lambda)$$
$$\left(\frac{\partial \mathcal{L}}{\partial \dot{\gamma}^m}\right)^{\cdot} = \left(\frac{1}{\sqrt{\dots}}\right)^{\cdot} g_{mj}(\gamma(\lambda)) \cdot \dot{\gamma}^j(\lambda) + \frac{1}{\sqrt{\dots}} (g_{mj}(\gamma(\lambda)) \ddot{\gamma}^j(\lambda) + \dot{\gamma}^s(\partial_s g_{mj}) \dot{\gamma}^j(\lambda))$$

Thought bubble: reparametrize  $g(\dot{\gamma}, \dot{\gamma}) = 1$  (it's a condition on my reparametrization)

By a clever choice of reparametrization  $(\frac{1}{\sqrt{\dots}})^{\cdot} = 0$

$$\frac{\partial \mathcal{L}}{\partial \gamma^m} = \frac{1}{2\sqrt{\dots}} \partial_m g_{ij}(\gamma(\lambda)) \dot{\gamma}^i(\lambda) \dot{\gamma}^j(\lambda)$$

putting this together as Euler-Lagrange equations:

$$g_{mj} \ddot{\gamma}^j + \partial_s g_{mj} \dot{\gamma}^s \dot{\gamma}^j - \frac{1}{2} \partial_m g_{ij} \dot{\gamma}^i \dot{\gamma}^j = 0$$

Multiply on both sides  $(g^{-1})^{qm}$

$$\ddot{\gamma}^q + (g^{-1})^{qm} (\partial_i g_{mj} - \frac{1}{2} \partial_m g_{ij}) \dot{\gamma}^i \dot{\gamma}^j = 0$$

$$\ddot{\gamma}^q + (g^{-1})^{qm} \frac{1}{2} (\partial_i g_{mj} + \partial_j g_{mi} - \partial_m g_{ij}) \dot{\gamma}^i \dot{\gamma}^j = 0$$

geodesic equation for  $\gamma$  in a chart.

$$(g^{-1})^{qm} \frac{1}{2} (\partial_i g_{mj} + \partial_j g_{mi} - \partial_m g_{ij}) =: \Gamma^q_{ij}(\gamma(\lambda))$$

Thought bubble:  $\left(\frac{\partial \mathcal{L}}{\partial \xi^{a+\dim M}}\right)^{\cdot}_{\sigma(x)} - \left(\frac{\partial \mathcal{L}}{\partial x^i_x}\right)_{\sigma(x)} = 0$

**Definition 108.** “Christoffel symbol”  $^{L.C.}\Gamma$  are the connection coefficient functions of the so-called Levi-Civita connection  $^{L.C.}\nabla$

We usually make this choice of  $\nabla$  if  $g$  is given.

$$(M, \mathcal{O}, \mathcal{A}, g) \rightarrow (M, \mathcal{O}, \mathcal{A}, g, ^{L.C.}\nabla)$$

abstract way:  $\nabla g = 0$  and  $T = 0$  (torsion)

$$\implies \nabla = ^{L.C.}\nabla$$

**Definition 109.** (a) The Riemann-Christoffel curvature is defined by

$$R_{abcd} := g_{am} R^m_{bcd}$$

(b) Ricci:  $R_{ab} = R^m_{amb}$

Thought bubble: with a metric,  $^{L.C.}\nabla$

(c) (Ricci) scalar curvature:

$$R = g^{ab} R_{ab}$$

Thought bubble:  $^{L.C.}\nabla$

**Definition 110.** Einstein curvature  $(M, \mathcal{O}, \mathcal{A}, g)$

$$G_{ab} := R_{ab} - \frac{1}{2} g_{ab} R$$

Convention:  $g^{ab} := (g^{“-1”})^{ab}$

F. Schuller: these indices are not being pulled up, because what would you pull them up with (student) Question: Does the Einstein curvature yield new information?

Answer:

$$g^{ab} G_{ab} = R_{ab} g^{ab} - \frac{1}{2} g_{ab} g^{ab} R = R - \delta_a^a R = R - \frac{1}{2} \dim M R = (1 - \frac{d}{2}) R$$

**Tutorial 9: Metric manifolds. Exercise 3: Levi-Civita Connection.** Suppose torsion-free  $T = 0$  and metric-compatible connection  $\nabla g = 0$

**Question Recall**  $T = 0$  on a chart.

$$\Gamma^c_{ba} = \frac{1}{2} (g^{-1})^{cm} \left( \frac{\partial g_{bm}}{\partial x^a} + \frac{\partial g_{ma}}{\partial x^b} - \frac{\partial g_{ab}}{\partial x^m} \right)$$

or

$$\Gamma^a_{bc} = \frac{1}{2} (g^{-1})^{am} \left( \frac{\partial g_{bm}}{\partial x^c} + \frac{\partial g_{mc}}{\partial x^b} - \frac{\partial g_{bc}}{\partial x^m} \right)$$

33. SYMMETRY

EY : 20150321 This lecture tremendously and lucidly clarified, for me at least, what a symmetry of the Lie algebra is, and in comparing structures  $(M, \mathcal{O}, \mathcal{A})$  vs.  $(M, \mathcal{O}, \mathcal{A}, \nabla)$ , clarified differences, and asking about differences is a good way to learn, the difference between  $\mathcal{L}$  and  $\nabla$ , respectively.

Feeling that the round sphere

$$(S^2, \mathcal{O}, \mathcal{A}, g^{\text{round}})$$

has rotational symmetry, while  
the potato

$$(S^2, \mathcal{O}, \mathcal{A}, g^{\text{potato}})$$

does not.

33.1.

33.2. Important

33.3. **Flow of a complete vector field.** Let  $(M, \mathcal{O}, \mathcal{A})$  smooth  $X$  vector field on  $M$

**Definition 111.** *A curve  $\gamma : I \subseteq \mathbb{R} \rightarrow M$  is called an integral curve of  $X$  if*

$$v_{\gamma, \gamma(\lambda)} = X_{\gamma(\lambda)}$$

**Definition 112.** *A vector filed  $X$  is **complete** if all integral curves have  $I = \mathbb{R}$*  EY: 20150321 (i.e. domain is all of  $\mathbb{R}$ )

Ex. minute 48:30 EY : reall good explanation by F.P.Schuller; take a pt. out for an incomplete vector field.

**Theorem 33.** compactly supported smooth vector field is complete.

**Definition 113.** The flow of a complete vector field  $X$  is a 1-parameter family

$$h^X = \mathbb{R} \times M \rightarrow M$$

where  $\gamma_p : \mathbb{R} \rightarrow M$  is the integral curve of  $X$  with

$$\begin{aligned} \gamma(0) &= p \\ \text{Then for fixed } \lambda \in \mathbb{R} \end{aligned}$$

$$h_{\lambda}^X : M \rightarrow M \text{ smooth}$$

picture  $h_{\lambda}^X(S) \neq S$ ( if  $X \neq 0$ )

33.4. **Lie subalgebras of the Lie algebra  $(\Gamma(TM), [\cdot, \cdot])$  of vector fields.**

- (a)  $\Gamma(TM) = \{ \text{ set of all vector fields } \}$   $C^\infty(M)$ -module =  $\mathbb{R}$ -vector space  
 $\implies [X, Y] \in \Gamma(TM)$   $[X, Y]f := X(Yf) - Y(Xf)$
- (i)  $[X, Y] = -[Y, X]$   
(ii)  $[\lambda X + Z, Y] = \lambda[X, Y] + [Z, Y]$   
(iii)  $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$   
 $(\Gamma(TM), [\cdot, \cdot])$  Lie algebra
- (b) Let  $X_1 \dots X_s$  for  $s$  (many) vector fields on  $M$ , such that

**Tutorial 11 Symmetry. Exercise 1. : True or false?**

- (a) 
  - 
  - $\phi^* : T^*N \rightarrow T^*M$  i.e.  $\phi^*\nu(X) = \nu(\phi_*X)$  for smooth  $\phi : M \rightarrow N$ , so the pullback of a covector  $\nu \in T^*N$  maps to a covector in  $T^*M$ .
  - 
  - 
  - 
  -
- (b)
- (c)

**Exercise 2. :** Pull-back and push-forward

**Question .** Let’s check this locally

$$\begin{aligned} \phi^*(df)(X) &= (df)(\phi_*X) = (df)(X^i \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}) = X^i \frac{\partial y^j}{\partial x^i} \frac{\partial f}{\partial y^j} \text{ where } \phi_*X = X^i \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j} \\ d(\phi^*f)(X) &= d(f(\phi))(X) = \frac{\partial f}{\partial y^j} \frac{\partial y^j}{\partial x^i} dx^i(X) = X^i \frac{\partial y^j}{\partial x^i} \frac{\partial f}{\partial y^j} \end{aligned}$$

So

$$\boxed{\phi^*(df) = d(\phi^*f)}$$

$$\forall p \in M, \forall X \in \mathfrak{X}(M)$$

The big idea is that this is a showing of the **naturality** of the pullback  $\phi^*$  with  $d$ , i.e. that this commutes:

$$\begin{array}{ccc} \Omega^1(M) & \xleftarrow{\phi^*} & \Omega^1(N) \\ d \uparrow & & d \uparrow \\ C^\infty(M) & \xleftarrow{\phi^*} & C^\infty(N) \end{array}$$

**Question .**

$$(\phi_*)^a_b := (dy^a)(\phi_*(\frac{\partial}{\partial x^b}))$$

Let  $g \in C^\infty(N)$

$$\begin{aligned} \phi_* \left( \frac{\partial}{\partial x^b} \right) g &= \frac{\partial x^b}{g} \phi(p) = \frac{\partial}{\partial x^b} g \phi x^{-1} x(p) = \frac{\partial}{\partial x^b} (g y y^{-1} \phi x^{-1})(x) = \\ &= \frac{\partial}{\partial x^b} (g y^{-1} (y \phi x^{-1}(x(p)))) = \frac{\partial g}{\partial y} \bigg|_y \frac{\partial y^a}{\partial x^b} \bigg|_x = \frac{\partial y^a}{\partial x^b} \frac{\partial g}{\partial y^a} \end{aligned}$$

Then

$$\phi_* \left( \frac{\partial}{\partial x^b} \right) = \frac{\partial y^a}{\partial x^b} \frac{\partial}{\partial y^a}$$

and so

$$(\phi_*)^a_b = \frac{\partial y^a}{\partial x^b}$$

**Question .**

**Exercise 3. :Lie derivative-the pedestrian way**

**Question .** While it is true that  $\forall p \in S^2$ , for  $x(p) = (\theta, \varphi)$ , and  $(yix^{-1})(\theta, \varphi) = (y^1, y^2, y^3) \in \mathbb{R}^3$  and that, at this point

$p$ ,  $(y^1)^2/a^2 + (y^2)^2/b^2 + (y^3)^2/c^3 = 1$ , this doesn’t imply (EY: 20150321 I think) that, globally, it’s an ellipsoid (yet). In the familiar charts given, spherical chart  $(U, x) \in \mathcal{A}$  and  $(\mathbb{R}^3, y = \text{id}_{\mathbb{R}^3}) \in \mathcal{B}$  it looks like an ellipsoid, but change to another choice of charts, and it could look something very different.

**Question .**

Equip  $(\mathbb{R}^3, \mathcal{O}_{\text{st}}, \mathcal{B})$  with the Euclidean metric  $g$ , and pullback  $g$ .  
Note that the pullback of the inclusion from  $\mathbb{R}^3$  onto  $S^2$  for the Euclidean metric is the following:

$$i^*g \left( \frac{\partial}{\partial \theta^i}, \frac{\partial}{\partial \theta^j} \right) = g \left( i_* \frac{\partial}{\partial \theta^i}, i_* \frac{\partial}{\partial \theta^j} \right) = g \left( \frac{\partial x^a}{\partial \theta^i} \frac{\partial}{\partial x^a}, \frac{\partial x^b}{\partial \theta^j} \frac{\partial}{\partial x^b} \right) = g_{ab} \frac{\partial x^a}{\partial \theta^i} \frac{\partial x^b}{\partial \theta^j}$$



With  $g_{ab} = \delta_{ab}$ , the usual Euclidean metric, this becomes the following:

$$g_{ij}^{\text{ellipsoid}} = \frac{\partial x^a}{\partial \theta^i} \frac{\partial x^a}{\partial \theta^j}$$

At this point, one should get smart (we are in the 21st century) and use some sort of CAS (Computer Algebra System). I like Sage Math (version 6.4 as of 20150322). I also like the Sage Manifolds package for Sage Math.

I like Sage Math for the following reasons:

- Open source, so it’s open and freely available to anyone, which fits into my principle of making online education open and freely available to anyone, anytime
- Sage Math structures everything in terms of Category Theory and Categories and Morphisms naturally correspond to Classes and Class methods or functions in Object-Oriented Programming in Python and they’ve written it that way

and I like Sage Manifolds for roughly the same reasons, as manifolds are fit into a category theory framework that’s written into the Python code. e.g.

```
sage: S2 = Manifold(2, 'S^2', r'\mathbb{S}^2', start_index=1) ; print S2
sage: print S2
2-dimensional manifold 'S^2'
sage: type(S2)
<class 'sage.geometry.manifolds.manifold.Manifold_with_category'>
```

With code (I’ve provided for convenience; you can make your own as I wrote it based upon to example of  $S^2$  on the sage-manifolds documentation website page), load it and do the following:

```
cf. https://github.com/ernestyalumni/diffgeo-by-sagemnfd/blob/master/S2.sage
http://sagemanifolds.obspm.fr/examples.html
sage: load("S2.sage")
sage: U_ep = S2.open_subset('U_{ep}')
sage: eps.<the,phi> = U_ep.chart()
sage: a = var('a')
sage: b = var('b')
sage: c = var('c')
sage: inclus = S2.diff_mapping(R3, {(eps, cart): [ a*cos(phi)*sin(the), b*sin(phi)*sin(the),c*cos(the) ]} , name="inc",latex_name=r'\mathcal{i}')
sage: inclus.pullback(h).display()
inc.*(h) = (c^2*sin(the)^2 + (a^2*cos(phi)^2 + b^2*sin(phi)^2)*cos(the)^2) dthe*dthe - (a^2 - b^2)*cos(phi)*cos(the)*sin(phi)*sin(the) dthe*dphi
- (a^2 - b^2)*cos(phi)*cos(the)*sin(phi)*sin(the) dphi*dthe + (b^2*cos(phi)^2 + a^2*sin(phi)^2)*sin(the)^2 dphi*dphi
sage: inclus.pullback(h)[2,2].expr()
(b^2*cos(phi)^2 + a^2*sin(phi)^2)*sin(the)^2
```

A new open subset  $U_{\text{ep}}$  was declared in  $S^2$ , a new chart  $(U_{\text{ep}}, (\theta, \phi))$  was declared, the constants,  $a, b, c$ , were declared, and the inclusion map given in the problem

$$y \circ \mathbf{i} \circ x^{-1} : (\theta, \phi) \mapsto (a \cos \phi \sin \theta, b \sin \phi \sin \theta, c \cos \theta)$$

Then the pullback of the inclusion map  $\rangle$  was done on the Euclidean metric  $h$ , defined earlier in the file

`S2.sage`

. Then one can access the components of this metric and do, for example,

`simplify_full()`, `full_simplify()`, `reduce_trig()`

on the expression.

In Python, I could easily do this, and give an answer quick in LaTeX:

```
sage: for i in range(1,3):
.....:     for j in range(1,3):
.....:         print inclus.pullback(h)[i,j].expr()
.....:         latex(inclus.pullback(h)[i,j].expr() )
.....:
c^2*sin(the)^2 + (a^2*cos(phi)^2 + b^2*sin(phi)^2)*cos(the)^2
```

(EY: I’ll suppress the LaTeX output but this sage math function gives you LaTeX code)

and so

$$\begin{aligned} i^*g = & c^2 \sin (the)^2 + \left(a^2 \cos (\phi)^2 + b^2 \sin (\phi)^2\right) \cos (the)^2 d\theta \otimes d\theta + \\ & -2\left(a^2 - b^2\right) \cos (\phi) \cos (the) \sin (\phi) \sin (the) d\theta \otimes d\phi + \\ & + \left(b^2 \cos (\phi)^2 + a^2 \sin (\phi)^2\right) \sin (the)^2 d\phi \otimes d\phi \end{aligned}$$

Question .

```
sage: polar_vees = eps.frame()
sage: X_1 = - sin(phi) * polar_vees[1] - cot( the ) * cos(phi) * polar_vees[2]
sage: X_2 = cos( phi ) * polar_vees[1] - cot( the ) * sin( phi ) * polar_vees[2]
sage: X_3 = polar_vees[2]
sage: X_2.lie_der(X_1).display()
(cos(the)^2 - 1)/sin(the)^2 d/dphi
sage: X_3.lie_der(X_1).display()
cos(phi) d/dthe - cos(the)*sin(phi)/sin(the) d/dphi
sage: X_3.lie_der(X_2).display()
sin(phi) d/dthe + cos(phi)*cos(the)/sin(the) d/dphi
Indeed, one can check on a scalar field  $f_{\text{eps}} \in C^\infty(S^2)$ :
sage: f_eps = S2.scalar_field({eps: function('f', the, phi ) }, name='f' )
sage: (X_1( X_2(f_eps)) - X_2(X_1(f_eps) ) ).display()
U_{ep} --> R
(the, phi) |--> -D[1](f)(the, phi)
sage: X_2.lie_der(X_1) == -X_3
True
sage: X_3.lie_der(X_1) == X_2
True
sage: X_3.lie_der(X_2) == -X_1
True
```

$$\implies \boxed{[X_i, X_j] = -\epsilon_{ijk} X_k}$$

So  $\text{span}_{\mathbb{R}}\{X_1, X_2, X_3\}$  equipped with  $[\ , \ ]$  constitute a Lie subalgebra on  $S^2$  (It’s closed under  $[\ , \ ]$

34. INTEGRATION

34.1.

34.2.

34.3. Volume forms.

**Definition 114.** *On a smooth manifold  $(M, \mathcal{O}, \mathcal{A})$*

*a  $(0, \dim M)$ -tensor field  $\Omega$  is called a volume form if*

- (a)  *$\Omega$  vanishes nowhere (i.e.  $\Omega \neq 0 \ \forall p \in M$ )*
- (b) *totally antisymmetric*

$$\Omega(\dots, \underbrace{X}_{ith}, \dots, \underbrace{Y}_{jth}, \dots) = -\Omega(\dots, \underbrace{Y}_{ith}, \dots, \underbrace{X}_{jth}, \dots)$$

*In a chart:*

$$\Omega_{i_1 \dots i_d} = \Omega_{[i_1 \dots i_d]}$$

Example  $(M, \mathcal{O}, \mathcal{A}, g)$  metric manifold  
construct volume form  $\Omega$  from  $g$   
In any chart:  $(U, x)$

$$\Omega_{i_1 \dots i_d} := \sqrt{\det(g_{ij}(x))} \epsilon_{i_1 \dots i_d}$$

where **Levi-Civita symbol**  $\epsilon_{i_1 \dots i_d}$  is defined as  $\epsilon_{123 \dots d} = +1$   
 $\epsilon_{1 \dots d} = \epsilon_{[i_1 \dots i_d]}$

*Proof.* (well-defined) Check: What happens under a change of charts

$$\begin{aligned} \Omega(y)_{i_1 \dots i_d} &= \sqrt{\det(g(y)_{ij})} \epsilon_{i_1 \dots i_d} = \\ &= \sqrt{\det(g_{mn}(x) \frac{\partial x^m}{\partial y^i} \frac{\partial x^n}{\partial y^j}) \frac{\partial y^{m_1}}{\partial x^{i_1}} \dots \frac{\partial y^{m_d}}{\partial x^{i_d}}} \epsilon_{[m_1 \dots m_d]} = \\ &= \sqrt{|\det g_{ij}(x)|} \left| \det \left( \frac{\partial x}{\partial y} \right) \right| \det \left( \frac{\partial y}{\partial x} \right) \epsilon_{i_1 \dots i_d} = \sqrt{\det g_{ij}(x)} \epsilon_{i_1 \dots i_d} \operatorname{sgn} \left( \det \left( \frac{\partial x}{\partial y} \right) \right) \end{aligned}$$

EY : 20150323  
Consider the following:

$$\begin{aligned} \Omega(y)(Y_{(1)} \dots Y_{(d)}) &= \Omega(y)_{i_1 \dots i_d} Y_{(1)}^{i_1} \dots Y_{(d)}^{i_d} = \\ &= \sqrt{\det(g_{ij}(y))} \epsilon_{i_1 \dots i_d} Y_{(1)}^{i_1} \dots Y_{(d)}^{i_d} = \\ &= \sqrt{\det(g_{mn}(x)) \frac{\partial x^m}{\partial y^i} \frac{\partial x^n}{\partial y^j} \epsilon_{i_1 \dots i_d} \frac{\partial y^{i_1}}{\partial x^{m_1}} \dots \frac{\partial y^{i_d}}{\partial x^{m_d}}} X^{m_1} \dots X^{m_d} = \\ &= \sqrt{\det(g_{mn}(x)) \frac{\partial x^m}{\partial y^i} \frac{\partial x^n}{\partial y^j}} \det \left( \frac{\partial y}{\partial x} \right) \epsilon_{m_1 \dots m_d} X^{m_1} \dots X^{m_d} = \\ &= \sqrt{\det(g_{mn}(x))} \left| \det \left( \frac{\partial x}{\partial y} \right) \right| \det \left( \frac{\partial y}{\partial x} \right) \epsilon_{m_1 \dots m_d} X^{m_1} \dots X^{m_d} = \\ &= \sqrt{\det(g_{mn}(x))} \epsilon_{m_1 \dots m_d} \operatorname{sgn} \left( \det \left( \frac{\partial x}{\partial y} \right) \right) X^{m_1} \dots X^{m_d} = \operatorname{sgn}(\det \left( \frac{\partial x}{\partial y} \right)) \Omega_{m_1 \dots m_d}(x) X^{m_1} \dots X^{m_d} \end{aligned}$$

If  $\det \left( \frac{\partial y}{\partial x} \right) > 0$ ,

$$\Omega(y)(Y_{(1)} \dots Y_{(d)}) = \Omega(x)(X_{(1)} \dots X_{(d)})$$

This works also if Levi-Civita symbol  $\epsilon_{i_1 \dots i_d}$  doesn't change at all under a change of charts. (around 42:43 <https://youtu.be/2XpnbvPy-Zg>)

Alright, let's require,  
restrict the smooth atlas  $\mathcal{A}$   
to a subatlas ( $\mathcal{A}^\uparrow$  still an atlas)

$$\mathcal{A}^\uparrow \subseteq \mathcal{A}$$

s.t.  $\forall (U, x), (V, y)$  have chart transition maps  $y \circ x^{-1}$   
 $x \circ y^{-1}$

s.t.  $\det \left( \frac{\partial y}{\partial x} \right) > 0$   
such  $\mathcal{A}^\uparrow$  called an **oriented** atlas

$$(M, \mathcal{O}, \mathcal{A}, g) \implies (M, \mathcal{O}, \mathcal{A}^\uparrow, g)$$

Note: associated bundles.  
Note also:  $\det \left( \frac{\partial y^b}{\partial x^a} \right) = \det(\partial_a(y^b x^{-1}))$   $\frac{\partial y^b}{\partial x^a}$  is an endomorphism on vector space  $V$ .

$\varphi : V \rightarrow V$   
 $\det \varphi$  independent of choice of basis  
 $g$  is a  $(0, 2)$  tensor field, not endomorphism (not independent of choice of basis)  $\sqrt{|\det(g_{ij}(y))|}$

**Definition 115.**  $\Omega$  be a volume form on  $(M, \mathcal{O}, \mathcal{A}^\uparrow)$  and consider chart  $(U, x)$

**Definition 116.**  $\omega_{(X)} := \Omega_{i_1 \dots i_d} \epsilon^{i_1 \dots i_d}$  same way  $\epsilon^{12 \dots d} = +1$   
 $\epsilon^{[\dots]}$

one can show

$$\omega_{(y)} = \det \left( \frac{\partial x}{\partial y} \right) \omega_{(x)}$$

scalar density

34.4. Integration on one chart domain  $U$ .

**Definition 117.**

(225)

$$\int_U f : \stackrel{(U, y)}{=} \int_{y(U)} d^d \beta \omega_{(y)}(y^{-1}(\beta)) f_{(y)}(\beta)$$

*Proof.* : Check that it's (well-defined), how it changes under change of charts

$$\begin{aligned} \int_U f : \stackrel{(U, y)}{=} \int_{y(U)} d^d \beta \omega_{(y)}(y^{-1}(\beta)) f_{(y)}(\beta) &= \stackrel{(U, y)}{=} \int_{x(U)} \int d^d \alpha \left| \det \left( \frac{\partial y}{\partial x} \right) \right| f_{(x)}(\alpha) \omega_{(x)}(x^{-1}(\alpha)) \det \left( \frac{\partial x}{\partial y} \right) = \\ &= \int_{x(U)} d^d \alpha \omega_{(x)}(x^{-1}(x)) f_{(x)}(\alpha) \end{aligned}$$

□

On an oriented metric manifold  $(M, \mathcal{O}, \mathcal{A}^\uparrow, g)$

$$\int_U f := \int_{x(U)} d^d \alpha \underbrace{\sqrt{\det(g_{ij}(x))(x^{-1}(\alpha))}}_{\sqrt{g}} f_{(x)}(\alpha)$$

34.5. Integration on the entire manifold.

35. LECTURE 13: RELATIVISTIC SPACETIME

Recall, from Lecture 9, the definition of Newtonian spacetime

$$(M, \mathcal{O}, \mathcal{A}, \nabla, t)$$

$$\begin{aligned} &\nabla \text{ torsion free} \\ &t \in C^\infty(M) \\ &dt \neq 0 \\ &\nabla dt = 0 \quad (\text{uniform time}) \end{aligned}$$

and the definition of relativistic spacetime (before Lecture )

$$(M, \mathcal{O}, \mathcal{A}^\uparrow, \nabla, g, T)$$

$$\begin{aligned} &\nabla \text{ torsion-free} \\ &g \text{ Lorentzian metric}(+ - - -) \\ &T \text{ time-orientation} \end{aligned}$$

35.1. Time orientation.

**Definition 118.**  $(M, \mathcal{O}, \mathcal{A}^\uparrow, g)$  a Lorentzian manifold. Then a time-orientation is given by a vector field  $T$  that

- (i) does **not** vanish anywhere

(ii)  $g(T, T) > 0$

Newtonian vs. relativistic

Newtonian

$X$  was called future-directed if

$$dt(X) > 0$$

$\forall p \in M$ , take half plane, half space of  $T_p M$

also stratified atlas so make planes of constant  $t$  straight

relativistic

half cone  $\forall p, q \in M$ , half-cone  $\subseteq T_p M$

This definition of spacetime

Question

I see how the cone structure arises from the new metric. I don’t understand however, how the  $T$ , the time orientation, comes in

Answer

$(M, \mathcal{O}, \mathcal{A}, g)$   $g \leftarrow (+ - - -)$

requiring  $g(X, X) > 0$ , select cones

$T$  chooses which cone

This definition of spacetime has been made to enable the following physical postulates:

- (P1) The worldline  $\gamma$  of a massive particle satisfies
- (i)  $g_{\gamma(\lambda)}(v_{\gamma, \gamma(lambda)}, v_{\gamma, \gamma(\lambda)}) > 0$

(ii)  $g_{\gamma(\lambda)}(T, v_{\gamma, \gamma(\lambda)}) > 0$
- (P2) Worldlines of massless particles satisfy
- (i)  $g_{\gamma(\lambda)}(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)}) = 0$

(ii)  $g_{\gamma(\lambda)}(T, v_{\gamma, \gamma(\lambda)}) > 0$
- picture: spacetime:

Answer (to a question)  $T$  is a smooth vector field,  $T$  determines future vs. past, “general relativity: we have such a time orientation; smoothness makes it less arbitrary than it seems” -FSchuller,

Claim: 9/10 of a metric are determined by the cone

spacetime determined by distribution, only one-tenth error

35.2. Observers.  $(M, \mathcal{O}, \mathcal{A}^\uparrow, \nabla, g, T)$

**Definition 119.** An observer is a worldline  $\gamma$  with

$$g(v_\gamma, v_\gamma) > 0$$

$$g(T, v_\gamma) > 0$$

together with a choice of basis

$$v_{\gamma, \gamma(\lambda)} \equiv e_0(\lambda), e_1(\lambda), e_2(\lambda), e_3(\lambda)$$

of each  $T_{\gamma(\lambda)} M$  where the observer worldline passes, if  $g(e_a(\lambda), e_b(\lambda)) = \eta_{ab} = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{bmatrix}_{ab}$

precise: observer = smooth curve in the frame bundle  $LM$  over  $M$

35.2.1. Two physical postulates.

(P3) A **clock** carried by a specific observer  $(\gamma, e)$  will measure a **time**

$$\tau := \int_{\lambda_0}^{\lambda_1} d\lambda \sqrt{g_{\gamma(\lambda)}(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)})}$$

between the two “events”

$$\gamma(\lambda_0)$$

“start the clock”

and

$$\gamma(\lambda_1)$$

“stop the clock”

Compare with Newtonian spacetime:

$$t(p) = 7$$

Thought bubble: proper time/eigentime  $\tau$

$$M = \mathbb{R}^4$$

Application/Example.  $\mathcal{O} = \mathcal{O}_{\text{st}}$

$$\mathcal{A} \ni (\mathbb{R}^4, \text{id}_{\mathbb{R}^4})$$

$$g : g_{(x)ij} = \eta_{ij} \quad ; \quad T_{(x)}^i = (1, 0, 0, 0)^i$$

$$\implies \Gamma_{(x) \ jk}^i = 0 \text{ everywhere}$$

$$\implies (M, \mathcal{O}, \mathcal{A}^\uparrow, g, T, \nabla) \quad \text{Riem} = 0$$

$\implies$  spacetime is flat

This situation is called special relativity.

Consider two observers:

$$\gamma : (0, 1) \rightarrow M$$

$$\gamma_{(x)}^i = (\lambda, 0, 0, 0)^i$$

$$\delta : (0, 1) \rightarrow M$$

$$\alpha \in (0, 1) : \delta_{(x)}^i = \begin{cases} (\lambda, \alpha \lambda, 0, 0)^i & \lambda \leq \frac{1}{2} \\ (\lambda, (1 - \lambda)\alpha, 0, 0)^i & \lambda > \frac{1}{2} \end{cases}$$

let’s calculate:

$$\tau_\gamma := \int_0^1 \sqrt{g_{(x)ij} \dot{\gamma}_{(x)}^i \dot{\gamma}_{(x)}^j} = \int_0^1 d\lambda 1 = 1$$

$$\tau_\delta := \int_0^{1/2} d\lambda \sqrt{1 - \alpha^2} + \int_{1/2}^1 \sqrt{1^2 - (-\alpha)^2} = \int_0^1 \sqrt{1 - \alpha^2} = \sqrt{1 - \alpha^2}$$

Note: piecewise integration

Taking the clock postulate (P3) seriously, one better come up with a realistic clock design that supports the postulate.  
idea.

- (P4) Postulate
- 2 little mirrors
- Let  $(\gamma, e)$  be an observer, and  $\delta$  be a *massive* particle worldline that is parametrized s.t.  $g(v_\gamma, v_\gamma) = 1$  (for parametrization/normalization convenience)
- Suppose the observer and the particle *meet* somewhere (in spacetime)

$$\delta(\tau_2) = p = \gamma(\tau_1)$$

*This* observer measures the 3-velocity (spatial velocity) of this particle as

(226)

$$v_\delta : \epsilon^\alpha(v_{\delta, \delta(\tau_2)})e_\alpha \qquad \alpha = 1, 2, 3$$

where  $\epsilon^0, \boxed{\epsilon^1, \epsilon^2, \epsilon^3}$  is the unique dual basis of  $e_0, \boxed{e_1, e_2, e_3}$

EY:20150407

There might be a major correction to Eq. (226) from the Tutorial 14 : Relativistic spacetime, matter, and Gravitation, see the second exercise, Exercise 2, third question:

(227)

$$v := \frac{\epsilon^\alpha(v_\delta)}{\epsilon^0(v_\delta)}e_\alpha$$

Consequence: An observer  $(\gamma, e)$  will extract quantities measurable in his laboratory from objective spacetime quantities always like that.

Ex:  $F$  Faraday (0, 2)-tensor of electromagnetism:

$$F(e_a, e_b) = F_{ab} = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{bmatrix}$$

observer frame  $e_a, e_b$

$$E_\alpha := F(e_0, e_\alpha)$$
$$B^\gamma := F(e_\alpha, e_\rho)\epsilon^{\alpha\beta\gamma} \text{ where } \epsilon^{123} = +1 \text{ totally antisymmetric}$$

35.3. **Role of the Lorentz transformations.** Lorentz transformations emerge as follows:

Let  $(\gamma, e)$  and  $(\tilde{\gamma}, \tilde{e})$  be observers with  $\gamma(\tau_1) = \tilde{\gamma}(\tau_2)$   
(for simplicity  $\gamma(0) = \tilde{\gamma}(0)$ )  
Now

$$\begin{array}{ll} e_0, \dots, e_1 & \text{at } \tau = 0 \\ \text{and } \tilde{e}_0, \dots, \tilde{e}_1 & \text{at } \tau = 0 \end{array}$$

both bases for the same  $T_{\gamma(0)}M$

Thus:  $\tilde{e}_a = \Lambda^b_a e_b \qquad \Lambda \in GL(4)$

Now:

$$\eta_{ab} = g(\tilde{e}_a, \tilde{e}_b) = g(\Lambda^m_a e_m, \Lambda^n_b e_n) = \Lambda^m_a \Lambda^n_b \underbrace{g(e_m, e_n)}_{\eta_{mn}}$$

i.e.  $\Lambda \in O(1, 3)$

Result: Lorentz transformations relate the *frames* of *any two observers* at the same point.  
“ $\tilde{x}^\mu - \Lambda^\mu_\nu x^\nu$ ” is utter nonsense

**Tutorial.** I didn’t see a tutorial video for this lecture, but I saw that the Tutorial sheet number 14 had the relevant topics. Go there.

- two types of matter
- point matter
- field matter
- point matter
- massive point particle
- more of a phenomenological importance
- field matter
- electromagnetic field
- more fundamental from the GR point of view
- both classical matter types

36.1. **Point matter.** Our postulates (P1) and (P2) already constrain the possible particle worldlines.

But what is their precise law of motion, possibly in the presence of “forces”,

- (a) without external forces

$$S_{\text{massive}}[\gamma] := m \int d\lambda \sqrt{g_{\gamma(\lambda)}(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)})}$$

with:

$$g_{\gamma(\lambda)}(T_{\gamma(\lambda)}, v_{\gamma, \gamma(\lambda)}) > 0$$

dynamical law Euler-Lagrange equation

similarly

$$S_{\text{massless}}[\gamma, \mu] = \int d\lambda \mu g(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)})$$

$$\begin{array}{ll} \delta_\mu & g(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)}) = 0 \\ \delta_\gamma & \text{e.o.m.} \end{array}$$

Reason for describing equations of motion by actions is that composite systems have an action that is the sum of the actions of the parts of that system, possibly including “interaction terms.”

Example.

$$S[\gamma] + S[\delta] + S_{\text{int}}[\gamma, \delta]$$

- (b) presence of external forces  
or rather presence of fields to which a particle “couples”  
Example

$$S[\gamma; A] = \int d\lambda m \sqrt{g_{\gamma(\lambda)}(v_{\gamma, \gamma(\lambda)}, v_{\gamma, \gamma(\lambda)})} + qA(v_{\gamma, \gamma(\lambda)})$$

where  $A$  is a **covector field** on  $M$ .  $A$  fixed (e.g. the electromagnetic potential)

Consider Euler-Lagrange eqns.  $L_{\text{int}} = qA_{(x)}\dot{\gamma}^m_{(x)}$

$$m(\nabla_{v_\gamma} v_\gamma)_a + \underbrace{\left( \frac{\partial L_{\text{int}}}{\partial \dot{\gamma}^m_{(x)}} \right) - \frac{\partial L_{\text{int}}}{\partial \gamma^m_{(x)}}}_* = 0 \implies \boxed{\begin{array}{l} m(\nabla_{v_\gamma} v_\gamma)^a = \underbrace{-qF^a_m \dot{\gamma}^m}_{\text{Lorentz force on a charged particle in an electromagnetic field}} \end{array}}$$

$$\frac{\partial L}{\partial \dot{\gamma}^a} = qA_{(x)a}, \qquad \left( \frac{\partial L}{\partial \dot{\gamma}^m_{(x)}} \right) = q \cdot \frac{\partial}{\partial x^m}(A_{(x)m}) \cdot \dot{\gamma}^m_{(x)}$$

$$\frac{\partial L}{\partial \gamma^a} = q \cdot \frac{\partial}{\partial x^a}(A_{(x)m}) \dot{\gamma}^m$$

$$* = q \left( \frac{\partial A_a}{\partial x^m} - \frac{\partial A_m}{\partial x^a} \right) \dot{\gamma}^m_{(x)} = q \cdot F_{(x)am} \dot{\gamma}^m_{(x)}$$

$F \leftarrow$  Faraday

$$S[\gamma]=\int (m\sqrt{g(v_\gamma,v_\gamma)}+qA(v_\gamma))d\lambda$$

36.2. Field matter.

**Definition 120.** *Classical (non-quantum) field matter is any tensor field on spacetime where equations of motion derive from an action.*

Example:

$$S_{\text{Maxwell}}[A]=\frac{1}{4}\int_M d^4x\sqrt{-g}F_{ab}F_{cd}g^{ac}g^{bd}$$

$A$   $(0,1)$ -tensor field  
= thought cloud: for simplicity one chart covers all of  $M$   
– for  $\sqrt{-g}$   $(+---)$

$$F_{ab}:=2\partial_{[a}A_{b]}=2(\nabla_{[a}A)_{b]}$$

Euler-Lagrange equations for fields

$$0=\frac{\partial \mathcal{L}}{\partial A_m}-\frac{\partial}{\partial x^s}\left(\frac{\partial \mathcal{L}}{\partial \partial_s A_m}\right)+\frac{\partial}{\partial x^s}\frac{\partial}{\partial x^t}\frac{\partial^2 \mathcal{L}}{\partial \partial_t \partial_s A_m}$$

Example ...

$$(\nabla_{\frac{\partial}{\partial x^m}}F)^{ma}=j^a$$

**inh**omogeneous Maxwell  
thought bubble  $j=qv_\gamma$

$$\partial_{[a}F_{b]}-(\hspace{0.1cm})$$

homogeneous Maxwell  
Other example well-liked by textbooks

$$S_{\text{Klein-Gordon}}[\phi]:=\int_M d^4x\sqrt{-g}[g^{ab}(\partial_a\phi)(\partial_b\phi)-m^2\phi^2]$$

$\phi$   $(0,0)$ -tensor field

36.3. **Energy-momentum tensor of matter fields.** At some point, we want to write down an action for the metric tensor field itself.

But then, this action  $S_{\text{grav}}[g]$  will be added to any  $S_{\text{matter}}[A,\phi,\dots]$  in order to describe the total system.

$$S_{\text{total}}[g,A]=S_{\text{grav}}[g]+S_{\text{Maxwell}}[A,g]$$

$$\delta A\qquad\Longrightarrow\quad\text{Maxwell's equations}$$

$$\delta g_{ab}\quad:\quad\boxed{\frac{1}{16\pi G}G^{ab}}+(-2T^{ab})=0$$

$G$  Newton's constant

$$G^{ab}=8\pi G_N T^{ab}$$

**Definition 121.**  $S_{matter}[\Phi,g]$  is a matter action, the ***so-called energy-momentum tensor*** is

$$T^{ab}:=\frac{-2}{\sqrt{-g}}\left(\frac{\partial \mathcal{L}_{matter}}{\partial g_{ab}}-\partial_s\frac{\partial \mathcal{L}_{matter}}{\partial \partial_s g_{ab}}+\dots\right)$$

– of  $\frac{-2}{\sqrt{g}}$  is Schrödinger minus (EY : 20150408 F.Schuller's joke? but wise)  
choose all sign conventions s.t.

$$T(\epsilon^0,\epsilon^0)>0$$

Example: For  $S_{\text{Maxwell}}$ :

$$T_{ab}=F_{am}F_{bn}g^{mn}-\frac{1}{4}F_{mn}F^{mn}g_{ab}$$

$$T_{ab}\equiv T_{\text{Maxwell}ab}$$

$$T(e_0,e_0)=\underline{E}^2+\underline{B}^2$$

$$T(e_0,e_\alpha)=(E\times B)_\alpha$$

Fact: One often does not specify the fundamental action for some matter, but one is rather satisfied to assume certain properties / forms of

$$T_{ab}$$

Example Cosmology: (homogeneous & isotropic)  
perfect fluid

of pressure  $p$  and density  $\rho$  modelled by

$$T^{ab}=(\rho+p)u^au^b-pg^{ab}$$

radiative fluid  
What is a fluid of photons:

$$T_{\text{Maxwell}}^{ab}g_{ab}=0$$

observe:  $T_{\text{p.f.}}^{ab}g_{ab}\stackrel{!}{=}0$

$$=(\rho+p)u^au^bg_{ab}-p\underbrace{g^{ab}g_{ab}}_4$$

$$\leftrightarrow \rho_p04p=0$$

$$\rho=3p$$

$p=\frac{1}{3}\rho$   
Reconvene at 3 pm? (EY : 20150409 I sent a Facebook (FB) message to the International Winter School on Gravity and Light: there was no missing video; it continues on Lecture 15 immediately)

**Tutorial 14: Relativistic Spacetime, Matter and Gravitation. Exercise 2: Lorentz force law.**

**Question** electromagnetic potential.

37. LECTURE 15: EINSTEIN GRAVITY

Recall that in Newtonian spacetime, we were able to reformulate the Poisson law  $\Delta\phi = 4\pi G_N\rho$  in terms of the Newtonian spacetime curvature as

$$R_{00} = 4\pi G_N\rho$$

$R_{00}$  with respect to  $\nabla_{\text{Newton}}$

$G_N$  = Newtonian gravitational constant

This prompted Einstein to postulate < 1915 that the relativistic field equations for the Lorentzian metric  $g$  of (relativistic) spacetime

$$R_{ab} = 8\pi G_N T_{ab} \text{✓}$$

However, this equation suffers from a problem

LHS:  $(\nabla_a R)^{ab} \neq 0$

generically

RHS:

$$(\nabla_a T)^{ab} = 0$$

thought bubble: = formulated from an action

Einstein tried to argue this problem away.

Nevertheless, the equations cannot be upheld.

37.1. **Hilbert.** Hilbert was a specialist for variational principles.

To find the appropriate left hand side of the gravitational field equations, Hibert suggested to start from an action

$$S_{\text{Hilbert}}[g] = \int_M \sqrt{-g} R_{ab} g^{ab}$$

thought bubble = “simplest action”

aim: varying this w.r.t. metric  $g_{ab}$  will result in some tensor

$$G^{ab} = 0$$

37.2. **Variation of  $S_{\text{Hilbert}}$ .**

$$0 \stackrel{!}{=} \underbrace{\delta}_{g_i} S_{\text{Hilbert}}[g] = \int_M [\underbrace{\delta\sqrt{-g}g^{ab}R_{ab}}_1 + \underbrace{\sqrt{-g}\delta g^{ab}R_{ab}}_2 + \underbrace{\sqrt{-g}g^{ab}\delta R_{ab}}_3]$$

$$\text{and 1 : } \delta\sqrt{-g} = \frac{-(\det g)g^{mn}\delta g_{mn}}{2\sqrt{-g}} = \frac{1}{2}\sqrt{-g}g^{mn}\delta g_{mn}$$

thought bubble

$$\delta\det(g) = \det(g)g^{mn}\delta g_{mn}$$

e.g. from

$$\det(g) = \exp \text{tr} \ln g$$

ad 2:  $g^{ab}g_{bc} = \delta^a_c$

$$\implies (\delta g^{ab})g_{bc} + g^{ab}(\delta g_{bc}) = 0$$

$$\implies \delta g^{ab} = -g^{am}g^{bn}\delta g_{mn}$$

ad 3:

$$\begin{aligned} \Delta R_{ab} &\stackrel{\text{normal coords at point}}{=} \delta\partial_b\Gamma^m_{am} - \delta\partial_m\Gamma^m_{ab} + \Gamma\Gamma - \Gamma\Gamma = \\ &= \partial_b\delta\Gamma^m_{am} - \partial_m\delta\Gamma^m_{ab} = \\ &= \nabla_b(\delta\Gamma^m_{am}) - \nabla_m(\delta\Gamma^m_{ab}) \\ &\implies \sqrt{-g}g^{ab}\delta R_{ab} = \sqrt{-g} \end{aligned}$$

“if you formulate the variation properly, you’ll see the variation  $\delta$  commute with  $\partial_b$ ” EY : 20150408 I think one uses the integration at the bounds, integration by parts trick

$\Gamma^i_{(x)jk} - \tilde{\Gamma}^i_{(x)jk}$  are the components of a (1,2)-tensor.

Notation:  $(\nabla_b A)^i{}_g =: A^i{}_{j;b}$

$$\begin{aligned} &\implies \sqrt{-g}g^{ab}\delta R_{ab} \\ &\stackrel{\nabla g=0}{=} \sqrt{-g}(g^{ab}\delta\Gamma^m_{am})_{;b} - \sqrt{-g}(g^{ab}\delta\Gamma^m_{ab})_{;m} = \sqrt{-g}A^b{}_{;b} - \sqrt{-g}B^m{}_{;m} \end{aligned}$$

Question: Why is the difference of coefficients a tensor?

Answer:

$$\Gamma^i_{(y)jk} = \frac{\partial y^i}{\partial x^m} \frac{\partial x^m}{\partial y^j} \frac{\partial x^q}{\partial y^k} \Gamma^m_{(x) ,nq} + \frac{\partial y^i}{\partial x^m} \frac{\partial^2 x^m}{\partial y^j \partial y^k}$$

Collecting terms, one obtains

$$\begin{aligned} 0 &\stackrel{!}{=} \delta S_{\text{Hilbert}} = \int_M [\frac{1}{2}\sqrt{-g}g^{mn}\delta g_{mn}g^{ab}R_{ab} - \sqrt{-g}g^{am}g^{bn}\delta g_{mn}R_{ab} + \underbrace{(\sqrt{-g}A^a)_{,a}}_{\text{surface}} - \underbrace{(\sqrt{-g}B^b)_{,b}}_{\text{surface term}}] \\ &= \int_M \sqrt{-g}\delta \underbrace{g_{mn}}_{\text{arbitrary variation}} [\frac{1}{2}g^{mn}R - R^{mn}] \implies G^{mn} = R^{mn} - \frac{1}{2}g^{mn}R \end{aligned}$$

Hence Hilbert, from this “mathematical” argument, concluded that one may take

$$R_{ab} - \frac{1}{2}g_{ab}R = 8\pi G_N T_{ab}$$

Einstein equations

$$S_{E-H}[g] = \int_M \sqrt{-g}R$$

37.3. **3. Solution of the  $\nabla_a T^{ab} = 0$  issue.** One can show ( $\rightarrow$  Tutorials) that the Einstein curvature

$$G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R$$

satisfy the so-called contracted differential Bianchi identity

$$(\nabla_a G)^{ab} = 0$$

37.4. **Variants of the field equations.**

(a) a simple rewriting:

$$R_{ab} - \frac{1}{2}g_{ab}R = 8\pi G_N T_{ab} = T_{ab}$$

$$G_N = \frac{1}{8\pi}$$

Contract on both sides  $g^{ab}$

$$R_{ab} - \frac{1}{2}g_{ab}R = T_{ab} || g^{ab}$$

$$R - 2R = T := T_{ab}g^{ab}$$

$$\implies R = -T$$

$$\implies R_{ab} + \frac{1}{2}g_{ab}T = T_{ab}$$

$$\iff R_{ab} = (T_{ab} - \frac{1}{2}Tg_{ab}) =: \hat{T}_{ab}$$



(b)

$$R_{ab} = \widehat{T}_{ab}$$

$$S_{E-H}[g] := \int_M \sqrt{-g}(R + 2\Lambda)$$

thought bubble:  $\Lambda$  cosmological constant

History:

1915:  $\Lambda < 0$  (Einstein) in order to get a non-expanding universe

>1915:  $\Lambda = 0$  Hubble

today  $\Lambda > 0$  to account for an accelerated expansion

$\Lambda \neq 0$  can be interpreted as a contribution

$-\frac{1}{2}\Lambda g$  to the energy-momentum

“dark energy”

Question: surface terms scalar?

Answer: for a careful treatment of the surface terms which we discarded, see, e.g. E. Poisson, “A relativist’s toolkit”

C.U.P. “excellent book”

Question: What is a constant on a manifold?

Answer:  $\int \sqrt{-g}\Lambda = \Lambda \int \sqrt{-g}1$

[back to dark energy]

[Weinberg, QCD, calculated]

idea: 1 could arise as the vacuum energy of the standard model fields

$\Lambda_{\text{calculated}} = 10^{120} \times \Lambda_{\text{obs}}$

“worst prediction of physics”

Tutorials: check that

- Schwarzschild metric (1916)
- FRW metric
- pp-wave metric
- Reisner-Nordstrom

$\implies$  are solutions to Einstein’s equations

in high school

$$m\ddot{x} + m\omega^2x^2 = 0$$

$$x(t) = \cos(\omega t)$$

ET: [elementary tutorials]

study motion of particles & observers in Schwarzschild S.T.

Satellite: Marcus C. Werner

Gravitational lensing

odd number of pictures Morse theory (EY:20150408 Morse Theory !!!)

Domenico Giulini

Hamiltonian form Canonical Formulations

Key to Quantum Gravity

TUTORIAL 13 SCHWARZSCHILD SPACETIME

EY : 20150408 I’m not sure which tutorial follows which lecture at this point.

The tutorial video is excellent itself. Here, I want to encourage the use of CAS to do calculations. There are many out there.

Again, I’m partial to the Sage Manifolds package for Sage Math which are both open-source and based on Python. I’ll use that here.

Exercise 1. Geodesics in a Schwarzschild spacetime

Question Write down the Lagrangian.

Load “Schwarzschild.sage” in Sage Math, which will always be available freely here <https://github.com/ernestyalumni/diffgeo-by-sagemnfd/blob/master/Schwarzschild.sage>:

```
sage: load("Schwarzschild.sage")
4-dimensional manifold 'M'
open subset 'U_sph' of the 4-dimensional manifold 'M'
Levi-Civita connection 'nabla_g' associated with the Lorentzian metric 'g' on the 4-dimensional manifold 'M'
```

and so on.

Look at the code and I had defined the Lagrangian to be

L

. To get out the coefficients of  $L$  of the components of the tangent vectors to the curve, i.e.  $t', r', \theta', \phi'$ , denoted

tp, rp, thp, php

in my .sage file, do the following:

```
sage: L.expr().coefficients(tp)[1][0].factor().full_simplify()
(2*G_N*M_0 - r)/r
sage: L.expr().coefficients(rp)[1][0].factor().full_simplify()
-r/(2*G_N*M_0 - r)
sage: L.expr().coefficients(php)[1][0].factor().full_simplify()
r^2
sage: L.expr().coefficients(thp)[1][0].factor().full_simplify()
r^2*sin(th)^2
```

Question There are 4 Euler-Lagrange equations for this Lagrangian. Derive the one with respect to the function  $t(\lambda)!$ .

```
sage: L.expr().diff(t)
0
```

This confirms that  $\frac{\partial L}{\partial t} = 0$

For  $\frac{d}{d\lambda} \frac{\partial L}{\partial t'}$ , then one needs to consider this particular workaround for Sage Math (computer technicality). One takes derivatives with respect to declared variables (declared with var) and then substitute in functions that are dependent upon  $\lambda$ , and then take the derivative with respect to the parameter  $\lambda$ . This does that:

```
sage: L.expr().diff( thp ).factor().subs( r == gamma1 ).subs( thp == gamma3.diff( tau ) ).subs( th == gamma3 ).diff(tau)\
....: .factor()
2*(2*cos(gamma3(tau))*gamma1(tau)*D[0](gamma3)(tau)^2 + 2*sin(gamma3(tau))*D[0](gamma1)(tau)*D[0](gamma3)(tau)
+ gamma1(tau)*sin(gamma3(tau))*D[0, 0](gamma3)(tau))*gamma1(tau)*sin(gamma3(tau))
```

Question Show that the Lie derivative of  $g$  with respect to the vector fields  $K_t := \frac{\partial}{\partial t}$ .

The first line defines the vector field by accessing the frame defined on a chart with spherical coordinates and getting the time vector. The second line is the Lie derivative of  $g$  with respect to this vector field.

```
sage: K_t = espher[0]
sage: g.lie_der(K_t).display() # 0, as desired
0
```

EY : 20150410 My question is this:  $\forall X \in \Gamma(TM)$  i.e.  $X$  is a vector field on  $M$ , or, specifically, a section of the tangent bundle, then does

$$\mathcal{L}_X g = 0$$

instantly mean that  $X$  is a symmetry for  $(M, g)$ ?  $\mathcal{L}_X g$  is interpreted geometrically as how  $g$  changes along the flow generated by  $X$ , and if it equals 0, then  $g$  doesn’t change.

38.

39.

40. CANONICAL FORMULATION OF GR I

Dynamical and Hamiltonian formulation of General Relativity.

Purpose

- (1) formulate and solve initial-value problems
- (2) integrate Einstein’s Equations by numerical codes
- (3) characterize degrees of freedom
- (4) characterize isolated systems, associated symmetry groups and conserved quantities, like Energy/Mass, Momenta (linear and angular), Poincaré charges
- (5) starting point for “canonical quantization” program.

How. We will rewrite Einstein’s Eq. in form of a *constrained Hamiltonian system*.

(− + ++)

$$T^{\mu\nu} = \begin{pmatrix} W & \frac{1}{c}S^m \\ cg^m & \mathbf{t}^{mn} \end{pmatrix}$$

- $W$  = Energy density (1 component)
- $g^m$  = Momentum density, (3 components)
- $S^m$  = Energy current-density (3 components)
- $\mathbf{t}^{mn}$  = Momentum current-density (6 components)

$$T^{\mu\nu} = T^{\nu\mu} \implies S^m = c^2 g^m$$

10 independent komp. (components)

Phys. dim.  $[T^{\mu\nu}] = \frac{J}{m^3}$   
 $[G^{\mu\nu}] = \frac{1}{m^2}$

$$k = \frac{\text{curvature}}{\text{Energy} \cdot \text{density}}$$

$$[k] = \frac{1}{m^2} / \frac{J}{m^3}, \quad {}^2k = \frac{\text{Curvature}}{\text{mass density}} = \left(\frac{1}{1.5 \text{ AU}}\right)^2 / \text{Density of water}$$
$$= \left(\frac{1}{10 \text{ km}}\right)^2 / \text{Nuclear density in core of neutron star} \simeq 5 \cdot 10^{17} \text{ kg}/m^3$$

If “Ein” for Einstein Tensor,  $G_{\mu\nu} = \text{Ein}\left(\frac{\partial}{\partial x^\mu}, \frac{\partial}{\partial x^\nu}\right)$

$$\text{Ein}(v, w) = \frac{1}{4}[\text{Ein}(v + w, v + w) - \text{Ein}(v - w, v - w)]$$
$$\text{Ein}(w, w) = -g(w, w) \sum_{\perp w} \text{Sec}$$

where  $\perp w$  is take the sum over any triple of mutually perp. 2-planes in  $\perp w$

$\text{Sec}(\text{Span}\{v, w\}) = \frac{\text{Riem}(v, w, v, w)}{[g(v, w)]^2 - g(v, v)g(w, w)}$   
“sectional curvature”

Identity:  $\nabla_\mu G^{\mu\nu} = 0$  (follows from twice-contracted II. Bianchi Identity  
 $\sum_{\lambda\mu\nu} \text{cycl} \nabla_\lambda R_{\alpha\beta\mu\nu} = 0$  )

$$\underbrace{\partial_0 G^{0\nu}}_{\text{contains at most 1st time der.}} + \underbrace{\partial_k G^{k\nu} + \Gamma G + \Gamma G}_{\text{contains at most 2nd. time derivatives}} \equiv 0$$
  
 $\implies$  4 out of 10 Einstein Eq. do not evolve the fields but rather constrain the initial data. The space-space components (6 Eqs.) are the evolution Eqs.  
10 Einstein Eq. - 4 constraints (underdetermined elliptic type)  
    \ - 6 evolution equations (undetermined hyperbolic type)

41.

42.

43.

44. LECTURE 22: BLACK HOLES

Only depends on Lectures 1-15, so does lecture on “Wednesday”  
Schwarzschild solution also vacuum solution (from tutorial EY : oh no, must do tutorial)  
Study the Schwarzschild as a vacuum solution of the Einstein equation:  
 $m = G_N M$  where  $M$  is the “mass”

$$g = \left(1 - \frac{2m}{r}\right) dt \otimes dt - \frac{1}{1 - \frac{2m}{r}} dr \otimes dr - r^2(d\theta \otimes d\theta + \sin^2 \theta d\varphi \otimes d\varphi)$$

in the so-called Schwarzschild coordinates

$t$	$r$	$\theta$	$\varphi$
$(-\infty, \infty)$	$(0, \infty)$	$(0, \pi)$	$(0, 2\pi)$

What staring at this metric for a while, two questions naturally pose themselves:

(i) What exactly happens  $r = 2m$ ?

$t$	$r$	$\theta$	$\varphi$
$(-\infty, \infty)$	$(0, 2m) \cup (2m, \infty)$	$(0, \pi)$	$(0, 2\pi)$

(ii) Is there anything (in the real world) beyond  $t \rightarrow -\infty$ ?

$$t \rightarrow +\infty$$

idea: Map of Linz, blown up  
Insight into these two issues is afforded by stopping to stare.  
Look at *geodesic* of  $g$ , instead.

44.1. **Radial null geodesics.** null -  $g(v_\gamma, v_\gamma) = 0$   
Consider null geodesic in “Schd”

$$S[\gamma] = \int d\lambda \left[ \left(1 - \frac{2m}{r}\right) \dot{t}^2 - \left(1 - \frac{2m}{r}\right)^{-1} \dot{r}^2 - r^2(\dot{\theta}^2 + \sin^2 \theta \dot{\varphi}^2) \right]$$

with  $[\dots] = 0$   
and one has, in particular, the  $t$ -eqn. of motion:

$$\left(\left(1 - \frac{2m}{r}\right) \dot{t}\right)' = 0$$

⇒

$$\left(1 - \frac{2m}{r}\right) \dot{t} = k = \text{const.}$$

Consider radial null geodesics

$$\theta \stackrel{!}{=} \text{const.} \qquad \varphi = \text{const.}$$

From  $\square$  and  $\square$

$$\implies \dot{r}^2 = k^2 \leftrightarrow \dot{r} = \pm k$$

$$\implies r(\lambda) = \pm k \cdot \lambda$$

Hence, we may consider

$$\tilde{t}(r) := t(\pm k \lambda)$$

Case A:  $\oplus$

$$\frac{d\tilde{t}}{dr} = \frac{\tilde{\dot{t}}}{\dot{r}} = \frac{k}{\left(1 - \frac{2m}{r}\right) k} = \frac{r}{r - 2m}$$

$$\implies \tilde{t}_+(r) = r + 2m \ln|r - 2m|$$

(**outgoing** null geodesics)

Case b.  $\pm$  (Circle around  $-$ , consider  $-$ ):

$$\tilde{t}_-(r) = -r - 2m \ln|r - 2m|$$

(**ingoing** null geodesics)

Picture

44.2. **Eddington-Finkelstein.** Brilliantly simple idea:

change (on the domain of the Schwarzschild coordinates) to different coordinates, s.t.

in those new coordinates,

*ingoing* null geodesics appear as straight lines, of slope  $-1$

This is achieved by

$$\bar{t}(t, r, \theta, \varphi) := t + 2m \ln|r - 2m|$$

Recall: ingoing null geodesic has

$$\tilde{t}(r) = -(r + 2m \ln|r - 2m|) \qquad (Schdcoords)$$

$$\iff \bar{t} - 2m \ln|r - 2m| = -r - 2m \ln|r - 2m| + \text{const.}$$

$$\therefore \bar{t} = -r + \text{const.}$$

(Picture)

*outgoing* null geodesics

$$\bar{t} = r + 4m \ln|r - 2m| + \text{const.}$$

Consider the new chart  $(V, g)$  while  $(U, x)$  was the Schd chart.

$$\bigcup_{\text{Schd}} \{ \text{horizon} \} = V$$

“chart image of the horizon”

Now calculate the *Schd metric*  $g$  w.r.t. Eddington-Finkelstein coords.

$$\bar{t}(t, r, \theta, \varphi) = t + 2m \ln|r - 2m|$$

$$\bar{r}(t, r, \theta, \varphi) = r$$

$$\bar{\theta}(t, r, \theta, \varphi) = \theta$$

$$\bar{\varphi}(t, r, \theta, \varphi) = \varphi$$

EY : 20150422 I would suggest that after seeing this, one would calculate the metric by your favorite CAS. I like the Sage Manifolds package for Sage Math.

[Schwarzschild\\_BH.sage on github](#)

[Schwarzschild\\_BH.sage on Patreon](#)

[Schwarzschild\\_BH.sage on Google Drive](#)

```
sage: load('Schwarzschild_BH.sage')
4-dimensional manifold 'M'
expr = expr.simplify_radical()
Levi-Civita connection 'nabla_g' associated with the Lorentzian metric 'g' on the 4-dimensional manifold 'M'
Launched png viewer for Graphics object consisting of 4 graphics primitives
```

Then calculate the Schwarzschild metric  $g$  but in Eddington-Finkelstein coordinates. Keep in mind to calculate the set of coordinates that uses  $\bar{t}$ , not  $\tilde{t}$ :

```
sage: gI.display()
gI = (2*m - r)/r dt*dt - r/(2*m - r) dr*dr + r^2 dth*dth + r^2*sin(th)^2 dph*dph
sage: gI.display( X_EF_I_null.frame())
gI = (2*m - r)/r dtbar*dtbar + 2*m/r dtbar*dr + 2*m/r dr*dtbar + (2*m + r)/r dr*dr + r^2 dth*dth + r^2*sin(th)^2 dph*dph
```

REFERENCES

[1] William Feller. **An Introduction to Probability Theory and Its Applications**, Vol. 1, 3rd Edition. 1968. ISBN-13: 978-0471257080

[2] Charles M. Grinstead. J. Laurie Snell. **Introduction to Probability**. Second Revised Edition. American Mathematical Society. 1997.

[3] Sheldon M. Ross. **Introduction to Probability and Statistics for Engineers and Scientists**. Second Edition. Academic Press. 2000.

[4] Jeffrey M. Lee. **Manifolds and Differential Geometry**, *Graduate Studies in Mathematics* Volume: 107, American Mathematical Society, 2009. ISBN-13: 978-0-8218-4815-9

[5] John Lee, **Introduction to Smooth Manifolds** (Graduate Texts in Mathematics, Vol. 218), 2nd edition, Springer, 2012, ISBN-13: 978-1441999818

[6] Victor Guillemin, Alan Pollack. **Differential Topology**. American Mathematical Society. 2010. ISBN-13: 978-0821851937 [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwjG96q9z63JAhWMLYgKHempDoMQFggmMAE&url=http%3A%2F%2Fwww.mat.unimi.it%2Fusers%2Fdedo%2Ftop%2520diff%2FGuillemin-Pollack\\_Differential%2520topology.pdf&usg=AFQjCNF5im0H5xeXRSK60qzM7zT97sdIsw](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwjG96q9z63JAhWMLYgKHempDoMQFggmMAE&url=http%3A%2F%2Fwww.mat.unimi.it%2Fusers%2Fdedo%2Ftop%2520diff%2FGuillemin-Pollack_Differential%2520topology.pdf&usg=AFQjCNF5im0H5xeXRSK60qzM7zT97sdIsw)

[7] Anant R. Shastri. **Elements of Differential Topology**. CRC Press. 2011. ISBN-13: 978-0415339209

[8] Morris W. Hirsch, **Differential Topology** (Graduate Texts in Mathematics), Graduate Texts in Mathematics (Book 33), Springer (September 16, 1997). ISBN-13: 978-0387901480

[9] Mikio Nakahara. **Geometry, Topology and Physics**. (Graduate Student Series in Physics) 2nd Edition. 2003. Taylor and Francis Group.

[10] P.-A. Absil, R. Mahony, and R. Sepulchre. **Optimization Algorithms on Matrix Manifolds**. Princeton University Press. 2008. ISBN 978-0-691-13298-3 <https://press.princeton.edu/titles/8586.html>

[11] John Baez, Javier P Muniain. **Gauge Fields, Knots and Gravity** (Series on Knots and Everything), World Scientific Publishing Company (October 24, 1994), ISBN-13: 978-9810220341

[12] Joseph J. Rotman, **Advanced Modern Algebra** (Graduate Studies in Mathematics), (Book 114), American Mathematical Society; 2 edition, (August 10, 2010), ISBN-13: 978-0821847411

[13] Andrew Baker, “Representations of Finite Groups”, 07112011 <http://www.maths.gla.ac.uk/~ajb/dvi-ps/groupreps.pdf>

[14] Yvette Kosmann-Schwarzbach, **Groups and Symmetries: From Finite Groups to Lie Groups**, Springer, 2010. e-ISBN 978-0-387-78866-1 <http://www.caam.rice.edu/~yad1/miscellaneous/References/Math/Groups/Groups%20and%20Symmetries.pdf>

[15] Agostino Prástaro. **Geometry of PDEs and Mechanics**. World Scientific Publishing Co. 1996. QC125.2.P73 1996 530.1’55353–dc20.  
Agostino Prástaro. *Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate, Università degli Studi di Roma “La Sapienza”*.

[16] Clemens Koppensteiner. *Complex Manifolds: Lecture Notes*. <http://www.caramdir.at/uploads/math/piii-cm/complex-manifolds.pdf>

[17] Stefan Vandoren. **Lectures on Riemannian Geometry, Part II: Complex Manifolds**. <http://www.staff.science.uu.nl/~vando101/MRIlectures.pdf>

[18] Kentaro Hori (Author, Editor), Sheldon Katz (Editor), Albrecht Klemm (Editor), Rahul Pandharipande (Editor), Richard Thomas (Editor), Cumrun Vafa (Editor), Ravi Vakil (Editor), Eric Zaslow (Editor). **Mirror Symmetry** (Clay Mathematics Monographs, V. 1). Clay Mathematics Monographs (Book 1). American Mathematical Society (August 19, 2003). ISBN-10: 0821829556 ISBN-13: 978-0821829554 <https://web.archive.org/web/20060919020706/http://math.stanford.edu/~vakil/files/mirrorfinal.pdf>

[19] Lawrence Conlon. **Differentiable Manifolds** (Modern Birkhäuser Classics). 2nd Edition. Birkhäuser; 2nd edition (October 10, 2008). ISBN-13: 978-0817647667

[20] Andrew Clarke and Bianca Santoro. *Holonomy Groups in Riemannian Geometry*. [arXiv:1206.3170](https://arxiv.org/abs/1206.3170) [math.DG]

[21] Urs Schreiber and Konrad Waldorf. *Parallel Transport and Functors*. [arXiv:0705.0452](https://arxiv.org/abs/0705.0452) [math.DG]

[22] Y. Eliashberg, N. Mishachev, *Introduction to the h-principle*, Grad. Studies in Math. Vol 48 (AMS 2002)

[23] Darryl D Holm. **Geometric Mechanics** - Part II: Rotating, Translating and Rolling. 2nd. Ed. Imperial College Publishing. (October 31, 2011)

[24] Tsutomu Kambe, **Geometrical Theory of Dynamical Systems and Fluid Flows**, Advanced Series in Nonlinear Dynamics: Volume 23, 2009, <http://www.worldscientific.com/worldscibooks/10.1142/7418> ISBN: 978-981-4282-24-6 (hardcover)

[25] Daniel Kleppner and Robert Kolenkow. **An Introduction to Mechanics**. Second Edition. Cambridge University Press. 2014.

[26] Louis N. Hand and Janet D. Finch. **Analytical Mechanics**. Cambridge University Press. 1998.

[27] Herbert Goldstein, Charles Poole, John Saffo. **Classical Mechanics**. Pearson; 3rd edition (June 15, 2001).

[28] G. J. Pert. **Introductory Fluid Mechanics for Physicists and Mathematicians**. 1st Edition. Wiley; 1 edition (May 28, 2013). ISBN-10: 1119944856. ISBN-13: 978-1119944850.

[29] Rolf H. Sabersky, Allen J. Acosta, Edward G. Hauptmann, E.M. Gates. **Fluid Flow: A First Course in Fluid Mechanics** (4th Edition). Prentice Hall. (August 22, 1998). ISBN-13: 978-0135763728

[30] L. D. Landau, E.M. Lifshitz, **Fluid Mechanics**, Second Edition: Volume 6 (Course of Theoretical Physics S) Butterworth-Heinemann, 1987, ISBN-13: 978-0750627672

[31] Walter Greiner. **Classical Mechanics: Point Particles and Relativity** (Classical Theoretical Physics). Springer New York; 2004 edition (December 4, 2003). ISBN-13: 978-0387955865

[32] Vladimir I. Arnold, Valery V. Kozlov, Anatoly I. Neishtadt. **Mathematical Aspects of Classical and Celestial Mechanics**. Third Edition. Springer. 2006.

[33] Ralph Abraham and Jerrold E. Marsden. **Foundations of Mechanics** (Ams Chelsea Publishing) Hardcover – May 21, 2008. American Mathematical Society; 2 edition (May 21, 2008). ISBN-10: 0821844385 ISBN-13: 978-0821844380

[34] Howard Georgi. **The Physics of Waves**. Benjamin Cummings; 1st edition. 1992. ISBN-13:978-0136656210.

[35] Sean Mauch. **Introduction to Methods of Applied Mathematics or Advanced Mathematical Methods for Scientists and Engineers**. January 24, 2004. Mauch Publishing Company, un-Incorporated.

[36] Robert M. Wald. **General Relativity**. The University of Chicago Press. 1984.

[37] [Frederic Schuller. Lectures on the Geometric Anatomy of Theoretical Physics](#). Sept. 2015.

[38] [The WE-Heraeus International Winter School on Gravity and Light 2015. Central Lecture Course](#).

[39] Yvonne Choquet-Bruhat. **General Relativity and the Einstein Equations** (Oxford Mathematical Monographs) 1st Edition. Oxford Mathematical Monographs (February 4, 2009) ISBN-13: 978-0199230723

[40] Jacob D. Bekenstein. "Universal upper bound on the entropy-to-energy ratio for bounded systems." **Physical Review D**. Volume 23, Number 2. 15 Jan. 1981.

[41] Jacob D. Bekenstein. "Holographic bound from second law of thermodynamics". 14 Apr 2000. <http://arxiv.org/abs/hep-th/0003058v2>

[42] T. Frankel, **The Geometry of Physics**, Cambridge University Press, Third Edition, 2012.

[43] Shigeyuki Morita. **Geometry of Differential Forms** (Translations of Mathematical Monographs, Vol. 201). American Mathematical Society (August 28, 2001). ISBN-10: 0821810456

[44] Jürgen Jost. **Riemannian Geometry and Geometric Analysis**. Seventh Edition. Springer. 2017.

[45] Tom M. Apostol. **Calculus, Vol. 1: One-Variable Calculus, with an Introduction to Linear Algebra**. Wiley; 2nd edition (January 16, 1991).ISBN-13: 978-0471000051

[46] Roger R. Bate, Donald D. Mueller, Jerry E. White. **Fundamentals of Astrodynamics** (Dover Books on Aeronautical Engineering). Publisher: Dover Publications; Revised ed. edition (June 1, 1971). ISBN-13: 978-0486600611

[47] I.E. Irodov. **Problems in General Physics**. Mir Publishers. Moscow. 1981. ISBN 5-03-000800-4

[48] The Committee Of Japan Physics Olympiad Japan. **Physics Olympiad - Basic To Advanced Exercises**. Wspc (January 15, 2014).