MAT 401: The Geometry of Physics

Spring 2009

Department of Mathematics
SUNY at Stony Brook

Welcome to The Geometry of Physics

In this class we will develop the mathematical language needed to understand Einstein's field equations. The infinitesimal structure of any space, even curved space, is Euclidean, and so is described with linear algebra. Calculus, in the form of continuity and differentiability properties of paths and surfaces, can express the connectedness of space. The synthesis of these points of view, of the infinitesimal with the global, of linear algebra with calculus, yields the powerful language of differential geometry, which Einstein used to express the physics of General Relativity.

Course Content

Before studying the field equations we must develop the language of geometry. We will try to integrate intuitive content with hard mathematics, and some of the topics will be partly review for many students. But hard work will be required... it took Einstein more than 2 years to understand the mathematics we will cover in a semester.

Homework assignment page Class notes Quiz Prep (including final exam info)

Announcements

- Final Exam: Wed May 13, from 11am-1:30pm, will take place in the usual classroom.
- There will be a makeup class on Monday (May 11) in P-131 in the math building, at 11am. We will go over the gravitational field equations.
- I'll be in my office on Tuesday the 12th, from 2-4pm and 5-7pm

Course Information:

Check out the topics we will cover...

Here is a link to the syllabus.

Textbook

A first Course in General Relavity by Bernard F. Schutz

Supplimentary books / Recommended reading

The Geometry of Physcis by Theodore Frankel, Second Edition
The Large Scale Structure of Space-Time by G. Ellis and S. Hawking
General Relativity by Robert Wald

Course Grading

One homework assugnment will be due each Wednesday.

Homeworks: 10% of total grade
Quizes: 10% of total grade
Test 1: 10% of total grade (Friday Feb 13)
Test 2: 20% of total grade (Friday Mar 6)
Test 3: 10% of total grade (Friday Mar 20)
Test 4: 10% of total grade (Friday April 17)
Final Exam: 30% of total grade

Your instructor is Brian Weber,

Office: 3-121 Math Tower

Course Prerequisites

Calculus IV, Math 305 or equivalent (differential equations)
Linear Algebra, Math 310 or equivalent

Americans with Disabilities Act

If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Disability Support Services, ECC (Educational Communications Center) Building, room 128, (631) 632-6748 or http://studentaffairs.stonybrook.edu/dss/**. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential. Students who requiring assistance during emergency evacuation are encouraged to discuss their needs with their professors and Disability Support Services. For procedures and information, go to the following web site: http://www.www.ehs.stonybrook.edu/fire/disabilities.asp**

Homework Assignments

Homework 1	Due Wed., Feb. 4
Homework 2	Due Wed., Feb. 11
Homework 3	Due Wed, Feb. 25
Homework 4	Due Wed, Mar. 4
Homework 5	Due Wed, Mar. 18
Homework 6	Due Wed, Mar 25
Homework 7	Due Wed, Apr 15

Spring Break Special: Notes on Electrodynamics Problems in Electrodynamics

Homework 8	Due Wed, Apr 22

The Hopf fibration and the Berger spheres

Homework 9	Due Wed, May 6

Class Notes

Lecture 1 - Algebraic Special Relativity (Mon, Jan 26)

Lecture 2 - Geometric Special Relativity (Wed, Jan 28)

Lecture 3 - Groups and Symmetry (Fri, Jan 30)

Lecture 4 - Orthogonal and Lorentz transformations (Mon, Feb 2)

Lecture 5 - Geometry of Minkowski space (Fri, Feb 6)

Lecture 6 - Vector spaces, linear maps, and dual spaces (Mon, Feb 9)

Lecture 7 - More on Dual Spaces (Mon, Feb 16)

Lecture 8 - More on Dual Spaces II (Wed, Feb 18)

Lecture 9 - Tensor Products (Fri, Feb 20)

Lecture 10 - The Tensor algebra (Mon, Feb 23)

Lecture 11 - Tensors as maps, dual spaces, transformation properties, alternating tensors, and wedge products (Wed, Feb 25)

Lecture 12 - Metric linear algebra (Mon, Mar 2)

Lecture 13 - Vectors as directional derivatives (Mon, Mar 9)

Lecture 14 - Covectors (Wed, Mar 11)

Lecture 15 - Tensor fields and the metric (Fri, Mar 13)

Lecture 16 - Lie brackets and the *d*-operator (Mon, Mar 16)

Lecture 17 - Relation between the classical vector operations and the *d*-operator, 4-velocity and 3-velocity, and 4-momentum

Lecture 18 - The Einstein equation and conservation of energy-momentum

Lecture 19 - Stereographic projection

Lecture 20 - The classical Maxwell equations, and the covariant derivative. (Mon, Mar 30)

Lecture 21 - Warped products (Wed, Apr 1)

Lecture 22 - Gauge invariance, wave equations in electrodynamics, and variation of pathlength (Fri, Apr 3)

Lecture 23 - The parallel transport equation, the Riemann curvature tensor, and the Jacobi equation (Wed, Apr 15)

Lecture 24 - The Riemann Curvature tensor in compents (Fri, Apr 17)

Lecture 25 - The Stress-Energy-Momentum tensor (Mon, Apr 20)

Lecture 26 - Traces and Norms (Mon, Apr 27)

Lecture 27 - Covariant Derivatives (Wed, Apr 29)

Lecture 28 - Curvature Identities (Fri, May 1)

Lecture 29 - Conservation laws (Mon, May 4)

Lecture 30 - Equations of motion for relativistic fluids, and Poisson's equaiton (Wed, May 6)

Lecture 31 - The relativistic Maxwell equations, and the gravitational field equations (Mon, May 11)

Quiz (and test) prep material

Quiz 1, Feb. 6

Quiz 2, Feb. 20

Quiz 3, Feb. 27

Test 2, March 6

Test 3, March 27

Test 4, April 24

Final Exam

The Geometry of Physics

Topics

Special Relativity (2 weeks)
Euclidean space and Minkowski space
Path integrals and worldlines
The Galilean, Orthogonal, and Poincare groups
Force, momentum, and Newtonian mechanics in Minkowski space
Curved spaces and intuitive GR: Equivalence principle; matter = divergence of space-time itself

Linear algebra (3 weeks)
Vector spaces, linear maps, matrices, and matrix groups
Dual spaces
Tensors, tensor products, wedge products
Vector fields and directional derivatives
Covectors
The d-operator
Electromagnetism in Minkowski space

Geometry I (2 weeks)
How can you tell if the space around you is curved?
Paths, parallel transport, and geodesics
Variation
The principle of least action

Geometry II (3 weeks)
The metric: the mathematical quantification of space
The connection: an absolute derivative
Curvature
Application: the geometry of surfaces, Gaussian curvature, the Theorema Egregium.

Physics (3 weeks)
The equivalence principle, general covariance
Newtonian gravity and curved space
Einstein's Field Equations
Special solutions to Einstein's equations
Principle of least action and the Einstein-Hilbert action (time permitting!)

Syllabus for Math 401, The Geometry of Physics

Spring 2009

Instructor Brian Weber, brweber at math dot sunysb dot edu
Office 3-121 Math Tower

Course Text *A First Course in General Relativity*, Bernard F. Schutz

Additional Texts
The Geometry of Physics, Theodore Frankel
General Relativity, Robert M. Wald
The Large-Scale Structure of Space-Time, S.W. Hawking and G.F.R. Ellis

Prerequisites
Grade of C or higher in Math 303 or 305 or equivalent, and Math 310 or equivalent. Unofficially, if you are comfortable with partial derivatives, path integrals, vector spaces, and linear transformations, you should be okay.

Course outline
Starting with special relativity, we will develop the mathematical language necessary for understanding General Relativity and the invariant Maxwell equations. Along the way we will learn enough math and physics that students can start understanding the modern research in these areas. The material will be divided into 5 topics:

- Special relativity
- Linear algebra, Tensor analysis
- Global geometry: geodesics, energy, and variation
- Infinitesimal geometry: metrics, connections, and curvature
- The mathematics of General Relativity

We will have regular quizzes and homework assignments to make sure everyone stays current with the material. We will have a test after we conclude each topic.

Exams We will have 4 in-class tests and a final exam.
Test 1: Friday Feb 13 (10% of grade)
Test 2: Friday Mar 6 (20% of grade)
Test 3: Friday Mar 20 (10% of grade)
Test 4: Friday April 17 (10% of grade)
Final: TBA (30% of grade)

Homework (10% of grade)
One problem set will be due each week. The problems will be turned in at the *beginning* of class each Wednesday. As a fair warning, you will have to work hard to be successful in this class. If you fall seriously behind on the homework, you will not be able to keep up in class and will not be prepared for the exams. You are encouraged to work in groups, but you must write up your own solutions.

Quizzes (10% of grade)
There will be a short quiz at the beginning of class each Friday (except the Fridays of scheduled tests). The purpose is to help everyone stay current with the mathematical techniques introduced during the prior week.

Makeup policy

All of your responsibilities for this class have been announced well ahead of time, namely in the first week of classes. Thus almost no requests for makeup homeworks or exams will be granted. The only exceptions, assuming evidence is provided, will be for serious illness, family emergency, or an unforeseeable catastrophe (tornado, car wreck, etc).

Academic Integrity

Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty are required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Technology & Management, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty, please refer to the academic judiciary website at http://www.stonybrook.edu/uaa/academicjudiciary/.

Course Withdrawals

The academic calendar, published in the Undergraduate Class Schedule, lists various dates that students must follow. Permission for a student to withdraw from a course after the deadline may be granted only by the Arts and Sciences Committee on Academic Standing and Appeals or the Engineering and Applied Sciences Committee on Academic Standing. The same is true of withdrawals that will result in an underload. A note from the instructor is not sufficient to secure a withdrawal from a course without regard to deadlines and underloads.

The Hopf Fibration and the Berger Spheres

Due —

Introduction $\mathbb{R}^4$ is the set of ordered quadruples of real numbers (x^1, x^2, x^3, x^4), along with the Euclidean distance function:

$$\text{dist}\left((x^1, x^2, x^3, x^4),\,(y^1, y^2, y^3, y^4)\right) \;=\; (x^1-y^1)^2 \,+\, (x^2-y^2)^2 \,+\, (x^3-y^3)^2 \,+\, (x^4-y^4)^2.$$

One can identify $\mathbb{R}^4$ with $\mathbb{C}^2$, the set of ordered pairs or complex numbers: a point

$$(z^1, z^2) = (x^1 + iy^1, x^2 + iy^2) \in \mathbb{C}^2$$

can be identified with the point

$$(x^1, y^1, x^2, y^2) \in \mathbb{R}^4.$$

If (z^1, z^2), (w^1, w^2) are two points in $\mathbb{C}^2$, the distance between then is

$$\text{dist}\left((z^1, z^2),\,(w^1, w^2)\right) \;=\; |z^1 - w^1|^2 \,+\, |z^2 - w^2|^2.$$

Recall that if $z \in \mathbb{C}$ then $|z|^2 = z\bar{z}$.

Let $\mathbb{S}^3 \subset \mathbb{R}^4$ denote the unit 3-sphere, defined to be

$$\begin{aligned}
\mathbb{S}^3 \;&\triangleq\; \left\{\,(x^1, x^2, x^3, x^4) \in \mathbb{R}^4 \; s.t. \; (x^1)^2 + (x^2)^2 + (x^3)^2 + (x^4)^2 \;=\; 1\,\right\} \\
&=\; \left\{\,(z^1, z^2) \in \mathbb{C}^2 \; s.t. \; |z^1|^2 + |z^2|^2 \;=\; 1\,\right\}.
\end{aligned}$$

1 The Hopf action

Problem 1) If $p, q \in \mathbb{C}^2$, the distance $\text{dist}(p, q)$ can be calculated in the $\mathbb{C}^2$ sense or the $\mathbb{R}^4$ sense. Prove that the distance is the same regardless of which distance function is used.

Def Given any $\theta \in \mathbb{R}$, let $\psi_\theta : \mathbb{C}^2 \to \mathbb{C}^2$ be the map

$$\psi_\theta(z^1, z^2) \;=\; \left(e^{i\theta} z^1, e^{i\theta} z^2\right).$$

Note that ψ_θ is the identity map if and only if θ is a multiple of 2π.

Problem 2) Given any $\theta \in \mathbb{R}$, prove that ψ_θ is an isometry that fixes the origin $(0, 0) \in \mathbb{C}^2$. Prove that, unless θ is a multiple of 2π, then $(0, 0)$ is the *only* fixed point of ψ_θ. Finally, prove that if $(z^1, z^2) \in \mathbb{S}^3$ then also $\psi_\theta(z^1, z^2) \in \mathbb{S}^3$.

Remark Thus $\psi_\theta : \mathbb{S}^3 \to \mathbb{S}^3$ is an isometric action: this is known as the Hopf action.

Problem 3) If $p = (z^1, z^2) \in \mathbb{S}^3$, the *orbit* of p under the Hopf action is defined to be the set of all $\psi_\theta(p)$ as θ varies. Prove that the orbit of any point $p \in \mathbb{S}^3$ is a circle of radius 1.

Remark Each orbit of ψ_θ is a circle, and of course each point of $\mathbb{S}^3$ lies in an orbit. Thus the union of the orbits (each a copy of $\mathbb{S}^1$) comprises $\mathbb{S}^3$. One says that $\mathbb{S}^3$ is *fibered* by $\mathbb{S}^1$; one calls the $\mathbb{S}^1$ orbits the *fibers*. The fibration of $\mathbb{S}^3$ by copies of $\mathbb{S}^1$ is called the *Hopf fibration*.

Problem 4) Considering $\mathbb{S}^3 \subset \mathbb{R}^4$ (instead of $\mathbb{S}^3 \subset \mathbb{C}^2$), prove that

$$\psi_\theta(x^1, x^2, x^3, x^4) \;=\; \big(x^1 \cos\theta \,-\, x^2 \sin\theta,\, x^1 \sin\theta \,+\, x^2 \cos\theta,\, x^3 \cos\theta \,-\, x^4 \sin\theta,\, x^3 \sin\theta \,+\, x^4 \cos\theta\big)\,.$$

The Hopf action ψ_θ produces an *action field*, which is just the velocity field of the rotation. Letting $\frac{d}{d\theta}$ denote the action field, compute $\frac{d}{d\theta}$ in terms of the coordinate fields $\left\{\frac{\partial}{\partial x^i}\right\}$.

2 The Hopf map

Def The *Hopf map* $\Psi : \mathbb{S}^3 \to \mathbb{S}^2$ is a continuous map defined as follows. Regard $\mathbb{S}^2$ to be $\mathbb{C} \cup \{\infty\}$ (via stereographic projection). If $p \in \mathbb{S}^3$ is an arbitrary point, then $p = (z^1, z^2)$ with $|z^1|^2 + |z^2|^2 = 1$. Define

$$\begin{aligned} \Psi(p) \;&\in\; \mathbb{C} \cup \{\infty\} \\ \Psi\left((z^1, z^2)\right) \;&=\; \frac{z^1}{z^2}. \end{aligned}$$

In contrast to the lower dimensional situation, there are NO (topologically nontrivial) continuous maps from $\mathbb{S}^2$ to $\mathbb{S}^1$.

Problem 5) Prove that $\Psi : \mathbb{S}^3 \to \mathbb{S}^2$ is onto. Which point in $\mathbb{S}^3$ maps to the 'point at infinity' on $\mathbb{S}^2$?

Problem 6) Prove that $p, q \in \mathbb{S}^3$ belong to the same Hopf fiber if and only if $\Psi(p) = \Psi(q)$.

3 The Berger spheres

Remark The Hopf map $\mathbb{S}^3 \to \mathbb{S}^2$ is an example of a *submersion*: a map from a higher dimensional space into (in this case onto) a lower dimensional space. As we have seen, Ψ takes the 3-sphere and collapses the Hopf circles (1-dimensional objects) to points. What is left is a 2-sphere (a 2-dimensional object). The purpose of this section is to see the process occurring dynamically: we will construct a family of metrics that shrinks the Hopf circles to points.

Problem 7) Let X, Y, and Z be vector fields given by

$$\begin{aligned} X = \frac{d}{d\theta} &= -x^2\frac{\partial}{\partial x^1} + x^1\frac{\partial}{\partial x^2} - x^4\frac{\partial}{\partial x^3} + x^3\frac{\partial}{\partial x^4} \\ Y &= -x^3\frac{\partial}{\partial x^1} + x^4\frac{\partial}{\partial x^2} + x^1\frac{\partial}{\partial x^3} - x^2\frac{\partial}{\partial x^4} \\ Z &= x^4\frac{\partial}{\partial x^1} + x^3\frac{\partial}{\partial x^2} - x^2\frac{\partial}{\partial x^3} - x^1\frac{\partial}{\partial x^4}. \end{aligned}$$

Prove that X, Y, and Z are all tangent to $\mathbb{S}^3$. Also prove that $|X|^2 = |Y|^2 = |Z|^2 = 1$, and that X, Y, and Z are mutually orthogonal.

Problem 8) Prove that

$$\begin{aligned} [X, Y] &= 2Z \\ [Y, Z] &= 2X \\ [Z, X] &= 2Y. \end{aligned}$$

This also proves that X, Y, Z cannot be considered coordinate fields.

<u>**Def**</u> We shall define the *Berger metric* on $\mathbb{S}^3$ as follows. Let $\eta^1 \triangleq X_\flat$, $\eta^2 \triangleq Y_\flat$, and $\eta^3 \triangleq Z_\flat$. Given $\alpha \in \mathbb{R}$, let

$$g_\alpha = \alpha^2\eta^1 \otimes \eta^1 + \eta^2 \otimes \eta^2 + \eta^3 \otimes \eta^3.$$

If $\alpha = 1$, then this is precisely the metric that $\mathbb{S}^3$ inherits from the ambient space $\mathbb{R}^4$.

Problem 9) Each of the metrics g_α has an associated covariant derivative ∇^α. Find

$$\nabla^\alpha_X Y \quad \nabla^\alpha_Y Z \quad \nabla^\alpha_Z X.$$

(It is best to use the Koszul formula directly). Note that the values of $\nabla^\alpha_Y X$, $\nabla^\alpha_Z Y$, $\nabla^\alpha_X Z$ are now automatic.

Problem 10) Compute the sectional curvatures

$$sec(X, Y) \quad sec(Y, Z) \quad sec(X, Z).$$

As measured in the metric g_α, the Hopf fibers are circles of radius α. As $\alpha \to 0$ and the fibers contract to points, the sectional curvatures remain bounded. This is a process known as *collapse with bounded curvature*.

Notes on Electrodynamics

1 The classical Maxwell equations

$$\begin{array}{ll}
\nabla \cdot \vec{B} \ = \ 0 & \text{no magnetic sources} \\
\nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} \ = \ 0 & \text{Faraday's law} \\
\nabla \times \vec{B} - \epsilon\mu \frac{\partial \vec{E}}{\partial t} \ = \ 4\pi\mu\vec{J} & \text{Ampere} - \text{Maxwell law} \\
\nabla \cdot \vec{E} \ = \ \frac{4\pi}{\epsilon}\rho & \text{Gauss' Law}
\end{array}$$

2 The Riemannian duality operator

Let M be an n-dimensional Riemannian space, with coordinates $\{x^1, \ldots, x^n\}$ and Riemannian metric g_{ij}. The duality operator $* : \Omega^p(M) \to \Omega^{n-p}(M)$ is defined (implicitly) according to the rule

$$\left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) \wedge * \left(dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) \ = \ \sqrt{\det(g_{ij})}dx^1 \wedge \cdots \wedge dx^n,$$

and extended linearly. The n-form $\sqrt{\det(g_{ij})}dx^1 \wedge \cdots \wedge dx^n$ is known as the *volume form*.

3 The Lorentzian duality operator

Let $\{x^0, \ldots, x^3\}$ be standard coordinates on Minkowski $(1+3)$-space. Without resorting to complex numbers, it is *not* possible to define the $*$ operator as naturally for Minkowsi space as it is for Riemannian spaces. The Lorentzian $*$ operator is defined, in a seemingly rather ad hoc way, to be the linear operator $* : \Omega^p(M) \to \Omega^{4-p}(M)$ given as follows:

$$* : \Omega^0(M) \to \Omega^4(M) \qquad *1 \ = \ \frac{1}{c^3} dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$$

$$\begin{array}{lrcl}
* : \Omega^1(M) \to \Omega^3(M) \qquad & *dx^0 & = & \frac{1}{c^3} dx^1 \wedge dx^2 \wedge dx^3 \\
& *dx^1 & = & \frac{1}{c} dx^0 \wedge dx^2 \wedge dx^3 \\
& *dx^2 & = & -\frac{1}{c} dx^0 \wedge dx^1 \wedge dx^3 \\
& *dx^3 & = & \frac{1}{c} dx^0 \wedge dx^1 \wedge dx^2
\end{array}$$

$$
\begin{array}{ll}
* : \Omega^2(M) \to \Omega^2(M) & \begin{aligned} *(dx^0 \wedge dx^1) &= -\frac{1}{c} dx^2 \wedge dx^3 \\ *(dx^0 \wedge dx^2) &= \frac{1}{c} dx^1 \wedge dx^3 \\ *(dx^0 \wedge dx^3) &= -\frac{1}{c} dx^1 \wedge dx^2 \\ *(dx^1 \wedge dx^2) &= c dx^0 \wedge dx^3 \\ *(dx^1 \wedge dx^3) &= -c dx^0 \wedge dx^2 \\ *(dx^2 \wedge dx^3) &= c dx^0 \wedge dx^1 \end{aligned}
\end{array}
$$

$$
\begin{array}{ll}
* : \Omega^3(M) \to \Omega^1(M) & \begin{aligned} *(dx^1 \wedge dx^2 \wedge dx^3) &= c^3 dx^0 \\ *(dx^0 \wedge dx^2 \wedge dx^3) &= c dx^1 \\ *(dx^0 \wedge dx^1 \wedge dx^3) &= -c dx^1 \\ *(dx^0 \wedge dx^1 \wedge dx^1) &= c dx^1 \end{aligned}
\end{array}
$$

$$
* : \Omega^4(M) \to \Omega^0(M) \qquad *(dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3) = -c^3.
$$

4 The divergence operator

Associated to the operator $d : \Omega^p(M) \to \Omega^{p+1}(M)$ is its *adjoint* operator, $\delta : \Omega^p(M) \to \Omega^{p-1}(M)$.

In the Riemannian setting, it is defined by $\delta = (-1)^{pn+p+1} * d*$.

In the Lorentzian setting, it is defined by

$$
\begin{array}{ll}
\delta : \Omega^1(M) \to \Omega^0(M) & \delta = *d* \\
\delta : \Omega^2(M) \to \Omega^1(M) & \delta = - * d * \\
\delta : \Omega^3(M) \to \Omega^2(M) & \delta = *d* \\
\delta : \Omega^4(M) \to \Omega^3(M) & \delta = *d*
\end{array}
$$

5 Gauge invariance

One of Maxwell's equations is $\nabla \cdot \vec{B} = 0$, which means $\vec{B}$ is a *pure curl*: $\vec{B} = \nabla \times \vec{A}$. Replacing $\vec{A}$ by $\vec{A} + \nabla f$ does not change the fact that $\vec{B} = \nabla \times \vec{A}$, so there is considerable freedom in choosing the vector potential $\vec{A}$.

Now consider Faraday's law, $\nabla \times \vec{E} + \frac{\partial}{\partial t}\vec{B} = 0$. Putting $\nabla \times \vec{A}$ in for $\vec{B}$ this reads

$$\nabla \times \left(\vec{E} + \frac{\partial \vec{A}}{\partial t} \right) \;=\; 0$$

so that $\vec{E} + \frac{\partial}{\partial t}\vec{A}$ is a *pure gradient*:

$$\vec{E} \,+\, \frac{\partial}{\partial t}\vec{A} \;=\; \nabla\varphi.$$

The quantity $\nabla\varphi$ is called the *generalized electrical potential* or the *electrical pseudopotential*. Once $\vec{A}$ is chosen, φ is defined up to a constant.

A particular choice of $(\varphi, \vec{A})$ is known as a *choice of gauge*.

6 The Coulomb Gauge

We prove that it is possible to choose $\vec{A}$ so that $\nabla \cdot \vec{A} = 0$; this is known as the *Coulomb gauge*. The problem is to choose a function f so that $\nabla \cdot \left(\vec{A} + \nabla f \right) = 0$. This is equivalent to finding a function f so that $\triangle f = -\nabla \cdot \vec{A}$. This is a Laplace equation for f, which is known to be solvable. (QED)

Due to Gauss' law $\nabla \cdot \vec{E} = \frac{4\pi}{\epsilon}\rho$, in the Coulomb gauge we have $\triangle\varphi = \frac{4\pi}{\epsilon}\rho$ (if ρ is given, this is known as the *Poisson equation* in the unknown φ).

The Coulomb gauge is useful in solving electrostatic problems, where it makes calculating $\vec{A}$ easy. It is not very useful in electrodynamic problems.

7 The Lorentz Gauge

Again let $\vec{A}$ be the magnetic vector potential: $\vec{B} = \nabla \times \vec{A}$. Again we have the electric pseudopotential φ, defined (up to constant addition) by

$$\vec{E} \,+\, \frac{\partial \vec{A}}{\partial t} \;=\; \nabla\varphi$$

This time, we choose $(\varphi, \vec{A})$ so that

$$\nabla \cdot \vec{A} \,-\, \frac{1}{c^2}\frac{\partial \varphi}{\partial t} \;=\; 0.$$

The Lorentz gauge is useful is solving electrodynamic problems.

8 Units for the classical quantities

SI base units are kilograms kg, meters m, seconds s, and Amperes A.

$$\begin{aligned} 1V &= \frac{1kg \cdot m^2}{A \cdot s^3} & 1C &= 1A \cdot s \\ \epsilon &\sim \frac{As}{Vm} = \frac{Farads}{m} & \mu &\sim \frac{Vs}{Am} = \frac{Henrys}{m} \\ \vec{E} &\sim \frac{V}{m} & \vec{B} &\sim \frac{Vs}{m^2} \\ \rho &\sim \frac{As}{m^3} = \frac{Coulombs}{m^3} & \vec{J} &\sim \frac{A}{m^2}. \end{aligned}$$

The symbol $\sim$ here means "has units".

Problems in Electrodynamics

Due —

1 Problems in Classical Electrodynamics

In this section, $\vec{E}$, $\vec{B}$, etc. will al be considered "classical" vector fields, that is, lists of 3 numbers attached to each point of 3-space.

Problem 1) Write out the Lorentz force law in components.

Problem 2) Write out the four Maxwell equations in components.

Problem 3) Using the charge- and current-free Maxwell equations, derive the following wave equations when $\vec{A}$ is the Lorentz gauge:

$$\triangle \vec{A} - \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} \;=\; 0 \qquad \triangle \phi - \frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} \;=\; 0$$

Hint: You will need the classical vector identity $\nabla \times \nabla \times \vec{A} \;=\; \nabla(\nabla \cdot \vec{A}) \;-\; \triangle \vec{A}$.

2 Problems in Minkowski Analysis

Problem 1) Given a 1-form $A = A_i dx^i$ $(0 \leq i \leq 3)$, find δA.

Problem 2) Given a function f, prove that $\delta df = \Box f$, where

$$\Box \;=\; \left(\frac{\partial}{\partial x^0}\right)^2 \;-\; \left(c\frac{\partial}{\partial x^1}\right)^2 \;-\; \left(c\frac{\partial}{\partial x^2}\right)^2 \;-\; \left(c\frac{\partial}{\partial x^3}\right)^2$$

is the *D'Alembertian* (the Lorentzian analog of the Laplacian).

Problem 3) Given an arbitrary 2-form F, find dF.

Problem 4) Given an arbitrary 2-form F, find δF.

3 Problems in Relativistic Electrodynamics

In this section, E, B, and A will be 1-forms: $E = E_1 dx^1 + E_2 dx^2 + E_3 dx^3$, $B = B_1 dx^1 + B_2 dx^2 + B_3 dx^3$.

Problem 1) Put $F = E \wedge dx^0 \, + \, c * \left(B \wedge dx^0 \right)$. This is called the Faraday tensor. Write out the tensor $F = F_{ij} = F_{ij} dx^i \otimes dx^j$ in matrix form.

Problem 2) Prove that the Faraday tensor has units $Volt \cdot s$.

Problem 3) Let $\gamma(\tau)$ be a path in Minkowski space. If $v = v^i = \frac{d}{d\tau}$ is its velocity vector, prove that the Lorentz force law is

$$\frac{dp_i}{d\tau} = qF_{ij}v^j,$$

where $p = -mc^2 v_\flat$ is the momentum. (Also, check that the units are consistent).

Problem 4) Using Faraday's law and $\nabla \cdot \vec{B} = 0$, prove that $dF = 0$.

Problem 5) If ρ is charge density and $\vec{J} = (J^1, J^2, J^3)$ is current density, define the current 4-vector to be

$$J \quad = \quad \rho \frac{\partial}{\partial x^0} \, + \, J^1 \frac{\partial}{\partial x^1} \, + \, J^2 \frac{\partial}{\partial x^2} \, + \, J^3 \frac{\partial}{\partial x^3}.$$

Show that the current 4-vector J has (consistent) units $Ampere/m^3$. Show that the 1-form $4\pi\epsilon^{-1}c^2 J_\flat$ has (consistent) units $Volt/s$.

Using the Maxwell-Ampere law and Gauss' law, prove that

$$\delta F \quad = \quad -4\pi\epsilon^{-1}c^2 J_\flat.$$

Recall that $\epsilon\mu = \frac{1}{c^2}$.

Problem 6) Since $dF = 0$, there is a 1-form A such that $F = dA$. As before, there is considerable freedom in choosing A: it can be modified into $A + df$, where f is any 0-form, without changing the equation $F = dA$. Prove that the choice of A so that $\delta A = 0$ is equivalent to working in the Lorentz gauge.

Problem 6) Prove that the source-free Maxwell equations: $dF = 0$, $\delta F = 0$ imply a wave equation for A: $\Box A = 0$. (That is, prove $\delta dA = \Box A$ when $\delta A = 0$).

Review for final

May 9, 2009

You must be familiar with the following concepts from Special Relativity:

- Euclidean space, Minkowski space, orthogonal transformation, Lorentz transformation
- Notation: $\mathbb{R}^n$, $\mathbb{R}^{k,n}$.
- Geometry of Lorentz space: Space-like, time-like, and null intervals. Pathlength. Path energy. Light cone. Pseudospheres.
- Orthogonal group, Euclidean group, Lorentz group, Poincare group.
- Notation: $O(n)$, $E(n)$, $O(k,n)$, $P(k,n)$.
- Relation between classical and relativistic velocity: $v = (\gamma, \gamma\vec{v})$.
- Momentum of a particle, and its relation with classical energy and momentum:

$$p = -mc^2 v_\flat \; = \; (-E, \gamma\vec{p}).$$

- Conservation of energy-momentum for interacting particles.
- The Cauchy stress tensor.
- The Stress-energy-momentum tensor, and the conservation law.

You must be familiar with the following concepts from linear algebra:

- Vector space, covector space, tensor spaces.
- Linear algebras, linear operators.
- Linear maps as tensors, and vice-versa.
- Wedge products.

- Notation: V, V^*, $\bigotimes^{r,s} V$, $\bigwedge^r V^*$.

You must be familiar with the following concepts from differential calculus:

- Coordinates, vectors, covectors, tensors, alternating tensors.
- The d-operator.
- The $*$-operator.
- Covariant derivatives of vector and tensor fields, and Christoffel symbols.
- The geodesic equation, and its derivation.
- The Riemann curvature operator.
- The Jacobi equation, and its derivation.
- Sectional curvature, Ricci curvature and scalar curvature.
- The two Bianchi identities.

You must be familiar with the following concepts from electromagnetics:

- The Classical Maxwell equations.
- The magnetic vector potential and the electrical pseudopotential.
- Wave equations for $\vec{E}$, $\vec{B}$, $\vec{A}$ and φ.
- The relativistic Maxwell equations.
- The electromagnetic 4-potential.
- The 4-potential wave equation.

You must be familiar with the following concepts from General relativity:

- The interpretation of gravity.
- Newton's law of gravitation, and the gravitational potential.
- The 'plausibility argument' for the field equations.
- The field equations.

Lecture 1 - Algebraic Special Relativity

January 27, 2009

From a mathematical point of view, special relativity is three postulates: 1) the existence of inertial reference frames, 2) the constancy of the speed of light in any reference frame, and 3) the conservation of energy-momentum. Today we will concern ourselves with deductions from the first two postulates, and leave the third for after we geometrize SR.

1 Inertial reference frames and Galilean relativity

We want to say that an inertial reference frame is the view of the universe (ie, the way of measuring space and time) of an unaccelerated observer. This is problematic however: an astronaut orbiting in a windowless spaceship would say he is unaccelerated, though most other observers would disagree. We will simply define, then accept as a postulate the existence of inertial reference frames, and leave aside the question of whether, and to what degree, any exist in the universe. We define an inertial reference frame to be a representation (more precisely, an isomorphism) of space-time as $\mathbb{R}^4$, the set of coordinates (x^0, x^1, x^2, x^3), so that any constant-time (ie constant x^0) slice is ordinary Euclidean 3-space, and so that time progresses uniformly (local clocks mark time, namely the progress of the x^0 coordinate, uniformly).

If you accept basic Newtonian concepts of conservation of energy and momentum, it is simple to argue that such a choice of coordinates is impossible for the astronaut in the example. Thus we must accept that in physical reality, inertial reference frames exist, at best, only approximately and only locally.

In Galilean relativity, we accept the notion of inertial reference frame, but postulate (as common-sense would have it) that time as measurements by different clocks coincide perfectly, and measurements of spatial displacements are identical regardless of the observer. This implies that light travels at different speeds in different frames. But that violates Maxwell's equations, which should hold regardless of frame. Possibly one could discard Maxwell's equations or discard the principle that physics be the same regardless of frame. Instead we discard the Galilean postulate that time and displacement measurements are frame-independent, and adopt the postulate of the constancy of the speed of light, as implied by Maxwell's equations.

2 Time Dilation

Consider the following thought experiment. You are in a spaceship, floating freely in space so that you perceive yourself motionless. Another spaceship is approaching on a collision course with speed v. It just misses you, but at the instant of closest approach a member of that ship's crew turns on a flashlight, whose beam hits the spaceship's ceiling a short time later.

Considering the two events: the release of a photon from the flashlight, and its subsequent absorption by the ceiling. From the crewmember's viewpoint, the light travels l meters with speed c, so the time between events is $\triangle\tau = l/c$. From your viewpoint the light travels a distance of $\sqrt{l^2 + v\triangle t}$ with speed c, so the time between events is $\triangle t = \sqrt{l^2 + v\triangle t}/c$. Eliminating l from these equations we get

$$\triangle t = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}\triangle\tau.$$

Thus the time interval between events is longer according to your view of the universe (your inertial reference frame), by a factor of

$$\gamma = \gamma_v \triangleq \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}},$$

than from the point of view of the whizzing-by crewmember.

3 Space dilation

This time imagine that the crewmember shone the flashlight toward the front of the ship. The two events under consideration are the emission of a photon from the flashlight and it subsequent absorption by the wall at a distance (measured by the crewmember) of $\triangle\xi$. The crewmember measures the time in transit to be $\triangle\tau = \triangle\xi/c$.

By time dilation, we measure the time between events to be $\triangle t = \gamma\triangle\tau$, so the distance traveled is

$$\begin{aligned}\triangle x &= c\triangle t \\ &= c\gamma\triangle\tau \\ &= \gamma\,\triangle\xi.\end{aligned}$$

4 Simultaneity

Assume we are standing at the center of an iron truss of length $2l$ with a light at either end. Each light emits a photon, which meet at precisely the center of the beam. We conclude that each photon has been traveling for $\triangle t = l/c$ seconds.

Now consider another observer, whizzing past at speed v, who happens to be directly overhead at the instant the two photons meet. The oncoming photon (photon A) has been in transit for a time of $\triangle\tau_A$ and so was emitted when the end of the beam was a distance of $\frac{l}{\gamma} + v\,\triangle\tau_A$ away, so the travel time was

$$\begin{aligned}
\triangle\tau_A &= \left(\frac{l}{\gamma} + v\,\triangle\tau_A\right)/c \\
\triangle\tau_A &= \frac{l/\gamma}{c + v}
\end{aligned}$$

The photon that caught up to us from behind (photon B) has been in transit for a time of $\triangle\tau_B$ and so was emitted when the end of the beam was a distance of $\frac{l}{\gamma} - v\triangle\tau_B$ away, so the travel time was

$$\begin{aligned}
\triangle\tau_B &= \left(\frac{l}{\gamma} - v\,\triangle\tau_B\right)/c \\
\triangle\tau_B &= \frac{l/\gamma}{c - v}
\end{aligned}$$

We are forced to conclude that the photons have been in transit for different periods of time, so were not released simultaneously.

5 Formulas

Assume (t, x^1, x^2, x^3) is a space-time coordinate system, and $(\tau, \xi^1, \xi^2, \xi^3)$ is a space-time coordinate system of an observer moving with speed $\vec{v} = (v_1, v_2, v_3)$ with respect to the original coordinate system.

If p and q are events in space-time, then the time between events as measured in the first system and the second system is related by

$$\triangle t \;=\; \gamma_v\,\triangle\tau \;=\; \frac{1}{\sqrt{1 - \frac{|\vec{v}|^2}{c^2}}}\,\triangle\tau$$

and the displacement between the two events as measured the two systems are related by

$$\begin{aligned}
\triangle x^1 &= \gamma_{v_1} \triangle \xi^1 = \frac{1}{\sqrt{1 - \frac{(v_1)^2}{c^2}}} \triangle \xi^1 \\
\triangle x^2 &= \gamma_{v_2} \triangle \xi^2 = \frac{1}{\sqrt{1 - \frac{(v_2)^2}{c^2}}} \triangle \xi^2 \\
\triangle x^3 &= \gamma_{v_3} \triangle \xi^3 = \frac{1}{\sqrt{1 - \frac{(v_3)^2}{c^2}}} \triangle \xi^3.
\end{aligned}$$

Lecture 2 - Geometric Special Relativity

Jan 28, 2009

1 Euclidean analytic geometry

Euclidean n-space has no natural origin and so is not naturally a vector space, despite how we are often taught to regard it. It makes no sense, intrinsically, to speak of a "position vector."

Nevertheless, when working in 3-space say, we often choose an origin and post x-, y-, and z-coordinate axes, and regard it as a vector space with basis $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. It is important to remember that Euclidean space is a purely geometric object and does not come equipped with axes. Supplying space with axes is a purely arbitrary construction.

Euclidean space does however come with notions of angles, distances, and lines. Once an orthonormal coordinate system is chosen, we may regard the geometric object of Euclidean n-space as the algebraic object $\mathbb{R}^n$, which does happen to be a vector space, and is easier to work with on paper.

1.1 Paths, pathlength, and path integrals

Consider Euclidean n-space, and choose an orthonormal coordinate system $(x^1, \ldots, x^n)$, thereby identifying it with $\mathbb{R}^n$.

Consider a path $\mathbf{\Gamma}(t) = (x^1(t), \ldots, x^n(t))$ through n-space. The path's velocity is

$$\frac{d\mathbf{\Gamma}}{dt}(t) \quad = \quad \left(\frac{dx^1}{dt}, \ldots, \frac{dx^n}{dt}\right)_{(x^1(t),\ldots,x^n(t))}.$$

It's speed is the velocity vector's length:

$$\begin{aligned}\left|\frac{d\mathbf{\Gamma}}{dt}\right| &= \sqrt{\left\langle\frac{d\mathbf{\Gamma}}{dt}, \frac{d\mathbf{\Gamma}}{dt}\right\rangle} \\ &= \sqrt{\left(\frac{dx^1}{dt}\right)^2 + \ldots + \left(\frac{dx^n}{dt}\right)^2}.\end{aligned}$$

If s is the arclength, we have $ds = |d\mathbf{\Gamma}| = |d\mathbf{\Gamma}/dt|\, dt$. To find the length of the path $\mathbf{\Gamma}(t)$ between $t = t_1$ and $t = t_2$, you integrate the path's arclength:

$$\int ds = \int_{t_1}^{t_2} \left|\frac{d\mathbf{\Gamma}}{dt}\right| dt.$$

Given a function $f(x^1, \ldots, x^n)$ defined on $\mathbb{R}^n$, it is possible to integrate f along the path $\mathbf{\Gamma}$:

$$\int_{\mathbf{\Gamma}} f\, ds = \int_{t_1}^{t_2} f\left(x^1(t), \ldots, x^n(t)\right) \left|\frac{d\mathbf{\Gamma}}{dt}\right| dt.$$

Finally, it is possible to take the derivative of f along the path Γ. Using the chain rule, we get

$$\frac{df}{dt} = \frac{dx^1}{dt}\frac{\partial f}{\partial x^1} + \frac{dx^2}{dt}\frac{\partial f}{\partial x^2} + \ldots + \frac{dx^n}{dt}\frac{\partial f}{\partial x^n}.$$

This indicates that we can write

$$\frac{d}{dt} = \frac{dx^1}{dt}\frac{\partial}{\partial x^1} + \frac{dx^2}{dt}\frac{\partial}{\partial x^2} + \ldots + \frac{dz^n}{dt}\frac{\partial}{\partial x^n}.$$

2 Minkowski analytic geometry

2.1 Minkowski's Pythagorean theorem

Consider the spaceships from the previous lecture. From our perspective, the crewmember traveled $\triangle x$ meters between when the light left the flashlight and when it struck the ceiling, and did so in $\triangle t$ seconds. The crewmember's own experience is that no distance was traveled, but that $\triangle\tau$ seconds passed. Using the time-dilation equation $\triangle t = \gamma_v \triangle\tau$, let's compute

$$\begin{aligned}(\triangle t)^2 - \frac{1}{c^2}(\triangle x)^2 &= (\triangle t)^2\left(1 - \frac{v^2}{c^2}\right) \\ &= \gamma_v^2(\triangle\tau)^2\gamma_v^{-2} \\ &= (\triangle\tau)^2.\end{aligned}$$

Graphically, $\triangle t$ and $\triangle x$ form the legs of a right triangle with hypotenuse $\triangle\tau$, so this gives a new pythagorean theorem. If p and $\overline{p}$ are points in space-time ("events"), then the "distance" (rather, proper time) between then is given by

$$\begin{aligned} p &= \left(x^0,\, x^1,\, x^2,\, x^3\right) \\ \overline{p} &= \left(\bar{x}^0,\, \bar{x}^1,\, \bar{x}^2,\, \bar{x}^3\right) \end{aligned}$$

$$\begin{aligned} \left|\overline{p\overline{p}}\right|^2 &= \left(\triangle x^0\right)^2 - \frac{1}{c^2}\left(\triangle x^1\right)^2 - \frac{1}{c^2}\left(\triangle x^2\right)^2 - \frac{1}{c^2}\left(\triangle x^3\right)^2 \\ &= \left(x^0 - \bar{x}^0\right)^2 - \frac{1}{c^2}\left(x^1 - \bar{x}^1\right) - \frac{1}{c^2}\left(x^2 - \bar{x}^2\right) - \frac{1}{c^2}\left(x^3 - \bar{x}^3\right)^2. \end{aligned}$$

Four-space that obeys this version of the Pythagorean theorem is called Minkowski space.

<u>The distinction between $\mathbb{R}^4$ and $\mathbb{R}^{1,3}$</u>. After choosing a coordinate system for space-time (a.k.a. inertial frame of reference) with coordinates (x^0, x^1, x^2, x^3), space-time becomes identified with $\mathbb{R}^4$. But notice that the time and space dimensions are treated differently due to our new Pythagorean theorem: the square of the time-dimension has a coefficient of $+1$, whereas the squared space-dimensions have coefficients $-1/c^2$ (giving these terms, by the way, units of time-squared). Because the geometry Minkowski space is so radically different from the geometry of Euclidean space, we usually call it $\mathbb{R}^{1,3}$ instead of $\mathbb{R}^4$, or sometimes "1+3 dimensional space" rather than "4 dimensional space."

Notice that proper time (i.e. distance in 1+3-space) can either be real and positive, zero, or imaginary. Respectively, these are called "time-like", "light-like" or "null", and "space-like" distances.

2.2 Paths in Minkowski space

Given an arbitrary path

$$\mathbf{\Gamma}(t) = \left(x^0(t),\, x^1(t),\, x^2(t),\, x^3(t)\right)$$

through space-time, it's *four-velocity* is

$$\frac{d\mathbf{\Gamma}}{dt} = \left(\frac{dx^0}{dt}, \frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right).$$

The meaning of this vector is not the same as for paths through Euclidean space: indeed, the 4-velocity is the path's velocity through proper time. Note that the path's speed through proper time is

$$\left|\frac{d\mathbf{\Gamma}}{dt}\right| = \sqrt{\left(\frac{dx^0}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx^1}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx^2}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx^3}{dt}\right)^2}$$

This could be well imaginary! At points where the speed is positive the path is called "time-like", at points where the speed is 0 the path is called "light-like" or "null", and at points where the speed is imaginary the path is called "space-like".

If s is the arclength parameter (perhaps the term "arctime parameter" is better) along the path $\mathbf{\Gamma}$ and ds the arclength element, the infinitesimal Pythagorean theorem for Minkowski space gives

$$(ds)^2 = \left(dx^0\right)^2 - \frac{1}{c^2}\left(dx^1\right)^2 - \frac{1}{c^2}\left(dx^2\right)^2 - \frac{1}{c^2}\left(dx^3\right)^2 .$$

Thus the Minkowski arclength is

$$\begin{aligned} \int_{\mathbf{\Gamma}} ds &= \int_{\mathbf{\Gamma}} \sqrt{(dx^0)^2 - \frac{1}{c^2}(dx^1)^2 - \frac{1}{c^2}(dx^2)^2 - \frac{1}{c^2}(dx^3)^2} \\ &= \int_{t_1}^{t_2} \sqrt{\left(\frac{dx^0}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx^1}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx^2}{dt}\right)^2 - \frac{1}{c^2}\left(\frac{dx^3}{dt}\right)^2}\, dt \end{aligned}$$

This is the total time measured along the path from $\mathbf{\Gamma}(t_1)$ to $\mathbf{\Gamma}(t_2)$.

Path of a physical particle In the case where the path $\mathbf{\Gamma}(t)$ represents the motion of a physical particle through space-time, we impose a physicality condition: that the rate of passage of proper time measured by the particle is unity. Namely, we impose the condition:

$$\left|\frac{d\mathbf{\Gamma}}{dt}(t)\right| = 1.$$

3 Formulas

Given a frame (x^0, x^1, x^2, x^3) for $\mathbb{R}^{1,3}$ and a line segment of length $\triangle s$, we have the following relationship between proper time and space-time measurements:

$$(\triangle s)^2 = \left(\triangle x^0\right)^2 - \frac{1}{c^2}\left(\triangle x^1\right)^2 - \frac{1}{c^2}\left(\triangle x^2\right)^2 - \frac{1}{c^2}\left(\triangle x^3\right)^2 .$$

Lecture 3 - Groups and Symmetry

Jan 30, 2009

1 Groups

It will be necessary for us to understand the basics of group theory, although we will not delve too deeply into this vast subject.

Def A group is a set G with an operation $*$, that satisfies the following three conditions:

- Existence of an identity element: there exists an element, commonly denoted 1 or e, so that whenever $a \in G$, we have $a * e = e * a = a$.
- Existence of inverses: Given any element $a \in G$, there is an element $b \in G$ (often denoted a^{-1}) so that $a * a^{-1} = a^{-1} * a = e$.
- Associativity: if $a, b, c \in G$, then $(a * b) * c = a * (b * c)$.

1.1 The integers

The simplest example of a group is the integers, with the operation being addition. This is denoted $(\mathbb{Z}, +)$. In this case, the identity element (the "1") is the number 0. Inverses clearly exist: given a number $n \in \mathbb{Z}$, it inverse element is the number $-n$. Addition is clearly associative. It is also commutative, so $(\mathbb{Z}, +)$ is what is known as an abelian group.

1.2 The cyclic groups

In this case, the underlying set is the set of the first n whole numbers, $\{0, 1, \ldots, n-1\}$. The operation, denoted $+_n$ or just $+$, is given by adding two numbers together in the ordinary way, then subtracting as many multiples of n as required to place the result back in the original set. If the group is $(\mathbb{Z}_7, +_7)$, we have, for example,

$$\begin{aligned}
1 +_7 1 &= 2 \\
1 +_7 6 &= 0 \\
2 +_7 6 &= 1 \\
5 +_7 6 &= 4.
\end{aligned}$$

As another example, the group $(\mathbb{Z}_{12}, +_{12})$ is the ordinary arithmetic of 12 hour clocks:

$$\begin{aligned}
&11\,(\text{o}'\text{clock}) + 2\,(hrs) = 1\,(\text{o}'\text{clock}) \\
&3 + 10 = 1 \\
&9 + 7 = 4 \\
&etc.
\end{aligned}$$

1.3 The Dihedral groups

Group are often used to encode geometric symmetries. The dihedral groups (denoted $(D_n, \circ)$) are the first examples: the set underlying the n^{th} dihedral group is defined to be the set of symmetries of the regular n-gon. The operation is on group elements is simply composition. We use exponential notation as a shorthand: $\alpha^2 = \alpha \circ \alpha$, for example.

The group D_n is generated by just two elements: let $\alpha \in D_n$ be rotation by $2\pi/n$, and let $\beta \in D_n$ be reflection about a pre-chosen axis of symmetry. The identity element, denoted 1 or e, is simply rotation by 0 radians. Clearly $\alpha^n = e$, so that α^{-1} is just the element α^{n-1}. Also, $\beta^2 = e$, so that $\beta^{-1} = \beta$ (β is its own inverse).

Reflections about other axes of symmetry can be constructed from these two elements. For example, reflection about the axis tilted at $2\pi/n$ from the original axis is given by $\alpha\beta\alpha^{-1}$. Reflection about the axis tilted at $2k\pi/n$ (k an integer) from the original axis is given by $\alpha^k\beta\alpha^{-k}$.

2 Symmetries of Euclidean space

Given a space with some notion of distance (for example Euclidean space or Minkowski space), an isometry is defined to be a transformation of that space which preserves distances between points. Given a space V, its isometries form a group, often denoted $Iso(V)$. The elements of the group are the isometries themselves, and the product is simply composition of isometries.

If V is n-dimensional Euclidean space, the group of isometries is called the Euclidean group. It is generated by translations, rotations, and reflections[1], and is denoted $E(n)$.

[1]Actually any rotation or translation can be decomposed into a pair of reflections, so the Euclidean group

Within the full Euclidean group is a distinguished subgroup, called the orthogonal group, denoted $\mathcal{O}(n)$. This is defined to be the group of isometries that fix a predetermined distinguished point (an "origin"). If o is the origin, the group $\mathcal{O}(n)$ is generated by the rotations that fix o and the reflection that fix o[2]

is actually generated by the reflections alone.

[2]Again, any rotation that fixes o can be decomposed into two reflections that fix o, so $O(n)$ is in fact generated by the reflections that fix o.

Lecture 4 - Lorentz transformations

Feb 2, 2009

1 Vector space transformations

If V is some vector space, a map $A : \mathbb{R}^n \to \mathbb{R}^n$ is called a linear operator, or a linear map, if it preserves the vector space operations of multiplication by constants and the addition/subtraction of vectors:

$$\alpha \in \mathbb{R},\ v,\, w\, \in\, \mathbb{R}^n \qquad \Longrightarrow \qquad A(\alpha v\, \pm\, w) \quad = \quad \alpha A(v)\, \pm\, A(w).$$

These transformations form a group called the linear group.

If V is an n-dimensional vecvtor space, then after a basis is chosen for V any operator $A : V \to V$ can be represented as an $n \times n$ matrix. It is also possible to express any vector $v \in V$ as an ordered n-tuple, usually arranged as a column-matrix of length n. Given a basis

$$\{\mathbf{e}_1,\, \mathbf{e}_2,\, \dots,\, \mathbf{e}_n\}$$

for V, any element $v \in V$ is given by

$$v \quad = \quad a^1\, \mathbf{e}_1\, +\, a^2\, \mathbf{e}_2\, +\, \dots\, +\, a^n\, \mathbf{e}_n$$

for some unique set of numbers $a^1, \dots, a^n$. It is typical to write v in column-vector notation: writing

$$v \quad = \quad \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{pmatrix}_{\{\mathbf{e}_i\}}$$

means precisely

$$v\ =\ a^1\mathbf{e}_1\, +\, \dots\, +\, a^n\mathbf{e}_n.$$

Once a basis has been chosen for a vector space V, it is possible to express any linear map $A : V \to V$ as a matrix. To do this, one must only record where A sends the basis vectors. We can define the numbers A^i_j implicitly as follows:

$$\begin{aligned} A(\mathbf{e}_j) &= A^1_j\,\mathbf{e}_1 + A^2_j\,\mathbf{e}_2 + \ldots + A^n_j\,\mathbf{e}_n \\ &= \sum_{i=1}^n A^i_j\,\mathbf{e}_i. \end{aligned}$$

It is then possible to regard A as a matrix

$$A \;=\; \left(A^i_j\right) \;=\; \begin{pmatrix} A^1_1 & A^1_2 & \ldots & A^1_n \\ A^2_1 & A^2_2 & \ldots & A^2_n \\ \vdots & & \ddots & \vdots \\ A^n_1 & A^n_2 & \ldots & A^n_n \end{pmatrix}_{\{\mathbf{e}_i\}}$$

1.1 Example - Rotation of the plane by θ

Consider Euclidean 2-space, with a distinguished point o. Let $A = A_\theta$ be the transformation that rotates the plane by θ about o.

<u>Exercise</u> Choose a basis, and express A as a matrix.
<u>Solution</u> Draw some coordinate axes, so we can consider Euclidean 2-space to be the coordinate plane. Let

$$\{\mathbf{e}_1,\, \mathbf{e}_2\}$$

be the usual unit vectors in the x and y directions, respectively. Using basic geometry, we know a rotation of θ has the following effects:

$$\begin{aligned} A_\theta(\mathbf{e}_1) &= \cos(\theta)\,\mathbf{e}_1 + \sin(\theta)\,\mathbf{e}_2 \\ A_\theta(\mathbf{e}_2) &= -\sin(\theta)\,\mathbf{e}_1 + \cos(\theta)\,\mathbf{e}_2 \end{aligned}$$

Therefore

$$A \;=\; \begin{pmatrix} A^1_1 & A^1_2 \\ A^2_1 & A^2_2 \end{pmatrix}_{\{\mathbf{e}_i\}} \;=\; \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}_{\{\mathbf{e}_i\}}$$

<u>Exercise</u>: Prove that indeed $A_\theta A_\alpha = A_{\theta+\alpha}$. You will have to use sum/difference formulas from trigonometry.

<u>Exercise</u> Choose another basis, and express A_θ as a matrix.
<u>Solution</u> Let

$$\begin{aligned} \mathbf{f}_1 &= \mathbf{e}_1 + \mathbf{e}_2 \\ \mathbf{f}_2 &= -\mathbf{e}_1 \end{aligned}$$

be a new basis for $\mathbb{R}^2$. Using the previous results, we get

$$\begin{aligned}
A(\mathbf{f}_1) &= A(\mathbf{e}_1 + \mathbf{e}_2) \;=\; (\cos(\theta) - \sin(\theta))\mathbf{e_1} \;+\; (\cos(\theta) + \sin(\theta))\mathbf{e}_2 \\
&= (\cos(\theta) - \sin(\theta))(-\mathbf{f_2}) \;+\; (\cos(\theta) + \sin(\theta))(\mathbf{f}_1 + \mathbf{f}_2) \\
&= (\cos\theta + \sin\theta)\,\mathbf{f_1} \;+\; 2\sin(\theta)\mathbf{f}_2 \\
A(\mathbf{f}_2) &= -A(\mathbf{e}_1) \;=\; -\cos(\theta)\mathbf{e}_1 \;-\; \sin(\theta)\mathbf{e}_2 \\
&= \cos(\theta)\mathbf{f}_2 \;-\; \sin(\theta)(\mathbf{f}_1 \;+\; \mathbf{f}_2) \\
&= -\sin(\theta)\mathbf{f}_1 \;+\; (\cos\theta - \sin\theta)\,\mathbf{f}_2
\end{aligned}$$

Thus we have

$$A_\theta \;=\; \begin{pmatrix} \cos\theta \,+\, \sin\theta & -\sin\theta \\ 2\sin\theta & \cos\theta \,-\, \sin\theta \end{pmatrix}$$

<u>Exercise</u>: Check that indeed $A_\theta A_\alpha = A_{\theta+\alpha}$.

2 Orthogonal transformations

Let's consider $\mathbb{R}^n$ with the usual orthonormal basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. Recall that the vector space $\mathbb{R}^n$ has an inner product. It is defined as follows: if $v, w \in \mathbb{R}^n$ are expressed as column vectors using the basis $\{\mathbf{e}_i\}$, we define $\langle v, w\rangle = v^T w$. A map $A : \mathbb{R}^n \to \mathbb{R}^n$ is called an orthogonal transformation if it preserves distances on $\mathbb{R}^n$. Due to the law of cosines:

$$\begin{aligned}
|\overline{pq}|^2 &= |\overline{op} - \overline{oq}|^2 \\
&= \langle \overline{op}, \overline{op}\rangle \;+\; \langle \overline{oq}, \overline{oq}\rangle \;-\; 2\,\langle \overline{op}, \overline{oq}\rangle
\end{aligned}$$

where p, q are arbitrary point in $\mathbb{R}^n$ and o is the origin, a transformation that fixes o will preserve distances if and only if it preserves the inner product. That is to say, A is in $\mathcal{O}(n)$ if and only if

$$\begin{aligned}
\langle Av,\, Aw\rangle \;&=\; \langle v,\, w\rangle \\
v^T\, A^T\, A\, w \;&=\; v^T\, w.
\end{aligned}$$

Since this must hold for any v, w, we have the following criterion:

$$A \in \mathcal{O}(n) \quad \text{if and only if} \quad A^T\, A \;=\; \mathrm{Id}\,.$$

3 The Lorentz group

Since Minkowski space $\mathbb{R}^{1,3}$ has a notion of distance, we can consider its group of isometries. Let

$$\{\mathbf{e}_0,\, \mathbf{e}_1,\, \mathbf{e}_2,\, \mathbf{e}_3\}$$

be an orthonormal basis for $\mathbb{R}^{1,3}$, meaning that they form an inertial reference frame. In particular,

$$|\mathbf{e}_0|^2 \;=\; 1 \quad |\mathbf{e}_1|^2 \;=\; -\frac{1}{c^2} \quad |\mathbf{e}_2|^2 \;=\; -\frac{1}{c^2} \quad |\mathbf{e}_3|^2 \;=\; -\frac{1}{c^2}.$$

If $v, w \in \mathbb{R}^{1,3}$ are expressed as column vectors using this basis, we can define the Minkowski inner product

$$\langle v,\, w\rangle_{1,3} \;=\; v^T\, I_{1,3}\, w$$

where

$$I_{1,3} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{c^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{c^2} & 0 \\ 0 & 0 & 0 & -\frac{1}{c^2} \end{pmatrix}.$$

Thus a transformation A is in $\mathcal{O}(1,3)$ if and only if

$$\begin{aligned} \langle Av,\, Aw\rangle_{1,3} \;&=\; \langle v, w\rangle \\ v^T\, A^T\, I_{1,3}\, A\, w \;&=\; v^T\, I_{1,3}\, w \end{aligned}$$

for any $v, w \in \mathbb{R}^{1,3}$. This means that

$$A \in \mathcal{O}(1,3) \quad \text{if and only if} \quad A^T\, I_{1,3}\, A \;=\; I_{1,3}.$$

The group $\mathcal{O}(1,3)$ is called the Lorentz group.

Similarly, this construction works for any of the spaces $\mathbb{R}^{k,n}$, using

$$I_{k,n} \;=\; \begin{pmatrix} 1 & & & & & & & \\ & 1 & & & & & 0 & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & -\frac{1}{c^2} & & & \\ & & & & & -\frac{1}{c^2} & & \\ & 0 & & & & & \ddots & \\ & & & & & & & -\frac{1}{c^2} \end{pmatrix}$$

where there are k many 1's and n many $-\frac{1}{c^2}$'s along the diagonal. We say that a transformation A is in $\mathcal{O}(n,k)$ if

$$A^T\, I_{k,n}\, A \;=\; I_{k,n}.$$

4 Generators of the Lorentz group

For convenience (and practicality) we shall only work with $\mathbb{R}^{1,3}$ (not $\mathbb{R}^{k,n}$), with orthonormal basis $\{\mathbf{e}_0, \ldots \mathbf{e}_3\}$.

There are three basic types of Lorentz transforms. First there are the boosts, which combine time with a space coordinate:

$$
\begin{aligned}
K_1(v) &= \begin{pmatrix} \gamma_v & -\frac{v}{c^2}\gamma_v & 0 & 0 \\ -v\gamma_v & \gamma_v & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
K_2(v) &= \begin{pmatrix} \gamma_v & 0 & -\frac{v}{c^2}\gamma_v & 0 \\ 0 & 1 & 0 & 0 \\ -v\gamma_v & 0 & \gamma_v & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
K_3(v) &= \begin{pmatrix} \gamma_v & 0 & 0 & -\frac{v}{c^2}\gamma_v \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v\gamma_v & 0 & 0 & \gamma_v \end{pmatrix}
\end{aligned}
$$

Second there are the spatial rotations, which do not involve the time coordinate:

$$
\begin{aligned}
P_{23}(\theta) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix} \\
P_{13}(\theta) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & -\sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta & 0 & \cos\theta \end{pmatrix} \\
P_{12}(\theta) &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned}
$$

These are just the orthogonal transformations of the the 3 spatial coordinates. Finally there are the space-reflection and time-reflection (ie, time reversing) transformations:

$$
\begin{aligned}
N_0 &= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
N_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
N_2 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
N_3 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
\end{aligned}
$$

5 The orthochronos group

Physically, one rotates one's space-time coordinates by changing one's spatial orientation, and performs a "boost" by changing speed. Thus is makes sense to restrict the Lorentz group, and consider just the boosts, rotations, and space-inversion matrices without the time-inversion matrix. This is called the orthochronous group, or $\mathcal{O}^+(1,n)$. Likewise, it is not physically possible to exchange left with right, forward with backwards, or up with down. The special orthochronous group, $\mathcal{SO}^+(1,n)$ is the group generated by just the boosts and the rotations, without the time- or space-inversion matrices.

6 The Poincare group

We have treated Minkowski space like a vector space with some fixed origin o, and considered points in Minkowski space to be "position vectors." But it is important to recognize that Minkowski space is not a vector space, and choosing an origin and a basis is an arbitrary choice.

The full set of isometries of Minkowski space consists of the the Lorenz group $\mathcal{O}(1,3)$ along with the translations. This group is called the Poincare group, denoted $P(1,3)$.

Lecture 5 - Geometry of Minkowski space

Feb 6, 2009

1 Geometry of $\mathbb{R}^{1,1}$

We define $\mathbb{R}^{1,1}$ as the set of ordered pairs of real numbers, equipped with the Minkowski distance: $\mathrm{dist}((a,b),(\alpha,\beta))^2 = (a-b)^2 - \frac{1}{c^2}(b-\beta)^2$. The set of isometries of $\mathbb{R}^{1,1}$, denoted $O(1,1)$, is generated by the boosts and the time- and space-reflections:

$$\begin{aligned} K(v) &= \begin{pmatrix} \gamma_v & -\frac{v}{c^2}\gamma_v \\ -v\gamma_v & \gamma_v \end{pmatrix} \\ N_0 &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ N_1 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

The principle invariant subsets are the light cone and the pseudospheres. The light cone is defined to be the points $x = (x^0, x^1)^T$ of distance 0 from the origin, given by solutions of

$$|x|^2 = 0 \qquad\qquad (x^0)^2 - \frac{1}{c^2}(x^1)^2 = 0.$$

This has the appearance of an "X", and is invariant under elements of $O(1,1)$. Sometimes it is useful to talk about the "future light cone," the part of the light cone with $x^0 \geq 0$, or the "past light cone," the part of the light cone with $x^0 \leq 0$.

The pseudosphere of radius $r > 0$ is defined to be the set of points $x = (x^0, x^1)^T$ of distance r from the origin, given by solutions of

$$|x|^2 = r^2 \qquad\qquad (x^0)^2 - \frac{1}{c^2}(x^1)^2 = r^2.$$

This has the appearance of a hyperbola, with two disconnected components. One may talk of the future and past parts of the pseudosphere.

2 Geometry of $\mathbb{R}^{1,2}$

We define $\mathbb{R}^{1,2}$ as the set of ordered triples of real numbers, equipped with the Minkowski distance: $\mathrm{dist}((a,b,c),(\alpha,\beta,\gamma))^2 = (a-b)^2 - \frac{1}{c^2}(b-\beta)^2 - \frac{1}{c^2}(c-\gamma)^2$. The set of isometries of $\mathbb{R}^{1,2}$, denoted $O(1,2)$, is generated by the boosts

$$
\begin{aligned}
K_1(v) &= \begin{pmatrix} \gamma_v & -\frac{v}{c^2}\gamma_v & 0 \\ -v\gamma_v & \gamma_v & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
K_2(v) &= \begin{pmatrix} \gamma_v & 0 & -\frac{v}{c^2}\gamma_v \\ & 1 & 0 \\ -v\gamma_v & 0 & \gamma_v \end{pmatrix}
\end{aligned}
$$

The (Euclidean) rotations in the $x^1 - x^2$ coordinates

$$
P_{12}(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}
$$

and the time- and space-reflections:

$$
\begin{aligned}
N_0 &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
N_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
N_2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}
\end{aligned}
$$

The principle invariant subsets are the light cone and the pseudospheres. The light cone is defined to be the points $x = (x^0, x^1, x^2)^T$ of distance 0 from the origin, given by solutions of

$$
|x|^2 = 0 \qquad (x^0)^2 - \frac{1}{c^2}(x^1)^2 - \frac{1}{c^2}(x^2)^2 = 0.
$$

This is a surface of revolution: it is the standard cone. It is invariant under elements of $O(1,2)$. Sometimes it is useful to talk about the "future light cone," the part of the light cone with $x^0 \geq 0$, or the "past light cone," the part of the light cone with $x^0 \leq 0$.

The pseudosphere of radius $r > 0$ is defined to be the set of points $x = (x^0, x^1, x^2)^T$ of distance r from the origin, given by solutions of

$$
|x|^2 = r^2 \qquad (x^0)^2 - \frac{1}{c^2}(x^1)^2 - \frac{1}{c^2}(x^2)^2 = r^2.
$$

This has the appearance of a two-sheeted hyperboloid of revolution. One may talk of the future and past parts of the pseudosphere.

3 Geometry of $\mathbb{R}^{1,3}$

We define $\mathbb{R}^{1,3}$ as the set of ordered quadruples of real numbers, equipped with the Minkowski distance: $\mathrm{dist}((a,b,c,d),(\alpha,\beta,\gamma,\delta))^2 \;=\; (a-b)^2 - \frac{1}{c^2}(b-\beta)^2 - \frac{1}{c^2}(c-\gamma)^2 - \frac{1}{c^2}(d-\delta)^2$. The set of isometries of $\mathbb{R}^{1,3}$, denoted $O(1,3)$, is generated by the boosts $K_1(v)$, $K_2(v)$, $K_3(v)$, by the Euclidean rotations $P_{12}(\theta)$, $P_{13}(\theta)$, $P_{23}(\theta)$, and by the time- and space-reflections N_0, N_1, N_2, N_3 (these matrices were given in lecture 4).

The principle invariant subsets are the light cone and the pseudospheres. The light cone is defined to be the points $x = (x^0, x^1, x^2, x^3)^T$ of distance 0 from the origin, given by solutions of

$$|x|^2 \;=\; 0 \qquad\qquad (x^0)^2 \,-\, \frac{1}{c^2}(x^1)^2 \,-\, \frac{1}{c^2}(x^2)^2 \,-\, \frac{1}{c^2}(x^2)^3 \;=\; 0.$$

This is a 3-dimensional surface in 4-space, and has spherical symmetry, which is to say it symmetric under any rotation of the $x^1 - x^2 - x^3$ coordinates. It is also invariant under the boosts and reflections, so is invariant under $O(1,3)$. Sometimes it is useful to talk about the "future light cone," the part of the light cone with $x^0 \geq 0$, or the "past light cone," the part of the light cone with $x^0 \leq 0$.

The pseudosphere of radius $r > 0$ is defined to be the set of points $x = (x^0, x^1, x^2, x^3)^T$ of distance r from the origin, given by solutions of

$$|x|^2 \;=\; r^2 \qquad\qquad (x^0)^2 \,-\, \frac{1}{c^2}(x^1)^2 \,-\, \frac{1}{c^2}(x^2)^2 \,-\, \frac{1}{c^2}(x^3)^2 \;=\; r^2.$$

This is again a two-sheeted hyperboloid, but of spherical symmetry (not just circular symmetry as in the $\mathbb{R}^{1,2}$ case). One may talk of the future and past parts of the pseudosphere.

Lecture 6 - Vector spaces, linear maps, and dual spaces

February 9, 2009

1 Vector spaces

A vector space V with scalars $\mathbb{F}$ is defined to be a commutative ring $(V, +)$ so that the scalars form a division ring with identity, and operate on the V in a way satisfying (here $\alpha, \beta \in \mathbb{F}$ and $\mathbf{v}, \mathbf{w} \in V$):

- $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$
- $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$
- $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$
- $1\mathbf{v} = \mathbf{v}$ where $1 \in \mathbb{F}$ is the identity element

If $o \in V$ is the identity of the group $(V, +)$ (ie, the 'origin' of the vector space V), it is an exercise to show that these axioms imply $0\mathbf{v} = o$ and $\alpha o = o$.

In our class, we will exclusively be concerned with real vector spaces, meaning $\mathbb{F}$ is the field $\mathbb{R}$.

2 Linear maps

If V and W are vector spaces with the same field of scalars, a *linear map* A is defined to be a map $A : V \to W$ satisfying

$$A(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = \alpha A(\mathbf{v}_1) + \beta A(\mathbf{v}_2)$$

where α, β are scalars and $\mathbf{v}_1, \mathbf{v}_2 \in V$. After bases $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ for V and $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ for W are chosen, it is possible to express A as a matrix. Specifically, we define the numbers A_i^j implicitly by

$$A(v_i) = A_i^1\mathbf{w}_1 + A_i^2\mathbf{w}_2 + \dots + A_i^m\mathbf{w}_m.$$

Then if $v = \alpha^1 \mathbf{v}_1 + \ldots + \alpha^n \mathbf{v}_n$, we have

$$
\begin{aligned}
A(v) &= A(\alpha^1 \mathbf{v}_1 + \alpha^2 \mathbf{v}_2 + \ldots + \alpha^n \mathbf{v}_n) \\
&= \alpha^1 A(\mathbf{v}_1) + \alpha^2 A(\mathbf{v}_2) + \ldots + \alpha^n A(\mathbf{v}_n) \\
&= \alpha^1 \left(A_1^1 \mathbf{w}_1 + A_1^2 \mathbf{w}_2 + \ldots + A_1^m \mathbf{w}_m \right) \\
&\quad + \alpha^2 \left(A_2^1 \mathbf{w}_1 + A_2^2 \mathbf{w}_2 + \ldots + A_2^m \mathbf{w}_m \right) \\
&\quad + \ldots \\
&\quad + \alpha^n \left(A_n^1 \mathbf{w}_1 + A_n^2 \mathbf{w}_2 + \ldots + A_n^m \mathbf{w}_m \right) \\
&= \left(\alpha^1 A_1^1 + \alpha^2 A_2^1 + \ldots + \alpha^n A_n^1 \right) \mathbf{w}_1 \\
&\quad + \left(\alpha^1 A_1^2 + \alpha^2 A_2^2 + \ldots + \alpha^n A_n^2 \right) \mathbf{w}_2 \\
&\quad + \ldots \\
&\quad + \left(\alpha^1 A_1^m + \alpha^2 A_2^m + \ldots + \alpha^n A_n^m \right) \mathbf{w}_m .
\end{aligned}
$$

Thus if we write $\mathbf{v}$ and $A(\mathbf{v})$ in vector notation, then by our calculations we have:

$$
\mathbf{v} = \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}_{\{\mathbf{v}_i\}} \qquad A(\mathbf{v}) = \begin{pmatrix} \alpha^1 A_1^1 + \alpha^2 A_2^1 + \ldots + \alpha^n A_n^1 \\ \alpha^1 A_1^1 + \alpha^2 A_2^1 + \ldots + \alpha^n A_n^1 \\ \vdots \\ \alpha^1 A_1^m + \alpha^2 A_2^m + \ldots + \alpha^n A_n^m \end{pmatrix}_{\{w_i\}}
$$

which means that A is an $n \times m$ matrix:

$$
A = \begin{pmatrix} A_1^1 & A_1^2 & \ldots & A_1^m \\ A_2^1 & A_2^2 & \ldots & A_2^m \\ \vdots & & \ddots & \vdots \\ A_n^1 & A_n^2 & \ldots & A_n^m \end{pmatrix}_{\{w_i\} \leftarrow \{v_i\}}
$$

and the action of A is given by matrix multiplication on the left.

<u>Example</u>

Let V be the vector space of quadratic polynomials with basis $\mathbf{e}_1 = 1$, $\mathbf{e}_2 = x$, $\mathbf{e}_3 = x^2$, and let W be the vector space of cubic polynomials with basis $\mathbf{f}_1 = 1$, $\mathbf{f}_2 = x$, $\mathbf{f}_3 = x^2$, and $\mathbf{f}_4 = x^3$. Let $A : V \to W$ be the map $A(P) = (1 + 2x)P$.

To express A as a matrix, we see where it send the basis vectors:

$$
\begin{aligned}
A(\mathbf{e}_1) &= (1 + 2x)1 = \mathbf{f}_1 + 2\mathbf{f}_2 \\
A(\mathbf{e}_2) &= (1 + 2x)x = \mathbf{f}_2 + 2\mathbf{f}_3 \\
A(\mathbf{e}_3) &= (1 + 2x)x^2 = \mathbf{f}_3 + 2\mathbf{f}_4
\end{aligned}
$$

Thus

$$
A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}_{\{f_i\} \leftarrow \{e_i\}} .
$$

3 Dual spaces

Assume V is a vector space with scalar field $\mathbb{F}$ (in our class, $\mathbb{F}$ will almost always be just the reals, $\mathbb{R}$). A *linear functional* on a vector space V is a linear map $f : V \to \mathbb{F}$. It is simple to prove that $A(o) = 0$ whenever A is a linear operator:

$$A(o) \;=\; A(0 \cdot \mathbf{v}) \;=\; 0 \cdot A(\mathbf{v}) \;=\; 0.$$

The space of linear operators on a vector space V is called its *dual vector space*, denoted V^*. If V is finite dimensional and a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for V has been chosen, there is a procedure for choosing a basis $\{\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*\}$ for V^*, called the *basis dual to* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. The procedure is very simple: define $\mathbf{v}_i^* : V \to \mathbb{R}$ by setting $\mathbf{v}_i^*(\mathbf{v}_j) = \delta_{ij}$ and extending linearly. To be more explicit, if $v = \alpha^1\mathbf{v}_1 + \cdots + \alpha^n\mathbf{v}_n$, then

$$\begin{aligned}
\mathbf{v}_i^*(v) &= \mathbf{v}_i^*\left(\alpha^1\mathbf{v}_1 + \ldots + \alpha^n\mathbf{v}_n\right) \\
&= \alpha^1\mathbf{v}_i^*(\mathbf{v}_1) + \ldots + \alpha^i\mathbf{v}_i^*(\mathbf{v}_i) + \ldots + \alpha^n\mathbf{v}_i^*(\mathbf{v}_n) \\
&= \alpha^1 \cdot 0 + \ldots + \alpha^i + \ldots + \alpha^n \cdot 0 \\
&= \alpha^i.
\end{aligned}$$

It is easy to verify that $\mathbf{v}_i^*$ is linear.

Theorem 3.1 *If* $\dim(V) \;=\; n \;<\; \infty$*, then also* $\dim(V^*) = n$*.*

<u>*Pf*</u> We only have to prove that what we called the "dual basis" (which consists of n many elements) is indeed a basis. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V, and $\{\mathbf{v}_1^*, \ldots, \mathbf{v}_n^*\}$ its "dual basis." We must prove that the $\mathbf{v}_i^*$ are linearly independent, and that they indeed span V^*. First, if $0 = \beta^1\mathbf{v}_1^* + \cdots + \beta^n\mathbf{v}_n^*$ for some constants β^i, then by plugging in $\mathbf{v}_j$ to both sides we get

$$0 \;=\; \beta^j.$$

Since j was arbitrary, this proves that all the coefficients are 0. Thus the $\mathbf{v}_i^*$ are independent.

To prove that the $\mathbf{v}_i^*$ span V^*, let $A \in V^*$. We can define the numbers A_i by

$$A(\mathbf{v}_i) \;=\; A_i.$$

It follows that $A = A_1\mathbf{v}_1^* + A_2\mathbf{v}_2^* + \ldots + A_n\mathbf{v}_n^*$: for let $v = \alpha^1\mathbf{v}_1 + \cdots + \alpha^n\mathbf{v}_n$ be a generic element in V; then

$$\begin{aligned}
&A(v) \;=\; A(\alpha^1\mathbf{v}_1 + \ldots + \alpha^n\mathbf{v}_n) \;=\; \alpha^1 A(\mathbf{v}_1) + \ldots + \alpha^n A(\mathbf{v}_n) \;=\; \alpha^1 A_1 + \ldots + \alpha^n A_n \\
&(A_1 e_1^* + \ldots + A_n e_n^*)\,(v) \;=\; A_1\mathbf{v}_1^*(v) + \ldots + A_n\mathbf{v}_n^*(v) \;=\; A_1\alpha^1 + \ldots + A_n\alpha^n.
\end{aligned}$$

$\square$

In the proof, note how we were able to write A as $A = A_1\mathbf{v}_1^* + \cdots + A_n\mathbf{v}_n^*$. This violates the usual motif of summing over upper-lower index pairs, indicating that the dual basis should probably be written with upper indices. From now on we will do this:

we will write $\mathbf{v}^i$, not $\mathbf{v}_i^*$.

Thus the basis dual to $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ will be written $\{\mathbf{v}^1, \ldots, \mathbf{v}^n\}$, with the same definition:

$$\begin{aligned}
&\mathbf{v}^i : V \to \mathbb{R} \\
&\mathbf{v}^i(\mathbf{v}_j) \;=\; \delta^i_j.
\end{aligned}$$

Note that this means dual vectors (elements of V^*) should be written in row form: if $\mathbf{v} = \alpha^1\mathbf{v}_1 + \cdots + \alpha^n\mathbf{v}_n$ and $A = A_1\mathbf{v}^1 + \cdots + A_n\mathbf{v}^n$, then we can write

$$v \;=\; \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}_{\{\mathbf{v}_i\}} \qquad\qquad A \;=\; (A_1\; A_2\; \ldots\; A_n)_{\{\mathbf{v}^i\}}\,.$$

As usual, we can express the action of A on v via matrix multiplication;

$$\begin{aligned}
A(v) &= A\left(\alpha^1\,\mathbf{v}_1 \,+\, \ldots \,+\, \alpha^n\,\mathbf{v}_n\right) \\
&= \alpha^1 A(\mathbf{v}_1) \,+\, \ldots \,+\, \alpha^n A(\mathbf{v}_n) \\
&= \alpha^1 A_1 \,+\, \alpha^2 A_2 \,+\, \ldots \,+\, \alpha^n A_n \;=\; \sum_{i=1}^{n} \alpha^i A_i \\
&= (A_1\, A_2\, \ldots\, A_n) \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}
\end{aligned}$$

4 The Einstein summation convention

In all cases so far considered, upper indices are summed over lower indices whenever a sum is required; two lower indices are never summed, likewise for two upper indices. For example, letting V be a vector space with basis $\{\mathbf{v}_i\}$, if $A = \left(A^i_j\right)$ is a linear operator and $\mathbf{v} = \alpha^1\mathbf{v}_1 + \alpha^2\mathbf{v}_2 + \cdots + \alpha^n\mathbf{v}_n$ a vector, we have

$$\begin{aligned}
A(v) &= \begin{pmatrix} A^1_1 & A^1_2 & \ldots & A^1_n \\ A^2_1 & A^2_2 & & A^2_n \\ \vdots & & \ddots & \vdots \\ A^n_1 & A^n_2 & \ldots & A^n_n \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}_{\{\mathbf{v}_i\}} \\
&= \begin{pmatrix} \alpha^1 A^1_1 + \alpha^2 A^1_2 + \cdots + \alpha^n A^1_n \\ \alpha^1 A^2_1 + \alpha^2 A^2_2 + \cdots + \alpha^n A^2_n \\ \vdots \\ \alpha^1 A^n_1 + \alpha^2 A^n_2 + \cdots + \alpha^n A^n_n \end{pmatrix}_{\{\mathbf{v}_i\}}
\end{aligned}$$

This is a lot of writing. But we can express the same information more compactly:

$$\mathbf{v} \;=\; \sum_{i=i}^{n} \alpha^i \mathbf{v}_i\,, \qquad A(\mathbf{v}_i) \;=\; \sum_{j=1}^{n} A_i^j \mathbf{v}_j\,,$$

$$A(\mathbf{v}) \;=\; A\left(\sum_{i=1}^{n} \alpha^i \mathbf{v}_i\right) \;=\; \sum_{i=1}^{n} \alpha^i A\left(\mathbf{v}_i\right) \;=\; \sum_{i=1}^{n} \alpha^i \left(\sum_{j=1}^{n} A_i^j \mathbf{v}_j\right) \;=\; \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha^i A_i^j \mathbf{v}_j.$$

If we just leave off the summation symbol, we can write this even more compactly:

$$\mathbf{v} \;=\; \alpha^i \mathbf{v}_i\,, \qquad A(\mathbf{v}) \;=\; A(\alpha^i \mathbf{v}_i) \;=\; \alpha^i A(\mathbf{v}_i) \;=\; \alpha^i A_i^j \mathbf{v}_j.$$

This is the **Einstein summation convention**: the summation symbol is left off, and any repeated upper and lower indices are summed over.

Lecture 7 - Vector spaces and their Duals

2/16/09

This lecture elaborates on some elements from lecture 6.

Throughout, we will use V to denote a vector space of dimension $n < \infty$, and we will use V^* to denote the dual of V. As always in this class, we assume that scalar field is $\mathbb{R}$. Recall that V^* is defined to to be the space of *linear functionals*, that is to say, $A \in V^*$ whenever A is a linear map $A : V \to \mathbb{R}$. It is important to note that V^* is not just a set, but is in fact a vector space.

1 Choosing bases for V and V^*

Given a vector space V, the choice of a basis is fundamentally an arbitrary procedure, though in some cases the choice is more-or-less natural.

In any case, in order to make calculations concrete, one must choose a basis by one means or another. Once a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ has been chosen, it is possible to express any vector $\mathbf{v} \in V$ as a linear combination of basis vectors:

$$\mathbf{v} \;=\; \alpha^1 \mathbf{v}_1 + \ldots + \alpha^n \mathbf{v}_n \qquad \textit{or} \qquad \mathbf{v} = \alpha^i \mathbf{v}_i \qquad \textit{or} \qquad \mathbf{v} \;=\; \left(\begin{array}{c} \alpha^1 \\ \vdots \\ \alpha^n \end{array} \right)_{\{\mathbf{v}_i\}}$$

(these three equations are precisely equivalent). Once a basis $\{\mathbf{v}_i\}$ for V has been chosen, it is possible to choose a basis for the vector space V^*. We define linear functionals $\mathbf{v}^i$ by requiring that

$$\mathbf{v}^i\left(\mathbf{v}_j\right) \;=\; \delta^i_j,$$

and then extending linearly. That is to say, $\mathbf{v}^i : V \to \mathbb{R}$ is the linear map

$$\begin{aligned}\mathbf{v}^i(\mathbf{v}) &= \mathbf{v}^i\left(\alpha^1\mathbf{v}_1 + \ldots + \alpha^i\mathbf{v}_i + \ldots + \alpha^n\mathbf{v}_n\right) \\ &= \alpha^1\mathbf{v}^i(\mathbf{v}_1) + \ldots + \alpha^i\mathbf{v}^i(\mathbf{v}_i) + \ldots + \alpha^n\mathbf{v}^i(\mathbf{v}_n) \\ &= \alpha^1 \cdot 0 + \ldots + \alpha^i \cdot 1 + \ldots + \alpha^n \cdot 0 \\ &= \alpha^i.\end{aligned}$$

Using Einstein notation, the same computation can be done with much less writing:

$$\begin{aligned}\mathbf{v}^i(\mathbf{v}) &= \mathbf{v}^i\left(\alpha^j\mathbf{v}_j\right) \\ &= \alpha^j\mathbf{v}^i(\mathbf{v}_j) \\ &= \alpha^j\delta^i_j \\ &= \alpha^i.\end{aligned}$$

2 The matrix δ^i_j

Just now we claimed that $\alpha^j\delta^i_j = \alpha^i$. Here we will prove this, and hopefully give some insight into the object δ^i_j. Of course δ^i_j can be expressed as a matrix:

$$\delta = \begin{pmatrix} \delta^1_1 & \delta^1_2 & \ldots & \delta^1_n \\ \delta^2_1 & \delta^2_2 & \ldots & \delta^2_n \\ \vdots & \vdots & \ddots & \vdots \\ \delta^n_1 & \delta^n_2 & \ldots & \delta^n_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} = I_n.$$

Given $\mathbf{v} \in V$, obviously $I_n(\mathbf{v}) = \delta(\mathbf{v}) = \mathbf{v}$. Letting $\mathbf{v} = \alpha^i\mathbf{v}_i$ and expressing this fact in matrix form, we have

$$\begin{aligned}\mathbf{v} = \alpha^i\mathbf{v}_i = \begin{pmatrix} \alpha^1 \\ \vdots \\ \alpha^n \end{pmatrix} &= \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix} = \begin{pmatrix} \delta^1_1 & \delta^1_2 & \ldots & \delta^1_n \\ \delta^2_1 & \delta^2_2 & \ldots & \delta^2_n \\ \vdots & \vdots & \ddots & \vdots \\ \delta^n_1 & \delta^n_2 & \ldots & \delta^n_n \end{pmatrix} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix} \\ &= \begin{pmatrix} \delta^1_1\alpha^1 + \delta^1_2\alpha^2 + \ldots + \delta^1_n\alpha^n \\ \delta^2_1\alpha^1 + \delta^2_2\alpha^2 + \ldots + \delta^2_n\alpha^n \\ \vdots \\ \delta^n_1\alpha^1 + \delta^n_2\alpha^2 + \ldots + \delta^n_n\alpha^n \end{pmatrix} = \begin{pmatrix} \delta^1_j\alpha^j \\ \delta^2_j\alpha^j \\ \vdots \\ \delta^n_j\alpha^j \end{pmatrix} = \delta^i_j\,\alpha^j\,\mathbf{v}_i.\end{aligned}$$

Therefore we have proven that $\delta^i_j\alpha^j = \alpha^i$.

3 Examples

If $\{\mathbf{v}_i\}$ is a basis of V, the basis $\{\mathbf{v}^i\}$ of V^* is called the basis dual to $\{\mathbf{v}_i\}$. The fact that the choice of the dual basis depends on the original choice of basis is illustrated by the two following examples. In the examples, we will use $V = \mathbb{R}^3$ with standard basis $\hat{\imath}$, $\hat{\jmath}$, $\hat{k}$.

Example 1 Let $V = \mathbb{R}^3$, and let $\mathbf{v}_1 = \hat{\imath}$, $\mathbf{v}_2 = \hat{\jmath}$, $\mathbf{v}_3 = \hat{k}$. be the standard basis. Determine the action of the dual basis $\mathbf{v}^1$, $\mathbf{v}^2$, $\mathbf{v}^3$ on a generic vector $\mathbf{v} = a\hat{\imath} + b\hat{\jmath} + c\hat{\jmath} \in \mathbb{R}^3$.

Solution. This is very simple: we can express $\mathbf{v} = a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3$, so

$$\begin{aligned}
\mathbf{v}^1(\mathbf{v}) &= \mathbf{v}^1(a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3) = a \\
\mathbf{v}^2(\mathbf{v}) &= \mathbf{v}^2(a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3) = b \\
\mathbf{v}^3(\mathbf{v}) &= \mathbf{v}^3(a\mathbf{v}_1 + b\mathbf{v}_2 + c\mathbf{v}_3) = c.
\end{aligned}$$

Example 2 Let $V = \mathbb{R}^3$ and let $\mathbf{w}_1 = \hat{\imath} + \hat{\jmath}$, $\mathbf{w}_1 = \hat{\imath} - \hat{\jmath}$, $\mathbf{w}_1 = \hat{\imath} + \hat{\jmath} + \hat{k}$ be a basis. Determine the action of the dual basis $\mathbf{w}^1$, $\mathbf{w}^2$, $\mathbf{w}^3$ on a generic vector $\mathbf{v} = a\hat{\imath} + b\hat{\jmath} + c\hat{\jmath} \in \mathbb{R}^3$.

Solution. You should check that we can express

$$\mathbf{v} = a\hat{\imath} + b\hat{\jmath} + c\hat{k} = \left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}_1 + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}_2 + c\mathbf{w}_3.$$

Therefore

$$\begin{aligned}
\mathbf{w}^1(\mathbf{v}) &= \mathbf{w}^1\left(\left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}_1 + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}_2 + c\mathbf{w}_3\right) \\
&= \left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}^1(\mathbf{w}_1) + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}^1(\mathbf{w}_2) + c\mathbf{w}^1(\mathbf{w}_3) \\
&= \frac{a}{2} + \frac{b}{2} - c \\
\mathbf{w}^2(\mathbf{v}) &= \mathbf{w}^2\left(\left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}_1 + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}_2 + c\mathbf{w}_3\right) \\
&= \left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}^2(\mathbf{w}_1) + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}^2(\mathbf{w}_2) + c\mathbf{w}^2(\mathbf{w}_3) \\
&= \frac{a}{2} - \frac{b}{2} \\
\mathbf{w}^3(\mathbf{v}) &= \mathbf{w}^3\left(\left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}_1 + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}_2 + c\mathbf{w}_3\right) \\
&= \left(\frac{a}{2} + \frac{b}{2} - c\right)\mathbf{w}^3(\mathbf{w}_1) + \left(\frac{a}{2} - \frac{b}{2}\right)\mathbf{w}^3(\mathbf{w}_2) + c\mathbf{w}^3(\mathbf{w}_3) \\
&= c.
\end{aligned}$$

Lecture 8 - Vector spaces and their Duals, II

2/18/09

This lecture completes our formal discussion of dual spaces.

1 The double dual, V^{**}

The space V^* is defined to be the space of linear operators on V. Of course, V^* is a vector space itself, so also has a dual, denoted V^{**}, called the "double-dual" of V.

But as a matter of fact, one can consider elements of V to acto on elements of V^*: there is an map

$$\mathcal{N} : V \hookrightarrow V^{**}$$

given by

$$\begin{aligned} &\mathcal{N}(\mathbf{v}) \in V^{**} \\ &\mathcal{N}(\mathbf{v})(f) \triangleq f(\mathbf{v}) \quad \textit{for any } f \in V^*. \end{aligned}$$

Theorem 1.1 *If V is a finite dimensional vector space, then the map $\mathcal{N} : V \to V^{**}$ is a a vector space isomorphism.*

Pf Homework problem 3.6. $\square$

Often we drop the "$\mathcal{N}$" from the notation, and just consider elements $\mathbf{v} \in V$ to act on elements $f \in V^*$ directly:

$$\begin{aligned} &\mathbf{v} \in V^{**} \quad \textit{acts on } V^* \textit{ by} \\ &\mathbf{v}(f) \triangleq f(\mathbf{v}) \quad \textit{for any } f \in V^*. \end{aligned}$$

2 Change-of-basis matrices

Consider two bases $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ for the vector space V. Any vector $\mathbf{v} \in V$ can be expressed as a column vector in either system, though the column vector will be different. If one knows the vector for $\mathbf{v}$ in the $\mathbf{e}_i$ system, how can $\mathbf{v}$ be expressed in the $\mathbf{f}_i$ system?

One has to know the relation between the two systems. Define the numbers A^i_j implicitly by

$$\mathbf{e}_j = A^i_j \mathbf{f}_i.$$

Then, for example,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{\{\mathbf{e}_i\}} = \begin{pmatrix} A^1_1 \\ A^2_1 \\ \vdots \\ A^n_1 \end{pmatrix}_{\{\mathbf{f}_i\}}$$

$$\mathbf{e}_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{\{\mathbf{e}_i\}} = \begin{pmatrix} A^1_j \\ \vdots \\ A^{j-1}_j \\ A^j_j \\ A^{j+1}_j \\ \vdots \\ A^n_j \end{pmatrix}_{\{\mathbf{f}_i\}}$$

Therefore, a vector $\mathbf{v} = \alpha^i \mathbf{e}_i$ can be expressed

$$\mathbf{v} = \alpha^j \mathbf{e}_j = \alpha^j A^i_j \mathbf{f}_i = \begin{pmatrix} A^1_j \alpha^j \\ A^2_j \alpha^j \\ \vdots \\ A^n_j \alpha^j \end{pmatrix}_{\{\mathbf{f}_i\}} .$$

The final vector is just the matrix multiplication

$$\begin{pmatrix} A^1_j \alpha^j \\ A^2_j \alpha^j \\ \vdots \\ A^n_j \alpha^j \end{pmatrix}_{\{\mathbf{f}_i\}} = \begin{pmatrix} A^1_1 & A^1_2 & \ldots & A^1_n \\ A^2_1 & A^2_2 & \ldots & A^2_n \\ \vdots & \vdots & \ddots & \vdots \\ A^n_1 & A^n_2 & \ldots & A^n_n \end{pmatrix}_{\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}} \begin{pmatrix} \alpha^1 \\ \alpha^2 \\ \vdots \\ \alpha^n \end{pmatrix}_{\{\mathbf{e}_i\}}$$

Notice the subscript $\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}$ on the matrix. It is used to indicate what bases A transitions between. The transformation from the $\mathbf{f}_i$ to the $\mathbf{e}_i$ basis is given by the inverse matrix:

$$A_{\{\mathbf{e}_i\} \leftarrow \{\mathbf{f}_i\}} = \left(A_{\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}} \right)^{-1} .$$

3 Active *vs.* Passive transformations

There are always two ways to think about an operator $A : V \to V$. A so-called *active transformation* uses a fixed coordinate system, and performs a transformation of the underlying space. A so-called *passive transformation* just changes the basis vectors and leaves the underlying space fixed. However these are conceptual differences only: any given operator can be interpreted in either way.

Let's illustrate this with an example. Let $V = \mathbb{R}^2$ with standard basis $\mathbf{e}_1 = \hat{\imath}$, $\mathbf{e}_2 = \hat{\jmath}$. Let A be given by

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Thought of as an active transformation, this is a rotation of space counterclockwise through an angle of θ.

On the other hand, consider another basis $\mathbf{f}_1 = \cos(\theta)\hat{\imath} - \sin(\theta)\hat{\jmath}$, $\mathbf{f}_2 = \sin(\theta)\hat{\imath} + \cos(\theta)\hat{\jmath}$. Then A is just the change-of-basis matrix from the $\mathbf{e}_i$ to the $\mathbf{f}_i$ bases.

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}_{\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}}.$$

Note that the new basis $\mathbf{f}_i$ is a rotation of the old basis $\mathbf{e}_i$ through a *clockwise* angle of θ.

Thus the matrix A can be considered to be either a transformation of the underlying space (an active transformation, in this case *counterclockwise* rotation by θ) or as a change of basis that leaves the vector space unchanged (a passive transformation, in this case a *clockwise* rotation of the basis vectors by θ).

4 Actions on the dual space

Let $A : V \to V$ be a linear operator. We have not defined any kind of action of A on the dual space V^*. But, as we shall see, there *should* be such an action.

To see this, consider A to be a passive transformation, changing from, say, the $\{\mathbf{e}_i\} \subset V$ to the $\{\mathbf{f}_i\} \subset V$ basis. Let $f \in V^*$ be a linear functional, and let $\mathbf{w} \in V$ be a vector. Let $\mathbf{w}_{\{\mathbf{e}_i\}}$, $f_{\{\mathbf{e}^i\}}$ be their expressions in the $\{\mathbf{e}_i\}$ (respectively $\{\mathbf{e}^i\}$) basis, and $\mathbf{w}_{\{\mathbf{f}_i\}}$, $f_{\{\mathbf{f}^i\}}$ be their expressions in the $\{\mathbf{f}_i\}$ (respectively $\{\mathbf{f}^i\}$) basis. We have

$$\mathbf{w}_{\{\mathbf{f}_i\}} = A_{\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}} \mathbf{w}_{\{\mathbf{e}_i\}}.$$

Now, since $f_{\{\mathbf{e}^i\}}$ and $f_{\{\mathbf{f}^i\}}$ are the same covector regardless of its expression in either basis, it must have the same action on $\mathbf{w}$, regardless of basis. Letting $A(f)$ indicate the action of

A on f, we therefore *must* have

$$\begin{aligned} A(f)(A(\mathbf{w})) &= f(\mathbf{w}). \\ (A(f))_{\{\mathbf{f}^i\}} \cdot (A\mathbf{w})_{\{\mathbf{f}_i\}} &= (A(f))_{\{\mathbf{f}^i\}} \cdot A_{\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}} \mathbf{w}_{\{\mathbf{e}_i\}}. \quad (\textit{Matrix multiplication}) \end{aligned}$$

Thus it *must* be the case that

$$(A(f))_{\{\mathbf{f}^i\}} = f_{\{\mathbf{e}^i\}} \cdot \left(A_{\{\mathbf{f}_i\} \leftarrow \{\mathbf{e}_i\}}\right)^{-1} \qquad (\textit{Matrix multiplication})$$

Expressing this abstractly (that is, without necessarily choosing a basis),

$$\begin{aligned} &\textit{Given } f : V \to \mathbb{R}, \textit{ we have } A(f) : V \to \mathbb{R}, \textit{ given by} \\ &A(f)(v) = f(A^{-1}\mathbf{v}) \quad \textit{for any } \mathbf{v} \in V. \end{aligned}$$

Lecture 9 - Tensor Products

Feb 18, 2009

1 Direct sum

If V and W are vector spaces and $v \in V$, $w \in W$ are vectors, the direct sum of v and w is defined to be their formal sum, denoted

$$v \oplus w.$$

This is subject to the linearity conditions

$$\begin{aligned} \alpha(v \oplus w) &= \alpha v \oplus \alpha w \\ (v \oplus w) + (v' \oplus w') &= (v + v') \oplus (w + w'), \end{aligned}$$

where $\alpha \in \mathbb{R}$ and $v, v' \in V$ and $w, w' \in W$. The space of all such sums is denoted $V \oplus W$, the direct sum of the vector spaces V and W. That is,

$$V \oplus W = \{v \oplus w \mid v \in V, w \in W\}.$$

Example Let $V = \mathbb{R}$ and $W = \mathbb{R}$. Describe the direct sum $V \oplus W$.

Solution We resort to choosing basis vectors. Let $v \in V$ and $w \in W$ be basis vectors. The space $V \oplus W$ is the set of "formal sums" of elements of V and W, meaning

$$x \in V \oplus W$$

if and only if

$$x = \alpha v \oplus \beta w.$$

Clearly, therefore, $V \oplus W$ is just a 2-dimensional vector space, so is isomorphic to $\mathbb{R}^2$.

Theorem 1.1 *$\mathbb{R}^n \oplus \mathbb{R}^m$ is isomorphic to $\mathbb{R}^{n+m}$.*

Pf Choose bases $\{\mathbf{e}_i\}_{i=1}^n$ for $\mathbb{R}^n$ and $\{\mathbf{f}_i\}_{i=1}^m$ for $\mathbb{R}^m$. Let $\{\mathbf{g}_i\}_{i=1}^{n+m}$ be a basis for $\mathbb{R}^{n+m}$. A generic element of $\mathbb{R}^n \oplus \mathbb{R}^m$ has the form

$$\mathbf{v} = \left(\alpha^1 \mathbf{e}_1 + \ldots + \alpha^n \mathbf{e}_n\right) \oplus \left(\beta^1 \mathbf{f}_1 + \ldots + \beta^m \mathbf{f}_m\right).$$

Let $\overline{\mathbf{v}} \in \mathbb{R}^{\mathbf{n}} \oplus \mathbb{R}^{\mathbf{m}}$ be another vector, given by

$$\overline{\mathbf{v}} \;=\; \left(\overline{\alpha}^1\mathbf{e}_1 \;+\; \ldots \;+\; \overline{\alpha^n}\mathbf{e}_n\right) \oplus \left(\overline{\beta}^1\mathbf{f}_1 \;+\; \ldots \;+\; \overline{\beta}^m\mathbf{f}_m\right).$$

Then the addition $\mathbf{v} + \overline{\mathbf{v}}$ is given by

$$\mathbf{v} \,+\, \overline{\mathbf{v}} \;=\; \left((\alpha^1 + \overline{\alpha}^1)\mathbf{e}_1 \;+\; \ldots \;+\; (\alpha^n + \overline{\alpha^n})\mathbf{e}_n\right) \oplus \left((\beta^1 + \overline{\beta}^1)\mathbf{f}_1 \;+\; \ldots \;+\; (\beta^m + \overline{\beta}^m)\mathbf{f}_m\right);$$

Let $\mathcal{A} : \mathbb{R}^n \oplus \mathbb{R}^m \to \mathbb{R}^{n+m}$ be defined by

$$\mathcal{A}(\mathbf{e}_i) = \mathbf{g}_i$$
$$\mathcal{A}(\mathbf{f}_i) = \mathbf{g}_{i+n}$$

and extending linearly. That is,

$$\begin{aligned}\mathcal{A}(\mathbf{v}) \;&=\; \mathcal{A}\left(\left(\alpha^1\mathbf{e}_1 \;+\; \ldots \;+\; \alpha^n\mathbf{e}_n\right) \oplus \left(\beta^1\mathbf{f}_1 \;+\; \ldots \;+\; \beta^m\mathbf{f}_m\right)\right)\\ &=\; \alpha^1\mathbf{g}_1 \;+\; \ldots \;+\; \alpha^n\mathbf{g}_n \;+\; \beta^1\mathbf{g}_{n+1} \;+\; \ldots \;+\; \beta^m\mathbf{g}_{n+m}.\end{aligned}$$

It is simple to verify that $\mathcal{A}$ is linear:

$$\begin{aligned}\alpha\mathcal{A}(\mathbf{v}) + \mathcal{A}(\overline{\mathbf{v}}) \;&=\; \alpha\alpha^1\mathbf{g}_1 \;+\; \ldots \;+\; \alpha\alpha^n\mathbf{g}_n \;+\; \alpha\beta^1\mathbf{g}_{i+n} \;+\; \alpha\beta^m\mathbf{g}_{n+m}\\ &\qquad +\overline{\alpha}^1\mathbf{g}_1 \;+\; \ldots \;+\; \overline{\alpha}^n\mathbf{g}_n \;+\; \overline{\beta}^1\mathbf{g}_{i+n} \;+\; \overline{\beta}^m\mathbf{g}_{n+m}\\ &=\; (\alpha\alpha^1 + \overline{\alpha}^1)\mathbf{g}_1 \;+\; \ldots \;+\; (\alpha\alpha^n + \overline{\alpha}^n)\mathbf{g}_n \;+\; (\alpha\beta^1 + \overline{\beta}^1)\mathbf{g}_{1+n} \;+\; \ldots \;+\; (\alpha\beta^m + \overline{\beta}^m)\mathbf{g}_{n+m}\\ &=\; \mathcal{A}(\alpha\mathbf{v} + \overline{\mathbf{v}}).\end{aligned}$$

It is also simple to verify that $Ker(\mathcal{A}) = \{0\}$:

$\mathcal{A}(\mathbf{v}) \;=\; 0$ *implies*
$\alpha^1\mathbf{g}_1 \;+\; \ldots \;+\; \alpha^n\mathbf{g}_n \;+\; \beta^1\mathbf{g}_{1+n} \;+\; \ldots \;+\; \beta^m\mathbf{g}_{n+m} \;=\; 0\mathbf{g}_1 \;+\; \ldots \;+\; 0\mathbf{g}_{n+m}$ *implies*
$\alpha^1 = 0, \ldots, \alpha^n = 0, \beta^1 = 0, \ldots, \beta^m = 0$ *implies*
$\mathbf{v} = 0.$

Finally, we can verify that $\mathcal{A}$ is onto: if $\mathbf{w} = \gamma^1\mathbf{g}_1 \;+\; \ldots \;+\; \gamma^{n+m}\mathbf{g}_{n+m}$ is an element of $\mathbb{R}^{n+m}$, then the element $\mathbf{v} \in \mathbb{R}^n \oplus \mathbb{R}^m$ given by

$$\mathbf{v} \;=\; \left(\gamma^1\mathbf{e}_1 \;+\; \ldots \;+\; \gamma^n\mathbf{e}_n\right) \;\oplus\; \left(\gamma^{1+n}\mathbf{f}_1 \;+\; \ldots \;+\; \gamma^{n+m}\mathbf{f}_m\right)$$

satisfies $\mathcal{A}(\mathbf{v}) \;=\; \mathbf{w}$. □

As a side note, the direct sum is also sometimes called the "cross product".

2 Tensor products

2.1 Definition of $V \otimes W$

The tensor product is formal multiplication of vectors, which is required to obey the linearity relations. If V, W are two vector spaces and $v \in V$, $w \in W$ are vectors, we denote their tensor product by

$$v \otimes w.$$

The linearity relations are the following:

$$\begin{aligned} v \otimes (\alpha w) &= \alpha(v \otimes w) \\ (\alpha v) \otimes w &= \alpha(v \otimes w) \\ v \otimes (w + w') &= v \otimes w + v \otimes w' \\ (v + v') \otimes w &= v \otimes w + v' \otimes w \end{aligned}$$

where $\alpha \in \mathbb{R}$ and $v, v' \in V$, $w, w' \in W$ are vectors. The tensor product $V \otimes W$ of two vector spaces is defined to be the linear span of elements of the form $v \otimes w$. That is,

$$V \otimes W = \quad \{v \otimes w \,|\, v \in V,\, w \in W\}.$$

2.2 A basis for $V \otimes W$

If bases for V and W are chose, it is possible to write down a basis for the vector space $V \otimes W$. Let $\{v_i\} \subset V$ be a basis for V and $\{w_i\} \subset W$ be a basis for W. The various tensor products $v_i \otimes w_j$ are elements of $V \otimes W$. A typical element of $V \otimes W$ is a linear combination of the $v_i \otimes w_j$:

$$\alpha^{ij} v_i \otimes w_j$$

where the various coefficients $\alpha^{ij} \in \mathbb{R}$.

Example Let $V = \mathbb{R}^2$ and $W = \mathbb{R}^2$, with bases $\{v_1, v_2\}$ and $\{w_1, w_2\}$. Find a basis for $V \otimes W$, and describe a typical element.

Solution $V \otimes W$ is the 4-dimensional space spanned by

$$v_1 \otimes w_1, \quad v_1 \otimes w_2, \quad v_2 \otimes w_1, \quad v_2 \otimes w_2.$$

A typical element $T \in V \otimes W$ can be written

$$T = \alpha^{11} v_1 \otimes w_1 + \alpha^{12} v_1 \otimes w_2 + \alpha^{21} v_2 \otimes w_1 + \alpha^{22} v_2 \otimes w_2,$$

where $\alpha^{11}, \alpha^{12}, \alpha^{21}, \alpha^{22} \in \mathbb{R}$.

Theorem 2.1 *The vector space $\mathbb{R}^n \otimes \mathbb{R}^k$ is isomorphic with $\mathbb{R}^{nk}$.*

Pf Homework assignment! $\square$

Lecture 10 - Tensor Products

Feb 23, 2009

1 The Tensor algebra over V

If V is a vector space with scalar field $\mathbb{R}$, we use the notation

$$\begin{aligned} \mathbb{R} &\triangleq \bigotimes^0 V \triangleq V^{\otimes 0} \\ V &\triangleq \bigotimes^1 V \triangleq V^{\otimes 1} \\ V \otimes V &\triangleq \bigotimes^2 V \triangleq V^{\otimes 2} \\ V \otimes V \otimes V &\triangleq \bigotimes^3 V \triangleq V^{\otimes 3} \\ &\textit{etc.} \end{aligned}$$

Elements of the space $\bigotimes^i V$ are called *(homogeneous) tensors of degree* i.

Theorem 1.1 *If $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis for V, then the n^i many elements of the form*

$$\mathbf{v}_{i_1} \otimes \mathbf{v}_{i_2} \otimes \cdots \otimes \mathbf{v}_{i_n}$$

constitute a basis for the vector space $\bigotimes^i V$.

□

The (infinite dimensional) algebra

$$\left(\bigotimes^0 V\right) \oplus \left(\bigotimes^1 V\right) \oplus \left(\bigotimes^2 V\right) \oplus \ldots.$$

is called the *tensor algebra* over V. A *tensor* is just an element of the tensor algebra. A tensor T is called *homogeneous of degree* i if $T \in \bigotimes^i V$. A tensor T is called *decomposable* if $T \in \bigotimes^i V$ can be written in the form

$$T = v^{(1)} \otimes \cdots \otimes v^{(i)},$$

where the $v^{(j)}$ are elements of V. Otherwise T is called *indecomposable*.

Examples. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis for V. A typical element $T \in \bigotimes^2 V$ is a linear combination of elements of the form $\mathbf{v}_i \otimes \mathbf{v}_j$, namely

$$T \;=\; T^{ij}\mathbf{v}_i \otimes \mathbf{v}_j,$$

where, of course, summation takes place in both the i and j indices.

A typical element $T \in \bigotimes^3 V$ is a linear combination of elements of the form $\mathbf{v}_i \otimes \mathbf{v}_j \otimes \mathbf{v}_k$, namely

$$T \;=\; T^{ijk}\mathbf{v}_i \otimes \mathbf{v}_j \otimes \mathbf{v}_k.$$

The tensor $T = \mathbf{v}_1 \otimes \mathbf{v}_1 + \mathbf{v}_2 \otimes \mathbf{v}_2 \otimes \mathbf{v}_2$ is not homogeneous. The tensor $S = \mathbf{v}_1 \otimes \mathbf{v}_2 + 2\mathbf{v}_1 \otimes \mathbf{v}_3$ is homogeneous and decomposable. The tensor $U = \mathbf{v}_1 \otimes \mathbf{v}_2 + 2\mathbf{v}_3 \otimes \mathbf{v}_4$ is homogeneous and indecomposable.

2 The bigraded tensor algebra

It is possible to tensor with the dual space V^*. We define

$$\bigotimes^{r,s} V \quad = \quad V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$$

(r many copies of V, s many copies of V^*). The tensor product is not commutative, meaning that $V \otimes W$ is not the same space as $W \otimes V$. However, these spaces are isomorphic in a natural way:

$$\begin{aligned} \mathcal{T} : V \times W &\to W \otimes V \\ \mathcal{T}(v \otimes w) \;&=\; w \otimes v. \end{aligned}$$

This is called the *transpose map*. Applying this as many times as necessary, we can see that

$$\begin{aligned} V \otimes V^* \otimes V^* \otimes V \quad &\approx \quad V \otimes V \otimes V^* \otimes V^* \\ V \otimes V^* \otimes V \otimes V^* \otimes V^* \otimes V \quad &\approx \quad V \otimes V \otimes V \otimes V^* \otimes V^* \\ &etc \quad , \end{aligned}$$

so that any tensor product of r many V's and s many V^*'s, no matter what the order, is isomorphic to $\bigotimes^{r,s} V$.

The (infinite dimensional) bigraded tensor algebra is commonly denoted

$$\bigotimes^{*,*} V.$$

An element $T \in \bigotimes^{*,*} V$ is called *homogeneous of bidegree* (r, s) if $T \in \bigotimes^{r,s} V$.

As examples, a typical element T of $\bigotimes^{1,1} V$ is given by

$$T \;=\; T^i_j \mathbf{v}_i \otimes \mathbf{v}^j.$$

A typical element T of $\bigotimes^{1,2} V$ is given by

$$T \;=\; T^i_{jk} \mathbf{v}_i \otimes \mathbf{v}^j \otimes \mathbf{v}^k.$$

a typical element T of $\bigotimes^{2,3} V$ is given by

$$T \;=\; T^{ij}_{klm} \mathbf{v}_i \otimes \mathbf{v}_j \otimes \mathbf{v}^k \otimes \mathbf{v}^l \otimes \mathbf{v}^m.$$

3 Tensors as bilinear maps and as operators

We give two concrete examples of uses for tensors.

Let $V = span\{\mathbf{v}_1, \mathbf{v}_2\}$ be a 2-dimensional vector space, with dual V^* and dual basis $\mathbf{v}^1, \mathbf{v}^2$. It is possible to consider elements of $\bigotimes^{0,2} V$ to be bilinear maps of the form $V \times V \to \mathbb{R}$, given by

$$v^* \otimes w^*(\tilde{v}, \tilde{w}) \quad = \quad v^*(\tilde{v})\, w^*(\tilde{w}),$$

were $v^*, w^* \in V^*$ and $\tilde{v}, \tilde{w} \in V$. If $T \in \bigotimes^{0,2} V$, is is simple to prove that T is *bilinear*, meaning it is linear in each entry:

$$\begin{aligned} T(\alpha v + \bar{v}, w) \;&=\; \alpha T(v, w) \,+\, T(\bar{v}, w) \\ T(v, \alpha w + \bar{w}) \;&=\; \alpha T(v, w) \,+\, T(v, \bar{w}). \end{aligned}$$

For example, let

$$T \in \bigotimes^{0,2} V \quad \textit{be given by} \quad T \;=\; \mathbf{v}^1 \otimes \mathbf{v}^1 \,-\, 2\mathbf{v}^1 \otimes \mathbf{v}^2$$

(notice that T is both homogeneous and decomposable). Then

$$
\begin{aligned}
T(\mathbf{v}_1, \mathbf{v}_1) &= \mathbf{v}^1 \otimes \mathbf{v}^1(\mathbf{v}_1, \mathbf{v}_1) - 2\mathbf{v}^1 \otimes \mathbf{v}^2(\mathbf{v}_1, \mathbf{v}_1) \\
&= \mathbf{v}^1(\mathbf{v}_1)\mathbf{v}^1(\mathbf{v}_1) - 2\mathbf{v}^1(\mathbf{v}_1)\mathbf{v}^2(\mathbf{v}_1) \\
&= 1 \\
T(\mathbf{v}_1, \mathbf{v}_2) &= \mathbf{v}^1 \otimes \mathbf{v}^1(\mathbf{v}_1, \mathbf{v}_2) - 2\mathbf{v}^1 \otimes \mathbf{v}^2(\mathbf{v}_1, \mathbf{v}_2) \\
&= \mathbf{v}^1(\mathbf{v}_1)\mathbf{v}^1(\mathbf{v}_2) - 2\mathbf{v}^1(\mathbf{v}_1)\mathbf{v}^2(\mathbf{v}_2) \\
&= -2 \\
T(\mathbf{v}_2, \mathbf{v}_1) &= \mathbf{v}^1 \otimes \mathbf{v}^1(\mathbf{v}_2, \mathbf{v}_1) - 2\mathbf{v}^1 \otimes \mathbf{v}^2(\mathbf{v}_2, \mathbf{v}_1) \\
&= \mathbf{v}^1(\mathbf{v}_2)\mathbf{v}^1(\mathbf{v}_1) - 2\mathbf{v}^1(\mathbf{v}_2)\mathbf{v}^2(\mathbf{v}_1) \\
&= 0 \\
T(\mathbf{v}_2, \mathbf{v}_2) &= \mathbf{v}^1 \otimes \mathbf{v}^1(\mathbf{v}_2, \mathbf{v}_2) - 2\mathbf{v}^1 \otimes \mathbf{v}^2(\mathbf{v}_2, \mathbf{v}_2) \\
&= \mathbf{v}^1(\mathbf{v}_2)\mathbf{v}^1(\mathbf{v}_2) - 2\mathbf{v}^1(\mathbf{v}_2)\mathbf{v}^2(\mathbf{v}_2) \\
&= 0
\end{aligned}
$$

Notice that T is *not* symmetric: for example $T(\mathbf{v}_1, \mathbf{v}_2) \neq T(\mathbf{v}_2, \mathbf{v}_1)$.

An *inner product* on a vector space V is a map $V \times V \to \mathbb{R}$ (commonly denoted $g(\cdot,\cdot)$ or $\langle\cdot,\cdot\rangle$) that satisfies

- Bilinearity: $g(\alpha v+\hat{v},w) \;=\; \alpha g(v,w)+g(\hat{v},w)$ and $g(v,\alpha w+\hat{w}) \;=\; \alpha g(v,w)+g(v,\hat{w})$
- Symmetry: $g(v,w)=g(w,v)$
- Nondegeneracy: given $v \in V$, there is at least one vector $w \in V$ so that $g(v,w) \neq 0$.

The Euclidean inner product on $\mathbb{R}^n$ is given by

$$g \;=\; \mathbf{v}^1 \otimes \mathbf{v}^1 \;+\; \mathbf{v}^2 \otimes \mathbf{v}^2 \;+\; \ldots \;+\; \mathbf{v}^n \otimes \mathbf{v}^n.$$

The Minkowski inner product is given by

$$g \;=\; \mathbf{v}^1 \otimes \mathbf{v}^1 \;-\; \frac{1}{c^2}\,\mathbf{v}^2 \otimes \mathbf{v}^2 \;-\; \ldots \;-\; \frac{1}{c^2}\,\mathbf{v}^n \otimes \mathbf{v}^n.$$

As another example, the tensor

$$g \;=\; \mathbf{v}^1 \otimes \mathbf{v}^2 \;+\; \mathbf{v}^2 \otimes \mathbf{v}^1$$

is an inner product. But the tensor

$$g \;=\; \mathbf{v}^1 \otimes \mathbf{v}^2$$

is not an inner product for two reasons, namely it is not symmetric, and it is degenerate: $g(\mathbf{v}_2,\cdot) \equiv 0$ no matter what goes in the second slot.

A second application of tensor products is to linear operators. After choosing a basis $\mathbf{v}_i$ and a dual basis $\mathbf{v}^i$, a tensor $A \in \bigotimes^{1,1} V$ is given by a linear combination of elements of the form $\mathbf{v}_i \otimes \mathbf{v}^j$:

$$A \;=\; A^i_j \mathbf{v}_i \otimes \mathbf{v}^j.$$

This tensor can be considered to be a linear map $V \to V$, as follows: given $v \in V$,

$$\begin{aligned} &A(v) \in V \\ &A(v) \;=\; A^i_j \mathbf{v}_i\, \mathbf{v}^j(v). \end{aligned}$$

Is is simple to verify that $A : V \to V$ is linear.

For example, let $V = \mathit{span}\{\mathbf{v}_1, \mathbf{v}_2\}$, and let A be the tensor

$$\begin{aligned} &A \in \bigotimes^{0,2} V \\ &A \;=\; \mathbf{v}_1 \otimes \mathbf{v}^1 \;-\; 2\mathbf{v}_2 \otimes \mathbf{v}^1 \;+\; 3\mathbf{v}_1 \otimes \mathbf{v}_2 \;+\; 10\mathbf{v}_2 \otimes \mathbf{v}^2 \end{aligned}$$

Then

$$\begin{aligned}
A(\mathbf{v}_1) &= \mathbf{v}_1\mathbf{v}^1(\mathbf{v}_1) - 2\mathbf{v}_2\mathbf{v}^1(\mathbf{v}_1) + 3\mathbf{v}_1\mathbf{v}_2(\mathbf{v}_1) + 10\mathbf{v}_2\mathbf{v}^2(\mathbf{v}_1) \\
&= \mathbf{v}_1 - 2\mathbf{v}_2 \\
A(\mathbf{v}_2) &= \mathbf{v}_1\mathbf{v}^1(\mathbf{v}_2) - 2\mathbf{v}_2\mathbf{v}^1(\mathbf{v}_2) + 3\mathbf{v}_1\mathbf{v}_2(\mathbf{v}_2) + 10\mathbf{v}_2\mathbf{v}^2(\mathbf{v}_2) \\
&= 3\mathbf{v}_1 + 10\mathbf{v}_2.
\end{aligned}$$

This is equivalent to our old notation, where take the numbers A_i^j to be implicitly defined by

$$A(\mathbf{v}_i) = A_i^j \mathbf{v}_j.$$

In either case, we have

$$A_1^1 = 1, \quad A_1^2 = -2, \quad A_2^1 = 3, \quad A_2^2 = 10.$$

Lecture 11 - Tensors as maps, dual spaces, transformation properties, alternating tensors, and wedge products

Feb 25, 2009

1 Tensors as maps

Let V be a vector space. We define V^* to be the vector space of linear maps $V \to \mathbb{R}$, but we also know that (in the finite dimensional case at least) the space V is the space of maps $V^* \to \mathbb{R}$. Likewise, we can consider elements of $\bigotimes^{0,k} V$ to be k-fold linear maps $V \times \cdots \times V \to \mathbb{R}$:

$$\begin{aligned} T \in \bigotimes^{0,k} V \quad given\ by \quad T \ &= \ T_{i_1 i_2 \ldots i_k} \mathbf{v}^{i_1} \otimes \mathbf{v}^{i_2} \otimes \cdots \otimes \mathbf{v}^{i_k} \\ T(v_{(1)}, \ldots, v_{(k)}) \ &= \ T_{i_1 i_2 \ldots i_k} \mathbf{v}^{i_1} \otimes \mathbf{v}^{i_2} \otimes \cdots \otimes \mathbf{v}^{i_k} \left(v_{(1)}, \ldots, v_{(k)} \right) \\ &= \ T_{i_1 i_2 \ldots i_k} \mathbf{v}^{i_1}(v_{(1)}) \mathbf{v}^{i_2}(v_{(2)}) \ldots \mathbf{v}^{i_k}(v_{(k)}), \end{aligned}$$

and elements of $\bigotimes^{k,0} V$ to be k-fold linear maps $V^* \times \cdots \times V^* \to \mathbb{R}$:

$$\begin{aligned} T \in \bigotimes^{k,0} V \quad given\ by \quad T \ &= \ T^{i_1 i_2 \ldots i_k} \mathbf{v}_{i_1} \otimes \mathbf{v}_{i_2} \otimes \cdots \otimes \mathbf{v}_{i_k} \\ T(v^{(1)}, v^{(2)}, \ldots, v^{(k)}) \ &= \ T^{i_1 i_2 \ldots i_k} \mathbf{v}_{i_1} \otimes \mathbf{v}_{i_2} \otimes \cdots \otimes \mathbf{v}_{i_k} \left(v^{(1)}, \ldots, v^{(k)} \right) \\ &= \ T^{i_1 i_2 \ldots i_k} \mathbf{v}_{i_1}(v^{(1)}) \mathbf{v}_{i_2}(v^{(2)}) \ldots \mathbf{v}_{i_k}(v^{(k)}). \end{aligned}$$

Finally, it is possible to regard any element $T \in \bigotimes^{r,s} V$ as a map $T : V^* \times \cdots \times V^* \times V \times \cdots \times V \to \mathbb{R}$. For example, an element $T \in \bigotimes^{1,2} V$, given by

$$T \ = \ T^i_{jk} \, \mathbf{v}_i \otimes \mathbf{v}^j \otimes \mathbf{v}^k$$

can be considered to be a map $V^* \times V \times V \to R$:

$$T(v^*, w, x) \ = \ T^i_{jk} \, \mathbf{v}_i(v^*) \mathbf{v}^j(w) \mathbf{v}^k(x),$$

where $v^* \in V^*$ and $w, x \in V$.

2 An addition to the Einstein notation

We introduce another feature of Einstein notation. Recall that we defined isomorphisms $V^* \otimes V \otimes V^* \approx V \otimes V^* \otimes V^*$, etc. However, it is sometimes important to preserve the order of the tensor products. As a point of fact,

$$T \;=\; v^* \otimes w \otimes x^* \quad \textit{and} \quad S \;=\; w \otimes v^* \otimes x^*$$

are different tensors. This is encoded in the Einstein notation by preserving the ordering of the indices:

$$T \;=\; T_i{}^j{}_k \, \mathbf{v}^i \otimes \mathbf{v}_j \otimes \mathbf{v}^k$$

and

$$S \;=\; S^j{}_{ik} \, \mathbf{v}_j \otimes \mathbf{v}^i \otimes \mathbf{v}^k$$

are in different tensor spaces. In fact,

$$T : V \times V^* \times V \to \mathbb{R},$$

whereas

$$S : V^* \times V \times V \to \mathbb{R}\,.$$

3 Dual spaces

If V^* is dual to V and V is dual to V^*, what is the dual to $\bigotimes^{r,s} V$? It is $\bigotimes^{s,r} V$.

Given a tensor $T^{r,s} \in \bigotimes^{r,s} V$, we can consider it to be a linear map $\bigotimes^{s,r} V \to \mathbb{R}$. On decomposable elements of $\bigotimes^{s,r} V$ we define this by

$$\begin{aligned} T &\left(v^{(i_1)} \otimes v^{(i_2)} \otimes \cdots \otimes v^{(i_s)} \otimes v_{(j_1)} \otimes v_{(j_2)} \otimes \cdots \otimes v_{(j_r)}\right) \\ &\triangleq\; T\left(v_{(j_1)},\, v_{(j_2)},\, \ldots,\, v_{(j_r)},\, v^{(i_1)},\, v^{(i_2)},\, \ldots,\, v^{(i_s)}\right), \end{aligned}$$

and extending linearly.

4 Transformation properties

Let $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_n\}$ be different bases for V, with $A = A_{\{f_i\} \leftarrow \{e_i\}}$ the the transition matrix between them. Let $\{e^i, \ldots, e^n\}$ and $\{f^1, \ldots, f^n\}$ be the respective dual bases, with transition maps $B = B_{\{f^i\} \leftarrow \{e^i\}}$. We have

$$f_i \;=\; A^j{}_i e_j \quad \textit{and} \quad f^i \;=\; B^i{}_j e^j.$$

As was discussed in Lecture 8, section 4, we have the relation $B = A^{-1}$. That is to say,

$$B^i{}_k A^k{}_j \;=\; \delta^i{}_j \quad and \quad A^i{}_k B^k{}_j \;=\; \delta^i{}_j.$$

A tensor $T \in \bigotimes^{1,1} V$, for instance, may have the expression

$$T \;=\; T^i{}_j \, e_i \otimes e^j$$

in the $\{e_i\} - \{e^i\}$ basis. Its expression in the $\{f_i\} - \{f^i\}$ basis is given by

$$\begin{aligned} T &= T^i{}_j \, e_i \otimes e^j \\ &= T^i{}_j \left(A^k{}_i f_k\right) \otimes \left(B^j{}_l f^l\right) \\ &= \left(A^k{}_i B^j{}_l T^i{}_j\right) \, f_k \otimes f^l. \end{aligned}$$

Likewise for elements of any of the spaces $\bigotimes^{r,s} V$.

5 Symmetric and alternating tensors

A tensor $T : V \times \cdots \times V \to \mathbb{R}$ (r many V's) is called a *symmetric tensor* if

$$T\left(v_{(1)}, \ldots, v_{(i)}, v_{(i+1)}, \ldots, v_{(r)}\right) \;=\; T\left(v_{(1)}, \ldots, v_{(i+1)}, v_{(i)}, \ldots, v_{(r)}\right).$$

That is, if you interchanging any two consecutive entries leaves the tensor unchanged. A tensor $T : V \times \cdots \times V \to \mathbb{R}$ (r many V's) is called an *alternating* or *antisymmetric tensor* if

$$T\left(v_{(1)}, \ldots, v_{(i)}, v_{(i+1)}, \ldots, v_{(r)}\right) \;=\; -T\left(v_{(1)}, \ldots, v_{(i+1)}, v_{(i)}, \ldots, v_{(r)}\right).$$

That is, if interchanging any two consecutive entries introduces a minus sign.

This leads us to two new definitions.

Definition The space $\bigodot^s V^* \subset \bigotimes^{0,s} V$ is the space of symmetric tensors of the type $V \times \cdots \times V \to \mathbb{R}$ (s many V's).

Definition The space $\bigwedge^s V^* \subset \bigotimes^{0,s} V$ is the space of alternating tensors of the type $V \times \cdots \times V \to \mathbb{R}$ (s many V's).

Example Let V be a vectors space with basis $\{v_i\}$ and dual basis $\{v^i\}$. Let

$$\begin{aligned} T &= v^1 \otimes v^2 + v^2 \otimes v^1 \\ S &= v^1 \otimes v^2 - v^2 \otimes v^1 \\ U &= v^1 \otimes v^2 \\ W &= v^1 \otimes v^1. \end{aligned}$$

Then T and W are symmetric tensors, S is an antisymmetric tensor, and U is neither symmetric nor antisymmetric.

6 The *Alt* map and the wedge product

There is a canonical way of transforming any tensor $T \in \bigotimes^{0,s} V$ into an alternating tensor, given by the *Alt* map:

$$\begin{aligned} Alt : \bigotimes^{0,s} V &\twoheadrightarrow \bigwedge^{s} V^* \\ Alt(T)(v_{(1)}, \ldots, v_{(s)}) \;&=\; \frac{1}{s!} \sum_{\pi \in \mathrm{Sym}(s)} (-1)^{|\pi|} T(v_{(\pi 1)}, \ldots, v_{(\pi s)}). \end{aligned}$$

For instance, if $T \in \bigotimes^{0,2} V$, then

$$Alt(T)(v,w) \;=\; \frac{1}{2} \left(T(v,w) \,-\, T(w,v) \right).$$

If $T \in \bigotimes^{0,3} V$, then

$$Alt(T)(v,w,x) \;=\; \frac{1}{6} \left(T(v,w,x) \,-\, T(v,x,w) \,-\, T(w,v,x) \,+\, T(w,x,v) \,+\, T(x,v,w) \,-\, T(x,w,v) \right).$$

Since $Alt(T)$ is itself a tensor, we should be able to express in terms of a basis. For instance if $T = v_1 \otimes v_2$ then

$$Alt(T) \;=\; \frac{1}{2} \left(v_1 \otimes v_2 \,-\, v_2 \otimes v_1 \right)$$

and if $T = v_1 \otimes v_2 \otimes v_3$, then

$$Alt(T) \;=\; \frac{1}{6} \left(v_1 \otimes v_2 \otimes v_3 \,-\, v_1 \otimes v_3 \otimes v_2 \,-\, v_2 \otimes v_1 \otimes v_3 \,+\, v_2 \otimes v_3 \otimes v_1 \,+\, v_3 \otimes v_1 \otimes v_2 \,-\, v_3 \otimes v_2 \otimes v_1 \right).$$

Theorem 6.1 *The map* $Alt : \bigotimes^{0,s} V \to \bigwedge^{s} V^*$ *is onto, and linear (meaning* $Alt(\alpha T + S) = \alpha Alt(T) + Alt(S)$*). If* $T \in \bigotimes^{0,s} V$*, then* $Alt(Alt(T)) = Alt(T)$*.*

□

Given two alternating tensors, $T \in \bigwedge^{n} V^*$ and $S \in \bigwedge^{m} V^*$, the *wedge product* $T \wedge S$ of T and S is defined to be

$$T \wedge S \;\triangleq\; Alt(T \otimes S).$$

Notice that $T \wedge S \in \bigwedge^{n+m} V^*$.

Theorem 6.2 *If* $T \in \bigwedge^{n} V^*$ *and* $S \in \bigwedge^{m} V^*$*, then* $T \wedge S \;=\; (-1)^{nm} S \wedge T$*.*

Example. Express $v^1 \wedge v^2$ as a tensor.
Solution:

$$v^1 \wedge v^2 \;=\; Alt(v^1 \otimes v^2) \;=\; \frac{1}{2} \left(v^1 \otimes v^2 \,-\, v^2 \otimes v^1 \right).$$

Lecture 12 - Metric linear algebra

March 2, 2009

1 Metrics

Let V be a vector space with basis $\{v_1, \ldots, \mathbf{v}_n\}$. Assume V is endowed with an *inner product* $g \in \bigotimes^2 V^*$. That is, g is given by

$$g \;=\; g_{ij}\mathbf{v}^i \otimes \mathbf{v}^j$$

where g satisfies

- Symmetry: $g(v,w) = g(w,v)$ for any $v, w \in V$. In other words, the matrix g_{ij} is symmetric: $g_{ij} = g_{ji}$.
- Nondegeneracy: if $0 \neq v \in V$, then there is some $\bar{v} \in V$ so that $g(v, \bar{v}) \neq 0$.

An inner products is often called a *metric*.

2 The musical isomorphisms

Given a basis $\{\mathbf{v}_i\} \subset V$, we have discussed the existence of a dual basis $\{\mathbf{v}^i\} \subset V^*$. One might be tempted to think that this leads to an isomorphism $V \to V^*$, but any such attempt to define such an isomorphism will be dependent on the basis that has been chosen.

If the vector space has a metric g, there is a *natural* (that is, basis-independent) isomorphism

$$\flat : V \to V^*.$$

This is defined by

$$\begin{aligned} &\flat(v) \in V^* \\ &\flat(v)(w) \;=\; g(v,w). \end{aligned}$$

Usually this is denoted more simply by

$$
\begin{aligned}
v \in V \quad &\mapsto \quad v_\flat \in V^* \\
v_\flat(\cdot) \;&=\; g(v, \cdot) \\
v_\flat(w) \;&=\; g(v, w) \quad for \quad w \in V.
\end{aligned}
$$

and the like. The fact that this is an isomorphism is equivalent to the nondegeneracy of the metric (homework problem). The inverse of the "$\flat$" isomorphism is the "$\sharp$" isomorphism

$$\sharp : V^* \to V \quad \textit{is given by} \quad \sharp \;=\; \flat^{-1}.$$

Given $f \in V^*$, we have

$$\sharp(f) \in V, \quad \textit{often denoted} \quad f^\sharp \in V.$$

It is easy to show (homework problem) that $f^\sharp \in V$ is characterized by

$$g(f^\sharp, v) \;=\; f(v).$$

3 The metric on the dual space

If $g = g_{ij}\mathbf{v}^i \otimes \mathbf{v}^j$ is a metric on V, we can define, in a natural (that is to say, basis-free) way, a metric on the dual space V^*. Given $f, g \in V^*$, we define

$$g(f,\, g) \;=\; g(f^\sharp,\, g^\sharp)$$

(recalling that $f^\sharp, g^\sharp \in V$ and $g : V \times V \to \mathbb{R}$). Considering g as a map $V^* \times V^* \to \mathbb{R}$, we can write

$$g \;=\; g^{ij}\mathbf{v}_i \otimes \mathbf{v}_j.$$

It is possible to prove that the matrix g^{ij} is the inverse of the matrix g_{ij} (homework problem). That is to say, it holds that

$$g^{ik}g_{kj} \;=\; \delta^i{}_j.$$

4 The musical isomorphisms in component form (raising and lowering indices)

Given a basis $\{\mathbf{v}_i\} \subset V$ and its dual basis $\{\mathbf{v}^i\} \subset V^*$, how can we express the musical isomorphisms? Assume

$$v \;=\; \alpha^i$$

is a vector (recall this is shorthand for $v = \alpha^i \mathbf{v}_i$). How can we find the components of the covector $v_\flat = \alpha_i$? (We are NOT assuming that the numbers α^i are the same as the numbers α_i.) By the definition of $\flat$, we have

$$\begin{aligned} v_\flat(\mathbf{v}_j) &= g(v, \mathbf{v}_j) \;=\; \alpha^i g(\mathbf{v}_i, \mathbf{v}_j) \\ &= \alpha^i g_{kl} \mathbf{v}^k \otimes \mathbf{v}^l (\mathbf{v}_i, \mathbf{v}_j) \;=\; \alpha^i g_{kl} \delta^k{}_i \delta^l{}_j \\ &= \alpha^i g_{ij} \end{aligned}$$

But of course also

$$\begin{aligned} v_\flat(\mathbf{v}_j) &= \alpha_i \mathbf{v}^i(\mathbf{v}_j) \\ &= \alpha_i \delta^i{}_j \;=\; \alpha_j. \end{aligned}$$

Therefore $\alpha_j \;=\; \alpha^i g_{ij}$.

This procedure is often called *lowering the index.*

Now we describe the $\sharp$ isomorphism in components. Let $f = f_i$ be a covector (recall that this means $f = f_i \mathbf{v}^i$). We define the numbers f^i by $f^\sharp = f^i$. Using the definition of $f^\sharp$, we have

$$\begin{aligned} f(\mathbf{v}_j) \;&=\; g(f^\sharp, \mathbf{v}_j) \;=\; g(f^i \mathbf{v}_i, \mathbf{v}_j) \;=\; f^i g_{ij} \\ f(\mathbf{v}_j) \;&=\; f_i \mathbf{v}^i(\mathbf{v}_j) \;=\; f_i \delta^i{}_j \;=\; f_j. \end{aligned}$$

Thus we can implicitly define f^i by the relationship

$$f^i g_{ij} \;=\; f_j.$$

Recalling that g^{ij} is the inverse of g_{ij}, we have

$$\begin{aligned} f^i g_{ij} g^{jk} \;&=\; f_j g^{jk} \\ f^i \delta^k{}_i \;&=\; f_j g^{jk} \\ f^k \;&=\; f_j g^{jk}. \end{aligned}$$

This means that $f^\sharp = f^k \mathbf{v}_i \in V$. This procedure is often called *raising the index.*

5 Raising and lowering tensor indices

Given an arbitrary tensor, for example $T = T^i{}_j \in \bigotimes^{1,1} V$, we can raise or lower its indices. For example, the corresponding $T_{ij} \in \bigotimes^{0,2} V$ is given by

$$T_{ij} \;=\; T^k{}_j g_{ki}$$

and the corresponding tensor $T^{ij} \in \bigotimes^{2,0} V$ is given by

$$T^{ij} \;=\; T^i{}_k g^{kj}.$$

Lecture 13 - Vectors as directional derivatives

March 9, 2009

1 Coordinates

Let M be some space, say Euclidean n-space, Minkowski $1+n$-space, or the like. Coordinates are functions on the space M that assign to each point some unique set of numbers. It is important to understand that a given space M is not a vector space, and coordinates are not basis vectors of any kind. Coordinates are functions, pure and simple.

2 Vectors, tangent spaces, and the tangent bundle

Intuitively, a vector is a magnitude and a direction. This is not a rigorous definition, however. A concept that can be made precise is the notion of the *derivative of a function along a curve*. To define this concept, let $p \in M$ be a point, let $f : M \to \mathbb{R}$ be a function, and let $\gamma : (-\epsilon, \epsilon) \to M$ be a curve parameterized by $\tau \in (-\epsilon, \epsilon)$ with $\gamma(0) = p$. Then the derivative of f along γ at p is defined to be

$$\frac{d}{d\tau}\Big|_p f \triangleq \lim_{h \to 0} \frac{f(\gamma(h)) - f(\gamma(0))}{h}.$$

One computes this expression using partial derivatives: if $\{x^1, \ldots, x^n\}$ are coordinates on M, we can write $f = f(x^1, \ldots, x^n)$ and compute

$$\frac{d}{d\tau} f = \frac{dx^1}{d\tau}\frac{\partial}{\partial x^1} f + \ldots + \frac{dx^n}{d\tau}\frac{\partial}{\partial x^n} f,$$

that is, the operator $\frac{d}{d\tau}$ is a linear combination of the operators $\frac{\partial}{\partial x^i}$

We have not defined the term "vector" yet, but intuitively two paths $\gamma(\tau)$ and $\widetilde{\gamma}(\tilde{\tau})$ which pass through the point p posses the same velocity vector at p if $\frac{dx^i}{d\tau} = \frac{dx^i}{d\tilde{\tau}}$, which is to say that $\frac{d}{d\tau} = \frac{d}{d\tilde{\tau}}$.

Our intuitive notion of vectors seems to coincide with the mathematically precise notion of directional derivatives. Thus we say v is a vector based at $p \in M$ if v_p is a linear combination of the directional derivatives $\partial/\partial x^i$:

$$v \;=\; v^i \, \frac{\partial}{\partial x^i}\Big|_p .$$

Note that we are justified in say that the partials $\partial/\partial x^i$ are directional derivatives: $\partial/\partial x^i$ is obtained by varying x^i and fixing all other coordinates.

The *tangent space* at p, denoted T_pM, is defined to be the vector space of all vectors based at p.

The *tangent bundle* of M, denoted TM, is defined to be the collection of all tangent spaces T_pM based at all points p of M.

3 Change of coordinates

If the coordinate functions are changed, it is important to know how to change the basis vectors of each tangent space T_pM. Let $\{x^1, \ldots, x^n\}$ and $\{y^1, \ldots, y^n\}$ be two coordinate systems on M. We have the relationship

$$\frac{\partial}{\partial x^i} \;=\; \frac{\partial y^j}{\partial x^i} \frac{\partial}{\partial y^j}.$$

For example, if r, θ are the so-called polar coordinates on the Euclidean plane and $x = r\cos\theta$, $y = r\sin\theta$ are the corresponding rectangular coordinates, we have

$$\begin{aligned}
\frac{\partial}{\partial r} &= \frac{\partial x}{\partial r}\frac{\partial}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial}{\partial y} \\
&= \cos(\theta)\,\frac{\partial}{\partial x} + \sin(\theta)\,\frac{\partial}{\partial y} \\
&= \frac{x}{\sqrt{x^2 + y^2}}\,\frac{\partial}{\partial x} + \frac{y}{\sqrt{x^2 + y^2}}\,\frac{\partial}{\partial y}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \theta} &= \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y} \\
&= -r\sin(\theta)\,\frac{\partial}{\partial x} + r\cos(\theta)\,\frac{\partial}{\partial y} \\
&= -y\,\frac{\partial}{\partial x} + x\,\frac{\partial}{\partial y}
\end{aligned}$$

Lecture 14 - Covectors

March 11, 2009

1 Covectors

To each point p of a space M is associated a tangent space, T_pM, which is a vector space. From our study of vector spaces, we know that for each of the tangent spaces T_pM there exists a dual space, called T_p^*M. This is an entirely abstract construction however.

It is possible to determine the nature of the dual space directly. We begin by defining the d-operator: if f is a function on M and $X \in T_pM$ is a vector, we can define the action of f on X by

$$df(X) \quad \triangleq \quad X(f).$$

Let's see how df operators on the basis vectors $\partial/\partial x^i$:

$$df\left(\frac{\partial}{\partial x^i}\right) \;=\; \frac{\partial f}{\partial x^i}.$$

Since the coordinates $\{x^i\}$ are functions, it makes sense to apply d to them as well:

$$dx^i\left(\frac{\partial}{\partial x^j}\right) \quad\triangleq\quad \frac{\partial x^i}{\partial x^j} \;=\; \delta^i{}_j.$$

Since

$$\left(\frac{\partial f}{\partial x^j}dx^j\right)\left(\frac{\partial}{\partial x^i}\right) \;=\; \frac{\partial f}{\partial x^j}\frac{\partial x^j}{\partial x^i} \;=\; \frac{\partial f}{\partial x^j}\delta^j{}_i \;=\; \frac{\partial f}{\partial x^i},$$

we can write

$$df \;=\; \frac{\partial f}{\partial x^j}\,dx^j.$$

Thus clearly $dx^1, \dots, dx^n$ is the basis dual to $\partial/\partial x^1, \dots, \partial/\partial x^n$. Thus we can take

$$T_p^*M \;=\; span\{dx^1, \dots, dx^n\}.$$

Lecture 15 - Tensor Fields and the Metric tensor

March 13, 2009

1 Tensor fields

Let M be a space. One may define vector fields, covector fields, and, more generally, tensor fields on M.

A vector field is the assignment of a vector to each point of M; likewise a covector field is the assignment of a covector to each point of M. For example, if M is Euclidean 2-space with standard x-y coordinates, then

$$X \;=\; X(x,y) \;=\; -y\,\frac{\partial}{\partial x} \;+\; (x\,-\,y^2)\,\frac{\partial}{\partial y}$$

is a vector field, and

$$\omega \;=\; \omega(x,y) \;=\; (x^2y\,-\,x)\,dx\;-\;xy\,dy$$

is a covector field.

There is no obstruction to having fields of higher order tensors. For instance

$$T \;=\; T^i{}_j{}^k\,\frac{\partial}{\partial x^i}\otimes dx^j\otimes\frac{\partial}{\partial x^k},$$

where each $T^i{}_j{}^k \;=\; T^i{}_j{}^k(x^1,\dots,x^n)$ is a function of the coordinates $\{x^1,\dots,x^n\}$.

2 The metric tensor

The most important tensor is the metric tensor. A metric on M is the assignment of an inner product to each tangent space T_pM of M. A metric gives a space its notion of distance. The *length* or *magnitude* of a vector $v\in T_pM$ is defined to be

$$|v| \;=\; \sqrt{g(v,v)}.$$

If $\gamma : [a, b] \to M$ is a path parameterized by τ (ie, $\gamma = \gamma(\tau)$, $a \le \tau \le b$), the vector tangent to γ is

$$\frac{d}{d\tau} \;=\; \frac{dx^i}{d\tau}\frac{\partial}{\partial x^i}.$$

The speed of γ is given by

$$\left|\frac{d}{d\tau}\right| \;=\; \sqrt{g\left(\frac{d}{d\tau}, \frac{d}{d\tau}\right)} \;=\; \sqrt{g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau}}.$$

The length of the path γ for $\tau \in [a, b]$ is given by

$$L_a^b(\gamma) \;=\; \int_a^b \left|\frac{d}{d\tau}\right| d\tau.$$

Example Let M be Euclidean 2-space with standard x-y coordinates. Define a tensor field g by

$$g \;=\; \frac{4}{(1 + x^2 + y^2)^2}\, dx \otimes dx \;+\; \frac{4}{(1 + x^2 + y^2)^2}\, dy \otimes dy.$$

It is simple to check that g is an inner product at each point of M (that is, it is symmetric and nondegenerate at each point). This metric is explored in the homework.

Example Let M be the following subset of $\mathbb{R}^2$:

$$M \;=\; \{(x, y) \,|\, x^2 + y^2 \;<\; 1\,\}.$$

Namely M is the interior of the unit ball. Let g be a metric defined on M by

$$g \;=\; \frac{4}{(1 - x^2 - y^2)^2}\, dx \otimes dx \;+\; \frac{4}{(1 - x^2 - y^2)^2}\, dy \otimes dy.$$

Note that g "blows up" (goes to infinity) on the boundary of M, so in particular it cannot be (continuously) continued beyond M. Let

$$\gamma(\tau) \;=\; (\tau, \tau), \quad 0 \le \tau \le b$$

be a path in M (it must be that $0 < b < 1/\sqrt{2}$ for the path to remain in M). The length of γ is

$$\begin{aligned}
\frac{d}{d\tau} &= \frac{dx}{d\tau}\frac{\partial}{\partial x} + \frac{dy}{d\tau}\frac{\partial}{\partial y} \;=\; \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \\
L_0^b(\gamma) &= \int_0^b \left|\frac{d}{d\tau}\right| d\tau \;=\; \int_0^b \left|\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right| d\tau \;=\; \int_0^b \sqrt{\frac{8}{(1 - x^2 - y^2)^2}}\, d\tau \\
&= \int_0^b \frac{2\sqrt{2}}{1 - 2\tau^2}\, d\tau \;=\; 2\tanh^{-1}(\sqrt{2}\tau)\Big|_0^b \;=\; 2\tanh^{-1}(\sqrt{2}b).
\end{aligned}$$

Notice that the pathlength $L_0^b(\gamma)$ approaches ∞ as b approaches $1/\sqrt{2}$, as expected.

Lecture 16 - Lie brackets and the d-operator

March 16, 2009

1 The Lie Bracket

Let X and Y be vector fields on a space M. We define the *Lie bracket* (sometimes called the *commutator* or just *the bracket*) $[X, Y]$ to be the operator

$$[X,\, Y](f) \;=\; X(Y(f)) \;-\; Y(X(f)).$$

Usually we just write $[X, Y] = XY - YX$. As it turns out, the bracket of two vector fields is again a vector field, meaning it is a first-order differential operator. In components, letting $X = X^i \frac{\partial}{\partial x^i}$ and $Y \;=\; Y^i \frac{\partial}{\partial x^i}$, we have

$$\begin{aligned}
[X,\, Y](f) \;&=\; X(Y(f)) \;-\; Y(X(f)) \\
&=\; X^i \frac{\partial}{\partial x^i}\left(Y^j \frac{\partial f}{\partial x^j}\right) \;-\; Y^i \frac{\partial}{\partial x^i}\left(X^j \frac{\partial f}{\partial x^j}\right) \\
&=\; X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial f}{\partial x^j} \;+\; X^i Y^j \frac{\partial^2 f}{\partial x^i \partial x^j} \;-\; Y^i \frac{\partial X^j}{\partial x^i}\frac{\partial f}{\partial x^j} \;-\; Y^i X^j \frac{\partial^2 f}{\partial x^i \partial x^j}.
\end{aligned}$$

By the commutativity of second partial derivatives, the two terms with second partials cancel out. We are left with

$$\begin{aligned}
[X,\, Y](f) \;&=\; \left(X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial}{\partial x^j} \;-\; Y^i \frac{\partial X^j}{\partial x^i}\frac{\partial}{\partial x^j}\right)(f) \\
&=\; \left(X(Y^j)\frac{\partial}{\partial x^j} \;-\; Y(X^j)\frac{\partial}{\partial x^j}\right)(f)
\end{aligned}$$

Thus $[X, Y]$ is the vector field:

$$[X,\, Y] \;=\; X^i \frac{\partial Y^j}{\partial x^i}\frac{\partial}{\partial x^j} \;-\; Y^i \frac{\partial X^j}{\partial x^i}\frac{\partial}{\partial x^j} \;=\; X(Y^j)\frac{\partial}{\partial x^j} \;-\; Y(X^j)\frac{\partial}{\partial x^j}$$

In a sense, the bracket is a kind of derivative operation, which measures how two vector fields mutually change with respect to one another.

2 Forms

Given a space M, let $\Omega^p(M)$ denote the collection of all fields of alternating p-tensors. An alternating p-tensor field is known as a *form*. For example,

$$\begin{array}{rcll}
f & = & f(x^1,\ldots,x^n) & \text{is a } 0-\text{form} \\
\omega & = & \omega_i dx^i & \text{where} \quad \omega_i = \omega_i(x^1,\ldots,x^n) \quad \text{is a } 1{-}\text{form} \\
\eta & = & \eta_{ij} dx^i \wedge dx^j & \text{where} \quad \eta_{ij} = \eta_{ij}(x^1,\ldots,x^n) \quad \text{is a } 2{-}\text{form} \\
& & etc. &
\end{array}$$

Example Let $\{x^1, x^2, x^3\}$ be coordinates on the 3-dimensional space M. Let

$$\omega \;=\; \left((x^1)^2 - x^2\right) dx^1 \wedge dx^2 + (x^1 - x^2)dx^1 \wedge dx^3 + (x^1x^3)dx^2 \wedge dx^3$$

be a 2-form, and let

$$X \;=\; x^1x^3\frac{\partial}{\partial x^1} \;-\; x^2\frac{\partial}{\partial x^3} \quad \text{and} \quad Y \;=\; (x^2)^2\frac{\partial}{\partial x^2}$$

be vector fields. Compute $\omega(X,Y)$.

Solution

$$\begin{aligned}
\omega(X,Y) \;&=\; \left((x^1)^2 - x^2\right) dx^1 \wedge dx^2(X,Y) \;+\; (x^1 - x^2)dx^1 \wedge dx^3(X,Y) \;+\; (x^1x^3)dx^2 \wedge dx^3(X,Y) \\
&=\; \left((x^1)^2 - x^2\right)\cdot\frac{1}{2}\cdot\left(x^1x^3(x^2)^2 \;-\; 0\right) \;+\; (x^1 - x^2)\cdot\frac{1}{2}\cdot(0 \;-\; 0) \;+\; (x^1x^3)\cdot\frac{1}{2}\cdot\left(0 \;+\; (x^2)^3\right) \\
&=\; \frac{1}{2}(x^1)^3(x^2)^2x^3 \;-\; \frac{1}{2}x^1(x^2)^3x^3 \;+\; \frac{1}{2}x^1(x^2)^3x^3
\end{aligned}$$

□

3 The d-operator

Given a function f, we have defined df to be the covector given by

$$df(X) \;=\; X(f).$$

If $\{x^i\}$ are coordinates on M, then we can express

$$df \;=\; \frac{\partial f}{\partial x^i}\, dx^i.$$

Since f is a 0-form and df is a 1-form, we can consider d to be a map

$$d : \Omega^0(M) \to \Omega^1(M).$$

It is possible to extend the d-operator to an operation

$$d : \Omega^p(M) \to \Omega^{p+1}(M)$$

as follows. In the $\{x^i\}$ coordinate system, any decomposable p-form ω can be written

$$\omega \;=\; f\, dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}.$$

Then we define $d\omega$ to be the $(p+1)$-form

$$\begin{aligned} d\omega &= df \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} \\ &= \frac{\partial f}{\partial x^j}\, dx^j \wedge dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p}. \end{aligned}$$

Ostensibly the definition of $d\omega$ depends on the coordinates used. If we had used different coordinates, say $\{y^i\}$ coordinates, how do we know we would end up with the same tensor?

As a point of fact, the definition of the d-operator is independent of the coordinate system, and is therefore and intrinsic operator. This is proved in the homework.

<u>Example</u> Consider the two coordinate systems $\{r, \theta\}$ (polar coordinates) and $\{x, y\}$ (rectangular coordinates) on Euclidean 2-space, where (as usual) $r^2 = x^2 + y^2$, $\tan\theta = y/x$. Let

$$\omega \;=\; r d\theta$$

be a 1-form. Simple computations give

$$\begin{aligned} dr &= \frac{\partial r}{\partial x}dx + \frac{\partial r}{\partial y}dy & \qquad d\theta &= \frac{\partial \theta}{\partial x}dx + \frac{\partial \theta}{\partial y}dy \\ dr &= \frac{x}{\sqrt{x^2+y^2}}dx + \frac{x}{\sqrt{x^2+y^2}}dy & \qquad d\theta &= \frac{-y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy \end{aligned}$$

and

$$\begin{aligned} dx &= \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta & \qquad dy &= \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta \\ dx &= \cos(\theta)\, dr - r\, \sin(\theta)\, d\theta & \qquad dy &= \sin(\theta)\, dr + r\, \cos(\theta)\, d\theta. \end{aligned}$$

Therefore we can express ω in rectangular coordinates as

$$\omega \quad = \quad r\, d\theta \quad = \quad \frac{-y}{\sqrt{x^2+y^2}}\, dx + \frac{x}{\sqrt{x^2+y^2}}\, dy$$

Prove that $d\omega$ is the same 2-form, regardless of how it is computed

<u>Solution</u> We compute $d\omega$ in the two different coordinate systems.
In the $\{r, \theta\}$ coordinate system:

$$\begin{aligned} d\omega &= d(r\, d\theta) \;=\; \frac{\partial r}{\partial r}dr \wedge d\theta + \frac{\partial r}{\partial \theta}d\theta \wedge d\theta \\ &= dr \wedge d\theta. \end{aligned}$$

In the $\{x, y\}$ coordinate system:

$$\begin{aligned}
d\omega &= d\left(\frac{-y}{\sqrt{x^2 + y^2}}\,dx \;+\; \frac{x}{\sqrt{x^2 + y^2}}\,dy\right) \\
&= \frac{\partial}{\partial y}\left(-y\left(x^2 + y^2\right)^{-1/2}\right) dy \wedge dx \;+\; \frac{\partial}{\partial x}\left(x\left(x^2 + y^2\right)^{-1/2}\right) dx \wedge dy \\
&= \left(-\left(x^2+y^2\right)^{-1/2} \;+\; y^2\left(x^2+y^2\right)^{-3/2}\right) dy \wedge dx \;+\; \left(\left(x^2+y^2\right)^{-1/2} \;-\; x^2\left(x^2+y^2\right)^{-3/2}\right) dx \wedge dy \\
&= \frac{1}{\sqrt{x^2+y^2}}\,dx \wedge dy.
\end{aligned}$$

Now, is $dr \wedge d\theta$ the same tensor as $(x^2+y^2)^{-1/2}dx \wedge dy$? We compute:

$$\begin{aligned}
dr \wedge d\theta &= \left(\frac{x}{\sqrt{x^2 + y^2}}dx \;+\; \frac{x}{\sqrt{x^2 + y^2}}dy\right) \wedge \left(\frac{-y}{x^2 + y^2}dx \;+\; \frac{x}{x^2 + y^2}dy\right) \\
&= \frac{x^2}{\left(x^2+y^2\right)^{3/2}}\,dx \wedge dy \;-\; \frac{y^2}{\left(x^2+y^2\right)^{3/2}}\,dy \wedge dx \\
&= \frac{1}{\sqrt{x^2+y^2}}\,dx \wedge dy.
\end{aligned}$$

Therefore, the tensor $d\omega$ is the same whether computed in the $\{r, \theta\}$ coordinate system of the $\{x, y\}$ coordinate system.

4 Four properties of the d-operator

In this section we demonstrate that the d-operator (which is defined in terms of a given coordinate system $\{x^i\}$) satisfies four properties. Assuming $\omega, \eta \in \Omega^*(M)$ and $f \in \Omega^0(M)$, then

- $d(\omega + \eta) \;=\; d\omega + d\eta$ (linearity)
- $df(X) \;=\; X(f)$
- $ddf \;\equiv\; 0$
- $d(\omega \wedge \eta) \;=\; (d\omega) \wedge \eta \;+\; (-1)^{|\omega|}\omega \wedge (d\eta)$ (generalized Leibnitz rule).

Now to prove these conditions: the first condition is implicit in the definition of $d : \Omega^p(M) \to \Omega^{p+1}(M)$; the second is the explicit definition of $d : \Omega^0(M) \to \Omega^1(M)$.

To prove that $ddf = 0$, we use two facts: 1-forms anticommute and second partial derivatives commute. In coordinates, we get

$$ddf \;=\; d\left(\frac{\partial f}{\partial x^i}\,dx^i\right) \;=\; \frac{\partial^2 f}{\partial x^j \partial x^i}\,dx^j \wedge dx^i.$$

Now we compute

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^j \partial x^i}\, dx^j \wedge dx^i \quad &= \quad \frac{\partial^2 f}{\partial x^i \partial x^j}\, dx^j \wedge dx^i \qquad \text{partials commute} \\
&= \quad -\frac{\partial^2 f}{\partial x^i \partial x^j}\, dx^i \wedge dx^j \qquad 1-\text{forms anticommute} \\
&= \quad -\frac{\partial^2 f}{\partial x^j \partial x^i}\, dx^j \wedge dx^i \qquad \text{relabel } i \text{ and } j
\end{aligned}$$

Thus $ddf = -ddf$ so therefore $ddf = 0$.

Next we prove the Leibniz rule. First note that the Leibniz rule is just the product rule, when dealing with 1-forms:

$$d(fg) \quad = \quad \frac{\partial(fg)}{\partial x^i} dx^i \;=\; g\frac{\partial f}{\partial x^i} dx^i \;+\; f\frac{\partial g}{\partial x^j} dx^j \;=\; g\,df \;+\; f\,dg.$$

By the linearity condition, without loss of generality we can assume ω and η are decomposable. Let $\omega \in \Omega^p(M)$ and $\eta \in \Omega^q(M)$ be given by

$$\begin{aligned}
\omega \;&=\; f\, dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_p} \\
\eta \;&=\; g\, dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_q}
\end{aligned}$$

then

$$\begin{aligned}
d(\omega \wedge \eta) \quad &= \quad d\left(fg\, dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}\right) \\
&= \quad g\, df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q} \;+\; f\, dg \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q} \\
&= \quad df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge (g dx^{j_1}) \wedge \cdots \wedge dx^{j_q} \;+\; (-1)^p\, f\, dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dg \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q} \\
&= \quad d\omega \wedge d\eta \;+\; (-1)^p \omega \wedge d\eta.
\end{aligned}$$

This completes the proof that the above four properties hold for the d-operator. In the homework, it is proven that these properties in fact characterize the d-operator completely. Since these properties are stated independent of any coordinates, the d-operator is therefore also independent of coordinates, despite the fact that the original definition appears to be coordinate-dependent.

Lecture 17 - Relation between the classical vector operations and the d-operator. Also 4-velocity and 3-velocity, and 4-momentum.

March 18, 2009

1 Classical Vector Analysis

In classical 3-dimensional vector analysis, we always work with standard orthonormal coordinates (x, y, z) (never polar or cylindrical coordinates) and we consider a *vector* at a point p to be an assignment of a list of 3 numbers to p. A vector field is an assignment of a list of 3 numbers to each point of a region of space. Typically vector fields are denoted in capitals: X, Y, W, etc.

Two algebraic operations

Let $X = (a, b, c)$ and $Y = (\alpha, \beta, \gamma)$ be two vectors based at a point p. We define their *inner product* and their *cross product* to be

$$\begin{aligned}
\langle X, Y\rangle &= a\alpha + b\beta + c\gamma \\
&= |X||Y|\cos\theta \\
X \times Y &= (b\gamma - c\beta,\ c\alpha - a\gamma,\ a\beta - b\alpha) \\
&= \hat{n}|X||Y|\sin\theta,
\end{aligned}$$

where θ is the angle between x and Y, and $\hat{n}$ is the unique unit vector normal to both X and Y obtained using the right-hand rule.

Three analytic operations

Let $X = (a, b, c)$ be a vector, and let f be a function. We define the *gradient* of f to be the vector field

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right).$$

It is interpreted to be the vector whose direction at each point is that of maximum increase of the function f, and whose length indicates the rate of that increase.

We define the *divergence* of the vector field X to be the function

$$\nabla \cdot X \;=\; \frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}.$$

The divergence of a vector field at a point is interpreted to be its tenancy to rarify or accumulate at that point.

We define the *curl* of the vector field X to be the vector field

$$\nabla \times X \quad = \quad \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}, \frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}, \frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} \right).$$

The curl of a vector field at a point is interpreted to be the tenancy of the vector field to rotate about a point, where the direction of $\nabla \times X$ is the axis of rotation (according to the right-hand rule) and the length of $\nabla \times X$ is the magnitude of the rotation.

Two identities

The following two identities are theorems of classical analysis:

$$\begin{aligned} \nabla \times (\nabla f) &= 0 \\ \nabla \cdot (\nabla \times X) &= 0. \end{aligned}$$

2 The d-operator and the classical vector operations

The classical vector operations all have equivalents in the language of forms. The three analytic operations can be recovered using the d-operator in various ways.

The dot product

This is just the inner product: if $X = X^i \frac{\partial}{\partial x^i}$ and $Y = Y^i \frac{\partial}{\partial x^i}$, then

$$\langle X, Y \rangle \;=\; g_{ij} X^i Y^j.$$

Note that if $g_{ij} = \delta_{ij}$ is the Euclidean inner product, then this is the classical dot product.

The cross product

Let $X = adx^1 + bdy + cdz$ and $Y = \alpha dx + \beta dy + \gamma dz$ be 1-forms. Then $X \wedge Y$ is the 2-form given by

$$\begin{aligned} X \wedge Y &= (adx + bdy + cdz) \wedge (\alpha dx + \beta dy + \gamma dz) \\ &= (b\gamma - c\beta)\, dy \wedge dz \;+\; (a\gamma - c\alpha)\, dz \wedge dx \;+\; (a\beta - b\alpha)\, dx \wedge dy \end{aligned}$$

The gradient

Let f be a 0-form (a function). Then

$$df \;=\; \frac{df}{dx}dx \;+\; \frac{df}{dy}dy \;+\; \frac{df}{dz}dz.$$

The curl

Let $\omega = adx + bdy + xdz$ be a 1-form. Then by the definition of $d\omega$ we have

$$\begin{aligned}
d\omega &= d(a) \wedge dx + d(b) \wedge dy + d(c) \wedge dz \\
&= \frac{da}{dx} dx \wedge dx + \frac{da}{dy} dy \wedge dx + \frac{da}{dz} dz \wedge dx \\
&\quad + \frac{db}{dx} dx \wedge dy + \frac{db}{dy} dy \wedge dy + \frac{db}{dz} dz \wedge dy \\
&\quad + \frac{dc}{dx} dx \wedge dz + \frac{dc}{dy} dy \wedge dz + \frac{dc}{dz} dz \wedge dz \\
&= \left(\frac{dc}{dy} - \frac{db}{dz} \right) dy \wedge dz + \left(\frac{da}{dz} - \frac{dc}{dx} \right) dz \wedge dx + \left(\frac{db}{dx} - \frac{da}{dy} \right) dx \wedge dy.
\end{aligned}$$

The divergence

Let $\omega = ady \wedge dz + bdz \wedge dx + cdx \wedge dy$ be a 2-form. We compute $d\omega$:

$$\begin{aligned}
d\omega &= \frac{da}{dx} dx \wedge dy \wedge dz + \frac{da}{dy} dy \wedge dy \wedge dz + \frac{da}{dz} dz \wedge dy \wedge dz \\
&\quad + \frac{db}{dx} dx \wedge dz \wedge dx + \frac{db}{dy} dy \wedge dz \wedge dx + \frac{db}{dz} dz \wedge dz \wedge dx \\
&\quad + \frac{dc}{dz} dz \wedge dx \wedge dy + \frac{dc}{dy} dy \wedge dx \wedge dy + \frac{dc}{dz} dz \wedge dx \wedge dy \\
&= \left(\frac{da}{dx} + \frac{db}{dy} + \frac{dc}{dz} \right) dx \wedge dy \wedge dz.
\end{aligned}$$

Two identities in one

In our new notation, the two vector identities $\nabla \times \nabla f = 0$ and $\nabla \cdot (\nabla \times X) = 0$ coalesce into the single fact that $dd = 0$.

Given a function f, the 1-form df is equivalent to the classical gradient. Given a 1-form ω, the 2-form $d\omega$ is equivalent to the classical curl. This means ddf is the classical analog of the curl of the gradient, and the fact that $ddf = 0$ is equivalent to the classical theorem $\nabla \times \nabla f = 0$.

Given a 2-form η, the 3-form $d\eta$ is equivalent to the classical divergence. If ω is a 1-form, then $dd\omega$ is equivalent to the divergence of the curl, and the fact that $dd\omega = 0$ is equivalent to the classical theorem $\nabla \cdot (\nabla \times X) = 0$.

3 A comment on notation

In the context of Minkowski space, indices i, j, k, etc will sum from 0 to 3. Indices a, b, c, etc will sum from 1 to 3 (ie, only over the space-dimensions, not time). For instance, in coordiantes a 4-vector v will be denoted

$$v = v^i \frac{\partial}{\partial x^i} = v^0 \frac{\partial}{\partial x^0} + v^1 \frac{\partial}{\partial x^1} + v^2 \frac{\partial}{\partial x^2} + v^3 \frac{\partial}{\partial x^3},$$

and a 3-vector $\vec{v}$ will be denoted

$$\vec{v} = \vec{v}^a \frac{\partial}{\partial x^a} = \vec{v}^1 \frac{\partial}{\partial x^1} + \vec{v}^2 \frac{\partial}{\partial x^2} + \vec{v}^3 \frac{\partial}{\partial x^3}.$$

4 4-velocity and 3-velocity

Let $\gamma(\tau)$ be a path through space-time, parameterized by τ. Typically γ represents a particle's *worldline*, its path through space-time. Given coordinates $\{x^i\}$, the velocity vector of γ is

$$v = \frac{d}{d\tau} = \frac{dx^i}{d\tau} \frac{\partial}{\partial x^i}.$$

If the coordinates $\{x^i\}$ constitute an inertial reference frame, we can define the particle's 3-velocity:

$$\begin{aligned} \vec{v} &= \frac{dx^a}{dx^0} \frac{\partial}{\partial x^a} \\ &= \frac{dx^1}{dx^0} \frac{\partial}{\partial x^1} + \frac{dx^2}{dx^0} \frac{\partial}{\partial x^2} + \frac{dx^3}{dx^0} \frac{\partial}{\partial x^3} \\ &= \frac{dx^1/d\tau}{dx^0/d\tau} \frac{\partial}{\partial x^1} + \frac{dx^2/d\tau}{dx^0/d\tau} \frac{\partial}{\partial x^2} + \frac{dx^3/d\tau}{dx^0/d\tau} \frac{\partial}{\partial x^3} \end{aligned}$$

This indicates the rate of change of the particle's position with respect to coordinate time.

Recall the Minkowski metric:

$$g_{11} = 1, \quad g_{22} = -\frac{1}{c^2}, \quad g_{22} = -\frac{1}{c^2}, \quad g_{33} = -\frac{1}{c^2}$$

Given a 4-vector $v = v^i \frac{\partial}{\partial x^i}$, we define its norm-square is

$$|v|^2 = g_{ij} v^i v^j = v^0 v^0 - \frac{1}{c^2} v^1 v^1 - \frac{1}{c^2} v^2 v^2 - \frac{1}{c^2} v^3 v^3$$

(recall that this can be positive, negative, or zero).

We will indicate 3-vectors by using an over-arrow: $\vec{v}$ means $\vec{v}$ is a 3-vector. If $\vec{v}$ is a 3-tensor, we define its norm-square to be the Euclidean norm-square:

$$|\vec{v}|^2 = \delta_{ab} v^a v^b = v^1 v^1 + \ldots + v^n v^n.$$

We impose the following *physicality condition* on timelike paths:

- If $\gamma(\tau)$ represents the path of a massive particle through space-time, its velocity vector $v = \frac{d}{d\tau}$ is time-like and $|v|^2 = +1$.

Using the physicality condition, we can compute the component $\frac{dx^0}{d\tau}$ for time-like paths:

$$\begin{aligned}
1 &= \left|\frac{d}{d\tau}\right|^2 \\
&= \left(\frac{dx^0}{d\tau}\right)^2 - \frac{1}{c^2}\left(\frac{dx^1}{d\tau}\right)^2 - \frac{1}{c^2}\left(\frac{dx^2}{d\tau}\right)^2 - \frac{1}{c^2}\left(\frac{dx^3}{d\tau}\right)^2 \\
\frac{1}{\left(\frac{dx^0}{d\tau}\right)^2} &= 1 - \frac{1}{c^2}\left(\frac{dx^1/d\tau}{dx^0/d\tau}\right) - \frac{1}{c^2}\left(\frac{dx^2/d\tau}{dx^0/d\tau}\right) - \frac{1}{c^2}\left(\frac{dx^3/d\tau}{dx^0/d\tau}\right) \\
&= 1 - \frac{1}{c^2}\left(\frac{dx^1}{dx^0}\right) - \frac{1}{c^2}\left(\frac{dx^2}{dx^0}\right) - \frac{1}{c^2}\left(\frac{dx^3}{dx^0}\right) \\
\left(\frac{dx^0}{d\tau}\right)^2 &= \frac{1}{1 - \frac{|\vec{v}|^2}{c^2}} \\
\frac{dx^0}{d\tau} &= \gamma_{|\vec{v}|}
\end{aligned}$$

Using this, we have

$$\begin{aligned}
v &= v^i \frac{\partial}{\partial x^i} \\
\vec{v} &= \vec{v}^a \frac{\partial}{\partial x^a}
\end{aligned}$$

where $v^0 = \gamma$, $v^a = \gamma \vec{v}^a$. A shorthand way to write this is to use "classical" vector notation:

$$\begin{aligned}
\vec{v} &\triangleq \left(\frac{dx^1}{dx^0}, \ldots, \frac{dx^n}{dx^0}\right) = \left(\vec{v}^1, \ldots, \vec{v}^n\right) \\
v &\triangleq \left(\frac{dx^0}{d\tau}, \frac{dx^1}{d\tau}, \ldots, \frac{dx^n}{d\tau}\right) \\
&= \left(v^0, v^1, \ldots, v^n\right) = \left(\gamma, \gamma\vec{v}^1, \ldots, \gamma\vec{v}^n\right) \\
&= (\gamma, \gamma\vec{v}).
\end{aligned}$$

5 Momentum

If $v = v^i = v^i \frac{\partial}{\partial x^i}$ is a velocity vector, we define the corresponding momentum (or *conjugate momentum*) covector to be

$$p = -mc^2 v_\flat,$$

or, in components,

$$p_i \quad = \quad -mc^2 g_{ij} v^j.$$

Recall the Minkowski metric

$$g_{00} \;=\; 1 \quad g_{11} \;=\; -\frac{1}{c^2} \quad g_{22} \;=\; -\frac{1}{c^2} \quad g_{33} \;=\; -\frac{1}{c^2}.$$

Then we have

$$\begin{aligned} v &= v^0 \frac{\partial}{\partial x^0} + v^1 \frac{\partial}{\partial x^1} + v^2 \frac{\partial}{\partial x^2} + v^3 \frac{\partial}{\partial x^3} \\ p &= -mc^2 v^0 dx^0 + mv^1 dx^1 + mv^2 dx^2 + mv^3 dx^3. \end{aligned}$$

Using vector notation

$$\begin{aligned} v &= (\gamma, \gamma\vec{v})_{\frac{\partial}{\partial x^i}} \\ p &= -mc^2 \left(\gamma, -\frac{1}{c^2}\gamma\,\vec{v}\right)_{dx^i} \\ &= \left(-m\,c^2\,\gamma, m\,\gamma\,\vec{v}\right)_{dx^i} = \left(-mc^2\gamma, \vec{p}\right). \end{aligned}$$

Note the space-components:

$$\vec{p} \;=\; m\gamma\vec{v}.$$

This closely resembles the classical 3-momentum $\vec{p} \;=\; m\vec{v}$.

Lecture 18 - Conservation of energy-momentum and the Einstein equation

March 20, 2009

1 Taylor expansion of γ

Let $\gamma = \gamma(v) = \left(1 - \frac{v^2}{c^2}\right)^{-1/2}$. This is a fairly complicated expression in terms of v, involving a square root and an inversion. A good way to simplify any function into a more workable form is to expand it into its Taylor series. Using the formula

$$\gamma(v) = \gamma(0) + \gamma'(0)v + \frac{1}{2!}\gamma''(0)v^2 + \ldots + \frac{1}{i!}\gamma^{(i)}v^i + \ldots$$

and the computations

$$\begin{aligned}
\gamma'(v) &= \frac{v}{c^2}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{3}{2}} \\
\gamma''(v) &= \frac{1}{c^2}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{3}{2}} + 3\frac{v^2}{c^4}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{5}{2}} \\
\gamma'''(v) &= 9\frac{v}{c^4}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{5}{2}} + 15\frac{v^3}{c^6}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{7}{2}} \\
\gamma''''(v) &= \frac{9}{c^4}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{5}{2}} + 90\frac{v^2}{c^6}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{7}{2}} + 105\frac{v^4}{c^8}\left(1 - \frac{v^2}{c^2}\right)^{-\frac{9}{2}} \\
& \quad etc
\end{aligned}$$

gives

$$\gamma(v) = 1 + \frac{1}{2}\frac{v^2}{c^2} + \frac{9}{24}\frac{v^4}{c^4} + \ldots.$$

2 The Relativistic Conservation Law

The Law of Conservation of Energy-Momentum states that the sum of all momenta remains constant in time.

For instance, assume one has k many particles, with momenta $p^{(1)}, \ldots, p^{(k)}$. After an interaction, there are l many particles, with momenta $\tilde{p}^{(1)}, \ldots, \tilde{p}^{(l)}$. The conservation law states that

$$\sum_{i=1}^{k} p^{(i)} \;=\; \sum_{i=1}^{l} \tilde{p}^{(i)}.$$

3 Energy

Letting $\gamma(\tau)$ be a particle's path through 4-space, its velocity is $v = \frac{\partial}{\partial \tau}$, and we have defined its momentum to be

$$p \;=\; -mc^2 v_\flat.$$

In components, we have

$$\begin{aligned} p &= -mc^2 \frac{dx^0}{d\tau} dx^0 \,+\, m\frac{dx^1}{d\tau}dx^1 \,+\, m\frac{dx^2}{d\tau}dx^2 \,+\, m\frac{dx^3}{d\tau}dx^3 \\ &= -mc^2\gamma dx^0 \,+\, m\gamma\vec{v}^1 dx^1 \,+\, m\gamma\vec{v}^2 dx^2 \,+\, m\gamma\vec{v}^3 dx^3 \\ &= \left(-mc^2\gamma,\, m\gamma\vec{v}\right). \end{aligned}$$

So the *3-momentum* can be written

$$\vec{p} \;=\; m\gamma\vec{v},$$

which approximates the classical momentum at speeds $v << c$.

So much for the p_1, p_2, p_3 components. How do we interpret the p_0 component? Using the Taylor series from above, we get

$$\begin{aligned} -p_0 &= mc^2\gamma \\ &= mc^2 \,+\, \frac{1}{2}m|\vec{v}|^2 \,+\, \frac{9}{24}m\frac{|\vec{v}|^4}{c^2} \,+\, \ldots. \end{aligned}$$

In the classical limit where $v << c$, all higher terms are negligible and we have

$$-p_0 \;\approx\; mc^2 \,+\, \frac{1}{2}m|\vec{v}|^2.$$

This is the classical kinetic energy plus the term mc^2. Following Einstein, we interpret p^0 as the particle's energy

$$E \;=\; -p_0.$$

Thus, the 4-momentum covector has components

$$\begin{aligned} p &= \left(-mc^2\gamma,\, m\gamma\vec{v}\right) \\ &= (-E,\, \vec{p}) \end{aligned}$$

where $E = mc^2\gamma$ is the particle's energy and $\vec{p} = m\gamma\vec{v}$ is the particle's 3-momentum.

4 The Einstein Equation

Using the physicality condition $|v|^2 = |v_\flat|^2 = 1$, we get

$$|p|^2 \;=\; |-mc^2 v_\flat|^2 \;=\; m^2c^4.$$

But on the other hand,

$$|p|^2 \;=\; |(-E, \vec{p})|^2 \;=\; E^2 \;-\; |\vec{p}|^2 c^2,$$

so we get the *Einstein equation*:

$$m^2c^4 \;=\; E^2 \;-\; |\vec{p}|^2 c^2.$$

This is the relativistic relationship between mass, energy, and momentum.

If the 3-velocity $\vec{v}$ is zero, this reads $E = mc^2$. On the other hand, in the case of massless particles (for instance photons), we have $E = |\vec{p}|c$.

Note that, in the case of light, we also have the equation $E = hf$ where f is the light's frequency.

Lecture 19 - Stereographic projection

March 25, 2009

1 The sphere

Coordinates on $\mathbb{R}^{n+1}$ will be denoted $\{x^1,\, x^2,\, \dots,\, x^n,\, x^{n+1}\}$ We define the n-sphere to be the subset of $\mathbb{R}^{n+1}$ given by

$$\mathbb{S}^n \quad = \quad \{\, (x^1,\, x^2,\, \dots,\, x^n) \;|\; (x^1)^2 \,+\, (x^2)^2 \,+\, \dots \,+\, (x^{n+1})^2 \,=\, 1\,\}.$$

2 Stereographic projection $\mathbb{S}^n \to \mathbb{R}^n$

We consider $\mathbb{R}^n \subset \mathbb{R}^{n+1}$ to be the plane given by $x^{n+1} = 0$. For convenience, we will let $(x^1, \dots, x^n, x^{n+1})$ be coordinates on $\mathbb{R}^{n+1}$ and $(\xi^1, \dots, \xi^n\}$ be coordinates on $\mathbb{R}^n \subset \mathbb{R}^{n+1}$.

We define the map $\phi : \mathbb{S}^n \to \mathbb{R}^n$ to be

$$\phi\left(x^1, \dots, x^n, x^{n+1}\right) \quad = \quad \left(\frac{x^1}{1 - x^{n+1}},\, \dots,\, \frac{x^n}{1 - x^{n+1}}\right).$$

This is defined only for $(x^1)^2 + \cdots + (x^{n+1})^2 = 0$, and when $x^{n+1} \neq 1$. The inverse map is given by

$$\phi^{-1}\left(\xi^1, \dots, \xi^n\right) \quad = \quad \left(\frac{2\xi^1}{(\xi^1)^1 + \cdots + (\xi^n)^2 + 1},\, \dots,\, \frac{2\xi^n}{(\xi^1)^1 + \cdots + (\xi^n)^2 + 1},\, \frac{(\xi^1)^2 + \cdots + (\xi^n)^2 \,-\, 1}{(\xi^1)^2 + \cdots + (\xi^n)^2 \,+\, 1}\right).$$

This sets us an association between coordinates $(x^1, \dots, x^n, x^{n+1})$ on $\mathbb{S}^n$ and coordinates $(\xi^1, \dots, \xi^n)$ on $\mathbb{R}^n$, given by

$$\begin{aligned}
\xi^1 \quad &= \quad \frac{x^1}{1 \,-\, x^{n+1}} \\
&\vdots \\
\xi^n \quad &= \quad \frac{x^n}{1 \,-\, x^{n+1}}
\end{aligned}$$

and

$$
\begin{aligned}
x^1 &= \frac{2\xi^1}{(\xi^1)^2 + \ldots + (\xi^n)^2 + 1} \\
&\vdots \\
x^n &= \frac{2\xi^n}{(\xi^1)^2 + \ldots + (\xi^n)^2 + 1} \\
x^{n+1} &= \frac{(\xi^1)^2 + \ldots + (\xi^n)^2 - 1}{(\xi^1)^2 + \ldots + (\xi^n)^2 + 1}.
\end{aligned}
$$

3 Examples

Example: $\mathbb{S}^1 \to \mathbb{R}^1$ We use coordinates (x, y) on $\mathbb{R}^2$ and a on $\mathbb{R}^1$. The 1-sphere (aka the circle) is defined by the equation $x^2 + y^2 = 1$. We have

$$
\begin{aligned}
\phi(x, y) &= \frac{x}{1 - y} \\
\phi^{-1}(a) &= \left(\frac{2a}{a^2 + 1}, \frac{a^2 - 1}{a^2 + 1} \right).
\end{aligned}
$$

That is to say, we have the association

$$
\begin{aligned}
a &= \frac{x}{1 - y} \\
x &= \frac{2a}{a^2 + 1} \\
y &= \frac{a^2 - 1}{a^2 + 1}
\end{aligned}
$$

(one easily checks that $x^2 + y^2 = \frac{(2a)^2}{(a^2+1)^2} + \frac{(a^2-1)^2}{(a^2+1)^2} = 1$, as required).

Example: $\mathbb{S}^2 \to \mathbb{R}^2$ We use coordinates (x, y, z) on $\mathbb{R}^3$ and (a, b) on $\mathbb{R}^2$. The 2-sphere is defined by the equation $x^2 + y^2 + z^2 = 1$. We define the map $\phi : \mathbb{S}^2 \to \mathbb{R}^2$ by

$$
\begin{aligned}
\phi(x, y, z) &= \left(\frac{x}{1 - z}, \frac{y}{1 - z} \right) \\
\phi^{-1}(a, b) &= \left(\frac{2a}{a^2 + b^2 + 1}, \frac{2b}{a^2 + b^2 + 1}, \frac{a^2 + b^2 - 1}{a^2 + b^2 + 1} \right).
\end{aligned}
$$

That is to say, we have the association

$$
\begin{aligned}
a &= \frac{x}{1 - z} \qquad b = \frac{y}{1 - z} \\
x &= \frac{2a}{a^2 + b^2 + 1} \qquad y = \frac{2b}{a^2 + b^2 + 1} \qquad z = \frac{a^2 + b^2 - 1}{a^2 + b^2 + 1}
\end{aligned}
$$

(one easily checks that $x^2+y^2+z^2 = \frac{(2a)^2}{(a^2+b^2+1)^2} + \frac{(2b)^2}{(a^2+b^2+1)^2} + \frac{(a^2+b^2-1)^2}{(a^2+b^2+1)^2} = 1$, as required).

Example: $\mathbb{S}^3 \to \mathbb{R}^3$ We use coordinates (x, y, z, w) on $\mathbb{R}^4$ and (a, b, c) on $\mathbb{R}^3$. The 3-sphere $\mathbb{S}^3 \subset \mathbb{R}^4$ is defined by the equation $x^2 + y^2 + z^2 + w^2 = 1$. We define the map $\phi : \mathbb{S}^3 \to \mathbb{R}^3$ by

$$\begin{aligned} \phi(x, y, z, w) &= \left(\frac{x}{1-w}, \frac{y}{1-w}, \frac{z}{1-w}\right) \\ \phi^{-1}(a, b, c) &= \left(\frac{2a}{a^2+b^2+c^2+1}, \frac{2b}{a^2+b^2+c^2+1}, \frac{2c}{a^2+b^2+c^2+1}, \frac{a^2+b^2+c^2-1}{a^2+b^2+c^2+1}\right). \end{aligned}$$

That is to say, we have the association

$$\begin{aligned} a &= \frac{x}{1-w} \qquad b = \frac{y}{1-w} \qquad c = \frac{z}{1-w} \\ x &= \frac{2a}{a^2+b^2+c^2+1} \qquad y = \frac{2b}{a^2+b^2+c^2+1} \qquad z = \frac{2c}{a^2+b^2+c^2+1} \qquad z = \frac{a^2+b^2+c^2-1}{a^2+b^2+c^2+1}. \end{aligned}$$

4 Vectors on $\mathbb{S}^2$

From now on, we specialize to the case of the 2-sphere.

Under the association $\mathbb{S}^2 \to \mathbb{R}^2$, there is an association between vectors on $\mathbb{R}^2$ (given generically by $v^1 \frac{\partial}{\partial a} + v^2 \frac{\partial}{\partial b}$) and vectors in $\mathbb{R}^3$ tangent to $\mathbb{S}^2$. We compute

$$\begin{aligned} \frac{\partial}{\partial a} &= \frac{\partial x}{\partial a}\frac{\partial}{\partial x} + \frac{\partial y}{\partial a}\frac{\partial}{\partial y} + \frac{\partial z}{\partial a}\frac{\partial}{\partial z} \\ &= \left(1 - z - x^2\right)\frac{\partial}{\partial x} - xy\frac{\partial}{\partial y} + (x - xy)\frac{\partial}{\partial z} \\ \frac{\partial}{\partial b} &= \frac{\partial x}{\partial b}\frac{\partial}{\partial x} + \frac{\partial y}{\partial b}\frac{\partial}{\partial y} + \frac{\partial z}{\partial b}\frac{\partial}{\partial z} \\ &= -xy\frac{\partial}{\partial x} + \left(1 - z - y^2\right)\frac{\partial}{\partial y} + (y - yz)\frac{\partial}{\partial z}. \end{aligned}$$

5 The metric

The plane $\mathbb{R}^2$ serves as a representation of $\mathbb{S}^2$ under stereographic projection. How should distances be measured on $\mathbb{R}^2$ in order that they correspond to distances on $\mathbb{S}^2$?

To answer this question, we must determine the inner product on $\mathbb{R}^2$ that corresponds to the inner product on $\mathbb{S}^2$.

The geometry of $\mathbb{S}^2$ is determined by its embedding in $\mathbb{R}^3$, on which the metric is $g = dx \otimes dx + dy \otimes dy + dz \otimes dz$. Thus we compute

$$
\begin{aligned}
g\left(\frac{\partial}{\partial a}, \frac{\partial}{\partial a}\right) &= g\left(\left(1 - z - x^2\right)\frac{\partial}{\partial x} - xy\frac{\partial}{\partial y} + (x - xy)\frac{\partial}{\partial z}, \left(1 - z - x^2\right)\frac{\partial}{\partial x} - xy\frac{\partial}{\partial y} + (x - xz)\frac{\partial}{\partial z}\right) \\
&= \left(1 - z + x^2\right)^2 + (-xy)^2 + (x - xz)^2 = (1 - z)^2 = \frac{4}{a^2 + b^2 + 1} \\
g\left(\frac{\partial}{\partial a}, \frac{\partial}{\partial b}\right) &= g\left(\left(1 - z - x^2\right)\frac{\partial}{\partial x} - xy\frac{\partial}{\partial y} + (x - xy)\frac{\partial}{\partial z}, -xy\frac{\partial}{\partial x} + \left(1 - z - y^2\right)\frac{\partial}{\partial y} + (y - yz)\frac{\partial}{\partial z}\right) \\
&= \left(1 - z - x^2\right)(-xy) + (-xy)\left(1 - z - y^2\right) + (x - xz)(y - yz) = 0 \\
g\left(\frac{\partial}{\partial b}, \frac{\partial}{\partial b}\right) &= g\left(-xy\frac{\partial}{\partial x} + \left(1 - z - y^2\right)\frac{\partial}{\partial y} + (y - yz)\frac{\partial}{\partial z}, -xy\frac{\partial}{\partial x} + \left(1 - z - y^2\right)\frac{\partial}{\partial y} + (y - yz)\frac{\partial}{\partial z}\right) \\
&= (-xy)^2 + \left(1 - z - y^2\right)^2 + (y - yz)^2 = (1 - z)^2 = \frac{4}{(a^2 + b^2 + 1)^2}.
\end{aligned}
$$

Thus

$$
\begin{aligned}
g\left(\frac{\partial}{\partial a}, \frac{\partial}{\partial a}\right) &= \frac{4}{(a^2 + b^2 + 1)^2} \\
g\left(\frac{\partial}{\partial a}, \frac{\partial}{\partial b}\right) &= h\left(\frac{\partial}{\partial a}, \frac{\partial}{\partial b}\right) = 0 \\
g\left(\frac{\partial}{\partial b}, \frac{\partial}{\partial b}\right) &= \frac{4}{(a^2 + b^2 + 1)^2},
\end{aligned}
$$

so that therefore the metric is

$$
g = \frac{4}{(a^2 + b^2 + 1)^2} da \otimes da + \frac{4}{(a^2 + b^2 + 1)^2} db \otimes db.
$$

Lecture 20 - The classical Maxwell equations, and the covariant derivative

March 30, 2009

1 The classical Maxwell equations

In classical electrodynamics, the electric field $\vec{E}$ and the magnetic field $\vec{B}$ (each being 3-vector fields) satisfy

$$\begin{array}{lll} \nabla\cdot\vec{B} \;=\; 0 & & \text{no magnetic sources} \\ \nabla\times\vec{E}+\frac{\partial\vec{B}}{\partial t} \;=\; 0 & & \text{Faraday's law} \\ \nabla\times\vec{B}-\epsilon\mu\frac{\partial\vec{E}}{\partial t} \;=\; 4\pi\mu\vec{J} & & \text{Ampere} - \text{Maxwell law} \\ \nabla\cdot\vec{E} \;=\; \frac{4\pi}{\epsilon}\rho & & \text{Gauss' Law} \end{array}$$

The equation $\nabla\cdot\vec{B}=0$ implies that $\vec{B}$ is a *pure curl*, meaning $\vec{B}=\nabla\times\vec{A}$. The vector field $\vec{A}$ is called the *vector potential*. It is defined up to a gradient, meaning that if $\vec{A}$ is replaced by $\vec{A}+\nabla f$ where f is any function, then the equation $\vec{B}=\nabla\times\vec{A}$ remains unchanged.

In the electrostatic case (where $\frac{\partial\vec{B}}{\partial t}=0$), we have $\nabla\times\vec{E}=0$, which implies $\vec{E}$ is a *pure gradient*. This means $E=\nabla\varphi$ for some function φ, called the *electrostatic potential*. Note that φ is defined up to a constant: replacing φ by $\varphi+c$ does not change the equation $\vec{E}=\nabla\phi$.

2 The Lorentz force law

A particle of charge q moving with velocity $\vec{v}$ experiences the force

$$\vec{\mathcal{F}} \;=\; q\left(\vec{E}+\vec{v}\times\vec{B}\right).$$

3 The covariant derivative and the Koszul formula

As we argued in class, there is a need for a new notion of the derivative of a vector field, which interacts with the space's metric.

Let X and Y be vector fields. The symbol $\nabla_X Y$ denotes the derivative of the vector field Y along trajectories of the vector field X. The operator ∇ is called the *covariant derivative*. How should we define this derivative? We require the following four properties:

- ∇ is linear in the first variable and additive in the second:
$$\nabla_{fX+hY} Z \;=\; f\nabla_X Z + h\nabla_Y Z$$
$$\nabla_X(Y+Z) \;=\; \nabla_X Y \;+\; \nabla_X Z$$
where f, h are functions and X, Y are vector fields.
- ∇ obeys the Leibnitz rule in the second variable:
$$\nabla_X(fY) \;=\; X(f)Y \;+\; f\nabla_X Y$$
- ∇ is *compatible with the metric* (a.k.a. Leibnitz rule for inner products):
$$X\langle Y,\, Z\rangle \;=\; \langle \nabla_X Y,\, Z\rangle \;+\; \langle Y,\, \nabla_X Z\rangle$$
where $\langle\cdot,\,\cdot\rangle$ means the same thing as $g(\cdot,\,\cdot)$
- ∇ is *torsion free*:
$$\nabla_X Y \;-\; \nabla_Y X \;=\; [X, Y]$$

But does there really exist an operator ∇ that satisfies these properties? Is it unique? The answer to both questions is yes: ∇ is uniquely (though implicitly) defined by the *Koszul formula*

$$2\langle \nabla_X Y,\, Z\rangle \;=\; X\langle Y,\, Z\rangle + Y\langle X,\, Z\rangle - Z\langle X,\, Y\rangle + \langle [X,Y], Z\rangle - \langle [Y,Z],\, X\rangle + \langle [Z,X],\, Y\rangle\,.$$

The operator ∇ is also called the *connection*.

4 The Christoffel symbols

Let M be a space with metric g. Let $\{x^i\}$ be coordinates, and $\frac{\partial}{\partial x^i}$ the coordinate fields. Since $\nabla_{\partial/\partial x^i} \frac{\partial}{\partial x^j}$ is a vector field, it can be expressed as a linear combination of the coordinate fields: we take

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} \;\triangleq\; \Gamma^k_{ij} \frac{\partial}{\partial x^k}$$

to be the (implicit) definition of the functions Γ^k_{ij}. These are called the *Christoffel symbols*.

Let's compute them. Before we start, notice that the brackets $\left[\partial/\partial x^i,\, \partial/\partial x^j\right]$ are zero:

$$\left[\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right] f \;=\; \frac{\partial^2 f}{\partial x^i \partial x^j} \;-\; \frac{\partial^2 f}{\partial x^j \partial x^i} \;=\; 0.$$

Therefore the Koszul formula gives

$$\begin{aligned}
2\left\langle \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k} \right\rangle \;&=\; \frac{\partial}{\partial x^i}\left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right\rangle + \frac{\partial}{\partial x^j}\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^k}\right\rangle - \frac{\partial}{\partial x^k}\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right\rangle \\
2\left\langle \Gamma^l_{ij} \frac{\partial}{\partial x^l}, \frac{\partial}{\partial x^k} \right\rangle \;&=\; \frac{\partial}{\partial x^i}\left\langle \frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right\rangle + \frac{\partial}{\partial x^j}\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^k}\right\rangle - \frac{\partial}{\partial x^k}\left\langle \frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right\rangle.
\end{aligned}$$

Since $\left\langle \partial/\partial x^i, \partial/\partial x^j \right\rangle \triangleq g(\partial/\partial x^i, \partial/\partial x^j) = g_{ij}$, this directly simplifies to

$$2\Gamma^l_{ij} g_{lk} \;=\; \frac{\partial}{\partial x^i} g_{jk} \;+\; \frac{\partial}{\partial x^j} g_{ik} \;-\; \frac{\partial}{\partial x^k} g_{ij}.$$

We can multiply both sides by g^{km} to eliminate the g_{lk} on the left side:

$$\begin{aligned}
2\Gamma^l_{ij} g_{lk} g^{km} \;&=\; \left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ik} - \frac{\partial}{\partial x^k} g_{ij}\right) g^{km} \\
2\Gamma^l_{ij} \delta^m_l \;&=\; \left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ik} - \frac{\partial}{\partial x^k} g_{ij}\right) g^{km} \\
2\Gamma^m_{ij} \;&=\; \left(\frac{\partial}{\partial x^i} g_{jk} + \frac{\partial}{\partial x^j} g_{ik} - \frac{\partial}{\partial x^k} g_{ij}\right) g^{km}.
\end{aligned}$$

Changing the index labels, we have the formula

$$\Gamma^k_{ij} \;=\; \frac{1}{2}\left(\frac{\partial g_{jl}}{\partial x^i} + \frac{\partial g_{il}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^l}\right) g^{lk}.$$

Lecture 21 - Warped Products

April 1, 2009

1 Warped Products

Consider the graph of a function $y = f(x)$, $\alpha < x < \beta$. Rotate around the x-axis, to obtain a surface of revolution in 3-space. Let M denote the surface; it has coordinates ξ, θ, where $\xi = x$ and θ is the rotation parameter.

Of course any point on M has a location in terms of (x, y, z) coordinates. The association is $(x, y, z) = (\xi,\ f(\xi)\cos\theta,\ f(\xi)\sin\theta)$.

2 First expression of the metric on M

Let's find expressions for the vector fields $\frac{\partial}{\partial \xi^i}$ and $\frac{\partial}{\partial \theta}$ in terms of the rectangular fields $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$ of the ambient space $\mathbb{R}^3$:

$$\begin{aligned}
\frac{\partial}{\partial \xi} &= \frac{\partial x}{\partial \xi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \xi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial \xi}\frac{\partial}{\partial z} \\
&= \frac{\partial}{\partial x} + f'(\xi)\cos\theta\frac{\partial}{\partial y} + f'(\xi)\sin\theta\frac{\partial}{\partial z} \\
\frac{\partial}{\partial \theta} &= \frac{\partial x}{\partial \theta}\frac{\partial}{\partial x} + \frac{\partial y}{\partial \theta}\frac{\partial}{\partial y} + \frac{\partial z}{\partial \theta}\frac{\partial}{\partial z} \\
&= -f(\xi)\sin\theta\frac{\partial}{\partial y} + f(\xi)\cos\theta\frac{\partial}{\partial z}.
\end{aligned}$$

Thus we can compute the inner products

$$\begin{aligned}
g\left(\frac{\partial}{\partial\xi},\frac{\partial}{\partial\xi}\right) &= g\left(\frac{\partial}{\partial x} + f'(\xi)\cos\theta\frac{\partial}{\partial y} + f'(\xi)\sin\theta\frac{\partial}{\partial z}, \frac{\partial}{\partial x} + f'(\xi)\cos\theta\frac{\partial}{\partial y} + f'(\xi)\sin\theta\frac{\partial}{\partial z}\right) \\
&= 1 + (f'(\xi))^2 \\
g\left(\frac{\partial}{\partial\theta},\frac{\partial}{\partial\theta}\right) &- g\left(-f(\xi)\sin\theta\frac{\partial}{\partial y} + f(\xi)\cos\theta\frac{\partial}{\partial z}, -f(\xi)\sin\theta\frac{\partial}{\partial y} + f(\xi)\cos\theta\frac{\partial}{\partial z}\right) \\
&= f^2(\xi).
\end{aligned}$$

Therefore the metric on M reads

$$g \;=\; \left(1 + f'(\xi)^2\right)\, d\xi\otimes d\xi \;+\; (f(\xi))^2\; d\theta\otimes d\theta.$$

3 Second expression of the metric on M

Consider again the graph $y = f(x)$, $\alpha < x < \beta$. We can express a location on this curve by giving the x-coordinate, but we can also express a location by giving the arclength between that point and the curve's starting position. That is, we can define the arclength variable r to be

$$r \;=\; \int_\alpha^\xi \sqrt{1 + (f'(t))}\, dt.$$

Notice that

$$dr \;=\; \sqrt{1 + (f'(\xi))^2}\, d\xi.$$

Therefore

$$\begin{aligned}
g &= \left(1 + f'(\xi)^2\right)\, d\xi\otimes d\xi \;+\; (f(\xi))^2\; d\theta\otimes d\theta \\
&= \left(\sqrt{1 + + (f'(\xi))^2}\, d\xi\right)\otimes\left(\sqrt{1 + + (f'(\xi))^2}\, d\xi\right) \;+\; (f(\xi))\; d\theta\otimes d\theta.
\end{aligned}$$

So we arrive at

$$g \;=\; dr\otimes dr \;+\; (\phi(r))^2\; d\theta\otimes d\theta,$$

where $\phi(r) \;=\; f(\xi(r))$. This is the most usable form of the warped product metric.

4 Covariant derivatives

As an example computation, let's show how to compute the covariant derivative $\nabla_{\partial/\partial\theta}\frac{\partial}{\partial r}$. In order to use index notation, put $x^1 = r$, $x^2 = \theta$. Then

$$\begin{aligned}\nabla_{\partial/\partial\theta}\frac{\partial}{\partial r} &= \nabla_{\partial/\partial x^2}\frac{\partial}{\partial x^1} = \Gamma^k_{12}\frac{\partial}{\partial x^k} \\ &= \Gamma^1_{12}\frac{\partial}{\partial x^1} + \Gamma^2_{12}\frac{\partial}{\partial x^2} = \Gamma^1_{12}\frac{\partial}{\partial r} + \Gamma^2_{12}\frac{\partial}{\partial\theta}.\end{aligned}$$

We use the formula

$$\Gamma^k_{ij} = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial x^j} + \frac{\partial g_{jl}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^l}\right) g^{lk}$$

and the metric components

$$g_{11} = 1,\ g_{12} = g_{21} = 0,\ g_{22} = (\phi(r))^2 \quad g^{11} = 1,\ g^{12} = g^{21} = 0,\ g^{22} = 1/(\phi(r))^2$$

to compute

$$\begin{aligned}\Gamma^1_{12} &= \frac{1}{2}\left(\frac{\partial g_{11}}{\partial x^2} + \frac{\partial g_{21}}{\partial x^1} - \frac{\partial g_{12}}{\partial x^1}\right) g^{11} + \frac{1}{2}\left(\frac{\partial g_{12}}{\partial x^2} + \frac{\partial g_{22}}{\partial x^1} - \frac{\partial g_{12}}{\partial x^2}\right) g^{21} = 0 \\ \Gamma^2_{12} &= \frac{1}{2}\left(\frac{\partial g_{11}}{\partial x^2} + \frac{\partial g_{21}}{\partial x^1} - \frac{\partial g_{12}}{\partial x^1}\right) g^{12} + \frac{1}{2}\left(\frac{\partial g_{12}}{\partial x^2} + \frac{\partial g_{22}}{\partial x^1} - \frac{\partial g_{12}}{\partial x^2}\right) g^{22} \\ &= 0 + \frac{1}{2}\left(0 + 2\phi(x^1)\phi'(x^1) - 0\right)\frac{1}{(\phi(x^1))^2} \\ &= \frac{\phi'(r)}{\phi(r)}.\end{aligned}$$

Thus

$$\nabla_{\frac{\partial}{\partial\theta}}\frac{\partial}{\partial r} = \frac{\phi'(r)}{\phi(r)}\frac{\partial}{\partial\theta}.$$

Lecture 22 - Gauge invariance, wave equations, and variation of curves

April 3, 2009

1 Wave equations

Consider again Maxwell's equations

$$\begin{array}{ll} \nabla\cdot\vec{B} \;=\; 0 & \text{no magnetic sources} \\ \nabla\times\vec{E} + \frac{\partial\vec{B}}{\partial t} \;=\; 0 & \text{Faraday's law} \\ \nabla\times\vec{B} - \epsilon\mu\frac{\partial\vec{E}}{\partial t} \;=\; 4\pi\mu\vec{J} & \text{Ampere} - \text{Maxwell law} \\ \nabla\cdot\vec{E} \;=\; \frac{4\pi}{\epsilon}\rho & \text{Gauss' Law} \end{array}$$

Recall the classical vector identity, valid for any vector field $\vec{A}$:

$$\nabla\times\nabla\times\vec{A} \;=\; \nabla(\nabla\cdot\vec{A}) - \triangle\vec{A}.$$

Now consider the charge-free (i.e. free space) situation, in which $\rho = 0$, $\vec{J} = 0$. Taking the curl of both side of the Ampere-Maxwell equation, we get

$$\begin{aligned} \nabla\times\nabla\times\vec{B} - \epsilon\mu\frac{\partial}{\partial t}\left(\nabla\times\vec{E}\right) \;&=\; 0 \\ \nabla(\nabla\cdot\vec{B}) - \triangle\vec{B} + \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{B} \;&=\; 0 \\ \triangle\vec{B} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\vec{B} \;&=\; 0 \end{aligned}$$

In the second line we used Faraday's law, and in the third line we used $\nabla\cdot\vec{B} = 0$. Note also that $c^2 = \frac{1}{\epsilon\mu}$. This is the classic wave equation. Similarly, had one taken the curl of both sides of the Faraday equation to start with, then used to Ampere-Maxwell equation to simplify the result, one would get the corresponding wave equation for the $\vec{E}$-field:

$$\triangle\vec{E} - \frac{1}{c^2}\frac{\partial\vec{E}}{\partial t^2} \;=\; 0.$$

2 Gauge invariance

Electromagnetics is an example of what is known as a gauge theory. We shall describe what this means in this section.

Maxwell's first equation $\nabla \cdot \vec{B} = 0$ implies that $\vec{B}$ is a pure curl, so that $\vec{B} = \nabla \times \vec{A}$ for some vector field $\vec{A}$. We have previously noted that there is considerable freedom in choosing $\vec{A}$, namely $\vec{A}$ can be replaced by $\vec{A} + \nabla f$ for any (differentiable) function f.

In the electrostatic case, $\nabla \times \vec{E} = 0$ implies that $\vec{E}$ is a pure gradient: $\vec{E} = \nabla\phi$ for some function ϕ, called the *static electric potential*.

In the general case, we have $\nabla \times \vec{E} + \frac{\partial}{\partial t}\nabla \times \vec{A} \;=\; 0$, so that $\vec{E} + \frac{\partial \vec{A}}{\partial t}$ is a pure gradient. Thus there exists a function φ so that

$$\vec{E} + \frac{\partial \vec{A}}{\partial t} \;=\; \nabla\varphi.$$

This function φ is called the *electric pseudopotential*.

Since there is some freedom in the choice of $\vec{A}$, there will be freedom in choosing φ as well. A particular choice of $(\vec{A}, \varphi)$ is called a *choice of gauge*. If f is any function, then replacing $(\vec{A}, \varphi)$ with $\left(\vec{A} + \nabla f, \varphi + \frac{\partial f}{\partial t}\right)$ does not change the equations $\vec{B} = \nabla \times \vec{A}$ or $\vec{E} + \frac{\partial \vec{A}}{\partial t} = \nabla\varphi$. This is known as *gauge invariance*.

To solve an electrodynamical problem, one usually chooses a propitious gauge (something that leads to a lot of cancelation), and then works with the more 'primitive' quantities $(\vec{A}, \varphi)$ instead of $\vec{E}$, $\vec{B}$.

3 Path variation

Let $\gamma(\tau)$, $0 < \tau < a$ be a path, with velocity vector $\frac{d}{d\tau}$. Our aim is to discover the conditions under which $\gamma(\tau)$ is a *geodesic*, which is to say, under what conditions it is the shortest path between its endpoints $p = \gamma(0)$ and $q = \gamma(a)$.

Recall the pathlength and the energy functionals

$$\begin{aligned} \mathcal{L}(\gamma) &= \int_0^a \sqrt{\left\langle \frac{d}{d\tau}, \frac{d}{d\tau} \right\rangle}\, d\tau \\ \mathcal{E}(\gamma) &= \int_0^a \left\langle \frac{d}{d\tau}, \frac{d}{d\tau} \right\rangle d\tau. \end{aligned}$$

These are closely related, although the energy functional $\mathcal{E}$ is easier to work with since there is no square-root.

Let $\gamma_s(\tau)$ be a smoothly family of paths, parametrized by $s \in (-b, b)$, with $\gamma_0(\tau) = \gamma(\tau)$ being the original path. Assume also that each path γ_s has the same endpoints as γ: $\gamma_s(0) = \gamma(0)$ and $\gamma_s(a) = \gamma(a)$ for all s. This is called a *variation* of γ.

There are two vector fields involved. First is $\frac{d}{d\tau}$, created by fixing s and varying τ is called the *direction field*. Second, it is possible to fix τ and vary s. This leads to the *variation field* $\frac{d}{ds}$.

If γ is indeed a geodesic, then is must be the case that $\mathcal{L}(\gamma_0) \leq \mathcal{L}(\gamma_s)$, or likewise, that $\mathcal{E}(\gamma_0) \leq \mathcal{E}(\gamma_s)$. Since $\mathcal{E}(\gamma_s)$, the energy the path γ_s, can be considered a function of s, this means that

$$\frac{d}{ds}\Big|_{s=0} \mathcal{E}(\gamma_s) \;=\; 0.$$

The key is this: *This must hold true not just for one variation, but for any conceivable variation of* γ. What property of γ could possibly lead to this?

We compute:

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=0} \mathcal{E}(\gamma_s) \;&=\; \frac{d}{ds}\int_0^a \left\langle \frac{d}{d\tau}, \frac{d}{d\tau} \right\rangle d\tau \\
&=\; \int_0^a \frac{d}{ds}\left\langle \frac{d}{d\tau}, \frac{d}{d\tau} \right\rangle d\tau \\
&=\; 2\int_0^a \left\langle \nabla_{\frac{d}{ds}} \frac{d}{d\tau}, \frac{d}{d\tau} \right\rangle d\tau.
\end{aligned}$$

The last line comes from the axiom that the connection ∇ is compatible with the metric. Now we work some switcheroo magic. From the connection's torsion-free axiom, we have

$$\nabla_{\frac{d}{ds}} \frac{d}{d\tau} \;=\; \nabla_{\frac{d}{d\tau}} \frac{d}{ds} \;+\; \left[\frac{d}{d\tau}, \frac{d}{ds}\right]$$

But the bracket in this case actually vanishes! (why?) Thus

$$\nabla_{\frac{d}{ds}} \frac{d}{d\tau} \;=\; \nabla_{\frac{d}{d\tau}} \frac{d}{ds}.$$

and we can continue our computation

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=0} \mathcal{E}(\gamma_s) \;&=\; 2\int_0^a \left\langle \nabla_{\frac{d}{d\tau}} \frac{d}{ds}, \frac{d}{d\tau} \right\rangle d\tau \\
&=\; 2\int_0^a \frac{d}{d\tau}\left\langle \frac{d}{ds}, \frac{d}{d\tau} \right\rangle d\tau \;-\; 2\int_0^a \left\langle \frac{d}{ds}, \nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} \right\rangle d\tau
\end{aligned}$$

The first term is a total derivative; by the fundamental theorem of calculus, we have

$$\begin{aligned}
\frac{d}{ds}\Big|_{s=0} \mathcal{E}(\gamma_s) \;&=\; 2\left\langle \frac{d}{ds}, \frac{d}{d\tau} \right\rangle\Big|_{\tau=0}^{\tau=a} \;-\; 2\int_0^a \left\langle \frac{d}{ds}, \nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} \right\rangle d\tau \\
&=\; -2\int_0^a \left\langle \frac{d}{ds}, \nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} \right\rangle d\tau.
\end{aligned}$$

If γ is indeed the minimizing path between its endpoints, we must therefore have

$$0 \quad = \quad -2\int_0^a \left\langle \frac{d}{ds}, \nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} \right\rangle d\tau$$

for ANY variation field $\frac{d}{ds}$. The only way this is possible is if $\nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} = 0$.

Therefore we can write down the geodesic equation:

$$\gamma \text{ is a geodesic} \quad \Longleftrightarrow \quad \nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} \; = \; 0$$

Lecture 23 - The parallel transport equation, the Riemann curvature tensor, and the Jacobi equation

April 15, 2009

1 Parallel transport

Given a path $\gamma(\tau)$ (not necessarily a geodesic), a vector field X is called *parallel*, or constant, along γ if

$$\nabla_{\frac{\partial}{\partial\tau}} X \;=\; 0.$$

In coordinates we can write

$$\frac{d}{d\tau} \;=\; \frac{dx^i}{d\tau}\frac{\partial}{\partial x^i} \qquad X \;=\; X^j\frac{\partial}{\partial x^j}.$$

Using the axioms of covariant differentiation, we compute

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial\tau}} X &= \nabla_{\frac{dx^i}{d\tau}\frac{\partial}{\partial x^i}}\left(X^j\frac{\partial}{\partial x^j}\right)\\
&= \frac{dx^i}{d\tau}\nabla_{\frac{\partial}{\partial x^i}}\left(X^j\frac{\partial}{\partial x^j}\right)\\
&= \frac{dx^i}{d\tau}\frac{\partial}{\partial x^i}\left(X^j\right)\frac{\partial}{\partial x^j} + \frac{dx^i}{d\tau}X^j\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j}\\
&= \frac{dX^j}{d\tau}\frac{\partial}{\partial x^j} + \frac{dx^i}{d\tau}X^j\Gamma_{ij}^k\frac{\partial}{\partial x^k}\\
&= \left(\frac{dX^k}{d\tau} + \frac{dx^i}{d\tau}X^j\Gamma_{ij}^k\right)\frac{\partial}{\partial x^k}.
\end{aligned}$$

The Christoffel symbols Γ_{ij}^k, the path γ, and the derivatives $\frac{dx^i}{d\tau}$ are known. Therefore the parallel transport equation is a system of n first order linear differential equations in the unknowns X^k:

$$\nabla_{\frac{d}{d\tau}} X \;=\; 0 \qquad \text{if and only if} \qquad \frac{dX^k}{d\tau} + X^j\frac{dx^i}{d\tau}\Gamma_{ij}^k \;=\; 0 \quad \text{for all } 1 < k < n.$$

This means that given a single vector $X \in T_{\gamma(0)}M$, it can be *transported* along the curve γ by solving this system of equations.

2 The Riemann tensor

Given three vector field X, Y, and Z, the Riemann tensor R is defined as follows:

$$R(X,Y)\,Z \;\triangleq\; \nabla_X\nabla_Y Z \;-\; \nabla_Y\nabla_X Z \;-\; \nabla_{[X,Y]}Z.$$

Essentially the Riemann tensor measures the **failure of commutativity of mixed second partials** when applied to vector fields. Note the obvious fact that R is anti-symmetric in the first two variables:

$$R(X,Y)\,Z \;=\; -R(Y,X)\,Z.$$

The Riemann tensor is also called the *curvature tensor*. Although at this point it is not immediately clear why, this tensor captures all of the geometric information of the space under consideration.

3 The Jacobi equation

Here we shall lay out the first geometric interpretation of the Riemann tensor: it governs variations of geodesics.

Let $p \in M$ be a point, let $\gamma(\tau)$ be a geodesic, and let $\gamma_s(\tau)$ be a variation consisting of a family of geodesics emanating from p. Again we have the directional field $\frac{d}{d\tau}$ and the variational field $\frac{d}{ds}$. We want to compute the second derivative of the variational field along the geodesic:

$$\nabla_{\frac{d}{d\tau}}\nabla_{\frac{d}{d\tau}}\frac{d}{ds}.$$

The strategy will be to move the $\frac{d}{ds}$ to the leftmost position. First use the torsion-free axiom to make the swap

$$\nabla_{\frac{d}{d\tau}}\nabla_{\frac{d}{d\tau}}\frac{d}{ds} \;=\; \nabla_{\frac{d}{d\tau}}\nabla_{\frac{d}{ds}}\frac{d}{d\tau} \;+\; \nabla_{\frac{d}{d\tau}}[\frac{d}{d\tau},\,\frac{d}{ds}].$$

However, $[d/d\tau,\, d/ds] = 0$, by the commutativity of partial derivatives. Now using the definition of the curvature tensor (and again using $[d/d\tau, d/ds] = 0$):

$$R\left(\frac{d}{d\tau},\,\frac{d}{ds}\right)\frac{d}{d\tau} \;=\; \nabla_{\frac{d}{d\tau}}\nabla_{\frac{d}{ds}}\frac{d}{d\tau} \;-\; \nabla_{\frac{d}{ds}}\nabla_{\frac{d}{d\tau}}\frac{d}{d\tau}$$

we can make another switch:

$$\begin{aligned}\nabla_{\frac{d}{d\tau}}\nabla_{\frac{d}{d\tau}}\frac{d}{ds} \;&=\; \nabla_{\frac{d}{d\tau}}\nabla_{\frac{d}{ds}}\frac{d}{d\tau}\\ &=\; \nabla_{\frac{d}{ds}}\nabla_{\frac{d}{d\tau}}\frac{d}{d\tau} \;+\; R\left(\frac{d}{d\tau},\,\frac{d}{ds}\right)\frac{d}{d\tau}\end{aligned}$$

However, we had assumed that the variation consists of geodesics, meaning $\nabla_{\frac{d}{d\tau}} \frac{d}{d\tau} = 0$. Thus

$$\nabla_{\frac{d}{d\tau}} \nabla_{\frac{d}{d\tau}} \frac{d}{ds} \;=\; R\left(\frac{d}{d\tau}, \frac{d}{ds}\right) \frac{d}{d\tau}.$$

This is most commonly written in the following form (recalling the anti-symmetry of R)

$$\nabla_{\frac{d}{d\tau}} \nabla_{\frac{d}{d\tau}} \frac{d}{ds} \;+\; R\left(\frac{d}{ds}, \frac{d}{d\tau}\right) \frac{d}{d\tau} \;=\; 0.$$

The tensor $R(\cdot,\, X)X$ is known as the *tidal curvature operator* in the direction X.

4 Examples of Jacobi equations

Example: Flat space. In flat (Euclidean) space, the Riemann tensor is precisely zero. Thus the Jacobi equation yields

$$\nabla_{\frac{d}{d\tau}} \nabla_{\frac{d}{d\tau}} \frac{d}{ds} \;=\; 0,$$

Which means that $\frac{d}{ds}$ is a linear field.

Example: Positively curved space. In a positively curved space, the tidal curvature

$$R\left(\frac{d}{ds}, \frac{d}{d\tau}\right) \frac{d}{d\tau} \quad\sim\quad \alpha^2 \frac{d}{ds} \;+\; \textit{other terms}$$

is, roughly speaking, proportional to a *positive* multiple of the variation field. Thus the Jacobi equation yields

$$\nabla_{\frac{d}{d\tau}} \nabla_{\frac{d}{d\tau}} \frac{d}{ds} \;+\; \alpha^2 \frac{d}{ds} \;+\; \textit{other terms} \;=\; 0$$

which is an equation of the form

$$f''(\tau) \;+\; \alpha^2 f(\tau) \;=\; 0,$$

the solution to which is $f(\tau) \;=\; \sin(\alpha\tau)$ (initial condition is $f(\tau) = 0$). Thus geodesics tend to curve in toward one another.

Example: Negatively curved space. In a negatively curved space, the tidal curvature

$$R\left(\frac{d}{ds}, \frac{d}{d\tau}\right) \frac{d}{d\tau} \quad\sim\quad -\alpha^2 \frac{d}{ds} \;+\; \textit{other terms}$$

is, roughly speaking, proportional to a *negative* multiple of the variation field. Thus the Jacobi equation yields

$$\nabla_{\frac{d}{d\tau}} \nabla_{\frac{d}{d\tau}} \frac{d}{ds} \;-\; \alpha^2 \frac{d}{ds} \;+\; \textit{other terms} \;=\; 0$$

which is an equation of the form

$$f''(\tau) \;-\; \alpha^2 f(\tau) \;=\; 0,$$

the solution to which is $f(\tau) \;=\; \sinh(\alpha\tau)$ (initial condition is $f(\tau) = 0$). The function sinh is initially nearly linear, but later is nearly exponential. Thus geodesics tend to bend away from one another.

5 Sectional Curvature

Given two vectors X and Y located at a point p, in the infinitesimal sense they span a plane. The *section curvature* of this plane is given by

$$\sec(X,\, Y) \;=\; \frac{\langle R(X,\, Y)Y,\, X\rangle}{\langle X,\, X\rangle\,\langle Y,\, Y\rangle \;-\; \langle X,\, Y\rangle^2}.$$

Lecture 24 - The curvature tensor in components

April 20, 2009

1 The Riemann curvature tensor in components

We have defined

$$R(X,Y)Z \;=\; \nabla_X\nabla_Y Z \,-\, \nabla_Y\nabla_X Z \,-\, \nabla_{[X,Y]}Z,$$

so that R is a map that absorbs three vector fields and spits out a vector field. Thus R is a $(1,3)$-tensor and can therefore be expressed in the form

$$R \;=\; R_{ijk}{}^l\, dx^i \otimes dx^j \otimes dx^k \otimes \frac{\partial}{\partial x^l}.$$

To compute the functions $R_{ijk}{}^l$ we have to plug in the coordinate fields:

$$R\left(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} \;=\; R_{ijk}{}^l\frac{\partial}{\partial x^l}.$$

Thus we compute

$$\begin{aligned}
R\left(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} \;&=\; \nabla_{\frac{\partial}{\partial x^i}}\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^k} \,-\, \nabla_{\frac{\partial}{\partial x^j}}\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^k} \\
&=\; \nabla_{\frac{\partial}{\partial x^i}}\left(\Gamma_{jk}^s\frac{\partial}{\partial x^s}\right) \,-\, \nabla_{\frac{\partial}{\partial x^j}}\left(\Gamma_{ik}^s\frac{\partial}{\partial x^s}\right) \\
&=\; \frac{\partial\Gamma_{jk}^s}{\partial x^i}\frac{\partial}{\partial x^s} \,+\, \Gamma_{jk}^s\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^s} \,-\, \frac{\partial\Gamma_{ik}^s}{\partial x^j}\frac{\partial}{\partial x^s} \,-\, \Gamma_{ik}^s\nabla_{\frac{\partial}{\partial x^j}}\frac{\partial}{\partial x^s} \\
&=\; \frac{\partial\Gamma_{jk}^s}{\partial x^i}\frac{\partial}{\partial x^s} \,+\, \Gamma_{jk}^s\Gamma_{is}^l\frac{\partial}{\partial x^l} \,-\, \frac{\partial\Gamma_{ik}^s}{\partial x^j}\frac{\partial}{\partial x^s} \,-\, \Gamma_{ik}^s\Gamma_{js}^l\frac{\partial}{\partial x^l}.
\end{aligned}$$

Changing the index labels somewhat, we have

$$R_{ijk}{}^l\frac{\partial}{\partial x^l} \;=\; R\left(\frac{\partial}{\partial x^i},\frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^k} \;=\; \left(\Gamma_{jk}^s\Gamma_{is}^l \,-\, \Gamma_{ik}^s\Gamma_{js}^l \,+\, \frac{\partial\Gamma_{jk}^l}{\partial x^i} \,-\, \frac{\partial\Gamma_{ik}^l}{\partial x^j}\right)\frac{\partial}{\partial x^l}.$$

Therefore

$$R_{ijk}{}^{l} \;=\; \Gamma^{s}_{jk}\Gamma^{l}_{is} \;-\; \Gamma^{s}_{ik}\Gamma^{l}_{js} \;+\; \frac{\partial \Gamma^{l}_{jk}}{\partial x^{i}} \;-\; \frac{\partial \Gamma^{l}_{ik}}{\partial x^{j}}.$$

Note the complexity of this formula. The Christoffel symbols are nonlinear first-order expressions in terms of the metric, and the curvature components $R_{ijk}{}^{l}$ involve derivatives and nonlinear combinations of the Christoffel symbols. Thus the curvature components $R_{ijk}{}^{l}$ are nonlinear second-order expressions in terms of the metric. In fact the curvature tensor is essentially as nonlinear as it is possible to be.

Lecture 25 - The stress-energy-momentum tensor

April 22, 2009

1 The Cauchy stress tensor

The Cauchy stress tensor $\vec{T}$ is a construct of classical physics. It is a $(0,2)$-tensor, and can be written essentially as a 3×3 matrix:

$$\vec{T} = \vec{T}_{ab} = \begin{pmatrix} \vec{T}_{11} & \vec{T}_{12} & \vec{T}_{13} \\ \vec{T}_{21} & \vec{T}_{22} & \vec{T}_{23} \\ \vec{T}_{31} & \vec{T}_{32} & \vec{T}_{33} \end{pmatrix}.$$

It's meaning is as follows: if $\hat{n}$ and $\hat{v}$ are unit vectors, then

$\vec{T}(\hat{n}, \hat{v})$ is the force communicated across the $\vec{n}$-plane in the $\vec{v}$-direction.

We can interpret the components $\vec{T}_{ab}$ as follows:

The diagonal elements are the *pressures*: $\vec{T}_{11}$, for instance, is the force communicated across the x^1-plane in the x^1-direction, that is, the pressure across the x^1-plane.

The off-diagonal elements are the *shear forces*: $\vec{T}_{12}$, for instance, is the x^2-force communicated across the x^1-plane.

The Cauchy stress tensor has many applications in classical physics and particularly in engineering applications. The notation $\vec{T}$ is meant to indicated that $\vec{T}$ is a classical (tensor) quantity, not meant to indicate that it is a vector.

2 The stress-energy-momentum tensor

The entries in the Cauchy stress tensor depend on the choice of reference frame, and is therefore inadequate for relativistic applications. We require a fully Lorentz-invariant version of this tensor.

First note that force is the same as the time-derivative of momentum:

$$\begin{aligned} f &= ma \\ &= m\frac{dv}{dt} \\ &= \frac{dp}{dt} \quad \text{where } p = mv \text{ is classical momentum} \end{aligned}$$

Thus force across a boundary can be regarded as momentum flux across a boundary.

The 4-dimensional *stress-energy-momentum tensor* (or *energy-momentum tensor* or *stress-energy tensor*) is

$$T = T_{ij} = \left(\begin{array}{cccc} T_{00} & T_{01} & T_{02} & T_{03} \\ T_{10} & & & \\ T_{20} & & \vec{T} & \\ T_{30} & & & \end{array} \right).$$

The space-components are just the components of the Cauchy stress tensor as seen in the observers rest-frame. The other components have the following interpretation:
T_{00} is energy (aka mass) density
T_{i0} $(i \neq 0)$ is momentum density
T_{0i} $(i \neq 0)$ is energy (aka mass) flux
$T_{ij} = \vec{T}_{ij}$ $(i, j \neq 0)$ are the various momenta fluxes (aka pressures and shear forces).

The stress-energy tensor can be computed from detailed knowledge of the distribution of matter, energy, and forces is some region of space-times.

Lecture 26 - Traces and norms

April 27, 2009

1 Traces and norms for classical vectors and matrices

If A^i_j is a square matrix, its trace $Tr(A)$ is $A^i_i = A^1_1 + \cdots + A^n_n$, the sum of the diagonal elements. The trace obeys the following properties: if A and B are $n \times n$ matrices and c is a scalar then

- Commutativity invariance: $Tr(AB) \;=\; Tr(BA)$
- Transpose invariance: $Tr(A^T) \;=\; Tr(A)$
- Linearity: $Tr(cA + B) \;=\; cTr(A) + Tr(B)$

As a consequence of the commutativity invariance, we can prove

Theorem 1.1 *If $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the $n \times n$ matrix A, then $Tr(A) = \lambda_1 + \cdots + \lambda_n$.*

Pf Let $A = C^{-1} J_A C$, where J_A is the Jordan canonical form for A. Recall that J_A is a matrix with the eigenvalues of A along its diagonal, has some 1's on the off-diagonal, and has 0's everywhere else. Clearly $Tr(J_A) = \lambda_1 + \cdots + \lambda_n$. Then using commutativity

$$Tr(A) \;=\; Tr(C^{-1}J_AC) \;=\; Tr(J_ACC^{-1}) \;=\; Tr(J_A) \;=\; \lambda_1 + \cdots + \lambda_n.$$

$\square$

It is possible to use the trace operation as an inner product. Consider vectors $v = v^i$ and $w = w^i$. We have

$$w^T v \;=\; \left(w^1, \ldots, w^n\right) \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \;=\; \begin{pmatrix} w^1v^1 & w^2v^1 & \ldots & w^nv^1 \\ w^1v^2 & w^2v^2 & & w^nv^2 \\ \vdots & & \ddots & \vdots \\ w^1v^n & w^2v^n & \ldots & w^nv^n \end{pmatrix}$$

$$Tr(w^Tv) \;=\; w^1v^1 \,+\, w^2v^2 \,+\, \ldots \,+\, w^nv^n \;=\; \langle w,\, v\rangle\,.$$

If $A = A^i_j$ is a matrix, how should the norm $|A|^2$ be defined? One the one hand, we want $|A|^2 = 0$ if and only if $A^i_j = 0$ for every i, j. On the other, it should be the case that $|cA|^2 = c^2|A|^2$ for any scalar c. The expressions

$$\begin{aligned} |A|^2 &= (A^1_1)^2 + (A^1_2)^2 + (A^2_1)^2 + \ldots + (A^n_n)^2 \\ &= \sum_{i,j=1}^{n} A^i_j A^i_j. \end{aligned}$$

meets both criteria. Now we compute

$$\begin{aligned} AA^T &= A^i_k (A^T)^k_j = \sum_{k=1} n A^i_k A^j_k \\ Tr(AA^T) &= Tr\left(\sum_k A^i_k A^j_k\right) = \sum_l \sum_k A^l_k A^l_k = |A|^2. \end{aligned}$$

It therefore makes sense to define an inner product

$$\langle A,\, B\rangle = Tr\left(AB^T\right).$$

One must check that this is bilinear (obvious), that it is symmetric ($Tr(AB^T) = Tr((AB^T)^T) = Tr(BA^T)$), and that it is nondegenerate ($Tr(AA^T) = 0$ iff A is the zero matrix).

2 Contractions of tensors

There is an analogy between classical matrix operations and tensor operations. If $v = v^i = v^i\frac{\partial}{\partial x^i}$ is a vector, the classical 'transpose' operation is analogous to the 'lowering of index' or 'flattening' operation:

$$\begin{aligned} v_\flat &= v_i dx^i \\ v_i &= v^j g_{ij}. \end{aligned}$$

There is another operation, known as the contraction, or C-operation. It is defined whenever a covector is tensored with a vector. It replaces tensoring with evaluation:

$$\mathsf{C}\,(\omega \otimes X) \triangleq \omega(X)$$

where ω is a covector and X a vector.

It turns out that C is the analog of the classical trace. Let $A = A_j{}^i dx^j \otimes \frac{\partial}{\partial x^i}$ be a $(1,1)$-tensor. Then

$$\begin{aligned} \mathsf{C}(A) &= \mathsf{C}\left(A_j{}^i dx^j \otimes \frac{\partial}{\partial x^i}\right) \\ &= A_j{}^i dx^j\left(\frac{\partial}{\partial x^i}\right) = A_j{}^i \delta^j_i \\ &= A_i{}^i. \end{aligned}$$

3 Contractions as norms

Just as traces can be used to define norms of matrices, contractions can be used to define norms of tensors. To start with, consider the norm of a vector $v = v^i$. Normally we have

$$|v|^2 \;=\; g(v,v) \;=\; g_{ij}v^iv^j \;=\; v^iv_i$$

recalling that $v_i = g_{ij}v^j$. We can get the same result another way however:

$$\begin{aligned}
&v \otimes v \;\to\; v_\flat \otimes v \;\to\; \mathsf{C}(v_\flat \otimes v)\\
&v^iv^j\frac{\partial}{\partial x^i}\otimes\frac{\partial}{\partial x^j} \;\to\; v_iv^j dx^i \otimes \frac{\partial}{\partial x^j} \;\to\; v_iv^j dx^i\left(\frac{\partial}{\partial x^j}\right) \;=\; v_iv^j\delta^i_j \;=\; v_iv^i.
\end{aligned}$$

This same procedure works with other kinds of tensors as well. For instance if $T = T^{ij}$ is a $(2,0)$-tensor, we can form the tensor product

$$T\otimes T \;=\; T^{ij}T^{kl}\frac{\partial}{\partial x^i}\otimes\frac{\partial}{\partial x^j}\otimes\frac{\partial}{\partial x^k}\otimes\frac{\partial}{\partial x^l},$$

then lower the indices on the first tensor

$$T_{\flat\flat}\otimes T \;=\; T_{ij}T^{kl}dx^i\otimes dx^j\otimes\frac{\partial}{\partial x^k}\otimes\frac{\partial}{\partial x^l}$$

(recall that $T_{ij} = g_{is}g_{jt}T^{st}$). Then it is possible to make two contractions

$$\mathsf{CC}\,(T_{\flat\flat}\otimes T) \;=\; T_{ij}T^{kl}dx^i\left(\frac{\partial}{\partial x^k}\right)dx^j\left(\frac{\partial}{\partial x^l}\right) \;=\; T_{ij}T^{kl}\delta^i_k\delta^j_l \;=\; T_{ij}T^{ij}.$$

Therefore we define

$$|T|^2 \;=\; T_{ij}T^{ij} \;=\; g_{ik}g_{jl}T^{kl}T^{ij}.$$

This can be done for any tensor of any type. If $T = T_{ab\ldots c}{}^{ij\ldots k}$, then

$$\begin{aligned}
|T|^2 \;&\triangleq\; T_{ab\ldots c}{}^{ij\ldots k}T_{\alpha\beta\ldots\gamma}{}^{\iota\upsilon\ldots\kappa}g^{a\alpha}g^{b\beta}\ldots g^{c\gamma}g_{i\iota}g_{j\upsilon}\ldots g_{k\kappa}\\
&=\; T_{ab\ldots c}{}^{ij\ldots k}T^{ab\ldots c}{}_{ij\ldots k}.
\end{aligned}$$

That is, corresponding indices are ‘contracted’ using the metric tensor.

4 Norms and contractions of the curvature tensor

The norm $|R|$ of the Riemann curvature tensor $R = R^l_{ijk}$ is given by

$$\begin{aligned}
|R|^2 \;&=\; R_{ijk}{}^lR_{mno}{}^pg^{im}g^{jn}g^{ko}g_{lp}\\
&=\; R_{ijk}{}^lR^{ijk}{}_l.
\end{aligned}$$

It is possible to contract the curvature tensor in other ways however. Contracting the first and fourth position, we get

$$\mathtt{C}_{1,4}\left(R_{ijk}{}^{l}dx^i\otimes dx^j\otimes dx^k\otimes\frac{\partial}{\partial x^l}\right) \;=\; R_{ijk}{}^{l}dx^i\left(\frac{\partial}{\partial x^l}\right)dx^j\otimes dx^k$$
$$=\; R_{ijk}{}^{l}\delta^i_l dx^j\otimes dx^k \;=\; R_{ljk}{}^{l}dx^j\otimes dx^k.$$

This is known as the Ricci curvature tensor:

$$\mathrm{Ric}_{jk} \;=\; R_{ljk}{}^{l}.$$

In the case where the two indices are the same, the interpretation is as follows:

$$\begin{aligned}\mathrm{Ric}_{jj} \;=\; R_{ljj}{}^{l} \quad &= \quad \left\langle R\left(\frac{\partial}{\partial x^l},\frac{\partial}{\partial x^j}\right)\frac{\partial}{\partial x^j},\frac{\partial}{\partial x^l}\right\rangle \\ &= \quad \sum_{l=1}^{n} const\cdot sec\left(\frac{\partial}{\partial x^l},\frac{\partial}{\partial x^j}\right).\end{aligned}$$

The rightmost expression is a multiple of the average of sectional curvatures of the planes that contain the vector $\frac{\partial}{\partial x^j}$.

Lecture 27 - Covariant derivatives

April 29, 2009

1 Leibnitz rules

Since covector fields can be nonconstant, there should be a way of taking their covariant derivatives. This is accomplished by imposing a new Leibnitz rule.

If X, Y are tensors and ω a covector, then $\omega(Y)$ is a function, so $X\left(\omega(Y)\right)$ is the derivative of this function in the X direction. We define $\nabla_X\omega$ implicitly by

$$X\left(\omega(Y)\right) \;=\; \left(\nabla_X\omega\right)(Y) + \omega\left(\nabla_X Y\right).$$

We know that $\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j} = \Gamma_{ij}^k\frac{\partial}{\partial x^k}$, and we know how to compute the symbols Γ_{ij}^k. From the above Leibnitz rule, we should be able to compute $\nabla_{\frac{\partial}{\partial x^i}}dx^j$. We have

$$\begin{aligned}
\frac{\partial}{\partial x^i}\left(dx^j\left(\frac{\partial}{\partial x^l}\right)\right) &= \left(\nabla_{\frac{\partial}{\partial x^i}}dx^j\right)\left(\frac{\partial}{\partial x^l}\right) + dx^j\left(\nabla_{\frac{\partial}{\partial x^i}}\frac{\partial}{\partial x^l}\right)\\
\frac{\partial}{\partial x^i}\left(\delta^j{}_l\right) &= \left(\nabla_{\frac{\partial}{\partial x^i}}dx^j\right)\left(\frac{\partial}{\partial x^l}\right) + dx^j\left(\Gamma_{il}^k\frac{\partial}{\partial x^k}\right)\\
0 &= \left(\nabla_{\frac{\partial}{\partial x^i}}dx^j\right)\left(\frac{\partial}{\partial x^l}\right) + \Gamma_{il}^j\\
-\Gamma_{il}^j &= \left(\nabla_{\frac{\partial}{\partial x^i}}dx^j\right)\left(\frac{\partial}{\partial x^l}\right).
\end{aligned}$$

This implies that

$$\nabla_{\frac{\partial}{\partial x^i}}dx^j \;=\; -\Gamma_{ik}^j dx^k.$$

If T and S are tensors, we define another Leibnitz rule by

$$\nabla_X\left(T\otimes S\right) \;=\; \left(\nabla_X T\right)\otimes S + T\otimes\nabla_X S.$$

This allows us to take covariant derivatives of any tensor. For example, let $T = T^i{}_j \frac{\partial}{\partial x^i} \otimes dx^j$. Then

$$\begin{aligned}
\nabla_X T &= \nabla_X \left(T^i{}_j \frac{\partial}{\partial x^i} \otimes dx^j \right) \\
&= X(T^i{}_j)\, \frac{\partial}{\partial x^i} \otimes dx^j \,+\, T^i{}_j \left(\nabla_X \frac{\partial}{\partial x^i} \right) \otimes dx^j \,+\, T^i{}_j \frac{\partial}{\partial x^i} \otimes \left(\nabla_X dx^j \right).
\end{aligned}$$

Another way to express this Leibnitz rule is as follows:

$$\nabla_X \left(T(Y_1, \ldots, Y_n, \omega_1, \ldots, \omega_n) \right) \;=\; \left(\nabla_X T \right) (Y_1, \ldots, Y_n, \omega_1, \ldots, \omega_n) \,+\, T\left(\nabla_X Y_1, \ldots, Y_n, \omega_1, \ldots, \omega_n \right) \,+\, \ldots \,+\, T\,(Y_1,$$

For example, if R is the Riemann tensor then

$$(\nabla_W R)(X, Y)Z \;=\; \nabla_W \left(R(X, Y)Z \right) \,-\, R(\nabla_W X, Y)Z \,-\, R(X, \nabla_W Y)Z \,-\, R(X, Y)\nabla_W Z.$$

2 Covariant derivatives as tensors

The covariant derivative $\nabla_X T$ is tensorial in the first variable, meaning $\nabla_{fX} T = f \nabla_X T$. If T is an (n, k)-tensor, the tensor ∇T accepts one additional vector field, meaning ∇T is an $(n, k+1)$-tensor.

A comma is used to indicate a derivative: if $T = T_{ij}$ is a $(0, 2)$-tensor for instance, then $\nabla T = T_{ij,k}$ is a $(0, 3)$-tensor.

For example, let $X = X^i$ be a vector field, ie a $(1, 0)$-tensor. Then $\nabla X = X^i{}_{,j}$ is a $(1, 1)$-tensor. Let's compute the components $X^i{}_{,j}$.

$$\begin{aligned}
\nabla X &= \nabla \left(X^i \frac{\partial}{\partial x^i} \right) \;=\; d(X^i) \otimes \frac{\partial}{\partial x^i} \,+\, X^i \nabla \frac{\partial}{\partial x^i} \\
&= \frac{dX^i}{dx^j} dx^j \otimes \frac{\partial}{\partial x^i} \,+\, X^i \, \Gamma^k_{li} \, dx^l \otimes \frac{\partial}{\partial x^l} \\
&= \left(\frac{dX^i}{dx^j} \,+\, X^s \, \Gamma^i_{js} \right) dx^j \otimes \frac{\partial}{\partial x^i}.
\end{aligned}$$

Therefore

$$X^i{}_{,j} \;=\; \frac{dX^i}{dx^j} + X^s \, \Gamma^i_{js}$$

As another example, let $\eta = \eta_i$ be a covector field, ie a $(0, 1)$-tensor. Then $\nabla \eta = \eta_{i,j}$

is a $(0,2)$-tensor. Let's compute the components $\eta_{,ij}$.

$$\begin{aligned}
\nabla\eta &= \nabla\left(\eta_i dx^i\right) \;=\; d(eta_i)\otimes dx^i \;+\; \eta_i \nabla dx^i \\
&= \frac{\partial \eta_i}{\partial x^j} dx^j \otimes dx^i \;-\; \eta_i \Gamma^i_{kl} dx^k \otimes dx^l \\
&= \left(\frac{\partial \eta_j}{\partial x^i} \;-\; \eta_s \Gamma^s_{ij}\right) dx^i \otimes dx^j .
\end{aligned}$$

Therefore

$$\eta_{i,j} \;=\; \frac{d\eta_i}{dx^j} \;-\; \eta_s\,\Gamma^s_{ij}.$$

Lecture 28 - Curvature identities

May 1, 2009

1 Four notions of curvature

Here we collate all our definitions of curvature operators into one place.

We have the Riemann curvature tensor

$$R\,(X,\,Y)\;Z \;\;=\;\; \nabla_X \nabla_Y Z \;-\; \nabla_Y \nabla_X Z \;-\; \nabla_{[X,\,Y]} Z,$$

an $(1,3)$-tensor. In a coordinate system $\{x^i\}$ the components are

$$\begin{aligned} R \;&=\; R_{ijk}{}^{l} dx^i \otimes dx^j \otimes dx^k \otimes \frac{\partial}{\partial x^l} \\ R_{ijk}{}^{l} \;&=\; \Gamma^s_{jk}\Gamma^l_{is} \;-\; \Gamma^s_{ik}\Gamma^l_{js} \;+\; \frac{\partial \Gamma^l_{jk}}{\partial x^i} \;-\; \frac{\partial \Gamma^l_{ik}}{\partial x^j} \end{aligned}$$

Of course it is possible to lower the final index to obtain a $(0,4)$-tensor

$$\begin{aligned} R(X,\,Y,\,Z,\,W) \;&=\; \langle R(X,Y)Z,\,W\rangle \\ R_{ijkl} \;&=\; R_{ijk}{}^{s} g_{sl}. \end{aligned}$$

The full curvature operator is extremely complex and its direct geometric meaning is difficult to interpret. The first simplification is the sectional curvature, which indicated the 'bending' of some specified 2-plane in an infinitesimal region around a point:

$$\sec(X,Y) \;\;=\;\; \frac{R(X,\,Y,\,Y,\,X)}{|X|^2|Y|^2 \,-\, \langle X,\,Y\rangle^2}$$

is the *sectional curvature* of the 2-plane spanned by X and Y. Note that the sectional curvature operator is NOT a tensor.

Contracting the first and final indices of the Riemann tensor gives the Ricci curvature tensor

$$\mathrm{Ric}_{ij} \;\;\triangleq\;\; R_{sij}{}^{s} \;=\; R_{kijl}g^{kl}.$$

If X is a unit vector, then $\mathrm{Ric}(X, X)$ is essentially a multiple of the average of sectional curvatures of all planes that contain X.

Contracting again, we get the *scalar curvature*

$$s \quad \triangleq \quad \mathrm{Ric}_{ij} g^{ij}.$$

The scalar curvature is essentially the sum of all sectional curvatures at a point.

2 Curvature Identities

In class and in the homeworks, we proved the following curvature identities

$$\begin{aligned}
&R(X, Y, Z, W) \;=\; -R(Y, X, Z, W) \\
&R(X, Y, Z, W) \;=\; -R(X, Y, W, Z) \\
&R(X, Y, Z, W) \;=\; R(Z, W, X, Y) \\
&R(X, Y)Z + R(Y, Z)X + R(Z, X)Y \;=\; 0 \qquad\qquad \text{(first Bianchi identity)} \\
&(\nabla_W R)(X, Y)Z + (\nabla_X R)(Y, W)Z + (\nabla_Y R)(W, X)Z \;=\; 0 \qquad \text{(second Bianchi identity)}
\end{aligned}$$

In components, these read

$$\begin{aligned}
&R_{ijkl} \;=\; -R_{jikl} \\
&R_{ijkl} \;=\; -R_{ijlk} \\
&R_{ijkl} \;=\; R_{klij} \\
&R_{ijkl} + R_{jkil} + R_{kijl} \;=\; 0 \\
&R_{ijkl,s} + R_{jskl,i} + R_{sikl,j} \;=\; 0.
\end{aligned}$$

Further identities can be derived by 'mixing and matching' these identities. For instance the second Bianchi identity is frequently presented

$$R_{ijkl,s} + R_{ijls,k} + R_{ijsk,l} \;=\; 0,$$

which can be derived from our second Bianchi identity and the rule $R_{ijkl} = R_{klij}$.

Lecture 29 - Conservation

May 4, 2009

1 The stress-energy tensor is symmetric

In class we gave an argument for the symmetry of the Cauchy stress tensor, which goes as follows. An "infinitesimal" volume element (side length l) has the following forces perpendicular to the z-axis:

$$-l^2\vec{T}_{12} - l^2\vec{T}_{12} + l^2\vec{T}_{21} + \vec{T}_{21}.$$

Thus the torque τ_z about the z-axis is $\frac{l}{2}\left(l^2\vec{T}_{21} - l^2\vec{T}_{12}\right)$ (torque=force×lever arm). Angular acceleration θ_z about the z-axis is related to torque about the z-axis τ_z and the second moment of inertia about the z-axis I_z as follows:

$$\tau_z = I_z\ddot{\theta}_z$$

(note the formal similarity to Newton's second law). The second moment of inertia is roughly a constant times mass times (linear dimension)^2, so $I_z = \alpha \cdot \rho l^3 \cdot l^2$ where ρ is the material's density. Thus

$$\ddot{\theta}_z = \frac{\vec{T}_{21} - \vec{T}_{12}}{\alpha\rho l^2}.$$

Since l was assumed to be very small, it follows that $\vec{T}_{12} = \vec{T}_{21}$. Repeating the argument using torque about the x- and y-axes gives $\vec{T}_{13} = \vec{T}_{31}$ and $\vec{T}_{23} = \vec{T}_{32}$. Thus the Cauchy stress tensor is symmetric.

The full stress-energy-momentum tensor is also symmetric, as can be seen by the equivalence of momentum density and energy flux density.

2 Conservation of energy

Imagine again the infinitesimal volume element of sidelength l. The total energy (aka mass) inside the volume element is l^3T_{00}. The energy (aka mass) leaving or entering the volume element must be related to the energy flux density (the elements T_{01}, T_{02}, and T_{03} of the stress-energy tensor).

The total energy leaving the volume element in the x-direction is

$$-l^2 T_{01}\big|_{x=l/2} + l^2 T_{01}\big|_{x=-l/2}.$$

The total energy leaving the volume element in the y-direction is

$$-l^2 T_{02}\big|_{y=l/2} + l^2 T_{02}\big|_{y=-l/2}.$$

The total energy leaving the volume element in the x-direction is

$$-l^2 T_{03}\big|_{z=l/2} + l^2 T_{03}\big|_{z=-l/2}.$$

The law of conservation of energy states that the time rate-of-change of energy in a volume equals the rate of flow of energy entering or leaving the volume. Restating this mathematically, we get

$$\begin{aligned}
l^3 \frac{\partial T_{00}}{\partial t} &= -l^2 \left(T_{01}\big|_{x=l/2} - T_{01}\big|_{x=-l/2} \right) - l^2 \left(T_{02}\big|_{y=l/2} - T_{02}\big|_{y=-l/2} \right) - l^2 \left(T_{03}\big|_{z=l/2} - T_{03}\big|_{z=-l/2} \right) \\
\frac{\partial T_{00}}{\partial t} &= -\frac{T_{01}\big|_{x=l/2} - T_{01}\big|_{x=-l/2}}{l} - \frac{T_{02}\big|_{y=l/2} - T_{02}\big|_{y=-l/2}}{l} - \frac{T_{03}\big|_{z=l/2} - T_{03}\big|_{z=-l/2}}{l}.
\end{aligned}$$

Since l is considered 'infinitesimally small" (which is essentially to say that we are taking a limit as l goes to zero), the right-hand side consists of derivatives. We get

$$\frac{\partial T_{00}}{\partial t} = -\frac{\partial T_{01}}{\partial x} - \frac{\partial T_{02}}{\partial y} - \frac{\partial T_{03}}{\partial z}.$$

3 Conservation of momentum

Thus conservation of energy can be restated

$$\frac{\partial T_{00}}{\partial x^0} + \frac{\partial T_{01}}{\partial x^1} + \frac{\partial T_{02}}{\partial x^2} + \frac{\partial T_{03}}{\partial x^3} = 0.$$

Conservation of momentum is as important to physics as conservation of energy. The above argument can be repeated from the point of view of total momentum contained in a volume element instead of total energy, and by looking at momentum fluxes instead of energy fluxes. One obtains the statement of conservation of momentum:

$$\frac{\partial T_{i0}}{\partial x^0} + \frac{\partial T_{i1}}{\partial x^1} + \frac{\partial T_{i2}}{\partial x^2} + \frac{\partial T_{i3}}{\partial x^3} = 0.$$

where i is either 1, 2, or 3.

4 Conservation of energy-momentum in arbitrary (ie curved) spacetimes

The arguments above are valid whenever it makes sense to take partial derivatives of tensors, namely, when considering flat space-time. If the space-time metric is not flat (which, according to the Einstein equations it essentially never is), then the operation of taking partial derivatives of tensors is dependent on the particular coordinate system.

To make differential expressions make sense in curved space-times, one must replace partial derivatives by covariant derivatives. Thus

$$\frac{\partial T_{ij}}{\partial x^k}$$

is replaced by

$$T_{ij,k} \triangleq \left(\nabla_{\frac{\partial}{\partial x^k}} T\right)\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right).$$

Thus the statement of conservation of energy-momentum is

$$g^{jk} T_{ij,k} = 0$$

or equivalently

$$T_{i,k}^{k} = 0 \quad \text{or even} \quad T^{ik}{}_{,k} = 0.$$

Lecture 30 - Equations of motion for relativistic fluids, and Poisson's equation

May 4, 2009

1 Explicit stress-energy-momentum tensors

First consider the example of dust. This is defined to be a 'fluid' comprised of noninteracting particles. Thus it can sustain no pressure and no shear. If ρ is the density and $U = U^i$ is the velocity of the dust, the stress-energy-momentum tensor is

$$\begin{aligned} T &= \rho U_\flat \otimes U_\flat \\ T_{ij} &= \rho U_i U_j. \end{aligned}$$

If a frame is chosen so that $U = \frac{\partial}{\partial x^0}$ (ie, if you are working in the frame in which the dust appears stationary) then

$$T_{ij} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now consider the case of 'ideal fluids', namely fluids with no viscosity. Such fluids can sustain pressure, but no shear (hence no viscosity). If ρ is density, P is pressure, and $U = U^i$ is the fluid's velocity, thee stress-energy-momentum tensor is

$$\begin{aligned} T &= (\rho + P)\, U_\flat \otimes U_\flat - Pg \\ T_{ij} &= \left(\rho + \frac{P}{c^2}\right) U_i U_j - \frac{P}{c^2} g_{ij} \\ T^{ij} &= \left(\rho + \frac{P}{c^2}\right) U^i U^j - \frac{P}{c^2} g^{ij}. \end{aligned}$$

Working in the frame in which the fluid appears stationary, then

$$T^{ij} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{pmatrix}.$$

2 Directional derivatives

If $X = X^i$ is a vector field, then it "gradient" is denoted

$$\nabla X = X^i{}_{,j} = X^i{}_{,j} \frac{\partial}{\partial x^i} \otimes dx^j.$$

If U^i is a vector, the derivative of X in the direction of U is

$$\begin{aligned} \nabla_U X &= (\nabla X)(U) = \left(X^i{}_{,j} \frac{\partial}{\partial x^i} \otimes dx^j \right) \left(U^k \frac{\partial}{\partial x^k} \right) \\ &= X^i{}_{,j} U^k \frac{\partial}{\partial x^i} \, dx^j \left(\frac{\partial}{\partial x^k} \right) = X^i{}_{,j} U^k \frac{\partial}{\partial x^i} \delta^j_k \\ &= X^i{}_{,k} U^k \frac{\partial}{\partial x^i}. \end{aligned}$$

3 Conservation laws imply the equations of motion

Recall that conservation of energy-momentum can be restated

$$g^{jk} T_{ij,k} = 0,$$

which can also be written $T_i{}^j{}_{,j} = 0$ or $T^{ij}{}_{,j} = 0$.

Consider the case of the perfect fluid from above. In addition to conservation of mass-energy, a fluid preserves the number of particles it contains. If n is the particle density, consider the flux number $N^i = nU^i$. The conservation of particle number is

$$N^i{}_{,i} = 0 \qquad \left(nU^i\right)_{,i} = 0.$$

Conservation of energy-momentum yields

$$
\begin{aligned}
0 \;&=\; T^{ij}{}_{,j} \;=\; \big((\rho+P)U^iU^j\big)_{,j} \;-\; \big(Pg^{ij}\big)_{,j} \\
&=\; \left(\frac{\rho+P}{n}U^i nU^j\right)_{,j} \;-\; \big(Pg^{ij}\big)_{,j} \\
&=\; \left(\frac{\rho+P}{n}U^i\right)_{,j} nU^j \;+\; \frac{\rho+P}{n}U^i\left(nU^j\right)_{,j} \;-\; P_{,j}g^{ij} \;-\; Pg^{ij}{}_{,j} \\
&=\; n\left(\frac{\rho+P}{n}U^i\right)_{,j} U^j \;-\; P^{,i}
\end{aligned}
$$

In invariant language, we cen state this

$$
n\nabla_U\left(\frac{\rho+P}{n}U\right) \;=\; \nabla P.
$$

This is precisely analogous to the classical Euler equation in Fluid mechanics:

$$
\rho\left(\frac{d\vec{v}}{dt} \;+\; \vec{v}\cdot\nabla\vec{v}\right) \;=\; \nabla P.
$$

4 Poisson's equation

Consider the equation

$$
\nabla\cdot\vec{E} \;=\; 4\pi\epsilon^{-1}\rho.
$$

In the electrostatic case, we have $E = \nabla\varphi$, so this equation reads

$$
\triangle\varphi \;=\; 4\pi\epsilon^{-1}\rho,
$$

where

$$
\triangle \;=\; \nabla\cdot\nabla \;=\; \frac{\partial}{\partial x}\frac{\partial}{\partial x} \;+\; \frac{\partial}{\partial y}\frac{\partial}{\partial y} \;+\; \frac{\partial}{\partial z}\frac{\partial}{\partial z}
$$

is called the Laplace operator, aka the Laplacian.

By physical reasoning, it is clear that a given distribution distribution of charge (as encoded by the charge density ρ) will lead to a definite electrical potential φ (at least up to an additive constant). Mathematically, this means that if the function ρ is specified (at least in a 'physically reasonable' fashion), then the equation

$$
\triangle\varphi \;=\; \alpha\rho
$$

can be solved for φ. The equation of the prescribed Laplacian is called *Poisson's equation.*

Notice that $\triangle$ is the trace of the Hessian. Given a function f, the Hessian is defined to be the matrix of second derivatives:

$$\nabla^2 f \;=\; f_{,ij} \;=\; \begin{pmatrix} f_{,11} & f_{,12} & \cdots & f_{,1n} \\ f_{,21} & f_{,22} & \cdots & f_{,2n} \\ \vdots & & \ddots & \vdots \\ f_{,n1} & f_{,n2} & \cdots & f_{,nn} \end{pmatrix}.$$

The Laplacian is the trace of the Hessian

$$tr\left(\nabla^2 f\right) \;=\; g^{ij} f_{,ij}.$$

When $g^{ij} = \delta^{ij}$, ie we are in Euclidean space, this is clearly just the classical Laplacian.

In Minkowski space, we have $g^{00} = 1$ but $g^{ii} = -c^2$ for $i \neq 1$. In this case

$$tr\left(\nabla^2 f\right) \;=\; g^{ij} f_{,ij} \;=\; f_{,00} \,-\, c^2 f_{,11} \,-\, c^2 f_{,22} \,-\, c^2 f_{,33}$$

so that $tr\left(\nabla^2 f\right)$ is the D'Alembertian, the Lorentzian analogue of the Laplacian.

Notice that this is a wave operator: $\Box f = 0$ is the wave equation for f.

Lecture 30 - The relativistic Maxwell equations, and the gravitational field equations

May 11, 2009

1 Maxwell's equations

In the Lorentzian context, the fundamental object of electromagnetics is the Faraday 2-form, F. Only after a frame is chosen is it possible to tease apart the $\vec{E}$ and $\vec{B}$ fields.

Let's first consider the classical equations

$$\begin{aligned} \nabla \cdot \vec{B} &= 0 \\ \nabla \times \vec{E} + \frac{\partial \vec{B}}{\partial t} &= 0. \end{aligned}$$

As we have described, these equations imply the existence of the two potentials of classical electrodynamics: the electrical pseudopotential φ and the magnetic vector potential $\vec{A}$.

Switching to the relativistic context, it was shown in the homeworks that these two equations lead to the single equation

$$dF = 0.$$

As we discussed in class, though we did not prove it, the equation $dF = 0$ implies that $F = dA$ for some 1-form $A = A_i dx^i$[1]. This is called the electromagnetic 4-potential. Note the following: if A is changed by adding df where f is any function, then the equation $F = dA$ is unchanged:

$$d(A + df) = dA + ddf = dA = F.$$

This is the new (and simpler!) statement of gauge invariance: the 4-potential (called the gauge) can by modified by adding df.

[1] Actually the assertion "$dF = 0$ implies $F = dA$" is valid only when a certain topological obstruction is not present. In the case of space-time, the obstruction is indeed not present.

Next let's consider the classical equations

$$
\begin{aligned}
\nabla \cdot \vec{E} &= \frac{4\pi}{\epsilon}\rho \\
\nabla \times \vec{B} - \epsilon\mu\frac{\partial \vec{E}}{\partial t} &= 4\pi\mu\vec{J}.
\end{aligned}
$$

As we have described, these lead to wave equations for the quantities $\vec{E}$ and $\vec{B}$, and, once an appropriate gauge has been chosen, for ϕ and $\vec{A}$ as well.

Switching to the relativistic context, it was shown in the homeworks that these two equations lead to the single equation

$$
\delta F = -\frac{4\pi}{\epsilon}c^2 J_\flat
$$

where J is the 4-current.

Before moving on, let's discuss a bit of mathematics. Recall from the homework that if f is a function, then $\delta df = \Box f$ where $\Box$ is the D'Alembertian. If ω is any k-form, then the D'Alembertian is no longer just δd. It turns out that $(d\delta + \delta d)\omega = \Box\omega$. Note that, for functions, $\delta f = 0$ so this equation holds for functions as well.

Returning to electromagnetics, we have $F = dA$ so that

$$
\begin{aligned}
\delta F &= -4\pi\epsilon^{-1}c^2 J_\flat \\
\delta dA &= -4\pi\epsilon^{-1}c^2 J_\flat.
\end{aligned}
$$

This is 'half' of a wave equation for A. If we work in the Lorentz gauge (meaning we choose A so that $\delta A = 0$), then

$$
\Box A = d\delta A + \delta dA = \delta dA = -4\pi\epsilon^{-1}c^2 J_\flat
$$

is a wave equation for A, with source $J_\flat$. Of course if $J_\flat = 0$, then we have just the wave equation: $\Box A = 0$.

2 Comments on electromagnetics

Classically one studies two (semi)independent objects: the $\vec{E}$-field and the $\vec{B}$-field. The relativistic field theory is a *unification* of electromagnetics: one studies the single field F.

The new Maxwell equations $dF = 0$ and $\delta F = \alpha J_\flat$ reduce back to the old equations only when a reference frame is chosen. However, if $\vec{E}$ and $\vec{B}$ are the fields as observed in the lab-frame, particles in different frames do not see the electric and magnetic fields the same way, and one will get incorrect results by applying the $\vec{E}$ and $\vec{B}$ fields to such moving particles.

In the low-energy situation, where velocities are small compared to c, the different particles under consideration all exist (approximately) in the same reference frame and observe (approximately) the same electrical and magnetic fields $\vec{E}$ and $\vec{B}$. In this situation we say that $\vec{E}$ and $\vec{B}$ *freeze out* from the Faraday 2-form, and the classical theory applies.

3 Traced Bianchi identities

The second Bianchi identity has consequences for Ricci and scalar curvatures. Recall that

$$R_{ijkl,m} + R_{ijlm,k} + R_{ijmk,l} = 0.$$

If we trace along jk (that means multiply both sides by g^{jk} and sum), we get

$$\begin{aligned} & g^{jk} R_{ijkl,m} - g^{jk} R_{jilm,k} + g^{jk} R_{ijmk,l} = 0 \\ & \mathrm{Ric}_{il,m} - R^{k}{}_{ilm,k} - \mathrm{Ric}_{im,l} = 0 \\ & \mathrm{Ric}_{il,m} - \mathrm{Ric}_{im,l} = R^{k}{}_{ilm,k}. \end{aligned}$$

This is the *traced second Bianchi identity*. Tracing again along il, we get

$$\begin{aligned} & g^{il} \mathrm{Ric}_{il,m} - g^{il} \mathrm{Ric}_{im,l} = g^{il} R^{k}{}_{ilm,k} \\ & \mathrm{Ric}^{l}{}_{l,m} - \mathrm{Ric}^{l}{}_{m,l} = \mathrm{Ric}^{k}{}_{m,k} \\ & s_{,m} = 2\mathrm{Ric}^{k}{}_{m,k}. \end{aligned}$$

In schematic form, this reads $grad\, Scal = 2\, div\, Ric$. This is known as the *doubly-traced second Bianchi identity*.

4 Gravity

Einstein argued from the *equivalence principle*. Varied 'interpretations' of the equivalence principle are possible, but one may consider it to be the principle that particles are not accelerated by gravity, rather mass curves space-time, and particles move along unaccelerated paths in a curved space-time.

Second, there is no true 'derivation' of the gravitational field equations; strictly speaking the field equations are axiomatic. However if relativistic gravity is to reduce to Newtonian gravity in the low-energy case, then we should be able to garner some 'hints' by carefully examining Newton's law of gravity.

In the Newtonian theory, one has a gravitational potential φ, which is related to force via $\vec{F} = -\nabla\varphi$. Or in Einstein notation

$$\vec{F}_\alpha = -\varphi_{,\alpha}.$$

The potential is 'generated' by mass via Newton's law of gravity

$$\triangle\varphi \;=\; 4\pi G\rho$$

where G is the gravitational constant and ρ is mass density.

Now consider a geometric 'thought experiment', where a number of particles situated different distances from a massive body are allowed free gravitational fall. The situation in 4-dimensional space-time shows we may consider this a variation of paths, with direction field $\frac{\partial}{\partial t}$ and variation field $\frac{\partial}{\partial s}$ (sorry I am not good enough with latex to be able to draw nice pictures here). Let us compute the Jacobi equation:

$$\begin{aligned}
\frac{\partial}{\partial s} &= \frac{\partial x^i}{\partial s}\frac{\partial}{\partial x^i}\\
\frac{\partial}{\partial t}\frac{\partial}{\partial t}\frac{\partial}{\partial s} &= \frac{\partial}{\partial t}\frac{\partial}{\partial t}\frac{\partial x^i}{\partial s}\frac{\partial}{\partial x^i}\\
&= \frac{\partial}{\partial s}\frac{\partial}{\partial t}\frac{\partial x^i}{\partial t}\frac{\partial}{\partial x^i},
\end{aligned}$$

because mixed partials commute. Now note that $\frac{\partial^2 x^i}{\partial t^2}$ is the i^{th} component of acceleration, which is related to force: $\vec{F}_i \;=\; -\frac{\partial\varphi}{\partial x^i}$. Thus

$$\begin{aligned}
\frac{\partial}{\partial t}\frac{\partial}{\partial t}\frac{\partial}{\partial s} &= -\frac{\partial}{\partial s}\frac{\partial\varphi}{\partial x^i}\frac{\partial}{\partial x^i}\\
&= -\frac{\partial x^j}{\partial s}\frac{\partial}{\partial x^j}\frac{\partial\varphi}{\partial x^i}\frac{\partial}{\partial x^i}\\
&= -\frac{\partial x^j}{\partial s}\frac{\partial^2\varphi}{\partial x^j\partial x^i}\frac{\partial}{\partial x^i}.
\end{aligned}$$

So much for the classical situation. Now consider the covariant situation, obtained by replacing partial derivatives with covariant derivatives. The Jacobi equation gives us

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial t}}\nabla_{\frac{\partial}{\partial t}}\frac{\partial}{\partial s} &= -R\left(\frac{\partial}{\partial s},\frac{\partial}{\partial t}\right)\frac{\partial}{\partial t}\\
&= -\frac{\partial x^j}{\partial s}R\left(\frac{\partial}{\partial x^j},\frac{\partial}{\partial t}\right)\frac{\partial}{\partial t}\\
&= -\frac{\partial x^j}{\partial s}R_{j00}{}^{i}\frac{\partial}{\partial x^i}.
\end{aligned}$$

Comparing this to the classical calculation, it should be the case that

$$\frac{\partial^2\varphi}{\partial x^i\partial x^j} \;=\; R_{i00}{}^{j}.$$

Now the Newton equation is

$$\sum_i \frac{\partial^2\varphi}{\partial x^i\partial x^i} \;\equiv\; \triangle\varphi \;=\; 4\pi G\rho,$$

so it should be the case that the covariant gravitational equation is something like

$$\mathrm{Ric}_{00} \;=\; R_{i00}{}^{i} \;=\; 4\pi G\rho,$$

that is, the gravitational field equations should be equations of *prescribed Ricci curvature*.

However $\mathrm{Ric}_{00} = 4\pi G\rho$ is not a lorentz-invariant expression: it depends on the choice of a time-vector. The Ricci tensor should not rely just the mass distribution, but on the entire stress-energy-momentum tensor T_{ij}, which is a Lorentz-invariant object.

Thus we are led to Einstein's first guess as to the field equations:

$$\mathrm{Ric}_{ij} \;=\; 4\pi G T_{ij} \qquad (??).$$

It turns out that this must be wrong, essentially for mathematical reasons: if one takes the divergence of both sides, then

$$\begin{aligned}\mathrm{Ric}_{ij,j} &= 4\pi G T_{ij,j}\\ \frac{1}{2}s_{,i} &= 0.\end{aligned}$$

(abusing Einstein notation here: one of the j^s on each side should be lifted). The left side was simplified using the doubly-traced second Bianchi identity, and the right side was simplified using conservation of energy-momentum. But $ds = s_{,i} = 0$ means that scalar curvature is constant, which is impossible since $s = g^{ij}\,\mathrm{Ric}_{ij} = g^{ij}T_{ij}$ is the trace of the stress-energy-momentum tensor, and this tensor, which represents the distribution of mass, energy, and forces in space, is by-and-large arbitrary (subject only to the conservation law).

Therefore our first 'guess' at the field equations is wrong. We therefore modify the equations:

$$\mathrm{Ric}_{ij} \;-\; \frac{s}{2}g_{ij} \;=\; 4\pi G T_{ij}.$$

Taking divergences of both sides, we get

$$\begin{aligned}\mathrm{Ric}_{ij,j} \;-\; \left(\frac{s}{2}g_{ij}\right)_{,j} &= 4\pi G T_{ij,j}\\ \frac{1}{2}s_i \;-\; \frac{s_{,j}}{2}g_{ij} \;-\; \frac{s}{2}g_{ij,j} &= 0\\ \frac{1}{2}s_i \;-\; \frac{s_{,i}}{2} &= 0,\end{aligned}$$

which checks out (recall that g is covariant-constant: $g_{ij,k} = 0$).

Notice what we have done. We 'guessed' at field equation $\mathrm{Ric}_{ij} = 4\pi G T_{ij}$ using a physical reasoning (ie comparing the Newtonian gravitation). It turned out to be demonstrably invalid, essentially for mathematical reasons. We then ad hoc appended a term that fixed the mathematical defect. In fact, this was more-or-less Einstein's process!

The question is, what other terms could one possibly append? We will not prove this, but there turns out to be only one further possible modification: adding a multiple of the metric. One gets

$$\mathrm{Ric}_{ij} \ - \ \frac{s}{2}g_{ij} \ + \ \Lambda g_{ij} \ = \ 4\pi G T_{ij}.$$

If we take a divergence of both sides, we get

$$\begin{aligned} &\left(\mathrm{Ric}_{ij} \ - \ \frac{s}{2}g_{ij}\right)_{,j} \ + \ (\Lambda g_{ij})_{,j} \ = \ 4\pi G T_{ij,j} \\ &0 \ + \ \Lambda_{,j} g_{ij} \ = \ 0 \end{aligned}$$

(we already computed the divergence of all terms except the new one). Thus $\Lambda_{,i} = 0$, so that Λ must be a constant. This is Einstein's famous *Cosmological Constant*.

Therefore we can write Einstein's Gravitational Field Equations

$$\mathrm{Ric}_{ij} \ - \ \frac{s}{2}g_{ij} \ + \ \Lambda g_{ij} \ = \ 4\pi G T_{ij},$$

where Λ is an arbitrary (that is, predetermined) constant. This is essentially an equation of *prescribed Ricci curvature*: the distribution of matter, energy, and forces (encoded in T_{ij}) determines the Ricci curvature of space. Since Ricci curvature is 'in principle' the D'Alembertian of the metric, it is reasonable to conclude that that one can solve for the metric once the Ricci curvature has been set.

MAT 401: The Geometry of Physics

Spring 2009

Department of Mathematics
SUNY at Stony Brook

Welcome to The Geometry of Physics

In this class we will develop the mathematical language needed to understand Einstein's field equations. The infinitesimal structure of any space, even curved space, is Euclidean, and so is described with linear algebra. Calculus, in the form of continuity and differentiability properties of paths and surfaces, can express the connectedness of space. The synthesis of these points of view, of the infinitesimal with the global, of linear algebra with calculus, yields the powerful language of differential geometry, which Einstein used to express the physics of General Relativity.

Course Content

Before studying the field equations we must develop the language of geometry. We will try to integrate intuitive content with hard mathematics, and some of the topics will be partly review for many students. But hard work will be required... it took Einstein more than 2 years to understand the mathematics we will cover in a semester.

Homework assignment page   Class notes   Quiz Prep (including final exam info)

Announcements

- There will be a makeup class on Monday (May 11) in P-131 in the math building. We will go over the gravitational field equations.
- I be in my office on Tuesday the 12th, from 2-4pm and 5-7pm

Course Information:

Check out the topics we will cover...

Here is a link to the syllabus.

Textbook

A first Course in General Relavity by Bernard F. Schutz

Supplimentary books / Recommended reading

The Geometry of Physcis by Theodore Frankel, Second Edition
The Large Scale Structure of Space-Time by G. Ellis and S. Hawking
General Relativity by Robert Wald

Course Grading

One homework assugnment will be due each Wednesday.

Homeworks: 10% of total grade
Quizes: 10% of total grade
Test 1: 10% of total grade (Friday Feb 13)
Test 2: 20% of total grade (Friday Mar 6)
Test 3: 10% of total grade (Friday Mar 20)
Test 4: 10% of total grade (Friday April 17)
Final Exam: 30% of total grade

Your instructor is Brian Weber,

Office: 3-121 Math Tower

Course Prerequisites

Calculus IV, Math 305 or equivalent (differential equations)
Linear Algebra, Math 310 or equivalent

Americans with Disabilities Act

If you have a physical, psychological, medical or learning disability that may impact your course work, please contact Disability Support Services, ECC (Educational Communications Center) Building, room 128, (631) 632-6748 or http://studentaffairs.stonybrook.edu/dss/**. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential. Students who requiring assistance during emergency evacuation are encouraged to discuss their needs with their professors and Disability Support Services. For procedures and information, go to the following web site: http://www.www.ehs.stonybrook.edu/fire/disabilities.asp**

Problem set 1

Due Feb 4

Problem 1) An electron flies by at half the speed of light. Mounted to the wall immediately behind it is a meter stick.

a) How much time does it take for the electron to traverse the meter stick?

b) From the electron's perspective, how fast is the meter stick moving?

c) From the electron's perspective, how much time does it take the meter stick to fly by?

d) From the electron's perspective, what is the length of the meter stick?

Problem 2) Consider the Minkowski plane $R^{1,1}$, and let p be the event $(2, -c)$. Given a massive particle located at space-time location p, which, if any, of the following events is it physically possible for it to reach? Justify your conclusions.

$$\begin{aligned} q_1 &= (2, 2c) \\ q_2 &= (4, c) \\ q_3 &= (7/2,\ -c/2) \end{aligned}$$

Problem 3) Let $\mathbf{\Lambda}(\tau) = \left(\tau,\ \frac{2}{3}c^2\tau^2\right)^T$ be a path through 2-space.

a) Find the pathlength, $0 < \tau < 1$, in Euclidean 2-space.

b) Find the pathlength, $0 < \tau < 1$, in Minkowski 1+1-space

Problem 4) Let $\mathbf{\Lambda}(\tau) = (\tau,\ c\ \sin(\tau),\ c\ \cos(\tau))^T$ through 3-space.

a) Find the pathlength, $0 < \tau < 1$, in Euclidean 3-space.

b) Find the pathlength, $0 < \tau < 1$, in Minkowski 1+2-space.

Problem 5) Consider an observer labeled o in the Minkowski plane, $\mathbb{R}^{1,1}$. Consider the following events:

- p_1 occurred 1s in the past, a distance of 2c to the left of o
- p_2 occurred 1s in the past, a distance of 5c to the right of o
- p_3 occurs 2s in the future, a distance of c to the right of o.

Draw a space-time diagram with p_1, p_2, and p_3 at the corners of a triangle. What are the lengths, in the Minkowski sense, of the sides of the triangle? Which of these events are separated by a space-like interval? a time-like interval?

Problem 6) Consider an observer o', located at the same space-time position as o, but traveling at $\frac{4}{5}c$ relative to o. Make a space-time diagram from the perspective of o' with the events p_1, p_2, p_3 again forming the corners of a triangle. Find and label the Minkowski lengths of the sides of the triangle. According to o', which event occurred first?

Problem set 2

Due Feb 11

Problem 1) Let $\{x^0, x^1\}$ be coordinates for $\mathbb{R}^{1,1}$; this will be the lab frame. A second observer is traveling at $\frac{3}{5}c$ with respect to the lab, and places $\{y^0, y^1\}$ coordinates on space-time (with the same origin as the $\{x^0, x^1\}$ coordinates). A third observer is traveling at $\frac{12}{13}c$ with respect to the lab, and places $\{z^0, z^1\}$ coordinates on space-time (with the same origin as the $\{x^0, x^1\}$ coordinates).

a) Make a sketch of space-time according using the $\{x^0, x^1\}$ coordinate system. As accurately and neatly as possible, superimpose the $\{y^1, y^2\}$ axes on your sketch. To do this, first find and label the points

$$p_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\{y^i\}} \qquad p_1 = \begin{pmatrix} 0 \\ c \end{pmatrix}_{\{y^i\}}.$$

b) On the same graph, superimpose the $\{z^1, z^2\}$ axes. To do this, first find and label the points

$$q_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_{\{z^i\}} \qquad q_1 = \begin{pmatrix} 0 \\ c \end{pmatrix}_{\{z^i\}}.$$

c) Superimpose the light cone and the pseudospheres of radius 1, 2, and 3.

**YOUR GRAPH MUST BE VERY NEAT AND VERY LARGE.

Problem 2) (*The Twin Paradox*) A traveler leaves the Earth traveling at speed v. After reaching a distance of d as measured by a stationary observer on Earth, the traveler immediately turns around and returns to Earth at the same speed. According to an Earth-bound observer, how long was the traveler gone? According to the traveler, how much time did the trip take?

Problem 3) (*Tachyons*) Tachyons are hypothetical particles that travel faster than light. Consider two spaceships leaving the space-time point o, traveling at velocity v relative to each other. Each spaceship is equipped with a tachyon emitter, which emits a tachyon of velocity $w > c$ as measured in the emitter's rest-frame. At time l, the first ship emits a tachyon, which is received and immediately re-emitted by the second ship, and then received again by the first ship. Let the event p be the emission of the tachyon from the first ship, let the event q be the reception and re-emission of the tachyon from the second ship, and let the event p' be its reception by the first ship.

a) Let $\{x^0, x^1\}$ be the space-time coordinates of the first ship. Express p and q in the x^0-x^1 coordinate system.

b) Let $\{\xi^0, \xi^1\}$ be the space-time coordinates of the second ship. Express p, q, and p' in the ξ^0-ξ^1 coordinate system.

c) Express p' in the x^0-x^1 coordinate system.

d) Show that if $w > \frac{v\,\gamma_v}{\gamma_v - 1}$, the first ship received the tachyon before it was emitted.

Problem 4) (*The velocity addition formula*) Assume an observer o' travels at speed v with respect to the lab. Of course o' sees the lab moving with speed $-v$. Assume, in addition, that o' sees a particle traveling in the direction opposite the lab's direction, with speed w. Prove that the lab observes the particle moving with speed

$$\frac{v + w}{1 + \frac{vw}{c^2}}.$$

(Hint: You may work in $\mathbb{R}^{1,1}$, instead of the full $\mathbb{R}^{1,3}$. First draw the situation from the perspective of o', then make a transformation from o''s frame back to the lab frame.)

Problem 5) ($\mathcal{O}(k,n)$ *is a group*) In this problem we will prove that the set $\mathcal{O}(k,n)$ is indeed a group. You may consider $\mathcal{O}(k,n)$ to be the matrix group consisting of $(k+n)\times(k+n)$ matrices A so that $A^T I_{k,n} A = I_{k,n}$. Recall that I_m is the $m \times m$ identity matrix, and that the matrix A^{-1} is the inverse of the matrix A, if it exists. Prove the following:

a) If $A, B \in \mathcal{O}(k,n)$, then $AB \in \mathcal{O}(k,n)$.
b) The identity matrix I_{k+n} is in $\mathcal{O}(k,n)$.
c) If $A \in \mathcal{O}(k,n)$, then A^{-1} exists and $A^{-1} \in \mathcal{O}(k,n)$.
d) If $A, B, C \in \mathcal{O}(k,n)$, then $(AB)C = A(BC)$.

Problem 6) Let V be the vector space of the 1-variable functions spanned by $f_1 = \sin(x)$, $f_2 = \cos(x)$, $f_3 = e^x$, and $f_4 = xe^x$. Let $A : V \to V$ be the linear operator $A(\alpha) = \int \alpha\, dx$, where the antiderivative is taken to have zero constant term.

a) Which of the following belong to V? Circle all that apply.

$$1 \qquad \cos(x) - e^x + 2xe^x \qquad \sin^2(x) \qquad e^{2x}$$

b) Express A as a matrix in the basis $\{f_1, f_2, f_3, f_4\}$.
c) Express A as a matrix in the basis

$$e_1 = \sin(x) + \cos(x) \qquad e_2 = \sin(x) - \cos(x) \qquad e_3 = e^x \qquad e_4 = 2e^x - xe^x$$

d) Express the function $2\sin(x) + \cos(x) - e^x + xe^x$ in both the $\{f_i\}$ and $\{e_i\}$ bases.

Problem set 3

Due Feb 25

Problem 1) Let A^i_j and B^i_j be the 3×3 matrices whose indices are given by the formulas

$$A^i_j \;=\; i + j \qquad B^i_j \;=\; i^2 - j^2.$$

a) Write down the matrices A and B in matrix form.
b) Write down the matrices $A^i_k B^k_j$ and $A^k_j B^i_k$ in matrix form.
c) If the vector v is given by $\mathbf{v} = \alpha^i \mathbf{v}_i$ where $\alpha^i = i - 3$, compute the number $A^1_j \alpha^j$.
d) Write the vector $A^i_j \alpha^j \mathbf{v}_i$ in column-vector form.
e) If the covector w is given by $\mathbf{w} = \beta_i \mathbf{v}^i$ where $\beta_i = i + 3$, compute the number $B^j_3 w_j$.
f) Write the covector $B^j_i w_j \mathbf{v}^i$ in row-vector form.

Problem 2) Prove the associativity of matrix multiplication: $(AB)C = A(BC)$, where A is an $n \times m$ matrix, B is an $m \times l$ matrix, and C is an $l \times p$ matrix.

Hint: If you are good at using Einstein notation, this is almost trivial.

Problem 3) Let V be the vector space of quadratic polynomials in one variable. Let $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = x$, and $\mathbf{v}_3 = x^2$; these 1-variable functions will constitute the "standard basis" of V. Let $\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3\} \subset V^*$ be the basis dual to $\{\mathbf{v}_i\}$. Let $\mathbf{p} = \alpha^i \mathbf{v}_i \in V$ where $\alpha^i = i^2 + 1$.

a) Compute $\mathbf{v}^1(\mathbf{p})$, $\mathbf{v}^2(\mathbf{p})$, and $\mathbf{v}^3(\mathbf{p})$.
b) Let $A, B, C \in V^*$ be the maps

$$\begin{aligned} A(\cdot) \;&=\; \int_{-1}^{1} \cdot \, dx \\ B(\cdot) \;&=\; \int_{-1}^{1} x \cdot \, dx \\ C(\cdot) \;&=\; \int_{-1}^{1} x^2 \cdot \, dx. \end{aligned}$$

Express A, B, and C in the $\{\mathbf{v}^i\}$ basis; that is, compute the numbers A_i, B_i, and C_i and write A, B, and C in the form $A = A_i \mathbf{v}^i$, $B = B_i \mathbf{v}^i$, $C = C_i \mathbf{v}^i$.

Problem 4) Let V be as above, this time with basis vectors $\mathbf{w}_1 = 1 + x + x^2$, $\mathbf{w}_2 = 1 - x + x^2$, $\mathbf{w}_3 = 1 + x - x^2$.

a) Compute the transition matrix $A_{\{\mathbf{w}_i\} \leftarrow \{\mathbf{v}_i\}}$.
b) If $\mathbf{p} = \alpha^i \mathbf{v}_i$ where $\alpha^i = i^2 + i - 3$, compute the numbers β^i where $\mathbf{p} = \beta^i \mathbf{w}_i$.
c) Compute the transition matrix $A_{\{\mathbf{v}_i\} \leftarrow \{\mathbf{w}_i\}}$.
d) If $\mathbf{p} = \gamma^i \mathbf{w}_i$ where $\gamma^i = i^3 - 3$, compute the numbers η^i where $\mathbf{p} = \eta^i \mathbf{v}_i$.
e) Express the maps $A, B, C \in V^*$ in covector form in the $\{\mathbf{w}^i\}$ basis. That is, compute the numbers $\bar{A}_i$, $\bar{B}_i$, and $\bar{C}_i$ and write A, B, and C in the form $A = \bar{A}_i \mathbf{w}^i$, $B = \bar{B}_i \mathbf{w}^i$, $C = \bar{C}_i \mathbf{v}^i$.

Problem 5) Let V be a 3-dimensional vector space with some basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \subset V$, and dual basis $\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3\} \subset V^*$. Let $f \in V^*$ be a linear functional, given by

$$f \;=\; 2\mathbf{v}^1 \,-\, \mathbf{v}^2 \,+\, \frac{1}{2}\mathbf{v}^3$$

and let $A : V \to V$ be a linear operator, given in the $\{\mathbf{v}_i\}$ basis by

$$A \;=\; \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Define new functionals $g, h \in V^*$ by

$$\begin{aligned} g(\cdot) \;&=\; f(A(\cdot)) \\ h(\cdot) \;&=\; A(f)(\cdot) \;\triangleq\; f(A^{-1}(\cdot)). \end{aligned}$$

a) Express g as a linear combination of basis elements of V^*: that is, find the numbers g_i so that $g = g_i\mathbf{v}^i$.

b) Express h as a linear combination of basis elements of V^*: that is, find the numbers h_i so that $h = h_i\mathbf{v}^i$.

Problem 6) In class we showed the existence of a map

$$\mathcal{N} : V \to V^{**},$$

which we call the *natural isomorphism*. Recall that this map is defined as follows: if $\mathbf{v} \in V$, then $\mathcal{N}$ associates to $\mathbf{v}$ a linear functional on V^* given by

$$\begin{aligned} &\mathcal{N}(\mathbf{v}) : V^* \to \mathbb{R} \\ &\mathcal{N}(\mathbf{v})(f) \;=\; f(\mathbf{v}). \end{aligned}$$

a) Prove that $\mathcal{N}$ is in fact a map from V to V^{**}. That is, given $\mathbf{v} \in V$, prove that the map $\mathcal{N}(\mathbf{v}) : V^* \to \mathbb{R}$ is indeed a linear functional.

Hint: This is equivalent to showing that $\mathcal{N}(\mathbf{v})(\alpha f \,+\, g) \;=\; \alpha\mathcal{N}(\mathbf{v})(f) \,+\, \mathcal{N}(\mathbf{v})(g)$ *for* $\alpha \in \mathbb{R}$, $f, g \in V^*$.

b) Prove that $\mathcal{N} : V \to V^{**}$ is a linear map. That is, show $\alpha \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w} \in V$ implies

$$\mathcal{N}(\alpha\mathbf{v} + \mathbf{w}) \;=\; \alpha\mathcal{N}(\mathbf{v}) \,+\, \mathcal{N}(\mathbf{w})$$

c) Prove that the kernel of $\mathcal{N}$ is the set $\{0\} \subset V$. That is, show that if $\mathbf{v} \in V$ and $\mathcal{N}(\mathbf{v})$ is the zero functional on V^*, then in fact $\mathbf{v} = 0$.

d) If V is finite dimensional, prove that $\mathcal{N} : V \to V^{**}$ is a vector space isomorphism.

Problem set 4

Due March 4

Problem 1) Let V be the space of quadratic polynomials, with basis $\mathbf{v}_1 = 1$, $\mathbf{v}_2 = x$ and $\mathbf{v}_3 = x^2$ and dual basis $\{\mathbf{v}^i\} \in V^*$. Let $A : V \to V$ be the operator

$$A(\cdot) \quad = \quad \frac{d}{dx}(\cdot) + \frac{1}{x}\int \cdot \, dx$$

where the antiderivative is taken to have zero constant term. We can regard A as an element of $\bigotimes^{1,1} V$. Express A in the basis $\{\mathbf{v}_i \otimes \mathbf{v}^j\}$. What is the number $A^2{}_3$?

Problem 2) Let V be a vector space with basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ and dual basis $\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3\} \subset V^*$. Let the numbers T_{ij} be given by

$$\begin{pmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{pmatrix} \quad = \quad \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & -1 \\ 1 & -2 & 0 \end{pmatrix}$$

Let $T = T_{ij}\mathbf{v}^i \otimes \mathbf{v}^j \in \bigotimes^{0,2} V$. We can regard T as a map $T : V \times V \to \mathbb{R}$, or in two different ways as maps $T : V \to V^*$. Let

$$v \;=\; \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}_{\{\mathbf{v}_i\}} \quad \in \; V.$$

Compute $T(v, v)$, and compute $T(v)$ in the two different possible ways.

Problem 3) Let V, W be finite dimensional vector spaces. Prove that any linear map $A : V \to W$ can be expressed as an element of $W \otimes V^*$.

Problem 4) Let V be a 1-dimensional vector space. Let W be a finite dimensional vector space. Prove that there exists an isomorphism $V \otimes W \approx W$.

Problem 5) Let V be a two-dimensional vector space, with basis $\mathbf{v}_1, \mathbf{v}^2$ and dual basis $\mathbf{v}^1$, $\mathbf{v}^2$. Let $\mathbf{w}_1 \;=\; \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$ $\mathbf{w}_2 \;=\; \frac{1}{2}(\mathbf{v}_1 - \mathbf{v}_2)$, and let $g \in \bigotimes^{0,2} V$ be given in the $\{\mathbf{v}^i\} \subset V^*$ basis by $g = \mathbf{v}^1 \otimes \mathbf{v}^2 + \mathbf{v}^2 \otimes \mathbf{v}^1$.

a) Express the elements of the dual basis $\mathbf{w}^1, \mathbf{w}^2 \in V^*$ in terms of the elements $\mathbf{v}^1, \mathbf{v}^2 \in V^*$.

b) Prove that g is an inner product.

c) Express g in terms of the $\{\mathbf{w}^i\}$ basis.

Problem 6) Let $V = span\{\mathbf{v}^1, \mathbf{v}^2, \mathbf{v}^3, \mathbf{v}^4\}$ be a 4-dimensional vector space. Let $T = \mathbf{v}^1 \otimes \mathbf{v}^2 \otimes \mathbf{v}^3 \otimes \mathbf{v}^4$. The alternating tensor $Alt(T)$ can be expressed as a linear combination of basis elements $\mathbf{v}^i \otimes \mathbf{v}^j \otimes \mathbf{v}^k \otimes \mathbf{v}^l \in \bigotimes^{0,4} V$, namely, $Alt(T) = \widetilde{T}_{ijkl} \mathbf{v}^i \otimes \mathbf{v}^j \otimes \mathbf{v}^k \otimes \mathbf{v}^l$. What are the numbers $\widetilde{T}_{1111}$, $\widetilde{T}_{1231}$, $\widetilde{T}_{1234}$, $\widetilde{T}_{2134}$, and $\widetilde{T}_{4132}$?

Problem 7) Assume V is an n-dimensional vector space. Prove that

$$dim\left(\bigwedge^i V^*\right) \quad = \quad \frac{n!}{i!(n-i)!}.$$

Problem set 5

Due March 18

Problem 1) Let $V = span\{\mathbf{v}_1, \mathbf{v}_2\}$ be a vector space with dual space V^*. Let

$$g \;=\; 4\mathbf{v}^1 \otimes \mathbf{v}^1 \;-\; 2\mathbf{v}^1 \otimes \mathbf{v}^2 \;-\; 2\mathbf{v}^2 \otimes \mathbf{v}^1$$

be an inner product on V.

a) Determine the corresponding inner product on V^*; explicitly write it as a linear combination of basis elements $\{\mathbf{v}_i \otimes \mathbf{v}_i\} \subset \bigotimes^{2,0} V$.

b) Let $v = 2\mathbf{v}_1 - 3\mathbf{v}_2 \in V$. Compute $v_\flat$.

c) Let $f = 2\mathbf{v}^1 - 3\mathbf{v}^2 \in V^*$. Compute $f^\sharp$.

Problem 2) Let $U \subset M$ be an open set. In class we defined an operator $d : \Omega^p(U) \to \Omega^{p+1}(U)$, assuming some coordinate system was given. In this problem we will show that the definition of the d-operator is actually independent of coordinates. Assume another operator $\tilde{d} : \Omega^p(U) \to \Omega^{p+1}(U)$ exists which obeys the following four properties: whenever $\omega, \eta \in \Omega^*(U)$ and $f \in \Omega^0(U)$, it holds that

- $\tilde{d}(\omega + \eta) \;=\; \tilde{d}\omega \;+\; \tilde{d}\eta$
- $\tilde{d}(\omega \wedge \eta) \;=\; \tilde{d}\omega \wedge \eta \;+\; (-1)^{|\omega|}\omega \wedge \tilde{d}\eta$
- $\tilde{d}f(X) \;=\; X(f)$
- $\tilde{d}\tilde{d}f \;=\; 0.$

Prove that $\tilde{d} \;=\; d$.

Problem 3) If $\omega = f_i dx^i$ is a 1-form, prove directly that

$$d\omega(X, Y) \;=\; X(\omega(Y)) \;-\; Y(\omega(X)) \;-\; \omega([X, Y]),$$

where $X = X^i \frac{\partial}{\partial x^i}, Y = Y^i \frac{\partial}{\partial x^i}$ are arbitrary vector fields.

Problem 4) Let M be Euclidean 2-space with standard (x, y)-coordinates. Define

$$g \;=\; \frac{4}{(1 \,+\, x^2 \,+\, y^2)^2}\, dx \otimes dx \;+\; \frac{4}{(1 \,+\, x^2 \,+\, y^2)^2}\, dy \otimes dy.$$

Let $r \in [0, \infty)$, $\theta \in [0, 2\pi)$ be new coordinates, defined so $r^2 = x^2 + y^2$ and $\tan(\theta) = y/x$. Compute g in the new coordinates.

Problem 5) Let M, g be as in Problem 4.

a) Let $\gamma : \mathbb{R} \to M$ be $\gamma(\tau) = (\tau, \tau)^T_{\{x^i\}}$. Compute that length of the path $\gamma([-\infty, \infty])$.

b) Let $\gamma_s : [0, 2\pi) \to M$ be $\gamma_s(\tau) = (s\cos\tau, s\sin\tau)^T_{\{x^i\}}$. Given a fixed $s \in [0, \infty)$, compute the length of the path $\gamma_s([0, 2\pi))$.

Problem set 6

Due March 25

Problem 1) Let $U \subset M$ be an open set. In class we defined an operator $d : \Omega^p(U) \to \Omega^{p+1}(U)$, assuming some coordinate system was given. In this problem we will show that the definition of the d-operator is actually independent of coordinates. Assume another operator $\tilde{d} : \Omega^p(U) \to \Omega^{p+1}(U)$ exists which obeys the following four properties: whenever $\omega, \eta \in \Omega^*(U)$ and $f \in \Omega^0(U)$, it holds that

- $\tilde{d}(\omega + \eta) \;=\; \tilde{d}\omega \,+\, \tilde{d}\eta$
- $\tilde{d}(\omega \wedge \eta) \;=\; \tilde{d}\omega \wedge \eta \,+\, (-1)^{|\omega|}\omega \wedge \tilde{d}\eta$
- $\tilde{d}f(X) \;=\; X(f)$
- $\tilde{d}\tilde{d}f \;=\; 0.$

Prove that $\tilde{d} \;=\; d$.

Problem 2) If $\omega = f_i dx^i$ is a 1-form, prove directly that

$$d\omega(X,Y) \;=\; X(\omega(Y)) \,-\, Y(\omega(X)) \,-\, \omega([X,Y]),$$

where $X = X^i\frac{\partial}{\partial x^i}, Y = Y^i\frac{\partial}{\partial x^i}$ are arbitrary vector fields.
Note: use the following definition: $v \wedge w = v \otimes w - w \otimes v$ *(no factor of* $\frac{1}{2}$*).*

Problem 3) If $dim(V) = n$, we know $dim \bigwedge^p V^* = \frac{n!}{p!(n-p)!}$. Thus $dim \bigwedge^p V^* = dim \bigwedge^{n-p} V^*$ and so, at least abstractly, $\bigwedge^p V^* \approx \bigwedge^{n-p} V^*$. The *duality operator* $* : \bigwedge^p V^* \to \bigwedge^{n-p} V^*$ (also called the *Hodge*-$*$ operator or the *Hodge duality* operator) realizes this isomorphism. On Euclidean 3-space, we define the $*$ operator by

$$\begin{aligned}
* : {\textstyle\bigwedge^0} \to {\textstyle\bigwedge^3} &\qquad *(1) \;=\; dx \wedge dy \wedge dz \\
* : {\textstyle\bigwedge^1} \to {\textstyle\bigwedge^2} &\qquad *dx \;=\; dy \wedge dz,\; *dy \;=\; -dx \wedge dz,\; *dz \;=\; dx \wedge dy \\
* : {\textstyle\bigwedge^2} \to {\textstyle\bigwedge^1} &\qquad *(dx \wedge dy) \;=\; dz,\; *(dx \wedge dz) \;=\; -dy,\; *(dy \wedge dz) \;=\; dx \\
* : {\textstyle\bigwedge^3} \to {\textstyle\bigwedge^0} &\qquad *(dx \wedge dy \wedge dz) \;=\; 1
\end{aligned}$$

and extended by linearity. Prove that, if $\omega = \omega_1 dx + \omega_2 dy + \omega_3 dz$, then $g(\omega,\omega) = *(\omega \wedge *\omega)$, where g is the Euclidean metric.

Problem 4) Assume a single stationary particle of mass m breaks into two particles of equal mass $\tilde{m}$. Prove that the two particles have the same 3-velocity $\tilde{\vec{v}}$, and find $\tilde{m}$ (it is not $\frac{m}{2}$).

Problem 5) Prove that a massive particle cannot spontaneously emit a photon.

Problem set 7

Due April 15

Problem 1) Verify the Koszul formula.

Problem 2) Let $f(x) = \sqrt{1-x^2}$, $x \in [-1, 1]$ be the equation of the upper unit semicircle. Rotate this curve around the x-axis to obtain the 2-sphere $\mathbb{S}^2$.

a) Compute g in the $\{\xi, \theta\}$ system.

b) Compute g in the $\{r, \theta\}$ system.

Problem 3) Let $\varphi^2 = \varphi^2(x^1)$ and let $g = dx^1 \otimes dx^1 + \varphi^2 \, dx^2 \otimes dx^2$ be the metric on a generic warped product (in more standard terminology, x^1 is r and x^2 is θ). Compute the eight Christoffel symbols: that is, compute

$$\Gamma^1_{11}, \quad \Gamma^2_{11}, \quad \Gamma^1_{12}, \quad \Gamma^2_{12}, \quad \Gamma^1_{21}, \quad \Gamma^2_{21}, \quad \Gamma^1_{22}, \quad \Gamma^2_{22}.$$

Problem 4) Given a generic warped product with metric $g = dr \otimes dr + \varphi(r)^2 \, d\theta \otimes d\theta$. Compute the covariant derivatives $\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r}$, $\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta}$, $\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial r}$, and $\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta}$.

Problem 5) Consider the example of the sphere from problem 2. Compute explicitly the covariant derivatives

$$\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r}, \quad \nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial r}, \quad \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta}, \quad \textit{and} \quad \nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta}$$

Problem 6) Let g be the metric on $\mathbb{S}^2$ computed in problem 2b. Let $\gamma(\tau) = \left(\frac{1}{2}\tau^2, \, \tau\right)$ be a path on the sphere, expressed in (r, θ) coordinates. Compute the acceleration

$$\nabla_{\frac{d}{d\tau}} \frac{d}{d\tau}.$$

Answers to Problem Set 7

2) $g = \frac{1}{1-\xi^2} d\xi^2 \otimes d\xi^2 + (1-\xi^2)^2 d\theta \otimes d\theta$, $g = dr \otimes dr + \sin^2(r) d\theta \otimes d\theta$.

3)

$$\begin{aligned}
&\Gamma^1_{11} = \Gamma^2_{11} = 0 \\
&\Gamma^1_{12} = 0 \quad \Gamma^2_{12} = \frac{\varphi'}{\varphi} \\
&\Gamma^1_{22} = -\varphi\varphi' \quad \Gamma^2_{22} = 0
\end{aligned}$$

4)

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} &= 0 \\
\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial r} &= \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta} = \frac{\varphi'(r)}{\varphi(r)} \frac{\partial}{\partial \theta} \\
\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta} &= -\varphi(r)\varphi'(r) \frac{\partial}{\partial r}
\end{aligned}$$

5)

$$\begin{aligned}
\nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial r} &= 0 \\
\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial r} &= \nabla_{\frac{\partial}{\partial r}} \frac{\partial}{\partial \theta} = -\cot(r) \frac{\partial}{\partial \theta} \\
\nabla_{\frac{\partial}{\partial \theta}} \frac{\partial}{\partial \theta} &= \cos(r)\sin(r) \frac{\partial}{\partial r}
\end{aligned}$$

Problem set 8

Due April 22

Problem 1) Do problem 1 from *Problems in Classical Electrodynamics*.

Problem 2) Do problem 3 from *Problems in Classical Electrodynamics*.

Problem 3) Do problem 1 from *Problems in Relativistic Electrodynamics*.

Problem 4) Do problem 3 from *Problems in Relativistic Electrodynamics*.

Problem 5) Given vector fields X and Y and functions f and h, prove that

$$[fX,\, hY] \;=\; fX(h)Y \;-\; hY(f)X \;+\; fh[X,\, Y].$$

Problem 6) Prove that the Riemann curvature tensor is indeed a tensor. That is, given vector fields X, Y, and Z and a function f, prove that

a) $R(f\,X,\, Y)Z \;=\; f\,R(X,\, Y)Z$
b) $R(X,\, f\,Y)Z \;=\; f\,R(X,\, Y)Z$
c) $R(X,\, Y)f\,Z \;=\; f\,R(X,\, Y)Z$

Problem 7) Let $g = dr \otimes dr \,+\, \varphi^2 d\theta \otimes d\theta$ be the metric on the sphere, where $\varphi = \varphi(r)$ is the function computed in the previous homework. Using results from that homework,

a) Compute $R\left(\frac{\partial}{\partial\theta},\, \frac{\partial}{\partial r}\right)\frac{\partial}{\partial r}$.
b) Compute the sectional curvature $\sec\left(\frac{\partial}{\partial r},\, \frac{\partial}{\partial\theta}\right)$.

Problem set 9

Due May 6

Problem 1) Do problem 2 from *Problems in Classical Electrodynamics*.

Problem 2) Do problem 2 from *Problems in Minkowski Analysis*.

Problem 3) Do problem 4 from *Problems in Relativistic Electrodynamics*.

Problem 4) Do problem 5 from *Problems in Relativistic Electrodynamics*.

Problem 5) Prove the *First Bianchi identity*:

$$R(X,Y)Z \;+\; R(Y,Z)X \;+\; R(Z,X)Y \;=\; 0.$$

Problem 5) Do problems 1-6 from *The Hopf Fibration and the Berger Spheres*

Quiz 1 Topics

Feb 4, 2009

You must be able to define the following:

Reference frame
Proper time
$\mathbb{R}^n$, $\mathbb{R}^{k,n}$
Space-like, time-like, and null intervals
Group
D_n
$\mathcal{O}(n)$, $E(n)$, $\mathcal{O}(k,n)$
The inner product on $\mathbb{R}^n$, $\mathbb{R}^{k,n}$
Linear transformation of a vector space

You must be able to carry out the following tasks:

Compute γ_v.

Given a vector space V and a linear transformation $A : V \to V$, determine the matrix representation of A after a basis has been specified.

Compute distances between points in $\mathbb{R}^n$ or in $\mathbb{R}^{k,n}$.

Compute the arclength of paths in $\mathbb{R}^n$ or in $\mathbb{R}^{k,n}$.

Compute boost and rotation matrices.

Quiz 2 Topics

Feb 20, 2009

You must be familiar with vector spaces and their duals.

You must be able to construct a basis for V^* dual to a chosen basis of V.

You must be able to express vectors and covectors as linear combinations of basis elements.

You must be familiar with Einstein summation notation.

You must be familiar with the action of operators on vectors and covectors.

Quiz 3 Topics

Feb 27, 2009

You must be familiar with the following:

The tensor product.

The spaces $\bigotimes^{r,s} V$.

The tensor algebra $\bigotimes^{*,*} V$.

Expressing tensors in the Einstein notation.

How tensors behave under change of bases.

Regarding tensors as linear maps.

The definition of inner products.

Test 2 Topics

Feb 27, 2009

1 Required Knowledge

If V is a vector space, what is V^*?

Given a basis, how can a basis for V^* be defined?

How can an element v of V be interpreted as a functional on V^* (that is, a map $V^* \to \mathbb{R}$)?

If $A : V \to V$ is an operator on V, how do we define the operation of A on V^*?

If A is a change-of-basis matrix for a vector space V, what is the corresponding change-of-basis matrix for V^*?

If V and W are vector spaces, how is $V \otimes W$ defined?

What is the definition of $\bigotimes^{r,s} V$? of $\otimes^{*,*}V$?

If T is a tensor (of whatever rank) and A is a change-of-basis, how is T expressed in the new basis?

In what different ways can a tensor T be interpreted as a map?

If A is an operator, say $A : V \to V$ (or $A : V \to V^*$ or $A : V \to V \otimes V^*$ or ...), how can A be expressed as a tensor?

What is a symmetric tensor? an antisymmetric tensor?

What is the definition of $\bigwedge^r V^*$? of $\bigwedge^* V^*$?

2 Practice Questions

1) Express the forms $\mathbf{v}^1 \wedge \mathbf{v}^2 - 2\mathbf{v}^3 \wedge \mathbf{v}^4$ and $12\mathbf{v}^1 \wedge \mathbf{v}^2 \wedge \mathbf{v}^3$ as elements of $\bigotimes^{0,2} V$ and $\bigotimes^{0,3} V$, respectively.

2) Let $V = \mathbb{R}^3$ with standard basis $\mathbf{v}_1 = \hat{\imath}$, $\mathbf{v}_2 = \hat{\jmath}$, $\mathbf{v}_3 = \hat{k}$. Let $\omega \in \bigotimes^{1,2} V$ be the cross product: $\omega(v, w) = v \times w$ for $v, w \in \mathbb{R}^3$. Prove that ω is a tensor. Express ω in terms of basis elements of $\bigotimes^{1,2} V$.

3) Let $V = \mathbb{R}^3$ again, with the standard basis. Let $\Omega \in \bigotimes^{0,3} V$ be the triple product: $\Omega(v, w, u) = (v \times w) \cdot u$. Prove that Ω is a tensor. Express Ω in a basis of $\bigotimes^{0,3} V$.

4) Let $\Omega : V \times V \times V \to \mathbb{R}$ be as above. Prove that Ω is a form. Express Ω in terms of the basis of $\bigwedge^3 V^*$.

5) Let V be the vector space of quadratic polynomials in the variable x, with the standard basis. Let $T_w : V \to \mathbb{R}$ be given by $T_w(v) = \int_0^1 w\, v\, dx$, where $w \in V$. If $w = \alpha^1 \mathbf{v}_1 + \alpha^2 \mathbf{v}_2 + \alpha^3 \mathbf{v}_3$, express T_w in terms of the dual basis $\{\mathbf{v}^i\} \subset V^*$.

Test 3 Topics

Mar 25, 2009

1 Required Knowledge

What are coordinates?

Given a space M and a point $p \in M$, what is T_pM? T_p^*M?

Given coordinates $\{x^1, \ldots, x^n\}$, what, precisely, does $\frac{\partial}{\partial x^i}\Big|_p$ mean?

Given a function f, what does df mean?

How are vector fields and covector fields expressed?

If X is a vector field and ω is a covector field, what is $\omega(X)$?

Given a metric g on V, how is the metric on V^* defined?

Given a metric g on V with components g_{ij}, what are the components g^{ij} of the corresponding metric g on V^*?

Let M be a space with metric g. If X is a vector field, what is $X_\flat$? If ω is a covector field, what is $\omega^\sharp$?

Let M be a space with metric g. If $\gamma(\tau)$, $a < \tau < b$ is a path, what is its length?

If ω is a p-form, how is $d\omega$ computed?

What are the four defining properties of the d operator?

2 Practice Problems

1) Consider the Minkowski metric on $\mathbb{R}^{1,1}$. Let $X = x^0 x^1 \frac{\partial}{\partial x^0} + \frac{\partial}{\partial x^1}$ be a vector field, and let $\omega = x^1 dx^0 + x^0 dx^1$ be a 1-form. Find $X_\flat$ and $\omega^\sharp$. Compute $\omega(X)$ and $X_\flat(\omega^\sharp)$.

2) Let $g = 4\left((x^1)^2 + 1\right) dx^1 \otimes dx^1 - x^1 x^2 dx^1 \otimes dx^2 - x^1 x^2 dx^2 \otimes dx^1 + (x^1 x^2)^2 dx^2 \otimes dx^2$ be a 2-tensor on $\mathbb{R}^2$. Is this a metric?

3) Let $g = e^{-(x^1)^2-(x^2)^2} dx^1 \otimes dx^1 + e^{-(x^1)^2-(x^2)^2} dx^2 \otimes dx^2$ be a 2-tensor on $\mathbb{R}^2$. Prove g is a metric. Compute g^{ij}. Let $X = \frac{\partial}{\partial x^1} + x^1 \frac{\partial}{\partial x^2}$ and compute $X_\flat$.

4) Let $\{x, y\}$ be coordinates on $\mathbb{R}^2$. Let $\omega = e^{-x^2-y^2} dx$. Compute $d\omega$, then convert $d\omega$ into polar coordinates. Next, convert ω into polar coordinates first, and then compute $d\omega$. Do you get the same answer?

5) Let $g = \frac{r^2}{(1+r^2)^2}\left(dr \otimes dr + r^2 d\theta \otimes d\theta\right)$ be a 2-tensor. Prove that this is a metric on $\mathbb{R}^2 - \{o\}$. Determine the distance from the origin to the "point at infinity."

Test 4 Topics

April 22, 2009

1 Things you must memorize

The defining properties of the connection.

The definition of the Christoffel symbols.

The geodesic equation.

The parallel transport equation.

The DERIVATION of the geodesic equation.

The definition of the Riemann curvature tensor.

The definition of sectional curvature.

The DERIVATION of the Jacobi equation.

2 Things you must know

What warped products are.

How to convert between the (ξ, θ) and (r, θ) coordinate systems on warped products.

What the classical vector potential $\vec{A}$ and electric pseudopotential φ are.

What a gauge is and what 'gauge invariance' means.

What the Coulomb gauge is.

What the Lorentz gauge is.

How to derive wave equations for $\vec{E}$ and $\vec{B}$, if you are given the Maxwell equations.

How to derive wave equations for φ and $\vec{A}$ in the Lorentz gauge, if you are given the Maxwell equations.

What the Faraday 2-form is.

3 Things you do not have to memorize

The Koszul formula.

The formula for Γ_{ij}^k in coordinates.

The formula for $R_{ijk}{}^l$ in coordinates.

The formula for the geodesic equation in coordinates.

The formula for the parallel transport equation in coordinates.

The Maxwell equations.