# Math 560 - Mathematical Physics I

HW 1 - Due Sept 17
HW 2 - Due Oct 15
HW 3 - Due Nov 5
HW 4 - Due Dec 10

Here is a link to my undergraduate class from last semester, The Geometry of Physics.

Here is some **backround/review material** on linear algebra:
  Vector spaces, linear maps, and dual spaces
  More on dual spaces, and even more still
  Tensor products, the tensor algebra, and uses of the tensor algebra
  The musical isomorphisms
  Vectors on manifolds, covectors on manifolds, and tensor fields

## Notes

Lagrangian formulation of classical mechanics

Hamiltonian formulation of classical mechanics

Symplectic Geometry

Quantum Mechanics

Nonrelativistic Spin and Lie Algebra Representations

Clifford Algebras

Relativistic Spin

# Homework 1

## Due Sept 17

From *The Geometry of Physics* by Frankel

2.1(2) 2.5(3) 2.8(3) 2.10(3) 2.10(4) 3.1(2) 3.1(3) 3.3(1) 3.3(2)

# Homework 2

Due Oct 8

## 1 Short Questions

Question 1 If $\omega = dp_i \wedge dq^i$ is the canonical 2-form on Euclidean space $\mathbb{R}^n \times \mathbb{R}^n$, and $A, B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ are dynamical variables, prove that

$$\omega(\nabla A,\, \nabla B) \;=\; \{A,\, B\}.$$

Question 2 Let $f(x) = c|x|$. Compute the Legendre transform $F = \mathcal{L}(f)$ (where $F = F(p)$). Show that $\mathcal{L}(\mathcal{L}(f)) = f$.

## 2 Motion in a central potential

Let $\mathbf{r} = (x^1, \ldots, x^n)$ be the coordinates of a point in Euclidean $n$-space and assume a potential of the form $U(r)$ where $r = |\mathbf{r}|$.

Prove that the motion is *planar*, so that after the plane of motion is determined, the motion can be described with two parameters, $r$ and $\theta$.

Assuming the point has mass 1, determine the Lagrangian and the Hamiltonian for this system. What are the momenta conjugate to $r$, $\theta$?

Prove that the dymanic variable $M = \dot{\theta} r^2$ is a constant of the motion. To do this, compute the Poisson bracket $\{H, M\}$ and note that $\partial M/\partial t = 0$.

Where velocity and acceleration are colloquially denoted $\dot{\mathbf{r}}(t)$ and $\ddot{\mathbf{r}}(t)$, we of course denote these by $\frac{d}{dt}$ and $\nabla_{d/dt}\frac{d}{dt}$. Using $\dot{r} = \frac{dr}{dt}$, $\dot{\theta} = \frac{d\theta}{dt}$, compute

$$\ddot{\mathbf{r}} \;=\; \left(\ddot{r} - r\,\dot{\theta}^2\right)\frac{\partial}{\partial r} \;+\; \left(\frac{2\,\dot{r}\,\dot{\theta}}{r} + \ddot{\theta}\right)\frac{\partial}{\partial \theta}$$

You may have to use

$$\nabla_{\partial/\partial r}\frac{\partial}{\partial r} \;=\; 0 \qquad \nabla_{\partial/\partial r}\frac{\partial}{\partial \theta} \;=\; \nabla_{\partial/\partial r}\frac{\partial}{\partial r} \;=\; \frac{1}{r}\frac{\partial}{\partial \theta} \qquad \nabla_{\partial/\partial \theta}\frac{\partial}{\partial \theta} \;=\; -r\,\frac{\partial}{\partial r}$$

It is possible to reduce this multidimensional problem to a 1-dimensional problem. Let $r$ be the coordinate on $\mathbb{R}^1$ and put $V(r) = U(r) - M^2/2r^2$ (the *effective potential*). Prove that the Euler-Lagrange equations for the 1-dimensional Lagrangian $L(r, v) = \frac{1}{2}v^2 - V$ give the same equation of motion for $r$ that the former Lagrangian gave.

Consider the case where the potential obeys a power law:$U(r) = r^\alpha$. The case $\alpha = 2$ is of course that of the harmonic oscillator, and $\alpha = -1$ yields the Kepler problem. Make graphs of the effective potential for $M = 1$ and $U(r) = r^\alpha$ for $\alpha = 2, -1, -2, -3$.

In this problem, we detetermine the angle made between an orbit's apocenter and the pericenter, assuming the orbit is closed. Since the Hamiltonian is a constant of the motion ($H = E$ where $E$ is the system's total energy), we have

$$\dot{r} \;=\; \sqrt{2E \,-\, 2V(r)}$$

Using $\dot{\theta} = Mr^{-2}$ and the chain rule, we can compute

$$\theta \;=\; \frac{1}{\sqrt{2}} \int \frac{Mr^{-2}}{\sqrt{E \,-\, V(r)}}\, dr,$$

which is valid as long as $\theta$ remains monotinic with respect to $r$. The angle we are concerned with is

$$\Theta \;=\; \frac{1}{\sqrt{2}} \int_{r_{min}}^{r_{max}} \frac{Mr^{-2}}{\sqrt{E \,-\, V(r)}}\, dr$$

Given $E$ and $M$, it is possible to compute $r_{min}$ and $r_{max}$. Do this, and compute $\Theta$ for $\alpha = 2, -1$. It is possible to prove that these are the only two numbers for which *all* orbits are closed (that is, given any orbit $\Theta$ will be a rational fraction of $\pi$).

Consider the case of Newtonian gravity in $n$-space. If the law of gravitation is still $\triangle U = -\rho$ where $U$ is the potential and $\rho$ is the mass distribution, prove that all bounded orbits decay, unless $n = 2, 3$. (First prove that if $\rho = \delta$ is the unit delta function at the origin, then $\triangle U = -\rho$ has the solution

$$U(q^1, \ldots, q^n) \;=\; \frac{1}{n\omega_n} \left((q^1)^2 + \cdots + (q^n)^2\right)^{\frac{2-n}{2}}$$

for $n \neq 2$, and

$$U(q^1, q^2) \;=\; \frac{1}{2\pi} \log |(q^1)^2 + \cdots + (q^n)^2|^{\frac{1}{2}}$$

for $n = 2$. Here $\omega_n = \frac{2\pi^{n/2}}{n\Gamma(n/2)}$ is the volume of the unit n-ball in Euclidean n-space.

# Homework 3

## Due Nov 5

Question 1. Let $\varphi_t : \mathbb{R}^2 \to \mathbb{R}^2$ be the family of diffeomorphisms $\varphi_t = (\varphi_t^1, \varphi_t^2)$ given by

$$\begin{aligned} \varphi_t^1(x^1, x^2) &= x^1 + t\,x^1 \\ \varphi_t^2(x^1, x^2) &= x^2 + t\,x^2. \end{aligned}$$

a) Find $\varphi_{-t}(x^1, x^2)$.

b) If

$$X_t(x^1, x^2) = \left.\frac{d\varphi_t^i}{dt}\right|_{\varphi_{-t}(x^1,x^2)} \frac{\partial}{\partial x^i}$$

is the direction field of the flow at time $t$, then prove that $X_t(x^1, x^2) = \frac{x^1}{1+t}\frac{\partial}{\partial x^1} + \frac{x^2}{1+t}\frac{\partial}{\partial x^2}$.

c) Let $Y = Y^i \frac{\partial}{\partial x^i}$ be the vector field given by $Y^1(x^1, x^2) = 1$, $Y^2(x^1, x^2) = x^1$. Compute $(\varphi_{t*}Y)(x^1, x^2)$.

d) Compute $L_X Y$ using the result from part (c) and the definition of the Lie derivative.

e) Compute $[X, Y]$ using the usual definition: $[X, Y](f) = X(Y(f)) - Y(X(f))$.

Question 2. Prove that the matrix group

$$\mathcal{H} = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \middle|\, x, y, z \in \mathbb{R} \right\}$$

has Lie algebra isomorphic to the Heisenberg algebra. This Lie group, which is diffeomorphic but not isomorphic to $\mathbb{R}^3$, is called the Heisenberg group.

In the natural frame $X$, $Y$, $Z$, compute $\nabla_X Y$, $\nabla_Y Z$, and $\nabla_Z X$.

Question 3. Recall the classical Maxwell equations

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

The first equation implies the existence of a vector field $\vec{A}$ so that

$$\vec{B} = \nabla \times \vec{A},$$

called the *magnetic vector potential*. Although $\vec{A}$ is not unique, any two choices $\vec{A}_1$, $\vec{A}_s$ must differ by a total differential: $\vec{A}_1 - \vec{A}_2 = \nabla f$ for some function $f$ (this is electromagnetic *guage invariance*). The second equation can therefore be written

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

which means that there is some function $\varphi$ so that

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

called the *electric pseudopotential*. Again it is not unique, although different choices differ only by an additive constant.

Consider a system consisting of a particle of charge $e$ and mass $m$ in the presence of an external electromagnetic field. The Lagrangian is

$$L(\vec{q},\, \vec{v},\, t) \;=\; \frac{m}{2} \left| \vec{v} \right|^2 + e \left( \vec{v} \cdot \vec{A} - \varphi \right).$$

Determine the momenta conjugate to $v^i$. What are the mechanical momenta, in terms of the canonical momenta?

Write down the Hamiltonian.

Prove that the Lorentz force law $F = e(\vec{v} \times \vec{B} + \vec{E})$ is recovered in the Euler-Lagrange equations.

Given $\vec{B} = \frac{\partial}{\partial z}$ and $\vec{E} = 0$, determine $\vec{A}$. If a particle at $(1, 0, 0)$ has initial velocity $(0, 1, 0)$, what is its initial canonical momenta? Detemine this particle's path through configuration space.

# Homework 4

## Due Dec 10

Question 1 Define the function

$$K(\vec{x}, t) \;=\; \frac{1}{(2\pi\hbar)^3}\int_{\mathbb{R}^3} \exp\left(\frac{i}{\hbar}\left(\vec{p}\cdot\vec{x} \,-\, Et\right)\right) d\vec{p},$$

where $E = E(\vec{p})$ is equal to the energy of a particle of momentum $p$ (ie, the Hamiltonian). This function is either called the *Schrödinger Kernel* or the *Schrödinger propagator*.

Prove that given any function $\psi_0$ the function

$$\psi(\vec{x},\, t) \;=\; \int_{\mathbb{R}^3} K(\vec{x} \,-\, \vec{y},\, t \,-\, t_0)\, \psi_0(\vec{y})\, d\vec{y}$$

is a solution of the Schrödinger equation with initial condition $\psi(\vec{x}, t_0) = \psi_0(\vec{x})$.

For the case of a free particle (no potentials), we have $E = \frac{1}{2m}|\vec{p}|^2$. Prove that

$$K(\vec{x},\, t) \;=\; e^{-\sqrt{-1}\frac{3}{4}\pi}\, \left(\frac{m}{2\pi\hbar t}\right)^{\frac{3}{2}} \exp\left(\sqrt{-1}\frac{m|\vec{x}|^2}{2\hbar t}\right).$$

Question 2. (Spreading of the wave function.) Consider the convolution above, applied to a particle constrained to move in one dimension. Assume the wave function at time 0 is

$$\psi(x,\, 0) \;=\; \frac{1}{(\pi\xi_0^2)^{-\frac{1}{4}}} \exp(ip_0x/\hbar)\, \exp(-x^2/2\xi_0^2).$$

This is a normalized Gaussian, with $\triangle\psi = \xi_0$. If time is allowed to vary, prove that

$$\triangle\psi \;=\; \xi_0\left(1 \,+\, \frac{\hbar^2 t^2}{m^2\xi_0^4}\right)^{\frac{1}{2}}.$$

Question 3. (Low dimensional Clifford algebras) Consider the real Clifford algebras $Cl(n)$. Prove that $Cl(1)$ is isomorphic to $\mathbb{C}$, $CL(2)$ is isomorphic to $\mathbb{H}$, and $Cl(3)$ is isomorphic to $\mathbb{H}\oplus\mathbb{H}$.

Question 4. (The Dirac Equation) Write down the Dirac equation for a particle is an externally applied electromagnetic field $\vec{B} = \partial/\partial z$. Solve this equation.

# 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**

---

# 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 $\alpha, \beta$ 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:

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

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* $\dim\left(V^*\right) = n$*.*

<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 $\beta^i$, 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

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

$\square$

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:

$$\mathbf{v} \;=\; \alpha^1\mathbf{v}_1 + \ldots + \alpha^n\mathbf{v}_n \qquad \textit{or} \qquad \mathbf{v} = \alpha^i\mathbf{v}_i \qquad \textit{or} \qquad \mathbf{v} \;=\; \begin{pmatrix} \alpha^1 \\ \vdots \\ \alpha^n \end{pmatrix}_{\{\mathbf{v}_i\}}$$

(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 $\delta^i_j$

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 $\delta^i_j$. Of course $\delta^i_j$ 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 $\theta$.

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 $\theta$.

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 $\theta$) 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 $\theta$).

## 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

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

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 $Ker(\mathcal{A}) = \{0\}$:

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

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

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,

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

## 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! □

# 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

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

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 $\bigotimes^i V$.*

$\square$

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

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

($r$ 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 $V \times \cdots \times V \to \mathbb{R}$:

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

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

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

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

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

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

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

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 $\alpha^i$ are the same as the numbers $\alpha_i$.) 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 $\gamma$ 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 $\partial/\partial x^i$:

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

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

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

# 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

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

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 $\tau$ (ie, $\gamma = \gamma(\tau)$, $a \le \tau \le b$), the vector tangent to $\gamma$ is

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

The speed of $\gamma$ 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 $\gamma$ 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

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

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

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

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 $\gamma$ 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 $\infty$ as $b$ approaches $1/\sqrt{2}$, as expected.

# Lagrangians

September 2009

## 1 Configuration Space, State Space

The set of possible configurations a (classical) physical system may adopt, called *configuration space* has a natural structure of a manifold, which we call $M$. Important in physics is not just the configuration of a system but also its instantaneous dynamics, that is, its velocity. The possible velocities of a system located at some point in configuration space naturally makes up a vector space (that is, velocities can be added, multiplied by constants, etc). The set of all possible configurations and possible velocities together makes up state space, which is identified with the tangent bundle, $TM$, of configuration space.

$$Configuration\, space\, is\, a\, manifold\; M$$
$$State\, space\, is\, its\, tangent\, bundle\; TM.$$

Example Consider an $n$-particle system in Euclidean 3-space. This system's configuration space is Euclidean $3n$-space. That is, if the $j^{th}$ particle is located at $(q^1_{(j)}, q^2_{(j)}, q^3_{(j)}) \in \mathbb{R}^3$, then the location of the system in $\mathbb{R}^{3n}$ is

$$(q^1, \ldots, q^{3n}) \;=\; (q^1_{(1)}, q^2_{(1)}, q^3_{(1)}, \ldots, q^1_{(j)}, q^2_{(j)}, q^3_{(j)}, \ldots, q^1_{(n)}, q^2_{(n)}, q^3_{(n)}).$$

The vectors of state space are, at each point $\mathbf{q} = (q^1, \ldots, q^n)$, spanned by

$$\frac{\partial}{\partial q^1}, \;\ldots,\; \frac{\partial}{\partial q^{3n}}$$

Example Consider the case of a pendulum: some weight attached to an arm of mass $m$, where the arm swings on a pivot. the configuration of the system is (after a zero angle has been chosen) naturally identified with the angular coordinate $\theta$. Since angles repeat themselves every $2\pi$, configuration space can be identified with the 1-sphere $\mathbb{S}^1$ (the circle). At each point of $\mathbb{S}^1$, the tangent space is spanned by the vector $\partial/\partial\theta$.

# 2 The Lagrangian

A *Lagrangian* is a function

$$L : TM \times \mathbb{R} \to \mathbb{R},$$

whose restriction to any tangent space $T_qM$ is a convex function. Given coordiates $q^i$ on some region of $M$, we can express any vector $\mathbf{v} \in T_qM$ with values $v^i$ as follows:

$$v \;=\; v^i \frac{\partial}{\partial q^i}.$$

Thus $(q^1, \ldots, q^n, v^1, \ldots, v^n)$ are local coordinates on the manifold $TM$. We can express the Lagrangian in these coordinates:

$$L(q^1, \ldots, q^n, v^1, \ldots, v^n, t).$$

In classical Newtonian physics, the Lagrangian is typically taken to be the difference between kinetic energy $T$ and potential energy $U$. Kinetic energy is $T(\mathbf{q}, \mathbf{v}) = \frac{1}{2} mg(\mathbf{v}, \mathbf{v})$, and potential energy is some function $U = U(q^1, \ldots, q^n)$ of configuration space (ie, does not involve 'velocity' terms). Such a Lagrangian, called a *natural Lagrangian* look like

$$L(q^1, \ldots, q^n, v^1, \ldots, v^n, t) \;=\; \frac{m}{2} g_{ij} v^i v^j \;-\; U(q_1, \ldots, q^n)$$

in the case that all the particle masses equal $m$.

If a Lagrangian can be written $L : TM \to \mathbb{R}$, that is if it does not explicitly depend on time, it is called *autonomous*.

## 2.1 Hamilton's principle of least action

If

$$\alpha : (t_1, t_2) \to M,$$

is a path on $M$, its *action* is the integral

$$\mathfrak{L}(\alpha) \;=\; \int_{t_1}^{t_2} L(\alpha(t),\, \dot{\alpha}(t),\, t)\, dt.$$

If $M$ is configuration space and $\alpha$ is the path a physical system actually takes between the configurations $\mathbf{q}_1 = \alpha(t_1)$ and $\mathbf{q}_2 = \alpha(t_2)$, then *Hamilton's principle of least action* states that $\mathfrak{L}(\alpha)$ is smallest amongst all paths $\beta : [t_1, t_2] \to \mathbb{R}$ that start and end at the points that $\alpha$ does:

$$\mathfrak{L}(\alpha) \;=\; \inf_{\beta}\, \mathfrak{L}(\beta)$$

where the infimum is taken over all paths $\beta$ where $\beta(t_1) = \alpha(t_1)$ and $\beta(t_2) = \alpha(t_2)$.

Hamilton's principle has an infinitesimal expression. Let $\alpha(t)$, $t \in [t_1, t_2]$ be the path taken by a physical system, and let $\alpha_s$ be a smooth family of paths parametrized by $s$, where $\alpha_0 = \alpha$, and $\alpha_s(t_1) = \alpha_0(t_1)$ and $\alpha_s(t_2) = \alpha_0(t_2)$ (that is, the endpoints of the variation are fixed). The *direction field* of such a variation is

$$\dot{\alpha}_s(t) \;\equiv\; \frac{d}{dt} \;=\; \frac{dq^i}{dt}\frac{\partial}{\partial q^i},$$

and the *variation field* is

$$\frac{d}{ds} \;=\; \frac{dq^i}{ds}\frac{\partial}{\partial q^i}.$$

Now, since $\mathfrak{L}(\alpha_0)$ is minimal, basic calculus says that

$$\left.\frac{d}{ds}\right|_{s=0} \mathfrak{L}(\alpha_s) \;=\; 0. \tag{1}$$

Therefore

$$\begin{aligned} 0 \;&=\; \frac{d}{ds}\int_{t_1}^{t_2} L(\alpha_s(t), \dot{\alpha}_s(t), t)\, dt \;=\; \int_{t_1}^{t_2} \frac{d}{ds} L(\alpha_s(t), \dot{\alpha}_s(t), t)\, dt \\ &=\; \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^i}\frac{dq^i}{ds} \;+\; \frac{\partial L}{\partial v^i}\frac{dv^i}{ds} \;+\; \frac{\partial L}{\partial t}\frac{dt}{ds} \right) dt \end{aligned}$$

But $\frac{dt}{ds} = 0$ (the coordinates parametrizing the variation and direction fields are independent), and along the path we have

$$\dot{\alpha}(t) \equiv \frac{d}{dt} \;=\; \frac{dq^i}{dt}\frac{\partial}{\partial q^i} \;=\; v^i \frac{\partial}{\partial q^i},$$

so therefore

$$\begin{aligned} 0 \;&=\; \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^i}\frac{dq^i}{ds} \;+\; \frac{\partial L}{\partial v^i}\frac{d}{ds}\left(\frac{dq^i}{dt}\right) \right) dt \\ &=\; \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^i}\frac{dq^i}{ds} \;+\; \frac{\partial L}{\partial v^i}\frac{d}{dt}\left(\frac{dq^i}{ds}\right) \right) dt \\ &=\; \left.\frac{\partial L}{\partial v^i}\frac{dq^i}{ds}\right|_{t_1}^{t_2} \;+\; \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q^i} \;-\; \frac{d}{dt}\frac{\partial L}{\partial v^i} \right) \frac{dq^i}{ds}\, dt. \end{aligned}$$

Since the variation is assumed to vanish at the endpoints, the first term is zero. Since the variation is arbitrary (meaning the functions $\frac{dq^i}{ds}$ may be arbitrarily chosen), it must be that the quantity in the parentheses *must vanish identically*. Thus,

$$\frac{\partial L}{\partial q^i} \;-\; \frac{d}{dt}\frac{\partial L}{\partial v^i} \;=\; 0.$$

These are called the Euler-Lagrange equations. They constitute a system of $n$ second-order, nonlinear ordinary differential equations, and furnish an infinitesimal version of Hamilton's least action principle.

Exercise Let $M$ be a Riemannian manifold with metric $g$, and let $L(q,v) = \frac{1}{2}g(v,v)$ (recall that the metric in general depends on $q$). Prove that the Euler Lagrange equations can be simplified to read

$$\nabla_{\frac{d}{dt}} \frac{d}{dt} \;=\; 0$$

where $\nabla$ is the Riemannian connection on $M$.

## 2.2 Canonical momenta, canonical forces, and equivalence of Lagrangian and Newtonian mechanics

The quantity

$$\frac{\partial L}{\partial v^i}$$

is called the *momentum conjugate to* $v^i$, or the $i^{th}$ *canonical momentum*. The motivation for this definition is due to the case of a natural Lagrangian, where

$$\begin{aligned}
\frac{\partial L}{\partial v^i} &= \frac{\partial}{\partial v^i}\left(\frac{m}{2}g_{kl}v^kv^l \;-\; U\right) \\
&= \frac{m}{2}g_{kl}\frac{\partial v^k}{\partial v^i}v^l \;+\; \frac{m}{2}g_{kl}v^k\frac{\partial v^l}{\partial v^i} \\
&= \frac{m}{2}g_{kl}\delta^k_iv^l \;+\; \frac{m}{2}g_{kl}v^k\delta^l_i \\
&= m\,g_{ik}v^k \;=\; m\,v_\flat
\end{aligned}$$

the the *mechanical momentum.*

The quantities

$$p_i \;=\; \frac{\partial L}{\partial q^i}$$

are called the *canonical forces.* Again this is motivated by the natural case. Classically, in the presence of a potential $U$, the force field is $\vec{F} = -\nabla U$, so the force in the $i^{th}$ direction is $-\frac{\partial U}{\partial q^i}$.

The Euler-Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial v^i} \;=\; \frac{\partial L}{\partial q^i}$$

are therefore precisely Newton's second law: $m\vec{a} \;=\; \frac{d}{dt}\vec{p} \;=\; \vec{F}$.

# 3 Invariance and conservation laws

## 3.1 Time invariance and conservation of energy

Given a physical path $\alpha(t)$ through configuration space, we have

$$\frac{dL}{dt} \;=\; \frac{\partial L}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial L}{\partial v^i}\frac{dv^i}{dt} + \frac{\partial L}{\partial t}.$$

Assuming $L$ does not depend explicitly on time, meaning $\partial L/\partial t = 0$, the Euler-Lagrange equations give

$$\begin{aligned}\frac{dL}{dt} &= \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right)\frac{dq^i}{dt} + \frac{\partial L}{\partial v^i}\frac{d}{dt}\left(\frac{dq^i}{dt}\right)\\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\frac{dq^i}{dt}\right).\end{aligned}$$

Thus

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\frac{dq^i}{dt} - L\right) \;=\; 0$$

and therefore the quantity $\frac{\partial L}{\partial v^i}\frac{dq^i}{dt} - L$ is a *constant of the motion*. Recalling that $v^i = \frac{dq^i}{dt}$ and that we have labeled $p_i = \frac{\partial L}{\partial v^i}$, this reads

$$p_i v^i - L \;=\; E$$

where $E$ is a constant that we call the system's *energy*. Again this is motivated by the classical definition: in the natural case, this is the sum of kinetic and potential energies. We have just proven the following theorem.

**Theorem 3.1 (Conservation of Energy)** *Assuming $\frac{\partial L}{\partial t} = 0$, the quantity $\frac{\partial L}{\partial v^i}\frac{dq^i}{dt} - L$ is constant on any path through state space that satisfies the Euler-Lagrange equations.*

## 3.2 Translation invariance and conservation of momentum

Assume the Lagrangian on $\mathbb{R}^{3n}$ is translation invariant (meaning that if all particles are displaced spatially by the same amount, the Lagrangian in unchanged). If this property holds, space is called *homogeneous* with respect to $L$.

Let $(q^1_{(a)}, q^2_{(a)}, q^3_{(a)})$ be the position of the $a^{th}$ particle. Assume all particles are translated in the direction $(T^1, T^2, T^3)$ by distance $s$, meaning that we chance each $q^i_{(a)}$ to

$q^i_{(a)} + sT^i$. Taking a derivative at $s=0$ we have $\frac{dL}{ds} = 0$ by assumption, and also $\frac{dq^i_{(a)}}{ds} = T^i$ and $\frac{dv^i_{(a)}}{ds} = 0$, so that

$$\begin{aligned} 0 &= \frac{dL}{ds} = \sum_a \sum_i \frac{\partial L}{\partial q^i_{(a)}} \frac{dq^i_{(a)}}{ds} + \frac{\partial L}{\partial v^i_{(a)}} \frac{dv^i_{(a)}}{ds} \\ &= \sum_a \left( \frac{\partial L}{\partial q^1_{(a)}} T^1 + \frac{\partial L}{\partial q^2_{(a)}} T^2 + \frac{\partial L}{\partial q^3_{(a)}} T^3 \right). \end{aligned}$$

Now the Euler-Lagrange equations give

$$\begin{aligned} 0 &= \sum_a \left( \frac{d}{dt}\left( \frac{\partial L}{\partial v^1_{(a)}} \right) T^1 + \frac{d}{dt}\left( \frac{\partial L}{\partial q^2_{(a)}} \right) T^2 + \frac{d}{dt}\left( \frac{\partial L}{\partial q^3_{(a)}} \right) T^3 \right) \\ &= T^1 \frac{d}{dt}\left( \sum_a \frac{\partial L}{\partial v^1_{(a)}} \right) + T^2 \frac{d}{dt}\left( \sum_a \frac{\partial L}{\partial v^2_{(a)}} \right) + T^3 \frac{d}{dt}\left( \sum_a \frac{\partial L}{\partial v^3_{(a)}} \right). \end{aligned}$$

Since $T^1$, $T^2$, $T^3$ are *arbitrary*, it must be that

$$\begin{aligned} 0 &= \frac{d}{dt} \sum_a \frac{\partial L}{\partial v^1_{(a)}} \\ 0 &= \frac{d}{dt} \sum_a \frac{\partial L}{\partial v^2_{(a)}} \\ 0 &= \frac{d}{dt} \sum_a \frac{\partial L}{\partial v^3_{(a)}}. \end{aligned}$$

Using again our labeling $\frac{\partial L}{\partial v^i_{(a)}} = p_{i,(a)}$ to indicate the $i$-component of the momentum of the $a^{th}$ particle, we can write this as

$$\begin{aligned} P_1 &= \sum_a p_{1,(a)} \\ P_2 &= \sum_a p_{2,(a)} \\ P_3 &= \sum_a p_{3,(a)} \end{aligned}$$

where the $P_i$ are *constants*. The quantity $P_i$ is called the system's *total linear momentum in the $i^{th}$ direction*; the vector $\mathbf{P} = (P_1, P_2, P_3)$ is called the system's *total linear momentum.*

## 3.3 Rotational invariance and conservation of angular momentum

Choose an origin and let $\varphi = (\varphi^1, \varphi^2, \varphi^3)$ be an axis of rotation in $\mathbb{R}^3$. If $\mathbf{q} = (q^1, q^2, q^3)$ is a position vector, and $\mathbf{q}(s)$ indicates a rotation of angle $s$ about $\varphi$, then

$$\frac{d}{ds}\Big|_{s=0} \mathbf{q} = \varphi \times \mathbf{q} \qquad \frac{d}{ds}\Big|_{s=0} \mathbf{v} = \varphi \times \mathbf{v}.$$

If we assume the Lagrangian is *rotationally invariant* in the sense that if each particle is rotated by the same amount about the same point, we have

$$0 \;=\; \frac{d}{ds} L(\mathbf{q}(s), \mathbf{v}_s, t).$$

Letting $(q^1_{(a)}, q^2_{(a)}, q^3_{(a)})$ be the position of the $a^{th}$ particle, we have

$$\begin{aligned}
0 &= \sum_a \sum_i \frac{\partial L}{\partial q^i_{(a)}} \frac{dq^i_{(a)}}{ds} \;+\; \frac{\partial L}{\partial v^i_{(a)}} \frac{dv^i_{(a)}}{ds} \;+\; \frac{\partial L}{\partial t} \frac{dt}{ds} \\
&= \sum_a \sum_i \frac{d}{dt} \left( \frac{\partial L}{\partial v^i_{(a)}} \right) \frac{dq^i_{(a)}}{ds} \;+\; \frac{\partial L}{\partial v^i_{(a)}} \frac{dv^i_{(a)}}{ds} \;+\; \frac{\partial L}{\partial t} \frac{dt}{ds} \\
&= \sum_a \big( \dot{\mathbf{p}}_{(a)} \cdot (\varphi \times \mathbf{q}_{(a)}) \;+\; \mathbf{p}_{(a)} \cdot (\varphi \times \dot{\mathbf{q}}_{(a)}) \big) \\
&= \sum_a \big( \varphi \cdot \big(\mathbf{q}_{(a)} \times \dot{\mathbf{p}}_{(a)}\big) \;+\; \varphi \cdot \big(\dot{\mathbf{q}}_{(\mathbf{a})} \times \mathbf{p}_{(a)}\big) \big) .
\end{aligned}$$

The last equality is due to the fact that the triple-product is invariant under cyclic permutation: $\mathbf{X} \cdot (\mathbf{Y} \times \mathbf{Z}) = \mathbf{Y} \cdot (\mathbf{Z} \times \mathbf{X})$. Continuing with our calculation,

$$\begin{aligned}
0 &= \varphi \cdot \sum_a \big( \mathbf{q}_{(a)} \times \dot{\mathbf{p}}_{(a)} \;+\; \dot{\mathbf{q}}_{(a)} \times \mathbf{p}_{(a)} \big) \\
&= \varphi \cdot \frac{d}{dt} \left( \sum_a \mathbf{q}_{(a)} \times \mathbf{p}_{(a)} \right) .
\end{aligned}$$

Since $\varphi$ is arbitrary, we conclude that the vector

$$\mathbf{M} \;=\; \sum_a \mathbf{q}_{(a)} \times \mathbf{p}_{(a)},$$

called the system's *total angular momentum*, is a constant of the motion.

### 3.4 Noether's Theorem

Noether's theorem provides a criterion for when we can find conservation laws. A *constant of the motion* is a function $I : TM \to \mathbb{R}$ such that $\frac{dI}{dt} = 0$ along any path that a physical system actually takes.

**Theorem 3.2 (Noether's theorem)** *If $L$ is invariant under a 1-parameter family of diffeomorphisms $\Phi_s : M \to M$ (parametrized by the variable $s$), then associated to this family is a constant of the motion, given by*

$$I(\mathbf{q}, \mathbf{v}) \;=\; \frac{\partial L}{\partial v^i} \frac{d\Phi_s(q^i)}{ds}\bigg|_{s=0}.$$

*Pf*

If $\alpha(t)$ is a physical path and $\alpha_s(t) = \Phi_s(\alpha(t))$ is its mapping under $\Phi_2$, then the fact that $L$ is invariant under the mapping $\Phi_s$ means that $\alpha_s(t)$ is also a physical path. We have

$$\begin{aligned} 0 &= \frac{d}{ds} L(\alpha_s(t), \dot{\alpha}_s(t), t) \;=\; \frac{\partial L}{\partial q^i}\frac{dq^i}{ds} + \frac{\partial L}{\partial v^i}\frac{dv^i}{ds} + \frac{\partial L}{\partial t}\frac{dt}{ds} \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right)\frac{dq^i}{ds} + \frac{\partial L}{\partial v^i}\frac{d}{dt}\left(\frac{dq^i}{ds}\right) \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\frac{dq^i}{ds}\right). \end{aligned}$$

□

# Hamiltonians

September 2009

## 1 The Floer map

Given a map $L : TM \to \mathbb{R}$ we can define another map

$$FL : TM \to T^*M,$$

called the Floer map. Given a vector $v \in T_qM$, we define the covector $FL(v) \in T_q^*M$ by defining its action on any other vector $w \in T_qM$ by

$$FL(v)(w) \;=\; \frac{d}{dt}\Big|_{t=0} L(q,\, v\,+\,tw).$$

In the case of a natural Lagragian on a manifold, we have

$$\begin{aligned}
L(q,\, v) &= \frac{m}{2}g(v,\, v) \,-\, U(q) \\
FL(v)(w) &\triangleq \frac{d}{dt}\Big|_{t=0} \Big(\frac{m}{2}g(v+tw,\, v+tw) \,+\, U(q)\Big) \\
&= \frac{d}{dt}\Big|_{t=0} \left(\frac{m}{2}g(v,\, v) \,+\, t\,m\,g(v,\, w) \,+\, \frac{1}{2}mt^2g(w,\, w)\right) \\
&= mg(v, w).
\end{aligned}$$

Therefore

$$FL(v) \;=\; m\,v_\flat.$$

Of course this is the classical momentum, which is, as usual, a covector.

Whether the Lagrangian is natural or not, we have

$$\begin{aligned}
FL(v)(w) &= \frac{d}{dt}L(q,\, v\,+\,tw) \\
&= \frac{\partial L}{\partial q^i}\frac{d\partial q^i}{dt} \,+\, \frac{\partial L}{\partial v^i}\frac{\partial v^i + tw^i}{\partial t} \\
&= \frac{\partial L}{\partial v^i}\,w^i,
\end{aligned}$$

so that the components of $FL(v)$ are just the components of the conjugate momenta.

We call $T^*M$ *phase space*. Covectors in phase space are called *generalized momenta* or *canonical momenta* associated to the system at that point.

# 2 The Legendre transform

## 2.1 Legendre transform of a 1-variable function

Given a 1-variable function $f$, a line $A(x) = mx + b$ is called a *supporting line for $f$ at $x_0$* if $f(x_0) = A(x_0)$, and for all $x$ in some interval $(x_0 - \epsilon, x_0 + \epsilon)$, we have $A(x) \leq f(x)$.

A 1-variable function $f$ is called *convex* if $f$ has a supporting line for all $x$ in its domain, and, whenever $A(x)$ is a supporting line, then $A(x) \leq f(x)$ globally. If $f$ has two or more derivatives and its domain is connected, it can be proved $f$ is convex iff $f''(x) \geq 0$ for all $x$.

Given a convex function $f(x)$, we define its Legendre transform $F(p)$ to be

$$F(p) \; = \; \sup_{x \in \mathbb{R}} \left( p\, x \; - \; f(x) \right),$$

whenever the right side is defined.

Example Find the Legendre transform of $f(x) = \frac{1}{2}cx^2$.
Solution Given $p$, we have to maximize the expression $xp - f(x)$. Taking a derivative we require $p - cx = 0$ so that $x = p/c$. Then $F(p) = \max_x px - f(x) = p^2/c - \frac{1}{2}c(p/c)^2$, so that $F(p) = \frac{1}{2}\frac{1}{c}p^2$.

## 2.2 Legendre transform of a multi-variable function

Likewise, if $f : \mathbb{R}^n \to \mathbb{R}$ is a multivariable function, then $f$ is called convex if all supporting hyperplanes remain $\leq f$ globally. In the case of a function with connected domain and two (continuous) derivatives, this is equivalent to the Hessian matrix

$$\left( \frac{\partial^2 f}{\partial x^i \partial x^j} \right)$$

being positive semi-definite (all eigenvalues $\geq 0$).

## 2.3 Legendre transform of a Lagrangian

On a general manifold, this notion of convexity does not make sense. However, given a function on any finite-dimensional vector space, it does. Recall that we required the

Lagrangian $L : TM \to \mathbb{R}$ to be convex when restricted to any tangent space $T_q M$.

We can define a function $H : T^*M \to \mathbb{R}$ to be the Legendre transform of $L$. Given $\mathbf{p} \in T_{\mathbf{q}}M$, we put

$$H(\mathbf{q}, \mathbf{p}) \;=\; \max_{\mathbf{v} \in T_{\mathbf{q}}M} \mathbf{p}(\mathbf{v}) \;-\; L(\mathbf{q}, \mathbf{v}).$$

If $L$ is continuously differentiable up to second order, then it can be shown that a vector $\mathbf{v}$ realizes his maximum if

$$\mathbf{p} \;=\; FL(\mathbf{v}).$$

That is,

$$H(q^i, p_i) \;=\; p_j v^j \;-\; L(q^i, v^i)$$

where $\mathbf{v} = FL^{-1}(\mathbf{q})$.

## 3 Phase space and Poisson brackets

A configuration space is a manifold $M$ with cotangent bundle $T^*M$. Locally, coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ exist on $T^*M$. However the cotangent bundle is itself a manifold, and has its own tangent bundle $TT^*M$, spanned at $T_{(q,p)}T^*M$ by

$$\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}, \frac{\partial}{\partial p^1}, \ldots, \frac{\partial}{\partial p^n},$$

and its own cotangent bundle $T^*T^*M$, spanned at $T^*_{(q,p)}T^*M$ by

$$dq^1, \ldots, dq^n,\; dp_1,\; \ldots, dp_n.$$

Given functions $f, g : T^*M \to \mathbb{R}$, there is a non-associative product between them, called the *Poisson bracket*, defined by

$$\{f,\, g\} \quad = \quad \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} \;-\; \frac{\partial g}{\partial p_i}\frac{\partial f}{\partial q^i}.$$

Easy-to-prove properties are

$$\begin{aligned}
\{f,\, c\} \;&=\; 0\\
\{f,\, g\} \;&=\; -\{g,\, f\}\\
\{f_1 \,+\, f_2,\, g\} \;&=\; \{f_1,\, g\} \,+\, \{f_2,\, g\}\\
\{f_1 f_2,\, g\} \;&=\; f_1\{f_2,\, g\} \,+\, f_2\{f_1,\, g\}.
\end{aligned}$$

Slightly more difficult to prove is the *Jacobi identity*:

$$\{f,\, \{g,\, h\}\} \,+\, \{g,\, \{h,\, f\}\} \,+\, \{h,\, \{f,\, g\}\} \;=\; 0$$

Two functions $f, g$ are said to *commute* if $\{f,\, g\} = 0$.

## 4 The canonical equations and dynamical variables

Given a Lagrangian system $\alpha(t)$ we compute

$$\begin{aligned}
\frac{\partial H}{\partial q^i} &= -\frac{\partial L}{\partial q^i} = -\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) = -\frac{dp_i}{dt} \\
\frac{\partial H}{\partial p_i} &= \frac{\partial p_j}{\partial p_i}v^j + p_j\frac{\partial v^j}{\partial p_i} - \frac{\partial L}{\partial v^j}\frac{\partial v^j}{\partial p_i} = v^i.
\end{aligned}$$

We arrive at *Hamilton's canonical equations*: any Lagrangian system satisfies

$$\boxed{\frac{dq^i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q^i}.}$$

A *dynamical variable* $A$ is a function $A : T^*M \times \mathbb{R} \to \mathbb{R}$. Given a Lagrangian system $\alpha(t)$ we have

$$\begin{aligned}
\frac{dA}{dt} &= \frac{\partial A}{\partial q^i}\frac{dq^i}{dt} + \frac{\partial A}{\partial p_i}\frac{dp_i}{dt} + \frac{\partial A}{\partial t}\frac{dt}{dt} \\
&= \frac{\partial A}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q^i} + \frac{\partial A}{\partial t},
\end{aligned}$$

which yields the elegant expression

$$\boxed{\frac{dA}{dt} = \{H, A\} + \frac{\partial A}{\partial t}.}$$

A dynamical variable $A$ is a constant of the motion iff $\partial A/\partial t + \{H, A\} = 0$. If $A$ does not explicitly depend on time, it is a constant of the motion iff it commutes with the Hamiltonian.

**Theorem 4.1 (Poisson's theorem)** *If $A, B$ are two dynamical variables and $\alpha(t)$ is a Lagrangian system, then*

$$\frac{d}{dt}\{A, B\} = \left\{\frac{dA}{dt}, B\right\} + \left\{A, \frac{dB}{dt}\right\}.$$

Pf

We compute

$$
\begin{aligned}
\frac{d}{dt}\{A,\, B\} &= \{H,\, \{A,\, B\}\} + \frac{\partial}{\partial t}\{A,\, B\} \\
&= -\{B,\, \{H,\, A\}\} - \{A,\, \{B,\, H\}\} + \left\{\frac{\partial A}{\partial t},\, B\right\} + \left\{A,\, \frac{\partial B}{\partial t}\right\} \\
&= \left\{\frac{\partial A}{\partial t} + \{H,\, A\},\, B\right\} + \left\{A,\, \frac{\partial B}{\partial t} + \{H,\, B\}\right\} \\
&= \left\{\frac{dA}{dt},\, B\right\} + \left\{A,\, \frac{dB}{dt}\right\}
\end{aligned}
$$

□

**Corollary 4.2** *The Poisson bracket of any two constants of the motion is again a constant of the motion.*

□

# 5 2-forms and Vortex lines

Let $\omega$ be a 2-form on some manifold. In coordinates, $\omega$ can be represented by a skew-symmetric matrix. Nonzero eigenvalues of a skew-symmetric matrix are always complex, and always come in complex-conjugate pairs (for a proof, consider the Jordan canonical form of the matrix).

If $\omega$ is a 2-form on an even dimensional manifold $M^{2n}$, we call it *nondegenerate* if there are no zero eigenvalues. If $\omega$ is a 2-form on an odd dimensional manifold $M^{2n+1}$, we call it nondegenerate if there is only one zero eigenvalue. Likewise, $\omega$ is nonsingular on $M^{2n+1}$ iff the space

$$\{\, X \mid \eta(X, \cdot) \;\equiv\; 0 \,\}$$

is spanned by just a single vector.

Now assume $\eta$ is a 1-form on $M^{2n+1}$ and that $d\eta$ is nonsingular. At each point there is a unique null direction for $d\eta$, called the *vortex direction.* Integrating this direction gives the *vortex lines.* Given a path $\gamma \subset M$, one can consider all the vortex lines emanating from $\gamma$; the resulting 2-dimensional object is called the *vortex tube* associated to $\eta$.

## 6 The phase flow and the Poincare 1-form

A differentiable map $T^*M \times \mathbb{R} \to T^*M$ can be defined by

$$(p, q, t) \to (p(t), q(t))$$

by solving the canonical equations. That is, if a Lagrangian system $\alpha$ starts with initial conditions $(\alpha(0), FL(\dot{\alpha}(0))) = (p, q)$, then we map $(p, q)$ to $(\alpha(t), FL(\dot{\alpha}(t)))$. This is called the *phase flow.* If $\alpha(t)$ is a Lagrangian system, it defines a curve in $T^*M \times \mathbb{R}$ by

$$t \hookrightarrow (\alpha(t),\, FL(\dot{\alpha}(t)),\, t)\,.$$

these are called the *flow lines.*

The 1-form

$$\eta \;=\; p_i\, dq^i \;-\; H\, dt$$

on $T^*M \times \mathbb{R}$ is called the Poincare-Cartan 1-form. Consider its differential

$$\omega \;=\; d\eta \;=\; dp_i \wedge dq^i \;-\; \frac{\partial H}{\partial q^i} dq^i \wedge dt \;-\; \frac{\partial H}{\partial p_i} dp_i \wedge dt,$$

which is called the *canonical 2-form.* As a matrix in the $(p, q, t)$-coordinate system, this has the form

$$\begin{aligned} \omega \;&=\; \omega_{ij} dp^i \wedge dq^j \;+ \\ \omega \;&=\; \begin{pmatrix} 0 & -I_n & H_p \\ I_n & 0 & H_q \\ -H_p & -H_q & 0 \end{pmatrix}. \end{aligned}$$

It is easy to check that the vector

$$\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} \;-\; \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i} \;+\; \frac{\partial}{\partial t} \tag{1}$$

is a null vector. It is also obvious that $\omega$ has at most one null vector. We have proven the theorem

**Theorem 6.1** *The lines of the phase flow are just the vortex lines of Poincare-Cartan 1-form.*

## 7 Canonical transformations

We consider changes of variables from $q^i, p_i$ coordinates on a cotangent bundle $T^*M$ to $Q^i, P_i$ coordinates, where $Q^i = Q^i(q^1, \ldots, q^n, p_1, \ldots, p_n)$ and $P_i = P_i(q^1, \ldots, q^n, p_1, \ldots, p_n)$. A *spatial* change of coordinates is where $Q^i = Q^i(q^1, \ldots, q^n)$ and the $P_i$ are defined (implicitly) by $p_i dq^i = p_i \frac{\partial q^i}{\partial Q^j} dQ^j$, that is, $P_j = p_i \frac{\partial q^i}{\partial Q^j}$.

Given a change of variables, the hamiltonian expressed in the new coordinates is sometimes denoted $K$ (the "Kamiltonian"). That is,

$$K(Q^i(q^i, p_i), P_i(q^i, p_i), t) \;=\; H(q^i, p_i, t).$$

The question arises, when does a change of variables preserve the canonical equations? For what kind of variable changes do we have

$$\frac{dQ^i}{dt} \;=\; \frac{\partial H}{\partial P_i} \qquad \frac{dP_i}{dt} \;=\; -\frac{\partial H}{\partial Q^i} \quad ?$$

**Theorem 7.1 (Canonical change of coordinate)** *Given new coordinates $Q^i$, $P_i$ on phase space, we retain the canonical equations*

$$\frac{dQ^i}{dt} \;=\; \frac{\partial H}{\partial P_i} \qquad \frac{dP_i}{dt} \;=\; -\frac{\partial H}{\partial Q^i}$$

*if there is a function $S : T^*M \to \mathbb{R}$ such that*

$$p_i\, dq^i \;=\; P_i\, dQ^i \;+\; dS.$$

Pf

If $p_i\, dq^i \;=\; P_i\, dQ^i \;+\; dS$ then $dp_i \wedge dq^i \;+\; dH \wedge dt = dP_i \wedge dQ^i \;+\; dK \wedge dt$, so the canonical 2-forms are equal. The vortex lines are therefore the same. Since the vortex lines are given by solutions to (1), the canonical equations are unchanged. □

The reverse implication is not in general true. A change of variables so that $p_i\, dq^i \;=\; P_i\, dQ^i \;+\; dS$ is called a *canonical change of variables.*

**Theorem 7.2** *The Poisson bracket is invariant under a canonical change of variables.*

Pf

The canonical 2-form is invariant under canonical coordinate changes, and we also have

$$\{A,\, B\} \;=\; \omega(dA,\, dB).$$

□

## 8 The Phase Flow and Poincare's Integral Invariants

**Theorem 8.1 (Poincare's integral invariant)** *Let $\gamma_0$ be a closed path in extended phase space, and let $\gamma_t$ be the image of the path, at time $t$, under the phase flow. Then*

$$\int_{\gamma_0} p_i \, dq^i \; - \; H \, dt \; = \; \int_{\gamma_t} p_i \, dq^i \; - \; H \, dt.$$

$\square$

**Corollary 8.2** *Let $p_i \wedge dq^i$ be the Poincare 1-form restricted to $T^*M$. If $\gamma$ is a closed path in $T^*M$ and $\gamma_t$ is its image under the phase flow at time $t$, then*

$$\int_{\gamma} p_i \wedge dq^i \; = \; \int_{\gamma_t} p_i \wedge dq^i.$$

**Corollary 8.3** *Given some surface $S \subset T^*M$, the phase flow preserves the oriented area of $S$.*

$\square$

**Corollary 8.4 (Liouville's theorem)** *The phase flow preserves volumes in phase space.*

**Corollary 8.5 (Poincare return)** *If $(p, q) \in T^*M$ is a system whose configuration space has bounded volume, and if $U \subset T^*(M)$ is a neighborhood of $(p, q)$ and $T > 0$, then at some time $t > T$ it will again hold that $(p(t), q(t)) \in U$ (here $(p(t), q(t))$ indicates the system's position at tile $t$).*

## 9 The Hamilton-Jacobi equation

The Lagrangian is a function on state space $TM$, but there is a way to define a corresponding function on configuration space $M$ itself. Fix a point $q_0 \in M$ and a time $t_0$. We define

$$S : TM \times \mathbb{R}^{>t_0} \to \mathbb{R}$$

to be

$$S(q, t) \; = \; \mathcal{L}(\gamma)$$

where $\gamma$ is the path traversed by a system starting at $q_0$ at time 0 and ending up at $q$ at time $t$.

Let us compute the differential $dS \in T^*(M \times \mathbb{R}^{>0})$. First assume $\gamma_u$ is a variation of Lagrangian paths, with $\gamma_u(t_0) = q_0$ for all $u$. Then

$$\begin{aligned}
\frac{dS}{du} &= \int_{t_0}^{t} \frac{dL}{du}\, dt \\
&= \int_{t_0}^{t} \left( \frac{\partial L}{\partial q^i} \frac{dq^i}{du} + \frac{\partial L}{\partial v^i} \frac{dv^i}{du} \right) dt \\
&= \frac{\partial L}{\partial v^i} \frac{dq^i}{du} \Big|_{t_0}^{t} + \int_{t_0}^{t} \left( \frac{\partial L}{\partial q^i} - \frac{d}{dt} \left( \frac{\partial L}{\partial v^i} \right) \right) \frac{dq^i}{du}\, dt \\
&= \frac{\partial L}{\partial v^i} \frac{dq^i}{du} \Big|_{t}.
\end{aligned}$$

Since

$$dS \left( \frac{d}{du} \right) = \frac{dS}{du},$$

we have

$$\frac{\partial S}{\partial q^i} = dS \left( \frac{\partial}{\partial q^i} \right) = \frac{\partial L}{\partial v^i} = p_i.$$

We have

$$\frac{dS}{dt} = L = p_i q^i - H$$

along an actual path. But

$$\begin{aligned}
\frac{dS}{dt} &= \frac{\partial S}{dt} + \frac{\partial S}{\partial q^i} \frac{\partial q^i}{\partial t} \\
&= \frac{\partial S}{dt} + p_i v^i.
\end{aligned}$$

Thus $\partial S / \partial t = H$. We have computed

$$dS = p_i\, dq^i - H\, dt.$$

# SYMPLECTIC GEOMETRY

## 1. What are the minimum requirements for a classical-type physics to exist?

1.1. **Our formulation of classical physics to date.** We have an extended phase space $T^*M \times \mathbb{R}$ with a Poincare-Cartan 1-form $\eta \;=\; p_i\,dq^i \;-\; H\,dt$. Notice that, when restricted to phase space $T^*M$, the form

$$p_i\,dq^i \;=\; \eta\big|_{T^*M}$$

is *intrinsic*, by which I mean, it does not depend on the existence of any Hamiltonian. We then constructed the *canonical 2-form*

$$\omega \;=\; d\eta,$$

which we proved was *nondegenerate*. We proved that the *phase flow* was in fact just the *vortex flow* of this 2-form on extended phase space. Again, notice that the form

$$dp_i \wedge dq^i \;=\; \omega\big|_{T^*M}$$

is *intrinsic to* $T^*M$, meaning is does not depend on the choice of a Hamiltonian. Also, if the phase flow, restricted to $T^*M$ is denoted by $\frac{d}{dt}$, then

$$i_{\frac{d}{dt}}\omega \;=\; -dH.$$

Another highly important feature of mechanics is the Poincare integral invariant. In global form this reads

$$\oint_{\gamma_t}\eta \;=\; const,$$

where $\gamma$ is some closed path and $\gamma_t$ is its image at time $t$ under the flow. In infinitesimal form this reads

$$\varphi_t^*\omega \;=\; \omega$$

where $\varphi_t$ is the flow itself. This means that the Lie derivative of $\omega$ vanishes

$$L_X\omega \;=\; 0$$

where $X = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} \;-\; \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i}$ is the flow field.

---

*Date*: September 2009.

1.2. **An attempt to reduce physics to its minimum.** It appears we need a cotangent bundle $T^*M$, a 2-form $\omega = dp_i \wedge dq^i$, and a Hamiltonian in order to define a physics. Note that the Hamiltonian itself is *a postiori*; the cotangent bundle and the 2-form constitute the substratum on which physics is built.

But can we do with even less? Instead of a cotangent bundle $T^*M$ and a canonical 2-form $\omega = dp_i \wedge dq^i$, maybe all we need is an even-dimensional manifold $N^{2n}$ and a nondegenerate 2-form $\omega$. We can still define the flow vector field $X$ (implicitly) by

$$i_X\omega \;=\; -dH.$$

The last question to ask is whether the flow defined by $X$ preserves the form $\omega$. We compute

$$\begin{aligned} L_X\omega &= d\,i_X\,\omega \;+\; i_X\,d\,\omega \\ &= -ddH \;+\; i_X\,d\,\omega \\ &= i_X\,d\,\omega. \end{aligned}$$

We want this to be 0 regardless of which Hamiltonian we chose. Thus we require $d\omega = 0$. It appears we have found the bare minimum for a physics to exist:

1.3. **Statement of the minimum required for a physics to exist.** For a reasonable "mechanics" to exist, we need an even-dimensional manifold $N^{2n}$ on which exists a nondegenerate 2-form $\omega$, which is *closed*: $d\omega = 0$. Any arbitrary function can be used as a Hamiltonian.

## 2. Vector fields and diffeomorphisms

2.1. **Diffeomorphisms define vector fields.** Let $\varphi_t : M \to M$ be a family of diffeomorphisms that are parameterized by $t$. For each $t$ it is possible to define a vector field $X_t$, which gives the "direction" of the diffeomorphism at each point. In coordinates, a diffeomorphism can be expressed

$$\varphi_t(x^1,\ldots,x^n) \;=\; \left(\varphi_t^1,\ldots,\varphi_t^n\right)$$

where each $\varphi_t^i$ is a function of the coordinates:

$$\varphi_t^i \;=\; \varphi_t^i(x^1,\ldots,x^n).$$

Fixing the time at $t = 0$, then given any point $p \in M$ we have

$$X(p) \;=\; \frac{d\varphi_t^i}{dt}\Big|_{t=0} \cdot \frac{\partial}{\partial x^i}.$$

If we fix the time at $t = t_0$, the formula is a little more complicated, due to the fact that the diffeomorphism has advanced in position

$$X_t(p) \;=\; \varphi_t^*\left(\frac{d\varphi_t^i}{dt}\Big|_{t=t_0}\right) \cdot \frac{d}{dx^i}.$$

Thus a family of diffeomorphism $\phi_t$ canonically gives rise, at time $t = 0$, to a vector field:

$$X \;=\; \frac{d\varphi_t^i}{dt}\Big|_{t=0} \frac{\partial}{\partial x^i}.$$

This is often called "differentiating" the family of diffeomorphisms, for obvious reasons. Can this be done in reverse? Does a vector field always give rise to a smooth family of diffeomorphisms?

2.2. **Vector fields define diffeomorphisms.** The answer is that if $X$ is a vector field of differentiability class $C^{0,1}$, then it does. First we state the theorem:

**Theorem 2.1** (Integration of vector fields)**.** *Let $X$ be a vector field of class $C^{0,1}$ on a manifold. If $p \in M$, then there is a path $\gamma : (-\epsilon, \epsilon) \to M$ so that $\gamma(0) = p$ and $\dot{\gamma}(t) = X(\gamma(t))$.*

<u>*Pf*</u> Let us use local coordinates $\{x^1, \ldots, x^n\}$ to examine the situation. Then if $\gamma$ is any path and $X$ the given vector field, we have

$$\gamma(t) = \left(\gamma^1(t), \ldots, \gamma^n(t)\right) \qquad X(p) = X^i(p)\frac{\partial}{\partial x^i}.$$

Also note that

$$\dot{\gamma}(t) \;=\; \frac{d\gamma^i}{dt}(t)\frac{\partial}{\partial x^i}.$$

Therefore solving $\dot{\gamma}(t) = X$ is therefore equivalent to solving the (nonlinear) system of ODES

$$\frac{d\gamma^i}{dt}(t) \;=\; X^i(\gamma^1(t), \ldots, \gamma^n(t)).$$

The ODE existence theorem states that this has a unique solution if the $X^i$ are $C^{0,1}$-differentiable. $\square$

## 3. The Lie Derivative

3.1. **Lie derivatives for objects that either push forward or pull back along diffeomorphisms.** Let $X$ be a vector field, and let $\phi_t : M \to M$ be its associated family of diffeomorphisms. If $Y$ is another vector field, we define

$$L_X Y(p) \;=\; \lim_{h\to 0} \frac{Y\big|_p \,-\, \varphi_{t*}\left(Y\big|_{\varphi_{-t}(p)}\right)}{h}.$$

That is, we push $Y$ forward along the flow, compare it to the $Y$ that was already defined there, and take a limit. This is called the Lie derivative of $Y$ along $X$.

There is no need to stop here: one can take a Lie derivative of any object either pushes forward or pulls back along a diffeomorphism. For instance, if $\eta$ is an object that pulls back under diffeomorphism, then for $p \in M$ we define the Lie derivative of $\eta$ along $X$, at the point $p$, to be

$$L_X\eta(p) \;=\; \lim_{h\to 0} \frac{\varphi_h^*\left(\eta\big|_{\varphi_h(p)}\right) \,-\, \eta\big|_p}{h}.$$

3.2. **Mixed objects.** Note the difference in treatment between objects that push forward and those that pull back. Before we unify our treatment of the Lie derivative, notice that, if $\eta$ pulls back, we have

$$L_X\eta \;=\; \lim_{h\to 0}\frac{\varphi_h^*\eta \,-\, \eta}{h} \;=\; \lim_{h\to 0}\varphi_{-h}^*\frac{\varphi_h^*\eta \,-\, \eta}{h}.$$

But then

$$L_X\eta \;=\; \lim_{h\to 0}\frac{\varphi_{-h}^*\varphi_h^*\eta \,-\, \phi_{-h}^*\eta}{h} \;=\; \lim_{h\to 0}\frac{\eta \,-\, \phi_{-h}^*\eta}{h}.$$

If we make the definition

$$\tilde{\varphi}_h = \varphi_{h*} \quad \textit{or} \quad \varphi_{-h}^*$$

as appropriate, then we can define

$$L_X A \;=\; \lim_{h\to 0}\frac{A \,-\, \tilde{\varphi}_h A}{h}.$$

Now it is possible to define a Lie derivative on mixed objects also. If $A \in \bigotimes^{k,l} M$ is some mixed tensor $A = V_1\otimes\cdots\otimes V_k\otimes\eta^1\otimes\cdots\otimes\eta^l$, then

$$L_X A \;=\; \lim_{h\to 0}\frac{A \,-\, \tilde{\varphi}_h A}{h}$$

where

$$\begin{aligned}\tilde{\varphi}_h(A) \;&=\; \tilde{\varphi}_h(V_1)\otimes\cdots\otimes\tilde{\varphi}_h(V_k)\otimes\tilde{\varphi}_h(\eta^1)\otimes\cdots\otimes\tilde{\varphi}_h(\eta^l)\\ &=\; \varphi_{h*}(V_1)\otimes\cdots\otimes\varphi_{h*}(V_k)\otimes\tilde{\varphi}_{-h}^*(\eta^1)\otimes\cdots\otimes\varphi_{-h}^*(\eta^l).\end{aligned}$$

3.3. **Properties of the Lie derivative.**

## 4. Introduction to Symplectic Geometry

4.1. **Definition of a symplectic manifold.** A 2-form $\omega$ with $d\omega = 0$ that is nondegenerate at every point of an even-dimensional manifold $N$ is called a *symplectic form*. A manifold $N$ that admits a symplectic form is called a *symplectic manifold*. The study of such pairs $(N,\omega)$ is the subject of *symplectic geometry*.

4.2. **Isomorphism between the tangent and cotangent spaces.** Similarities between Riemannian and symplectic geometry abound. One similarity is that a canonical isomorphisms $T_pM \to T_qM$ exist in both cases. We have already discussed this in the Riemannian case (the $\sharp$ and $\flat$ maps). In the symplectic case, we can interpret $\omega$ as a map $\omega : T_pM \to T_p^*M$ by

$$Y \in T_pM \quad \textit{maps to} \quad i_Y\omega \,\in\, T_p^*M.$$

One can prove that this map is an isomorphism, and therefore has an inverse. Unfortunately there is no widely accepted notation for these maps: the $\sharp$ and $\flat$ symbols are reserved for the Riemannian isomorphisms only.

4.3. **Hamiltonian flows.** Let $H : M \to \mathbb{R}$ be any smooth function. Such a function defines a kind of symplectic gradient vector field, usually given by $X_H$ which is defined (implicitly) by

$$i_{X_H}\omega \ = \ -dH.$$

The vector field $X_H$ is called the *Hamiltonian vector field* defined by $H$. Since any vector field defines a flow, we now have a *Hamiltonian flow*, $\varphi_t$, associated to $H$.

Let us examine how a function $f : M \to \mathbb{R}$ changes along the Hamiltonian flow.

$$L_{X_H}f \ = \ X_H(f) \ = \ -\omega^{-1}(dH)(\nabla f)$$

# Quantum Mechanics

November 2009

## 1 Basics of wave mechanics

Historically Schrödinger's wave mechanics was the second mathematical quantum theory, following Heisenberg's "matrix mechanics." The starting point for wave mechanics is the assumption that elementary particles are not point-like objects, but distributed, wave-like objects. The basic wave-like object is the *plane wave* (or *monochromatic wave*), given by

$$\Psi(\vec{q},\, t) \;=\; e^{i\left(\vec{k}\cdot\vec{q}\,-\,\omega t\right)}.$$

The number $\omega$ is the wave's radian frequency, and the vector $\vec{k}$ is called the *wave number*. The direction of the wave number indicates the wave's direction of travel, and its magnitude indicates the wave's "spatial frequency." The quantity $\omega/|\vec{k}|$ is the wave's speed, in the sense that this is speed at which the crests of either $Re(\Phi(\vec{q},t))$ or $Im(\Phi(\vec{q},t))$ will travel.

Introducing some physics, if $\Phi$ represents electromagnetic radiation for instance, we have the relationship between energy and frequency $E=\hbar\omega$, and and between energy and momentum $|\vec{p}|=E/c$, from which we derive $|\vec{k}|=|\vec{p}|$. Clearly also the direction of $\vec{k}$ and $\vec{p}$ are the same, so we have

$$\Psi(\vec{q},\, t) \;=\; e^{\frac{i}{\hbar}(\vec{p}\cdot\vec{q}\,-\,Et)}$$

and $E=|\vec{p}|c$. Any complex wave can then be expressed as a superposition of plane waves:

$$\Psi(\vec{q},\, t) \;=\; \int_{\mathbb{R}^3} f(p_1,p_2,p_3)\, e^{\frac{i}{\hbar}(\vec{p}\cdot\vec{q}\,-\,Et)}\, dp_1dp_2dp_3.$$

In general we assume any particle is also a superposition of plane waves, so let us examine plane waves in more detail. Taking derivatives, we get

$$\begin{aligned}\frac{\partial}{\partial q^i}\Psi \;&=\; -\frac{1}{i\hbar}\,p_i\,\Psi\\ \frac{\partial}{\partial t}\Psi \;&=\; \frac{1}{i\hbar}\,E\,\Psi.\end{aligned}$$

It stands to reason that we should make the identification

$$p_i = -i\hbar\frac{\partial}{\partial q^i} \tag{1}$$

$$E = i\hbar\frac{\partial}{\partial t}. \tag{2}$$

Any classical physical particle obeys the relationship $E = \frac{|\vec{p}|^2}{2m} + V(\vec{q})$ where $V$ is potential energy. Making the identifications (1) and (2) in this equation, we get the quantum relationship

$$i\hbar\frac{\partial}{\partial t} = \frac{-\hbar}{2m}\left(\left(\frac{\partial}{\partial q^1}\right)^2 + \left(\frac{\partial}{\partial q^2}\right)^2 + \left(\frac{\partial}{\partial q^3}\right)^2\right) + V.$$

Thus any physical quantum particle is given by a wave function $\Psi$ that satisfies

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2m}\triangle\Psi + V\Psi.$$

This is the *Schrödinger equation*. We discuss its physical interpretation below.

## 2 Hilbert spaces and Hermitian and unitary operators

A Hilbert space is a vector space $\mathcal{H}$ which is topologically complete, and which possesses a sesquilinear function $(\cdot,\,\cdot) : \mathcal{H} \to \mathcal{H}$ that obeys the following axioms: If $v, w \in \mathcal{H}$ and $\alpha \in \mathbb{C}$ then

- Linearity in the second variable: $(v,\,\alpha w) = \alpha\,(v,\,w)$.
- Hermitian-symmetry: $(v,\,w) = \overline{(w,\,v)}$. This implies two additional things: first that $(\alpha v,\,w) = \bar{\alpha}\,(v,\,w)$, and second that $(v,\,v)$ is always real.
- Nondegeneracy: $(v,\,v) \geq 0$, with equality if an only if $v = 0$.

Such a function $(\cdot,\,\cdot)$ is called an *Hermitian inner product*.

If $V$ is a vector space, the set of all linear operators is also a vector space. Is subspace of *bounded* linear operators is denoted $V^*$. If $V$ is a general vector space or even a Banach space, $V$ and $V^*$ usually little resemble on another (except in the finite-dimensional case). However any Hilbert spaces is *self-adjoint*, meaning that $\mathcal{H}^*$ is isomorphic (technically skew-isomorphic) to $\mathcal{H}$. The skew-isomorphism is given by

$$v \mapsto (v,\,\cdot)\,,$$

whenever $v \in \mathcal{H}$. Notice that this is strictly analogous to the $\flat$-map from Riemannian geometry.

If $A : \mathcal{H} \to \mathcal{H}$ is any linear operator, we can define its *Hermitian adjoint* $A^\dagger$ implicitly by

$$\left(A^\dagger v,\, w\right) \;=\; (v,\, Aw)\,.$$

In the finite dimensional case, any operator $A$ can be expressed as an $n \times n$ matrix. In this case $A^\dagger = \overline{A}^T$, the conjugate-transpose.

A linear operator is called *Hermitian* if $A = A^\dagger$, and anti-Hermitian if $A = -A^\dagger$. In other words, $A$ is Hermitian if

$$\langle Av,\, w\rangle \;=\; \langle v,\, Aw\rangle\,,$$

and anti-Hermitian if

$$\langle Av,\, w\rangle \;+\; \langle v,\, Aw\rangle \;=\; 0.$$

Note that if $A$ is Hermitian, then $iA$ is anti-Hermitian.

A linear operator $U : \mathcal{H} \to \mathcal{H}$ is called *unitary* if

$$\langle Uv,\, Uw\rangle \;=\; \langle v,\, w\rangle$$

(that is, unitary operators play the same role for complex vector spaces that orthogonal matrices play for real vector spaces).

We prove one final fact: if $A$ is a bounded anti-Hermitian operator, then the operator

$$e^A \;=\; I \;+\; A \;+\; \frac{1}{2!}A^2 \;+\; \ldots$$

is unitary. First note that $e^A$ exists, since $A$ is bounded. Obviously $e^0 = I$ is unitary, and also $\frac{d}{dt}e^{tA} \;=\; Ae^{tA}$. Thus we have

$$\begin{aligned}
\frac{d}{dt}\left\langle e^{tA}v,\, e^{tA}w\right\rangle \;&=\; \left\langle \frac{d}{dt}e^{tA}v,\, e^{tA}w\right\rangle \;+\; \left\langle e^{tA}v,\, \frac{d}{dt}e^{tA}w\right\rangle \\
&=\; \left\langle Ae^{tA}v,\, e^{tA}w\right\rangle \;+\; \left\langle e^{tA}v,\, Ae^{tA}w\right\rangle \\
&=\; 0.
\end{aligned}$$

The final equality holds because $A$ is anti-Hermitian. This proves that $\langle v,\, w\rangle \;=\; \left\langle e^{tA}v,\, e^{tA}w\right\rangle$ for all $t$.

# 3 Quantum Mechanics

## 3.1 Further consideration of the Schrödinger equation

If $H(\vec{q}, \vec{p}, t)$ is the classical Hamiltonian, we can construct a quantum Hamiltonian by using the rule (1) above (there could be ambiguities in this conversion, due to the noncommutativity of operators, but common sense can usual dispels of these). The classical identification of $E$ with the $H$ and the identification of $E$ with $i\hbar\partial/\partial t$ leads to the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi \;=\; H\Psi.$$

The abstract solution (assuming $H$ does not depend explicitly on time) is

$$\Psi(\vec{q},t) \;=\; e^{-\frac{i}{\hbar}(t-t_0)H}\Psi(\vec{q},t_0).$$

Before doing more work in quantum mechanics, we have to consider the meaning of this equation a little more thoroughly.

In the Schrödinger picture we considered particles to be wave functions, that is, complex valued function. In addition, we have to assume that $\int_{\mathbb{R}^3}|\Psi|^2 = 1$, or else the probability-density interpretation would fail. Since each wave function is in $L^2$ and $L^2$ is a Hilbert space, this suggests that the set of physical states (wave functions) should occupy a Hilbert space. Notice also that what had been dynamical variables in the classical setting, namely $E$ and $p_i$, are now operators. Whereas multiplying dynamical variable is a commutative operation (eg $p_1q^1 = q^1p_1$), composing operators is often not commutative. For instance

$$[p_1,q^1]f \;\;=\;\; -i\hbar[\frac{\partial}{\partial q^1},q^1]f \;=\; -i\hbar\left(\frac{\partial}{\partial q^1}(q^1f) \;-\; q^1\frac{\partial}{\partial q^1}f\right) \;=\; -i\hbar\, f,$$

so that $[p_1,q^1] = -i\hbar$.

## 3.2 Definition of Quantum Mechanics

We define quantum mechanics to be a complex Hilbert space $\mathcal{H}$, also called state space, equipped with a set of Hermitian operators. Elements of $\mathcal{H}$ are called *kets*, usually dented $\big|v\big\rangle$ where $v$ is some index. In addition there is a distinguished Hermitian operator $H:\mathcal{H}\to\mathcal{H}$, called the Hamiltonian, and so that any state $\big|v\big\rangle \in \mathcal{H}$ evolves with time according to the equation

$$i\hbar\frac{\partial}{\partial t}\big|v\big\rangle \;=\; H\big|v\big\rangle.$$

Normally the set of operators are generated by two sets of operators: $\{p_i\}_{i=1}^m$, $\{q^i\}_{i=1}^m$, where $p_i:\mathcal{H}\to\mathcal{H}$ and $q^i:\mathcal{H}\to\mathcal{H}$, and the commutators are

$$[q^i,\, p_j] \;=\; i\hbar\delta^i_j \qquad [q^i,\, q^j] \;=\; 0 \qquad [p_i,\, p_j] \;=\; 0.$$

Note that a set of Hermitian operators never form a Lie algebra, as the bracket of two Hermitians is an anti-Hermitian.

### 3.3 Bras and kets

If $|v\rangle$ and $|w\rangle$ are kets and since ket-space $\mathcal{H}$ is a Hilbert space, there is an inner product between them:

$$(|v\rangle, |w\rangle) \in \mathbb{C}\,.$$

Note that

$$(|v\rangle, \cdot)$$

is an element of $\mathcal{H}^*$. Any element of $\mathcal{H}^*$ is called a *bra*, and is denoted $\langle v|$. Obviously any bra operates on the kets, as follows

$$\begin{aligned} \langle v| \;&\triangleq\; (\langle v|, \cdot) \\ \langle v|w\rangle \;&\triangleq\; (|v\rangle, |w\rangle)\,. \end{aligned}$$

If $A : \mathcal{H} \to \mathcal{H}$ is an operator, then the symbol

$$\langle v|A$$

is interpreted to mean act by $A$, then act by $\langle v|$. That is to say,

$$\langle v|A|w\rangle \;=\; \langle v|\Big(A|w\rangle\Big) \;=\; (\langle v|, A|w\rangle)\,.$$

Using this definition, it is easy to prove that the bra corresponding to the ket $A|v\rangle$ is $\langle v|A^\dagger$.

### 3.4 Quantum numbers and probabilities

If $A$ is an observable (an Hermitian operator $\mathcal{H} \to \mathcal{H}$) and $|v\rangle$ is a quantum state (an element of $\mathcal{H}$), then $|v\rangle$ is said to be in the pure $A$-state $a$ if $a \in \mathbb{C}$ and

$$A|v\rangle \;=\; a|v\rangle.$$

The following terminology is also used: $|v\rangle$ has quantum number $a$ of type $A$ if $A|v\rangle = a|v\rangle$.

The *expected value* of $A$ on the quantum state $|v\rangle$ is defined to be the quantity

$$\langle A\rangle \;\triangleq\; \langle v|A|v\rangle$$

(recall that we assume $\langle v|v\rangle = 1$). We can also define the standard deviation $\triangle A$ of the observable $A$ on the state $|v\rangle$:

$$(\triangle A)^2 \;\triangleq\; \Big\langle (A - \langle A\rangle)^2 \Big\rangle \;=\; \langle A^2\rangle - \langle A\rangle^2\,.$$

## 3.5 Unitary Evolution

A system $\big|v\big\rangle$ evolves according to the equation $i\hbar\big|v\big\rangle = H\big|v\big\rangle$ where the Hamiltonian $H : \mathcal{H} \to \mathcal{H}$ is some distinguished Hermitian operator. The solution of this equation is

$$\big|v(t)\big\rangle \;=\; U(t)\big|v(0)\big\rangle$$

where

$$U(t) \;=\; e^{\frac{i}{\hbar}tH}$$

is a unitary operator, called the evolution operator. The map $U(t) : \mathcal{H} \to \mathcal{H}$ defines a unitary evolution of Hilbert space, called the Hamiltonian flow.

## 3.6 Finding explicit solutions of the Schrödinger equation

If $\big|E\big\rangle$ is an eigenvector of the Heisenberg equation

$$H\big|E\big\rangle \;=\; E\big|E\big\rangle$$

where $E$ is some number, then

$$e^{\frac{i}{\hbar}Et}\big|E\big\rangle$$

solves the Schrödinger equation.

In principle at least, the Hamiltonian is diagonalizable and any physical state is a superposition of its eigenstates. However in many of not most quantum systems, one must contend with the problem of non-normalizability of eigenstates.

## 3.7 Interpretation of the wave function

If $\Psi$ is the Schrödinger wave function for some particle, the quantity $\rho(\vec{q},t) = |\Psi(\vec{q},t)|^2$ is interpreted to be the probability density of finding the particle at position $\vec{q}$ at time $t$. The probability of finding the particle in the region of space $\Omega$ is

$$\int_\Omega \overline{\Psi}\Psi\, dq^1 dq^2 dq^3,$$

so that the probability of finding it at *some* location is

$$\int_{\mathbb{R}^3} \overline{\Psi}\Psi\, dq^1 dq^2 dq^3 \;=\; \big\langle\Psi\big|\Psi\big\rangle.$$

This probability is of course unity, so that we require

$$\big\langle\Psi\big|\Psi\big\rangle \;=\; 1.$$

Since the wave function $\Psi(\vec{q},t)$ at time $t$ is associated to the state vector $\big|\Psi\big\rangle(t) = U(t)\big|\Psi\big\rangle(0)$ at time $t$, we have another interpretation of the fact that the evolution operator is unitary:

$$\Big|U(t)\big|\Psi\big\rangle\Big| \;=\; \big\langle\Psi|U(t)^{\dagger}U(t)\big|\Psi\big\rangle \;=\; \big\langle\Psi\big|\Psi\big\rangle \;=\; 1,$$

which is conservation of total probability.

Now if $\rho = \overline{\Psi}\Psi$ is the probability density, we can find the probability current density as follows

$$\frac{\partial\rho}{\partial t} \;=\; \overline{\frac{\partial\Psi}{\partial t}}\Psi \,+\, \overline{\Psi}\frac{\partial\Psi}{\partial t} \;=\; \overline{\frac{i}{\hbar}H\Psi}\Psi \,+\, \overline{\Psi}\frac{i}{\hbar}H\Psi \;=\; -2\hbar^{-1}Im\left(\overline{\Psi}H\Psi\right).$$

For instance consider the case of a natural Hamiltonian, $H = -\frac{\hbar^2}{2m}\triangle \,+\, V$. Then

$$\begin{aligned}
-2\hbar^{-1}Im\left(\overline{\Psi}H\Psi\right) \;&=\; \frac{i\hbar}{2m}\left(\overline{\Psi}H\Psi \,-\, \Psi\overline{H\Psi}\right)\\
&=\; \frac{i\hbar}{2m}\left(\overline{\Psi}\triangle\Psi \,-\, \Psi\triangle\overline{\Psi}\right)\\
&=\; \frac{i\hbar}{2m}\nabla\cdot\left(\overline{\Psi}\nabla\Psi \,-\, \Psi\nabla\overline{\Psi}\right)\\
&=\; -\frac{\hbar}{m}\nabla\cdot Im\left(\overline{\Psi}\nabla\Psi\right),
\end{aligned}$$

where $\nabla\cdot\vec{X}$ indicates the divergence of the vector field $\vec{X}$. Thus we can write

$$\frac{\partial\rho}{\partial t} \,+\, \frac{\hbar}{m}\nabla\cdot Im\left(\overline{\Psi}\nabla\Psi\right) \;=\; 0.$$

Of course this is the typical "continuity equation" (aka conservation of mass) from fluid mechanics, and indicates that the probability current density of a quantum system with a natural Hamiltonian must be

$$\vec{J} \;=\; \frac{\hbar}{m}Im\left(\overline{\Psi}\nabla\Psi\right).$$

# 4 Heisenberg's wave mechanics

In addition to the Schrödinger picture of quantum mechanics, there are other ways of making the abstract formulation of QM concrete. One of these is Heisenberg's "matrix mechanics". The physical system is considered to be static, but the observables change according to the formula

$$\frac{dA}{dt} \;=\; \frac{i}{\hbar}\left[H,\, A\right] \,+\, \frac{\partial A}{\partial t}.$$

It is possible to derive the Heisenberg equation from the Schrödinger equation. To do so, let $A, B, H, \ldots$ be operators on Hilbert space. In the Schrödinger representation they are interpreted to be fixed, and the physical system (which is a point of Hilbert space) changes. In the Heisenberg representation, the physical system is considered to be fixed, and it is the operators which change. To distinguish the Heisenberg operators from the Schrödinger operators we will denote the Heisenberg operators with a lower-vase $h$, namely $A_h, B_h, H_h, etc.$ Let $U(t) = e^{\frac{i}{\hbar}Ht}$ be the Schrödinger evolution operator. We assert that

$$A_h \;=\; U(t)AU(-t),$$

and prove that the operators $A_h, etc$ so defined satisfy the Heisenberg equation.

$$\begin{aligned}
\frac{dA_h}{dt} &= \frac{dU}{dt}AU^{-1} \,+\, UA\frac{dU^{-1}}{dt} \,+\, U\frac{dA}{dt}U^{-1} \\
&= \frac{i}{\hbar}UHAU^{-1} \,-\, \frac{i}{\hbar}UAHU^{-1} \,+\, U\frac{\partial A}{\partial t}U^{-1} \\
&= \frac{i}{\hbar}UHU^{-1}UAU^{-1} \,-\, \frac{i}{\hbar}UAU^{-1}UHU^{-1} \,+\, U\frac{\partial A}{\partial t}U^{-1} \\
&= \frac{i}{\hbar}\left(H_hA_h \,-\, A_hH_h\right) \,+\, \frac{\partial A_h}{\partial t},
\end{aligned}$$

which is the Heisenberg equation.

# 5 The uncertainty relations

Let $A$ and $B$ be Hermitian operators on $\mathcal{H}$. Using Hermiticity, we compute

$$\langle [A,B]\rangle \;\triangleq\; \langle v|AB|v\rangle \,-\, \langle v|BA|v\rangle \;=\; \langle v|AB|v\rangle \,-\, \overline{\langle v|AB|v\rangle} \;=\; 2i\,\mathrm{Im}\,\langle v|AB|v\rangle.$$

The Schwartz inequality says $\langle v|w\rangle^2 \leq \langle v|v\rangle\langle w|w\rangle$, with equality if and only if $|v\rangle$ is a (complex) multiple of $|w\rangle$. The same inequality can be used with the kets $A|v\rangle$ and $B|v\rangle$ to give

$$\Big|\langle [A,B]\rangle\Big| \;\leq\; 2\Big|\langle v|AB|v\rangle\Big| \;\leq\; 2\langle v|A^2|v\rangle\langle v|B^2|v\rangle \;=\; 2\sqrt{\langle A^2\rangle}\sqrt{\langle B^2\rangle}.$$

Replacing $A$, $B$ by $A-a$, $B-b$, respectively, where $a$, $b$ are any constants, we have $[A,B] = [A-a, B-b]$, so we have

$$\Big|\langle [A,B]\rangle\Big| \;\leq\; 2\sqrt{\langle (A-a)^2\rangle}\sqrt{\langle (B-b)^2\rangle}.$$

Putting $a = \langle A\rangle$ and $b = \langle B\rangle$ we have

$$\Big|\langle [A,B]\rangle\Big| \;\leq\; 2\sqrt{\langle (A-\langle A\rangle)^2\rangle}\sqrt{\langle (B-\langle B\rangle)^2\rangle} \;=\; 2\triangle A\triangle B.$$

This is called the Heisenberg inequality. It says that the product of variances of any two quantities associated to a quantum system is greater than a predetermined minimum, namely the average of the commutator. On the intuitive level, if Hermitian operators represent measurements, then the uncertainty in the measurements of two variables can never be simultaneously reduced to an arbitrarily small value, unless the operators commute.

As an application, recall that $\left[p_i, q^j\right] = -i\hbar\delta_i^j$. Thus the Heisenberg inequality gives

$$\delta_i^j \hbar/2 \quad \leq \quad \triangle p_i \triangle q^j.$$

A second type of Heisenberg inequality is the time-energy relation. To prove this, we use the Heisenberg equation $\frac{dA}{dt} = \frac{i}{\hbar}[H, A]$, to get

$$\hbar \Big| \left\langle \frac{dA}{dt} \right\rangle \Big| \quad = \quad \big| \langle [H, A] \rangle \big| \ \leq \ 2 \triangle H \triangle A.$$

Using $\langle dA/dt \rangle = d \langle A \rangle / dt$ (proved below), let us formally define

$$\triangle \tau_A \quad \triangleq \quad \triangle A \, / \, \Big| \frac{d \langle A \rangle}{dt} \Big|.$$

This quantity is interpreted to be the amount of time for the statistics of $A$ to change appreciably, or more specifically, for the A-average of the system to move by an amount equal to the standard distribution of $A$. The Heisenberg inequality now reads

$$\hbar/2 \quad \leq \quad \triangle H \triangle \tau_A.$$

# 6 The classical limit

In the limit where $\hbar \to 0$ (i.e. in the physics where $\hbar$ is inconsequentially small), the laws of quantum mechanics must reduce to the laws of classical mechanics; the success or failure of this reduction provides the first theoretical test of the physical validity of quantum mechanics. We examine how this occurs in the two representations of quantum mechanics we have studied so far.

## 6.1 $\hbar \to 0$ in the Heisenberg representation

In this section we always assume our observables be Heisenberg observables. We prove two theorems first.

**Theorem 6.1** *If $A : \mathcal{H} \to \mathcal{H}$ is Hermitian then*

$$\left\langle \frac{dA}{dt} \right\rangle \quad = \quad \frac{d \langle A \rangle}{dt}.$$

Pf

$$\frac{d\langle A\rangle}{dt} \;=\; \frac{d}{dt}\langle v|A|v\rangle \;=\; \left\langle v\middle|\frac{dA}{dt}\middle|v\right\rangle \;=\; \left\langle\frac{dA}{dt}\right\rangle.$$

□

**Theorem 6.2** *If the Hermitian operator $A$ is a composition of various $p_i$ and $q^i$ then*

$$\begin{aligned}\left[q^i, A\right] &= i\hbar\frac{\partial A}{\partial p_i}\\ [p_i, A] &= -i\hbar\frac{\partial A}{\partial q^i}.\end{aligned}$$

Pf
Since $[p_i, q^i] = -i\hbar$ and all other commutators are zero, we compute

$$\begin{aligned}[p_i, (q^i)^n] &= p_i q^i \dots q^i \,-\, q^i \dots q^i p_i \;=\; p_i q^i \dots q^i \,-\, q^i \dots q^i p_i q^i \,-\, q^i \dots q^i [q^i, p_i]\\ &= p_i q^i \dots q^i \,-\, q^i \dots q^i p_i q^i \,-\, i\hbar (q^i)^{n-1}.\end{aligned}$$

continuing to apply commutators, we eventually get

$$\begin{aligned}[p_i, (q^i)^n] &= p_i q^i \dots q^i \,-\, q^i p_i q^i \dots q^i p_i q^i \,-\, (n-1) i\hbar (q^i)^{n-1}\\ &= -ni\hbar \left(q^i\right)^{n-1}.\end{aligned}$$

Similarly we compute $[q^i, (p_i)^n] = in\hbar(p_i)^{n-1}$. □

Now we can derive the promised reduction of quantum-to-classical:

$$\frac{d\langle q^i\rangle}{dt} \;=\; \left\langle\frac{dq^i}{dt}\right\rangle \;=\; \left\langle\frac{i}{\hbar}[H, q^i]\right\rangle \;=\; \left\langle\frac{\partial H}{\partial p_i}\right\rangle$$

and likewise for $p_i$. We therefore have the *Ehrenfest Equaitons*

$$\frac{d\langle q^i\rangle}{dt} \;=\; \left\langle\frac{\partial H}{\partial p_i}\right\rangle \qquad \frac{d\langle p_i\rangle}{dt} \;=\; -\left\langle\frac{\partial H}{\partial q^i}\right\rangle.$$

In the limit where $\hbar \to 0$, the obstruction represented by the Heisenberg inequalities to both the position and momentum distributions from being arbitrarily $L^1$-close to delta functions disappears. In this limit, the observables $p_i$ and $q^i$, and therefore the observables $H$, $\partial H/\partial q^i$ and $\partial H/\partial p_i$, have well-defined values. We can denote these values simply by $p_i$, $q_i$, $H$, $\partial H/\partial q^i$, and $\partial H/\partial p_i$, respectively. Using this notation, we see that in the classical limit we recover Hamilton's canonical equations

$$\frac{dq^i}{dt} \;=\; \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} \;=\; -\frac{\partial H}{\partial q^i}.$$

## 6.2 $\hbar \to 0$ in the Schrödinger representation

Let $\Psi$ be a wave function that evolves according to the Schrödinger equation. Assume the Hamiltonian is natural, meaning $H = -\frac{\hbar^2}{2m}\triangle + V$. Writing

$$\Psi = A\, e^{\frac{i}{\hbar}S}$$

where $A$ and $S$ are real-valued function, and plugging into the Schrödinger equation

$$i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\triangle\Psi + V\Psi,$$

we get

$$\begin{aligned} & i\hbar\frac{\partial A}{\partial t}e^{\frac{i}{\hbar}S} - A\frac{\partial S}{\partial t}e^{\frac{i}{\hbar}S} \\ & \quad = -\frac{\hbar^2}{2m}\left(\triangle A\, e^{\frac{i}{\hbar}S} + 2\frac{i}{\hbar}\langle \nabla A, \nabla S\rangle\, e^{\frac{i}{\hbar}S} - \frac{1}{\hbar^2}A|\nabla S|^2 e^{\frac{i}{\hbar}S} + \frac{i}{\hbar}A\triangle S\, e^{\frac{i}{\hbar}S}\right) + VAe^{\frac{i}{\hbar}S} \end{aligned}$$

Canceling the $e^{iS/\hbar}$ factor and separating the real and imaginary parts,

$$\begin{aligned} \frac{\partial S}{\partial t} &= \frac{\hbar^2}{2m}\frac{1}{A}\triangle A - \frac{1}{2m}|\nabla S|^2 - V \\ \frac{\partial A}{\partial t} &= -\frac{1}{2m}\left(2\langle \nabla A, \nabla S\rangle + A\triangle S\right). \end{aligned}$$

Now $A^2 = |\Psi|^2 = \rho$ has physical meaning, so multiplying the second equation by $A$ gives

$$\frac{\partial S}{\partial t} = \frac{\hbar^2}{2m}\frac{1}{A}\triangle A - \frac{1}{2m}|\nabla S|^2 - V \qquad (3)$$

$$\frac{\partial A^2}{\partial t} = -\frac{1}{m}\nabla\cdot\left(A^2\nabla S\right). \qquad (4)$$

Equation (4) is precisely the continuity equation $\frac{\partial \rho}{\partial t} = -\frac{\hbar}{m}\nabla\cdot Im\left(\overline{\Psi}\nabla\Psi\right)$ from section 3.7.

Equation (3) can be seen as a perturbation of the Hamilton-Jacobi equation, where $S$ is equated with the classical action. After taking $\hbar \to 0$, (3) is precisely the Hamilton-Jacobi equation. This allows the interpretation that, in the limit where $\hbar$ is infinitesimally small, the quantum wave function is an ensemble of non-interacting particles, with some initial distribution of positions and momenta, that evolves according to the classical laws of motion.

# Spinors and the Dirac Equation

Brian Weber

*December 26, 2009*

## 1 Symmetries of the Universe

The Dirac equation can be derived using the relativistic relation between mass and energy $E = m$ (normalized units) and the symmetries of the universe, so it is necessary to have a firm grasp on the latter. The relativistic symmetry group is the orthochronous Poincare group $SO^+(1,3) \ltimes \mathbb{R}^4$, though we shall mainly use just the orthochronous Lorentz group $SO^+(1,3)$. Its Lie algebra is $\mathfrak{so}(1,3)$. We first study the Euclidean case of three spatial dimensions.

## 2 Spinors over 3-space

First we shall study the orthogonal group $SO(3)$. It's Lie algebra $\mathfrak{so}(3)$ has basis

$$\mathrm{j}^x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \qquad \mathrm{j}^y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \qquad \mathrm{j}^z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with brackets

$$[\mathrm{j}^x, \mathrm{j}^y] = \mathrm{j}^z \qquad [\mathrm{j}^y, \mathrm{j}^z] = \mathrm{j}^x \qquad [\mathrm{j}^z, \mathrm{j}^x] = \mathrm{j}^y .$$

The group $SU(2)$ has the same Lie algebra: $\mathfrak{so}(3) \approx \mathfrak{su}(2)$. The Lie algebra $\mathfrak{su}(2)$ is the vector space of trace-free anti-hermitian matrices. The modified Pauli matrices $\frac{1}{2i}\sigma^1$, $\frac{1}{2i}\sigma^2$, $\frac{1}{2i}\sigma^3$ are a basis, where

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

One computes

$$\left[\frac{1}{2i}\sigma^1, \frac{1}{2i}\sigma^2\right] = \frac{1}{2i}\sigma^3 \qquad \left[\frac{1}{2i}\sigma^2, \frac{1}{2i}\sigma^3\right] = \frac{1}{2i}\sigma^1 \qquad \left[\frac{1}{2i}\sigma^3, \frac{1}{2i}\sigma^1\right] = \frac{1}{2i}\sigma^2.$$

In fact the anti-Hermitian matrices $\frac{1}{i}\sigma^i$ form a basis for the purely imaginary quaternions. Adjoining $\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ we have

$$span_{\mathbb{R}}\{\sigma^0,\ \frac{1}{i}\sigma^1,\ \frac{1}{i}\sigma^2,\ \frac{1}{i}\sigma^3\} \ \approx \ \mathbb{H}.$$

It is possible to obtain an orthogonal action of $SU(2)$ on $\mathbb{R}^3$ as follows. First map $\mathbb{R}^3$ onto the three-space of trace-free antihermitian matrices, meaning we map a covector $\eta = (x, y, z) \in \mathbb{R}^3$ to

$$\begin{aligned} H_\eta \ &= \ \begin{pmatrix} -iz & -ix \, - \, y \\ -ix \, + \, y & iz \end{pmatrix} \\ &= \ -ix\sigma^1 \, - \, iy\sigma^2 \, - \, iz\sigma^3 \ = \ -i\eta\vec{\sigma} \end{aligned}$$

where

$$\vec{\sigma} \ = \ \begin{pmatrix} \sigma^1 \\ \sigma^2 \\ \sigma^3 \end{pmatrix}.$$

Notice that $\det(X_\eta) = |\eta|^2$.

Assigning a vector to a trace-free antihermitian matrix defines a representation of $\mathbb{R}^3$ on $\mathbb{C}^2$: the action of an element $\eta = (x, y, z)$ on an element $v = \binom{v_1}{v_2} \in \mathbb{C}^2$ is by

$$\eta \cdot v \ \triangleq \ \begin{pmatrix} -iz & -ix \, + \, y \\ -ix \, - \, y & iz \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

In fact this is a *Clifford representation.* To see this, we check the Clifford relation $\eta \cdot \eta + |\eta|^2 \ = \ 0$:

$$\begin{aligned} \eta \cdot \eta \cdot v \ = \ H_\eta H_\eta v \ &= \ \begin{pmatrix} -iz & -ix \, + \, y \\ -ix \, - \, y & iz \end{pmatrix} \begin{pmatrix} -iz & -ix \, + \, y \\ -ix \, - \, y & iz \end{pmatrix} v \\ &= \ -\left(x^2 + y^2 + z^2\right) \, v \ = \ -|\eta|^2 v \end{aligned}$$

There is also an orthogonal action of $SU(2)$ on $\mathbb{R}^3$. If the covector $\eta$ is identified with the matrix $H_\eta$ then we define $U(\eta)$ by identifying it with the matrix $UH_\eta\overline{U}^T$. Namely

$$\begin{aligned} \eta \ &\leftrightarrow H_\eta \\ U(\eta) \ &\leftrightarrow \ UH_\eta\overline{U}^T. \end{aligned}$$

It is easy to check that this operation preserves antihermiticity and the determinant, so the action is well-defined and orthogonal. Equally importantly, it commutes with Clifford multiplication: if $\eta, \mu \in \mathbb{R}^3$ and $v \in \mathbb{C}^2$ then

$$\begin{aligned} U(\eta) \cdot U(\mu) &= UH_\eta \overline{U}^T UH_\mu \overline{U}^T = UH_\eta H_\mu \overline{U}^T = U(\eta \cdot \mu) \\ U(\eta) \cdot U(v) &= UH_\eta \overline{U}^T Uv = UH_\eta v = U(\eta \cdot v). \end{aligned}$$

Note that $SU(2)$ is connected and simply connected. One can prove that if $U \in SU(2)$ acts as the identity (meaning $U(\eta) = \eta$ for all $\eta$), then $U \in \{I, -I\}$. Therefore $SU(2)$ is a nontrivial double cover and in fact the universal cover of $SO(3)$. Incidentally, since $SU(2)$ is topologically the 3-sphere, $SO(3)$ is topologically $\mathbb{R}\mathrm{P}^3$

The Lie group $\mathrm{Spin}(r, s)$ is by definition the nontrivial double cover of $SO^+(r, s)$, so we have that $\mathrm{Spin}(3) = SU(2)$. A space on which $\mathrm{Spin}(r, s)$ acts irreducibly is called a *Spinor space of signature* $(r, s)$, or $S_{r,s}$. Thus $S_3 = \mathbb{C}^2$. A *Pauli spinor* is an element of $\mathbb{C}^2$ on which $SU(2)$ acts as the symmetry group. Note there is a canonical Hermitian inner product on the space of Pauli spinors, and a canonical volume form.

# 3 Spinors over Minkowski space

## 3.1 $SO(1, 3)$, $Spin(1, 3)$, and $SL(2, \mathbb{C})$

Now we consider the symmetries of Minkowski space. The orthochronous Lorentz group, $SO^+(1, 3)$, is defined to be those rigid transformations of $\mathbb{R}^{1,3}$ that preserve both orientation and the direction of time. Its Lie algebra has a basis of the three boost generators $\mathcal{K}^x$, $\mathcal{K}^y$, and $\mathcal{K}^z$, representing infinitesimal changes in velocity along the three axes, and the three rotations $\mathrm{j}^x$, $\mathrm{j}^y$, $\mathrm{j}^z$. Together, these are

$$\mathcal{K}^x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \mathcal{K}^y = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\mathcal{K}^z = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \qquad \mathrm{j}^x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathrm{j}^y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \qquad \mathrm{j}^z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The brackets are

$$\begin{array}{lll}
[\mathcal{K}^x, \mathcal{K}^y] \;=\; -\,\mathfrak{j}^z & [\mathcal{K}^y, \mathcal{K}^z] \;=\; -\,\mathfrak{j}^x & [\mathcal{K}^z, \mathcal{K}^x] \;=\; -\,\mathfrak{j}^y \\
[\mathfrak{j}^x, \mathfrak{j}^y] \;=\; \mathfrak{j}^z & [\mathfrak{j}^y, \mathfrak{j}^z] \;=\; \mathfrak{j}^x & [\mathfrak{j}^z, \mathfrak{j}^x] \;=\; \mathfrak{j}^y \\
[\mathcal{K}^x, \mathfrak{j}^x] \;=\; 0 & [\mathcal{K}^y, \mathfrak{j}^y] \;=\; 0 & [\mathcal{K}^z, \mathfrak{j}^z] \;=\; 0 \\
[\mathcal{K}^x, \mathfrak{j}^y] \;=\; \mathcal{K}^z & [\mathcal{K}^y, \mathfrak{j}^z] \;=\; \mathcal{K}^x & [\mathcal{K}^z, \mathfrak{j}^x] \;=\; \mathcal{K}^y \\
[\mathfrak{j}^x, \mathcal{K}^y] \;=\; \mathcal{K}^z & [\mathfrak{j}^y, \mathcal{K}^z] \;=\; \mathcal{K}^x & [\mathfrak{j}^z, \mathcal{K}^x] \;=\; \mathcal{K}^y\,.
\end{array}$$

Note that the boosts do not generate an algebra; this is the "Thomas procession." The algebra $\mathfrak{so}(1,3)$ is simple, but the complexification splits: putting $\mathfrak{a}^x = \frac{1}{2}(\mathfrak{j}^x - i\,\mathcal{K}^x)$, etc and $\mathfrak{b}^x = \frac{1}{2}(\mathfrak{j}^x + i\,\mathcal{K}^x)$, etc, then the $\mathfrak{a}'\,s$ and $\mathfrak{b}'\,s$ commute, and $[\mathfrak{a}^x, \mathfrak{a}^y] = \mathfrak{a}^z$ and cyclic permutations, and $[\mathfrak{b}^x, \mathfrak{b}^y] = \mathfrak{b}^z$ and cyclic permutations. Notice that the $\mathfrak{a}'\,s$ and $\mathfrak{b}'\,s$ are antihermitian.

Now consider the three dimensional complex Lie group $SL(2,\mathbb{C})$, defined to be the $2\times 2$ complex matrices with unit determinant. Its Lie algebra consists of the trace-free complex matrices, a (complex) basis of which is

$$\mathfrak{c}^x \;=\; \frac{i}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \mathfrak{c}^y \;=\; \frac{i}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \mathfrak{c}^z \;=\; \frac{i}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

(the Pauli matrices again). Therefore $\mathfrak{sl}(2,\mathbb{C}) = \mathfrak{su}(2)\otimes\mathbb{C}$, the complexification of $\mathfrak{su}(2)$. We have just seen that $\overline{\text{the group}}$ $\mathfrak{sl}(2,\mathbb{C})$ bears a relationship to $\mathfrak{so}(1,3)$ as well: $\mathfrak{so}(1,3)\otimes\mathbb{C} \approx \mathfrak{sl}(2,\mathbb{C})\oplus\overline{\mathfrak{sl}(2,\mathbb{C})}$. Let $\lambda, \rho : \mathfrak{so}(1,3)\otimes\mathbb{C} \to \mathfrak{sl}(2,\mathbb{C})$ be the projections onto $span_{\mathbb{C}}\{\mathfrak{a}^x, \mathfrak{a}^y, \mathfrak{a}^z\}$ and $span_{\mathbb{C}}\{\mathfrak{b}^x, \mathfrak{b}^y, \mathfrak{b}^z\}$, respectively. These yield inequivalent representations of the real algebra $\mathfrak{so}(1,3)$ on $\mathbb{C}^2$, called the $+$ and $-$ representations. To check the inequivalence, notice that if $\pi \in \mathfrak{so}(1,3)$ then $\lambda(\pi) = \overline{\rho(\pi)}$, so that for instance $\lambda(\mathfrak{j}^x\,\mathfrak{j}^y\,\mathfrak{j}^z) = -\rho(\mathfrak{j}^x\,\mathfrak{j}^y\,\mathfrak{j}^z)$.

The algebra $\mathfrak{so}(1,3)\otimes\mathbb{C} = \mathfrak{sl}(2,\mathbb{C})\oplus\overline{\mathfrak{sl}(2,\mathbb{C})}$ consists of $4\times 4$ complex matrices

$$\mathfrak{U} \;=\; \begin{pmatrix} \mathfrak{u} & 0 \\ 0 & \overline{\mathfrak{v}} \end{pmatrix},$$

where $\mathfrak{u}, \mathfrak{v} \in sl(2,\mathbb{C})$. The real subalgebra $\mathfrak{so}(1,3) \subset \mathfrak{so}(1,3)\otimes\mathbb{C}$ sits inside as follows

$$\begin{aligned}
\mathfrak{j}^x \;&=\; \begin{pmatrix} \mathfrak{a}^x & 0 \\ 0 & \mathfrak{b}^x \end{pmatrix} \;=\; \begin{pmatrix} \mathfrak{a}^x & 0 \\ 0 & \overline{\mathfrak{a}^x} \end{pmatrix}, \quad etc, \\
\mathcal{K}^x \;&=\; \begin{pmatrix} i\,\mathfrak{a}^x & 0 \\ 0 & -i\,\mathfrak{b}^x \end{pmatrix} \;=\; \begin{pmatrix} i\,\mathfrak{a}^x & 0 \\ 0 & \overline{i\,\mathfrak{a}^x} \end{pmatrix}, \quad etc.
\end{aligned}$$

In other words, it is the diagonal subalgebra

Exponentiation $\mathfrak{so}(1,3)\otimes\mathbb{C}$ gives $SL(2,\mathbb{C})\times SL(2,\mathbb{C})$, the group of $4\times 4$ matrices of the form

$$\begin{pmatrix} U & 0 \\ 0 & \overline{V} \end{pmatrix},$$

where $U, V \in \mathfrak{sl}(2, \mathbb{C})$. Therefore the exponentiation of the real (diagonal) subalgebra is the diagonal subgroup consisting of matrices of the form

$$\mathcal{U} \;=\; \begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix}, \quad U \in SL(2, \mathbb{C}).$$

The spin group $Spin(1, 3)$ is therefore $SL(2, \mathbb{C})$:

$$\begin{array}{ccc} Spin(1,3) & \xrightarrow{\;\approx\;} & SL(2,\mathbb{C}) \overset{diag}{\subset} SL(2,\mathbb{C}) \oplus \overline{SL(2,\mathbb{C})} \\ \big\downarrow & & \\ SO(1,3) & & \end{array}$$

The space on which $Spin(1, 3)$ acts as the symmetry group is (by definition) the spin space $\mathcal{S} = \mathcal{S}_{1,3} \approx \mathbb{C}^4$.

## 3.2 The symplectic form on half spin-spaces, and the four spin spaces

Consider the Clifford multiplication $\mathbb{R}^{1,3} : \mathcal{S} \to \mathcal{S}$. If $\omega$ is the complex volume form, then we have a way of splitting spin space $\mathcal{S}$ into two irreducible components

$$\begin{aligned} \mathcal{S}^+ &= \pi^+ \mathcal{S} \;=\; \frac{1}{2} (Id + \omega) \mathcal{S} \\ \mathcal{S}^- &= \pi^- \mathcal{S} \;=\; \frac{1}{2} (Id - \omega) \mathcal{S}\,. \end{aligned}$$

Clifford multiplication of vectors in $\mathbb{R}^{1,3}$ anti-commute with $\omega$ so that $\eta \in \mathbb{R}^{1,3}$ acts as $\eta : \mathcal{S}^\pm \to \mathcal{S}^\mp$. Since the factors of $Spin(1, 3)$ that act on $\mathcal{S}^+$ and $\mathcal{S}^-$ are complex conjugates of each other, we regard $\mathcal{S}^+ \;=\; \overline{\mathcal{S}^-}$. Spin space is therefore

$$\mathcal{S} \;=\; \mathcal{S}^+ \oplus \mathcal{S}^-.$$

Let $\epsilon$ be the unit antisymmetric $2 \times 2$ matrix:

$$\epsilon \;=\; \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

If $A$ is any $2 \times 2$ matrix then $A^T \epsilon A \;=\; \det(A)\epsilon$. Therefore $SL(2, \mathbb{C})$ is characterized as the group of $2 \times 2$ matrices $U$ with $U^T \epsilon U = \epsilon$. We can therefore define a symplectic form on $\mathcal{S}^+$, which is preserved by the symmetry group $SL(2, \mathbb{C})$:

$$\begin{aligned} \langle \psi,\, \varphi \rangle &= \psi^T \epsilon \varphi. \\ \langle U\psi,\, U\varphi \rangle &= (U\psi)^T \epsilon U \varphi \;=\; \psi^T U^T \epsilon U \varphi \;=\; \psi^T \epsilon \varphi \;=\; \langle \psi,\, \varphi \rangle\,. \end{aligned}$$

This symplectic form is nondegenerate, and therefore determines natural isomorphisms from the space $\mathcal{S}^+$ to its dual $\mathcal{S}^{+*}$, and from the space $\mathcal{S}^-$ to its dual $\mathcal{S}^{-*}$.

$$\begin{array}{rcl} \psi \in \mathcal{S}^+ & \mapsto & \langle \psi, \cdot \rangle \, \in \, \mathcal{S}^{+*} \\ \psi \, = \, \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} & \mapsto & \psi^T \epsilon \, = \, (\psi_2, \, -\psi_1). \end{array}$$

We now have four spin spaces

$$\mathcal{S}^+, \qquad \mathcal{S}^-, \qquad \mathcal{S}^{+*}, \qquad \mathcal{S}^{-*},$$

and natural isomorphisms amongst them: Letting $\psi = \binom{\psi_1}{\psi_2} \in \mathcal{S}^+$,

$$\begin{array}{ll} \mathcal{S}^+ \to \mathcal{S}^- & \psi \, \mapsto \, \overline{\psi} \\ \mathcal{S}^+ \to \mathcal{S}^{+*} & \psi \, \mapsto \, \psi^T \epsilon \\ \mathcal{S}^+ \to \mathcal{S}^{-*} & \psi \, \mapsto \, \overline{\psi}^T \epsilon \end{array}$$

and so forth. A *natural spinor* is an element of $\mathcal{S} \; = \; \mathcal{S}^+ \oplus \mathcal{S}^-$, in contrast to a *Dirac spinor*, which we below to be an element of $\mathcal{S}^+ \oplus \mathcal{S}^{-*}$. More on this presently.

### 3.3 Clifford representations

Now we study the action of a covector $\eta \in \mathbb{R}^{1,3}$ on $\mathcal{S}^+$. There are several isomorphisms between the 4-covectors (with the Minkowski norm) and $2 \times 2$ matrices. To give a few of many possible examples, if $\eta = (t, x, y, z) \in \mathbb{R}^{1,3}$ we can set

$$\begin{array}{rcl} M_\eta & = & \begin{pmatrix} -x - iy & -t + z \\ t + z & x - iy \end{pmatrix} \\ \tilde{M}_\eta & = & \begin{pmatrix} x - iy & t - z \\ -t - z & -x - iy \end{pmatrix} \\ H_\eta & = & \begin{pmatrix} t + z & x - iy \\ x + iy & t - z \end{pmatrix} \; = \; t\sigma^0 \, + \, x\sigma^1 \, + \, y\sigma^2 \, + \, z\sigma^3 \\ \tilde{H}_\eta & = & \begin{pmatrix} t - z & -x + iy \\ -x - iy & t + z \end{pmatrix} \; = \; t\sigma^0 \, - \, x\sigma^1 \, - \, y\sigma^2 \, - \, z\sigma^3. \end{array}$$

We are using $\sigma^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. These matrices satisfy

$$\det(M_\eta) \; = \; \det(\tilde{M}_\eta) \; = \; \det(H_\eta) \; = \; \det(\tilde{H}_\eta) \; = \; 0 \tag{1}$$

Notice that the matrices $H_\eta$ and $\tilde{H}_\eta$ are related by *quaternionic conjugation.* Also notice that none of the maps $\eta \mapsto M_\eta$, etc, constitutes a Clifford representation, as the Clifford relation always fails: $M_\eta M_\eta \neq -|\eta|^2 Id$, $\tilde{H}_\eta \tilde{H}_\eta \neq -|\eta|^2 Id$, etc. However $M_\eta \tilde{M}_\eta = -|\eta|^2$ and $H_\eta \tilde{H}_\eta = -|\eta|^2$, which gives a way of creating (at least) two Clifford representations:

$$\begin{aligned} \eta \quad &\mapsto \mathcal{M}_\eta \quad = \quad \begin{pmatrix} 0 & \tilde{M}_\eta \\ M_\eta & 0 \end{pmatrix} \\ \eta \quad &\mapsto \mathcal{H}_\eta \quad = \quad \begin{pmatrix} 0 & \tilde{H}_\eta \\ H_\eta & 0 \end{pmatrix}. \end{aligned}$$

But which spin spaces do these representations act on? To answer this question we examine how the symmetry group $SL(2,\mathbb{C})$ acts on these $4 \times 4$ matrices. If $M_{4\times 4} : \mathcal{S}^+ \oplus \mathcal{S}^- \to \mathcal{S}^+ \oplus \mathcal{S}^-$ is any $4 \times 4$ matrix then $U \in SL(2,\mathbb{C})$ must act on $M_{4\times 4}$ by

$$U(M_{4\times 4}) \; = \; \begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix} M_{4\times 4} \begin{pmatrix} U^{-1} & 0 \\ 0 & \overline{U}^{-1} \end{pmatrix}.$$

So we ask, is $U(\mathcal{M}_\eta)$ of the form $\mathcal{M}_\mu$? Is $U(\mathcal{H}_\eta)$ of the form $\mathcal{H}_\mu$? The answer is affirmative for the first question and negative for the second. Therefore

$$\underset{\eta}{\mathcal{M}} : \overset{+}{\mathcal{S}} \oplus \overset{-}{\mathcal{S}} \to \overset{+}{\mathcal{S}} \oplus \overset{-}{\mathcal{S}}.$$

Now notice that $H_\eta = -\epsilon M_\eta$ and $\tilde{H}_\eta = \tilde{M}_\eta \epsilon$, which means that the following diagram commutes

$$\begin{tikzcd} \mathcal{S}^+ \arrow[r, "M\cdot"] \arrow[d, "Id\cdot"'] & \mathcal{S}^- \arrow[r, "\tilde{M}_\eta\cdot"] \arrow[d, "\cdot^T\epsilon"] & \mathcal{S}^+ \arrow[d, "Id\cdot"] \\ \mathcal{S}^+ \arrow[r, "(H_\eta\cdot)^T"] & \mathcal{S}^{-*} \arrow[r, "\tilde{H}_\eta\cdot^T"] & \mathcal{S}^+ \end{tikzcd} \tag{2}$$

Therefore we should regard $H_\eta : \mathcal{S}^+ \to \mathcal{S}^{-*}$, $\tilde{H}_\eta : \mathcal{S}^{-*} \to \mathcal{S}^+$, and so

$$\mathcal{H}_\eta : \mathcal{S}^+ \oplus \mathcal{S}^{-*} \to \mathcal{S}^+ \oplus \mathcal{S}^{-*}.$$

We can express the $4 \times 4$ matrix $\mathcal{H}_\eta$ as

$$\mathcal{H}_\eta \;\; = \;\; \begin{pmatrix} 0 & \tilde{H}_\eta \\ H_\eta & 0 \end{pmatrix} \; = \; t\gamma^0 + x\gamma^1 + y\gamma^2 + z\gamma^3$$

where the $\gamma^i$ are the *Dirac matrices*

$$\gamma^0 \; = \; \begin{pmatrix} 0 & \sigma^0 \\ \sigma^0 & 0 \end{pmatrix} \quad \text{and} \quad \gamma^i \; = \; \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} \quad \text{for } i = 1, 2, 3.$$

This is a Clifford representation of $Cl(1,3) \approx Cl(\mathbb{R}^{1,3})$ on $\mathcal{S}^+ \oplus \mathcal{S}^{-*}$. In this representation the element traditionally denoted $\gamma^5$ plays the role of the volume element:

$$\gamma^5 \; \triangleq \; -i\gamma^0\gamma^1\gamma^2\gamma^3 \; = \; \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}.$$

This gives projections

$$\begin{aligned}\pi^{+} &= \frac{1}{2}\left(1 + \gamma^{5}\right)\\ \pi^{-*} &= \frac{1}{2}\left(1 - \gamma^{5}\right)\end{aligned}$$

A *Dirac spinor* is a pair $\Phi = (\varphi', \varphi) \in \mathcal{S}^{+} \oplus \mathcal{S}^{-*}$ where

$$\begin{aligned}\varphi' &= \begin{pmatrix}\varphi^{1'}\\ \varphi^{2'}\end{pmatrix} \in \mathcal{S}^{+}\\ \varphi &= (\varphi_1, \varphi_2) \in \mathcal{S}^{-}.\end{aligned}$$

An element $\varphi' \in \mathcal{S}^{+}$ is called a *left-handed spinor*, and an element $\varphi \in \mathcal{S}^{-}$ is called a *right handed spinor*. Typically we use 4-vector notation, and write

$$\Phi = \begin{pmatrix}\varphi'\\ \varphi^{T}\end{pmatrix} = \begin{pmatrix}\varphi^{1'}\\ \varphi^{2'}\\ \varphi_1\\ \varphi_2\end{pmatrix}$$

so we have the Clifford action of an element $\eta \in \mathbb{R}^{1,3}$ on an element of $\mathcal{S}^{+} \oplus \mathcal{S}^{-*}$

$$\eta.\Phi = \begin{pmatrix}0 & \tilde{H}_\eta\\ H_\eta & 0\end{pmatrix}\begin{pmatrix}\varphi^{1'}\\ \varphi^{2'}\\ \varphi_1\\ \varphi_2\end{pmatrix}$$

Finally we discuss is the action of the symmetry group $SL(2,\mathbb{C})$ on $H_\eta$ and $\tilde{H}_\eta$. We have

$$\begin{aligned}U(H_\eta) &= -U(\epsilon M_\eta) = -U(\epsilon)U(M_\eta)\\ &= -\epsilon\overline{U}M_\eta U^{-1} = -\overline{U}^{T}\epsilon M_\eta U^{-1} = \overline{U}^{T}H_\eta U^{-1}.\end{aligned}$$

Similarly $U(\tilde{H}_\eta) = U\tilde{H}_\eta\overline{U}^{T^{-1}}$. Therefore

$$U(\mathcal{H}_\eta) = \begin{pmatrix}U & 0\\ 0 & \overline{U}^{T}\end{pmatrix}\begin{pmatrix}0 & \tilde{H}_\eta\\ H_\eta & 0\end{pmatrix}\begin{pmatrix}U^{-1} & 0\\ 0 & \overline{U}^{T^{-1}}\end{pmatrix}.$$

It is clear that $U(\Phi) = (U\varphi', \phi\overline{U}^{T})$, and so we have $U(\mathcal{H}_\eta)U(\Phi) = U(\mathcal{H}_\eta\Phi)$, as we must.

## 3.4 The symplectic form and Hermitian inner product on $\mathcal{S}^+ \oplus \mathcal{S}^{-*}$

The symplectic form on $\mathcal{S}^+$ carries over to Dirac spinors. Given $\Phi = (\varphi', \varphi)$, $\Psi = (\psi', \psi)$ we define

$$\langle \Phi,\, \Psi\rangle \;=\; \Phi^T \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} \Psi.$$

The Clifford action is symmetric with respect to this form:

$$\langle \eta.\Phi,\, \Psi\rangle \;=\; \Phi^T \mathcal{H}_\eta^T \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} \Psi \;=\; \Phi^T \begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon \end{pmatrix} \mathcal{H}_\eta \Psi \;=\; \langle \Phi,\, \eta.\Psi\rangle$$

The space $\mathcal{S}^{-*}$ acts on $\mathcal{S}^+$ after complex conjugation, and vice-versa. This gives a way for a Dirac spinor $\Phi = (\varphi', \varphi)$ to act on itself. Given $\Phi = (\varphi', \varphi)$ we define the Hermitian functional

$$|\Phi|^2 \;\;=\;\; \overline{\varphi}\varphi' \;+\; \varphi\overline{\varphi'}.$$

This is real, but not positive definite. Polarizing we get

$$(\Phi,\, \Psi) \;=\; \overline{\varphi}\psi' \;+\; \psi\overline{\varphi'}.$$

In the 4-vector notation

$$\Phi \;=\; \begin{pmatrix} \varphi^{1'} \\ \varphi^{2'} \\ \varphi_1 \\ \varphi_2 \end{pmatrix} \qquad \Psi \;=\; \begin{pmatrix} \psi^{1'} \\ \psi^{2'} \\ \psi_1 \\ \psi_2 \end{pmatrix}$$

$$\begin{aligned} \langle \Phi,\, \Psi\rangle &= \overline{\Phi}^T \gamma^0 \Psi \\ &= \overline{\varphi_1}\psi^{1'} \;+\; \overline{\varphi_2}\psi^{2'} \;+\; \overline{\varphi^{1'}}\psi_1 \;+\; \overline{\varphi^{2'}}\psi_2. \end{aligned}$$

We can prove that the action of $\eta \in \mathbb{R}^{1,3}$ is Hermitian with respect to this Hermitian inner product. It is easy to prove that $\overline{\mathcal{H}}_\eta^T \gamma^0 = \gamma^0\, \mathcal{H}_\eta$. Therefore

$$(\eta.\Phi,\, \Psi) \;=\; (\Phi,\, \eta.\Psi)\,.$$

## 3.5 Penrose notation

We can use the Penrose notation to help us wade through the spin spaces. An element $\varphi' \in \mathcal{S}^+$ is denoted with primed upper indices $\varphi^{A'}$, and an element $\varphi \in \mathcal{S}^-$ is denoted with unprimed upper indices: $\varphi^A$ (the "+" indicates *with prime*, the "−" indicates *without prime*). The isomorphism between them is

$$\begin{aligned} \mathcal{S}^- &\;\to\; \mathcal{S}^+ \\ \varphi &\;\mapsto\; \overline{\varphi}, \end{aligned}$$

or in components

$$\varphi^A \quad \mapsto \quad \overline{\varphi}^{A'} \;\triangleq\; \overline{\varphi^A}.$$

Similarly elements of dual spaces are indicated with primed and unprimed lower indices respectively: an element $\psi \in \mathcal{S}^{+*}$ is denoted $\psi_{A'}$ and an element $\psi \in \mathcal{S}^{-*}$ is denoted $\psi_A$. Now consider the isomorphism $\mathcal{S}^+ \to \mathcal{S}^{+*}$, given by transposition and matrix multiplication by $\epsilon$. This indicates that $\epsilon$ should be indexed by two primed lower indices: $\epsilon_{A'B'}$. The stated isomorphism is

$$\varphi^{A'} \mapsto \varphi_{A'} \;\triangleq\; \epsilon_{A'B'}\varphi^{B'}.$$

Similarly the isomorphism $\mathcal{S}^- \to \mathcal{S}^{-*}$ is

$$\varphi^A \mapsto \varphi_A \;\triangleq\; \epsilon_{AB}\varphi^B.$$

Covectors are denoted $\eta_a$, which are mapped to the matrices $\eta_{AA'}$, which are the maps $\mathcal{S}^+ \to \mathcal{S}^{-*}$. This mapping is achieved by means of the $\sigma$-symbol, also called the mixed spinor-tensor:

$$\eta_{AA'} \;=\; \eta_a\, \sigma^a{}_{AA'}.$$

Of course the matrix $\eta_{AA'}$ is just the matrix $H_\eta$ from above, so we have

$$\sigma^0{}_{AA'} \;=\; \sigma^0, \quad \sigma^1{}_{AA'} \;=\; \sigma^1, \quad \sigma^2{}_{AA'} \;=\; \sigma^2, \quad \sigma^3{}_{AA'} \;=\; \sigma^3,$$

where the sigmas on the right sides of the equations are the appropriate Pauli matrices. One can verify directly that the matrix $\eta^A{}_{A'} = \eta_{BA'}\epsilon^{AB}$ is the matrix $M_\eta$.

The four main properties of the $\sigma$-symbol are

$$\sigma^a{}_{AA'} \;=\; \overline{\sigma}^a{}_{AA'} \tag{3}$$

$$\sigma^a{}_{AA'}\sigma_b{}^{AA'} \;=\; \delta^a_b \tag{4}$$

$$\sigma_a{}^{AA'}\sigma^a{}_{BB'} \;=\; \epsilon_B{}^A\epsilon_{B'}{}^{A'} \tag{5}$$

$$\sigma^{[a}{}_{AA'}\sigma^{b]AB'} \;=\; -\frac{i}{2}\varepsilon^{abcd}\sigma_{cAA'}\sigma_d{}^{AB'} \tag{6}$$

where $\varepsilon$ is the antisymmetric symbol. Property (**??**) is called the *reality condition*; it restates the fact that the $\sigma$ matrices are Hermitian.

The Penrose notation lets us easily express the symplectic forms and inner products on the spin spaces. On $\mathcal{S}^+$, $\mathcal{S}^-$, etc the symplectic form is

$$\langle\psi, \varphi\rangle \;=\; \psi^{B'}\varphi^{A'}\epsilon_{A'B'} \;=\; \psi_{A'}\varphi^{A'} \;=\; -\psi^{A'}\varphi_{A'}$$

and so forth. A Dirac spinor $\Phi$ is a pair $\Phi = (\varphi^{A'}, \varphi_A)$. The Hermitian inner product can be expressed

$$\begin{aligned}
(\Phi,\, \Psi) \;&=\; \Big((\varphi^{A'}, \varphi_A),\, (\psi^{A'}, \psi_A)\Big) \\
&=\; \overline{\varphi^{A'}}\psi_A \;+\; \overline{\varphi_A}\psi^{A'} \\
&=\; \overline{\varphi}^A\psi_A \;+\; \overline{\varphi}_{A'}\psi^{A'}
\end{aligned}$$

# 4 Dirac's Equation

Einstein's Equation $E = mc^2$ in properly relativistic form is $g^{ij}p_ip_j = m^2c^4$, where $p = -mc^2 v_\flat = (-E, p_1, p_2, p_3) = (-E, \vec{p})$ is the energy-momentum 4-vector and the norm is taken using the Lorentz metric. Using the standard quantum-mechanical identifications $p_i \mapsto -i\hbar\partial_i$ and $E \mapsto i\hbar\partial_t$, we get the Klein-Gordon equation

$$\left(\Box + \frac{m^2c^4}{\hbar^2}\right)\phi \;=\; 0$$

where

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

for scalar fields $\phi$. This is inadequate as a physical evolution equation for three reasons. First, since it is of second order, the state of a quantum system is not given solely by the wavefunction $\phi$. Second, the probability 4-current density is given by

$$J_i \;=\; \frac{i\hbar}{2m}\left(\phi\frac{\partial\bar{\phi}}{\partial x^i} - \bar{\phi}\frac{\partial\phi}{\partial x^i}\right) \;=\; \frac{\hbar}{m}Im(\bar{\phi}\frac{\partial\phi}{\partial x^i}).$$

It is easy to check that

$$\frac{\partial}{\partial x^i}J^i \;=\; 0$$

so that the conserved quantity is $J_0 = \frac{\hbar}{m}Im(\bar{\phi}\frac{\partial}{\partial x^0}\phi)$. But since $\phi$ and $\partial\phi/\partial x^0$ can be chosen simultaneously, the probability density need not be positive everywhere. This presents a serious problem of interpretation. Third, there is the problem of negative energy states.

We would like to take the 'square root' of the Einstein equation and write $p = \pm mc$, interpreted as an operator equation. But on what space would $p$ and $mc$ operate? Of course they must operate on the space of Dirac spinors $\mathcal{S}\otimes\overline{\mathcal{S}}^*$ via the Clifford representation. Given a Dirac spinor $\Phi$ we have $p\cdot p\cdot\Phi = -|p|^2\Phi$, so that

$$(p - mc)\cdot(p - mc)\cdot\Phi \;=\; (p\cdot p - 2mcp + m^2c^2)\cdot\Phi \;=\; (-|\mathbf{p}|^2 - m^2)\Phi.$$

Replacing $\mathbf{p}$ by $-i\hbar\nabla$ as usual, we see that $p - m$ (as well as the antiparticle operator $p + m$) squares to the opposite of the Klein-Gordon operator $\Box + m^2$. Letting $D$ denote differentiation acting via the Clifford representation, we have the Dirac equation for the electron in free space

$$(\hbar D - im)\Phi \;=\; 0.$$

# 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 gravitiational 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.