MAT 561: Mathematical Physics, Spring 2007

MAT 560
References
Homeworks
Contact William
Contact Jerry

Lectures: Mondays 9:45 -- 11:30 AM and 1 -- 2 PM, in the Math Tower, P-131.

Instructors: Jerry Jenquin and William D. Linch III.

Office Hours: Jerry's are TBA
William's are TBA.

Suggested Prerequisites We'll be assuming familiarity with the material covered in MAT 530, 531 and 560. Some previous exposure to physics, while helpful, is by no means necessary.

Course Content: In this second semester we will cover Thermodynamics, Statistical Mechanics, and Quantum Mechanics. The latter will be divided into two ``sections''. One more mathematical and one more physical.

Texts and Online Notes: Although there are no official texts for this course here's a list of references for some of the topics we'll be covering and some of the prerequisite topics. In particular we recommend the following texts to complement the lectures.

1. *Mathematical Methods of Classical Mechanics* by V.I. Arnold.
2. *A Course in Mathematics for Students of Physics* by Paul Bamberg and Shlomo Sternberg.

Homework: We will provide four problem sets throughout the semester: one for each ``section'' of this semester's material. There will be four corresponding homework sessions and, at every session, each student must present a solution to one of the problems. Before each presentation the student is expected to give a ``practice run'' in front of either William or Jerry for the sake of quality control. Also a TeX'd solution must be turned in sometime before the presentation. The students are not required to arrive at the solutions independently but must present them independently.

Solutions: Below are solutions to the Thermodynamics problem set as done by the students.

1. Nate Round
2. Josh Rembaum
3. Gabe Drummond-Cole
4. Joe Walsh
5. Andrew Stimpson
6. Ki Song
7. Chris Bay

Grades: The course grade will be determined solely by the homework.

DSS advisory: If you have a physical, psychiatric, medical, or learning disability that could adversely affect your ability to carry out assigned course work, we urge you to contact the Disabled Student Services office (DSS), Educational Communications Center (ECC) Building, room 128, (631) 632-6748..

MAT 560: Mathematical Physics, Fall 2006

Lectures: Tuesdays and Thursdays 11:20 AM -- 12:40 PM in the Math Tower, P-131.

Instructors: Jerry Jenquin and William D. Linch III.

Office Hours: Jerry's are 10 -- 11 AM on Tuesdays and Thursdays. William's are 1:30 -- 2:30 PM on Tuesdays and Thursdays.

Suggested Prerequisites We'll be assuming familiarity with the material covered in MAT 530 and 531. Some previous exposure to physics, while helpful, is by no means necessary.

Course Content: In this first semester we will cover Classical Newtonian Mechanics, Classical Relativistic Mechanics, and Electromagnetism. Specifically we hope to cover the following:

Classical Newtonian Mechanics

- Paths in Euclidean space and Newton's 2nd Law.
- Phase space and symplectic geometry.
- Hamiltonian mechanics in the Newtonian setting.
- Variational principles, Lagrangian mechanics.
- Symmetry and Noether's theorem.
- The Euclidean group, symmetry, and conserved charges (a.k.a. Newtonian kinematics).
- Time translation, energy, and dynamics.
- Hamiltonian mechanics from Lagrangian mechanics.
- Gravitational potentials and solvable systems.

Classical Relativistic Mechanics

- Geometry on Minkowski space.
- Lagrangian for paths on Minkowski space.
- The Poincare group, symmetry, and relativistic kinematics.
- Reparameterization invariance.

Electromagnetism

- Differential forms, Stoke's theorem, currents, flows.
- Hodge star for Euclidean, Lorentzian signatures and duality.
- Electromagnetic fields and Maxwell's equations
- PDE's on Minkowski space and Poincare symmetry
- Laplace and wave equations, Green's operators, boundary conditions
- Exact solutions: propagating waves, monopoles, instantons, ...
- Lagrangian formulation of electromagnetism
- Hamiltonian theory of electromagnetism
- Gauge symmetry and connections on principal R-bundles
- Magnetic sources, Dirac charge quantization, and principal U(1)-bundles
- Generalizations in various directions.

Texts and Online Notes: Although there are no official texts for this course here's a list of references for some of the topics we'll be covering and some of the prerequisite topics. In particular we recommend the following texts to complement the lectures.

1. *Mathematical Methods of Classical Mechanics* by V.I. Arnold.
2. *A Course in Mathematics for Students of Physics* by Paul Bamberg and Shlomo Sternberg.
3. *Overview of Selected Topics in Physics* by William D. Linch III. These notes offer a treatment closer to what one would find in a physics text. It's a work in progress.

We are also fortunate to have Gabriel Drummond-Cole's TeX'd course notes, annotated with physics commentary by William.

- Lecture 1
- Lecture 3
- Lecture 4

Homework: We will provide several problem sets throughout the semester. The best way to learn the material is to attempt these problems and even come up with and solve some problems on your own.

Grades: Throughout the semester students will be expected to present homework solutions in class. The course grade will be determined solely by these presentations.

DSS advisory: If you have a physical, psychiatric, medical, or learning disability that could adversely affect your ability to carry out assigned course work, we urge you to contact the Disabled Student Services office (DSS), Educational Communications Center (ECC) Building, room 128, (631) 632-6748..

MAT 561: Mathematical Physics, Spring 2007

Here are the problem sets. They will be assigned on a first-come-first-serve basis. If you have any questions or want some help feel free to stop by anytime or make an appointment. We also encourage you to seek out each other for help, especially if your respective assigned problems are related.

- Thermodynamics Homework
- Statistical Mechanics Homework
- Jerry's Quantum Homework
- William's Quantum Homework

Solutions: Below are solutions to the Thermodynamics problem set as done by the students.

1. Nate Round
2. Josh Rembaum
3. Gabe Drummond-Cole
4. Joe Walsh
5. Andrew Stimpson
6. Ki Song
7. Chris Bay

Below are solutions to Jerry's Quantum Mechanics problems as done by the students.

1. Gabe Drummond-Cole
2. Nate Round
3. Joe Walsh
4. Josh Rembaum
5. Andrew Stimpson
6. Ki Song
7. Chris Bay

Jerry Jenquin

I can be contacted in the following ways:

Email:	jjenquin@math.sunysb.edu
Phone:	(631) 632 - 8262
Office:	Math Tower Room 3-115
Smail:	Mathematics Department Stony Brook University Stony Brook, NY 11794-3651

Problem 1. *Let V be a finite-dimensional vector space and let $(\,,\,) : V \times V \to V$ be a skew-symmetric bilinear pairing on V. Then V has a basis of the form* $\{x_i, y_i\}_{i=1}^m, \{z_i\}_{i=1}^n$ *such that* $\begin{cases} (x_i, y_i) &= 1 \\ (x_i, y_j) &= 0 \\ (z_i, \alpha) &= 0 \quad \forall \alpha \in V \end{cases}$

If $(\,,\,)$ is identically zero, then any basis for V is of the above form, with $n = \dim(V)$ and $m = 0$. If not, then $(x, y) \neq 0$ for some $x, y \in V$. Furthermore $x \neq y$ by skew-symmetry, so after rescaling if necessary we have $(x, y) = 1$.

Now we can write $V = \operatorname{span}\{x, y\} \oplus V'$, where $V' = \{\alpha \in V | (\alpha, x) = (\alpha, y) = 0\}$, and apply the same arguement recursively to V'. Since V is finite dimensional, the recursion terminates and we arrive at a basis of the desired form.

Corollary 1. *If $(\,,\,)$ is nondegenerate, then the dimension of V is even.*

Problem 2. *If ω_1 and ω_0 are symplectic forms on $\mathbb{R}^{2n}$ and $\omega_1 = \omega_0$ at 0, then there exits a local diffeomorphism f defined in some neighborhood of 0 so that $f^*\omega_1 = \omega_0$.*

We will solve this problem using time-dependent vector-fields. Given a time dependent vector field $\xi : \mathbb{R} \times M \to TM$, let $\phi : \mathbb{R} \times M \to M$ be the solution to the initial value problem

$$\begin{cases} \xi_s(\phi_s(x)) = d\phi_{(s,x)}(\frac{\partial}{\partial t}) \\ \phi_0(x) = x \end{cases}$$

Lemma 1. *For any time-dependent form α_t, time-dependent vector field ξ_t and one parameter family of diffeomorphisms ϕ_t generated by ξ_t, we have*

$$\frac{\partial}{\partial t}(\phi_t^* \alpha_t) = \frac{\partial}{\partial t}\alpha_t + \mathrm{Lie}_{\xi_t}\alpha_t$$

Proof: First define $\bar{\phi} : \mathbb{R} \times M \to \mathbb{R} \times M$ by extending ϕ by the identity, so that $\bar{\phi}(s, x) = (s, \phi_s(x))$. Then $\phi_t^*\alpha_t = \bar{\phi}^*\alpha_t$.

First consider the case where α_t is a time-dependent function. At a point $(s, x) \in \mathbb{R} \times M$, we have

$$\begin{aligned}
\frac{\partial}{\partial t}(\phi_t^*\alpha_t)_{(s,x)} &= \frac{\partial}{\partial t}(\alpha \circ \bar{\phi}_t)_{(s,x)} \\
&= d\alpha_{(s,\phi_s(x))} \circ d\bar{\phi}_{(s,x)}(\frac{\partial}{\partial t}) \\
&= d\alpha_{(s,\phi_s(x))} \circ (d(\mathrm{ID}_{\mathbb{R}}) + d\phi)_{(s,x)}(\frac{\partial}{\partial t}) \\
&= d\alpha_{(s,\phi_s(x))}(\frac{\partial}{\partial t} + \xi_s(\phi_s(x))) \\
&= \frac{\partial}{\partial t}\alpha_{(s,\phi_s(x))} + \xi_s\alpha(\phi_s(x)) \\
&= \phi_t^*(\frac{\partial}{\partial t}\alpha_t + \mathrm{Lie}_{\xi_t}\alpha_t)(s, x)
\end{aligned}$$

Next consider the case where $\alpha_t = df_t$ is an exact 1-form. Then

$$\begin{aligned}\frac{\partial}{\partial t}(\phi_t^* df_t)_{(s,x)} &= d(\frac{\partial}{\partial t}\phi_t^* f_t) \\ &= d(\phi_t^*(\frac{\partial}{\partial t} f_t + \mathrm{Lie}_{\xi_t} f_t)) \\ &= (\phi_t^*(\frac{\partial}{\partial t} df_t + \mathrm{Lie}_{\xi_t} df_t))\end{aligned}$$

because d commutes with $\frac{\partial}{\partial t}$ and Lie_{ξ_t}.

Now, every form α_t can be written locally as a wedge product of fuctions and exact 1-forms. Since pullback commutes with wedge and $\frac{\partial}{\partial t}$ is a derivation of the wedge product, the lemma follows. ■

We proceed with the proof of Problem 2. Define:

$$\omega_t = t\omega_1 + (1-t)\omega_0$$

Observe that ω_t is a linear combination of closed forms, and thus is closed and locally exact. Then there exists a 1-form β such that:

$$d\beta = \omega_1 - \omega_0$$

At the point 0, $\omega_1 = \omega_0$ and thus $\omega_t = t\omega_1 + (1-t)\omega_0 = \omega_0$. Therefore ω_t is nondegenerate at 0. Nondegeneracy is an open condition, so ω_t is nondegenerate in a neighborhood of 0 as well. Thus it is possible to define a time-dependent vector field ξ_t by the relation:

$$\iota_{\xi_t}\omega_t = -\beta$$

Let ϕ_t be the one parameter family generated by ξ_t. Then by our Lemma,

$$\begin{aligned}\frac{\partial}{\partial t}(\phi_t^*\omega_t) &= \phi_t^*(\frac{\partial}{\partial t}\omega_t + \mathrm{Lie}_{\xi_t}\omega_t) \\ &= \phi_t^*(\omega_1 - \omega_0 + d(\iota_{\xi_t}\omega_t) + \iota_{\xi_t} d\omega_t) \\ &= \phi_t^*(d\beta - d\beta + 0) \\ &= 0\end{aligned}$$

Thus $\phi_1^*\omega_1 = \phi_0^*\omega_0$. But ϕ_0^* is the identity, so $\phi_1^*\omega_1 = \omega_0$. ϕ_1 is the desired diffeomorphism.

MAT 561 - HW 1 Problem 2

Josh Rembaum

March 6, 2007

2 Assuming question 1.

(a) Let $\omega \in \Omega^2(M)$ be closed, non-degenerate. Claim: there exist local coordinates in a neighborhood of $m \in M$, (U, ϕ), $\phi = (x^1, \ldots, x^{2n})$ such that

$$\omega = dx^1 \wedge dx^2 + \ldots dx^{2n-1} \wedge dx^{2n}$$

Proof: For $m \in M$, pick a chart (U, ϕ), $\phi : U \to \mathbb{R}^{2n}$, and $\phi(m) = 0$. Consider $(\phi^{-1})^*\omega$.

It is a non-degenerate 2-form,[1] $\omega \in \Lambda^2(\mathbb{R}^{2n})^*$, so by problem 1a),
$\exists\,(V, y),\ 0 \in V$ such that

$$(\phi^{-1})^*\omega\Big|_0 = dy^1 \wedge dy^2 + \ldots + dy^{2n-1} \wedge dy^{2n}\Big|_0$$

Further, $(\phi^{-1})^*\omega$ is also closed,[2] so, as a 2-form on $\mathbb{R}^{2n}$, by problem 1b),
$\exists\, f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, a diffeomorphism, defined on a neighborhood of 0, such that:

$$(f^* \circ (\phi^{-1})^*)\omega = dy^1 \wedge dy^2 + \ldots + dy^{2n-1} \wedge dy^{2n}$$

on the neighborhood. Then,

$$\begin{aligned} \omega &= (f^{-1} \circ \phi)^* dy^1 \wedge dy^2 + \ldots + dy^{2n-1} \wedge dy^{2n} \\ &= d(y^1 \circ f^{-1} \circ \phi) \wedge d(y^2 \circ f^{-1} \circ \phi) + \ldots + \\ &= \ + d(y^{2n-1} \circ f^{-1} \circ \phi) \wedge d(y^{2n} \circ f^{-1} \circ \phi) \end{aligned}$$

Let $x^i = y^i \circ f^{-1} \circ \phi$ and we have the coordinates desired:

$$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2n-1} \wedge dx^{2n}. \quad \square$$

[1] This is because ω is non-degenerate and ϕ is a bijection, so their composition remains injective.

[2] Since pull-back commutes with external derivative.

(b) Suppose $\omega \in \Omega^2(M)$ is a *closed* form with constant rank $k \leq n$, i.e.:

$$\forall m \in M,\ \omega^k\Big|_m \neq 0,\ \text{and}\ \omega^{k+1}\Big|_m = 0$$

Claim: $\exists$ local coordinates such that

$$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2k-1} \wedge dx^{2k}$$

Proof: by inducting on the dimension of M.
Base case: $\dim M = 2k$.
Then $\omega^k \neq 0 \Rightarrow \omega$ is a non-vanishing top form, so it is non-degenerate (and closed by assumption).
Thus, we can apply the results of 2a) to get local coordinates such that:

$$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2k-1} \wedge dx^{2k}$$

Suppose that for $\dim M = 2k + i$ there exist local coordinates such that:

$$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2k-1} \wedge dx^{2k}$$

Consider the case for $\dim M = 2k + i + 1$.
$\omega^{k+1} = 0 \quad \Rightarrow \quad \omega$ is degenerate,[3] so:

$$\exists X \in \mathfrak{X}(M) \quad \text{s.t.} \quad \iota_X \omega = 0$$

Further, we can choose local coordinates $\{z^1, \ldots, z^{2n-1}, y\}$ such that $X = \frac{\partial}{\partial y}$.
Then

$$\iota_X \omega = 0 \quad \Rightarrow \quad \omega = \sum_{i<j} a_{ij} dz^i \wedge dz^j$$

Want to show that a_{ij} is independent of y, so that ω will be defined on a dimension $2k + i$ submanifold and we can apply the inductive hypothesis.
ω closed $\quad \Rightarrow \quad d\omega = 0$, i.e.:

$$0 = \sum_{i<j} \left(\frac{\partial a_{ij}}{\partial y} dy \wedge dz^i \wedge dz^j + \text{stuff not involving } y \right)$$

The linear combination of basis vectors is 0 if and only if the coefficients are all 0, so we have that
$\forall i \forall j\ \frac{\partial a_{ij}}{\partial y} = 0$.
Thus the a_{ij}'s are independent of y, so that ω is a closed 2-form on a dimension $2n + i$ submanifold of M and we can apply the induction hypothesis to get that there are local coordinates such that:

$$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2k-1} \wedge dx^{2k} \quad \square$$

[3] In the case $i = 0$, ω is degenerate, since it's a skew-symmetric form on an odd-dimensional manifold, so in local coordinates its matrix, ω_{ij} is skew-symmetric and the odd dimension gives $\omega_{ij} = -\omega_{ji}$, but $det(\omega_{ij}) = det(\omega_{ji})$, thus $det(\omega_{ij}) = 0$, hence it's a degenerate matrix. This holds for every point, so ω is degenerate as a form.

[3.] Assume that α is a constant rank $2k+1$ one-form on M^N. Prove that there exists a basis $\{x_1, y_1, \ldots, x_k, y_k, z, \ldots\}$ with respect to which

$$\alpha = dz + \sum_{i=1}^{k} x^i dy^i.$$

Proof. Since α is rank $2k+1$, $\alpha \wedge (d\alpha)^k$ is nowhere zero, so $(d\alpha)^k$ is nowhere zero, but $(d\alpha)^{k+1}$ is identically zero, so $(d\alpha)$ is an exact (hence closed) two-form of constant rank k.

Then by Darboux's theorem, there exists a set of coordinates $\{x^1, y^1, \ldots, x^k, y^k, z^1, \ldots, z^{N-2k}\}$ with respect to which

$$d\alpha = \sum_{i=1}^{k} dx^i \wedge dy^i.$$

Then

$$d(\alpha - \sum_{i=1}^{k} x^i dy^i) = d\alpha - \sum_{i=1}^{k} dx^i \wedge dy^i = 0$$

So

$$\alpha - \sum_{i=1}^{k} x^i dy^i$$

is closed, hence locally exact, so that locally there exists a function z such that

$$\alpha - \sum_{i=1}^{k} x^i dy^i = dz.$$

Now

$$\alpha \wedge (d\alpha)^k = dz \wedge \bigwedge_{i=1}^{k} dx^i \wedge dy^i$$

because all the other terms have multiples of some dx^i or dy^i. Writing dz locally and omitting those terms which are functional multiples of dx^i or dy^i, this is

$$\left(\sum_{j=1}^{N-2k} \frac{\partial z}{\partial z^j} dz^j \right) \wedge \bigwedge_{i=1}^{k} dx^i \wedge dy^i.$$

Because α is of constant rank $2k+1$, this form is nowhere zero, so in particular,

$$\sum_{j=1}^{N-2k} \frac{\partial z}{\partial z^j} dz^j$$

is nowhere zero. This means that if the range of z (as a function $M \to \mathbb{R}^N\}$ is restricted to $\langle z^1, \ldots, z^{N-2k} \rangle$, its derivative is rank one, so it is a coordinate.

Then on $\langle x^1, y^1, \ldots, x^k, y^k, z^1, \ldots, z^{N-2k} \rangle$ it is still a coordinate, and in particular a coordinate independent of the set $\{x^1, y^1, \ldots, x^k, y^k\}$, as desired.

(a) Start with a one-form $\alpha \in \Omega^1(\mathbb{R}^{2k})$ of constant rank $2k$. Show there exist local coordinates $x^1, \ldots, x^{2k}$ such that

$$\alpha \wedge (d\alpha)^{k-1} = g dx^2 \wedge dx^3 \wedge \cdots \wedge dx^{2k}$$

for some positive g. Define a function f so that $f^k = g$ and define a one-form σ so that $\alpha = f\sigma$. Show that σ has rank $2k-1$.

Proof. By definition of a rank $2k$ one-form, $(d\alpha)^k$ is nonvanishing, and $\alpha \wedge (d\alpha)^k \equiv 0$. (In this case, this second property is trivial since $(d\alpha)^k$ is a top form.) We can assume $\alpha \wedge (d\alpha)^{k-1}$ does not vanish locally, since its derivative is $(d\alpha)^k$. Since $\alpha \wedge (d\alpha)^{k-1}$ is a nonvanishing $2k-1$-form, we can find a locally-defined nonvanishing vector field ξ such that $i(\xi)[\alpha \wedge (d\alpha)^{k-1}] \equiv 0$. Therefore, around any point, we can find local coordinates $t, x^2, x^3, \ldots, x^{2k}$ such that $\xi = \frac{\partial}{\partial t}$ in this coordinate system.

Hence, $i(\xi)[\alpha \wedge (d\alpha)^{k-1}] = i(\xi)[dx^2 \wedge \cdots \wedge dx^{2k}] = 0$. Since $\alpha \wedge (d\alpha)^{k-1}$ is nonvanishing, at each point p there is a unique one dimensional subspace of the tangent space at p where $\xi(p)$ must reside. Ergo, $\alpha \wedge (d\alpha)^{k-1}$ and $dx^2 \wedge \cdots \wedge dx^{2k}$ are linearly dependent. So $\alpha \wedge (d\alpha)^{k-1} = g dx^2 \wedge \cdots \wedge dx^{2k}$ for some continuous, nonvanishing g. If g is negative, we can switch x^2 to $-x^2$ to make g positive.

Let f be such that $f^k = g$ and let $\sigma = \frac{1}{f}\alpha$.

$$d\sigma = -\frac{1}{f^2} df \wedge \alpha + \frac{1}{f} d\alpha$$

Since $\alpha \wedge \alpha = 0$, we can quickly calculate:

$$\begin{aligned}
(d\sigma)^{k-1} &= -(k-1)\frac{1}{f^k} df \wedge \alpha \wedge (d\alpha)^{k-2} + \frac{1}{f^{k-1}}(d\alpha)^{k-1} \\
\sigma \wedge (d\sigma)^{k-1} &= \frac{1}{f^k}\alpha \wedge (d\alpha)^{k-1} = \frac{1}{g}(g dx^2 \wedge \cdots \wedge dx^{2k}) = dx^2 \wedge \cdots \wedge dx^{2k} \\
(d\sigma)^k &= d(\sigma \wedge (d\sigma)^{k-1}) = d(dx^2 \wedge \cdots \wedge dx^{2k}) = 0
\end{aligned}$$

Hence, $\sigma \wedge (d\sigma)^{k-1}$ vanishes nowhere and $(d\sigma)^k$ is identically 0; i.e. σ has constant rank $2k-1$. □

(b) Prove that the above still holds when $n \geq k$ and $\alpha \in \Omega^1(\mathbb{R}^{2n})$ still has rank $2k$.

Proof. We know that $(d\alpha)^k$ is nowhere vanishing, and $\alpha \wedge (d\alpha)^k \equiv 0$. Hence, $(d\alpha)^{k+1} = d(\alpha \wedge (d\alpha)^k) \equiv 0$. Thus, $d\alpha$ is a 2-form of constant rank k. Therefore, by Darboux's Theorem (Problem 2), there exist local coordinates $x^1, x^2, \ldots, x^{2n}$ such that $d\alpha = dx^1 \wedge dx^2 + dx^3 \wedge dx^4 + \cdots + dx^{2k-1} \wedge dx^{2k}$.

Then $(d\alpha)^k = k!(dx^1 \wedge dx^2 \wedge \cdots \wedge dx^{2k-1} \wedge dx^{2k})$, and since α has rank $2k$, $\alpha \wedge (d\alpha)^k = 0$. Expressed in this coordinate system, $\alpha = \sum_{i=1}^{2n} a^i dx^i$ for some functions a^i. $\alpha \wedge (d\alpha)^k = 0$ implies that $a^i \equiv 0$ for $i > 2k$. So

$$\alpha = \sum_{i=1}^{2k} a^i(x^1, \ldots, x^{2n}) dx^i.$$

$d\alpha = \sum_{i=1}^{2k} \sum_{j=1}^{2j} \frac{\partial a^i}{\partial x^j} dx^j \wedge dx^i$. Therefore, for any $1 \leq i \leq 2k$ and for any $j > 2k$, $\frac{\partial a^i}{\partial x^j} = 0$, since the only term of the sum that carries a factor of $dx^j \wedge dx^i$ is $\frac{\partial a^i}{\partial x^j} dx^j \wedge dx^i$, and this term does not appear at all in $d\alpha$ by choice of coordinates. Hence, each a^i is independent of x^j for $j > 2k$. Thus,

$$\alpha = \sum_{i=1}^{2k} a^i(x^1, \ldots, x^{2k}) dx^i.$$

is completely independent of $x^{2k+1}, \ldots, x^{2n}$. Therefore, we can restrict to the $2k$-dimensional submanifold generated by $\{x^1, \ldots, x^{2k}\}$, and apply part (a).

$\square$

(c) Use part (b) to show that if $\alpha \in \Omega^1(M)$ has constant rank $2k$ then we can find a positive function f and a one-form σ of constant rank $2k-1$ such that

$$\alpha = f\sigma.$$

Use this fact and question (3) to find coordinates that verify the claim.

Proof. Part (b) allows us to define σ and f locally as follows: Given a coordinate chart $x : U \to \mathbb{R}^{2n}$, $x_*\alpha$ is a one form of constant rank $2k$ on $\mathbb{R}^{2n}$. Thus, there exist a positive function f' and a rank $2k-1$ 1-form σ' on $\mathbb{R}^{2n}$ such that $f'\sigma' = x_*\alpha$. Let $f = x^*f'$ and $\sigma = x^*\sigma'$.

By question 3, there are local coordinates $z^1, z^2, \ldots, z^n, y^1, y^2, \ldots, y^n$ such that $\sigma = dy^1 + \sum_{i=2}^k z^i dy^i$. Then $\alpha = f dy^1 + \sum_{i=2}^k f z^i dy^i$. Define $x^1 = f$ and $x^i = f z^i$ for $2 \leq i \leq k$. $(d\alpha)^k = dx^1 \wedge dy^1 \wedge \cdots \wedge dx^k \wedge dy^k \neq 0$, so $x^1, \ldots, x^k, y^1, \ldots y^k$ are linearly independent. Hence, the x^i, y^i are part of a coordinate system such that $\alpha = \sum_{i=1}^k x^i dy^i$.

$\square$

Consider a thermodynamic system whose only configurational variable is volume V so that the equilibrium submanifold M is 2-dimensional. **NOTE:** For any of these results to make sense, we must assume that $Q \neq 0$.

(a) Let Q denote the heat 1-form and let p denote pressure. Explain how we can interpret the ratio of two forms as a function

$$f = \frac{Q \wedge dp}{Q \wedge dV}.$$

Proof: Because 2-forms are top forms on a 2-manifold, in local coordinates (y^1, y^2) we can write

$$Q \wedge dp = f_p(y)\, dy^1 \wedge dy^2 \quad \text{and} \quad Q \wedge dV = f_V(y)\, dy^1 \wedge dy^2.$$

Because V is a configurational variuable, $Q \wedge dV$ is nowhere-vanishing, so we can provisionally define $f \equiv f_p/f_V$. This definition is coordinate-free because in any other coordinates, the top form corresponding to $dy^1 \wedge dy^2$ will be related to $dy^1 \wedge dy^2$ by a multiplication by a non-vanishing function[1], which will cancel out in the quotient. ■

(b) Fix a point $x \in M$ and any adiabatic vector $\xi \in T_xM$. Show that

$$f(x) = \frac{dp(\xi)}{dV(\xi)} \quad \text{which justifies the expression} \quad f = \left(\frac{dp}{dV}\right)_{\text{adiabatic}}$$

where the right hand side is just the ratio of 1-forms evaluated on some adiabatic vector.

Proof: Pick $\eta \in T_xM$ such that $Q(\eta) \neq 0$. Part (a) showed that

$$Q \wedge dp = f\; Q \wedge dV.$$

So at the point x,

$$Q(\eta)\, dp(\xi) = (Q \wedge dp)(\eta, \xi) = f(x)(Q \wedge dV)(\eta, \xi) = f(x)\, Q(\eta)\, dV(\xi)$$

because $Q(\xi) = 0$. So because $Q(\eta) \neq 0$, $dp(\xi) = f(x)\, dV(\xi)$. Note that since f is independent of the choice of ξ, the ratio $\frac{dp(\xi)}{dV(\xi)}$ is also independent of ξ. ■

(c) Let T denote temperature. Based on the discussion above, state what is meant by

$$\frac{dT \wedge dp}{dT \wedge dV} = \left(\frac{dp}{dV}\right)_{\text{isothermal}}.$$

Statement: To calculate, $\frac{dp}{dV}$ along infinitesimal isothermal directions, we can just take the ratio of dp and dV evaluated on isothermal tangent vectors. But just as in part (b), this can be calculated by taking the ratio of the two 2-forms $dT \wedge dp$ and $dT \wedge dV$ because an isothermal tangent vector is in the kernel of dT by definition.

[1] the determinant of the Jacobian of the transition function

(d) Since the differentials of V, T and p, T are linearly independent, we may write

$$Q = \Lambda_V dV + C_V dT \quad \text{or} \quad Q = \Lambda_p dp + C_p dT$$

where the Λs and Cs are functions on M. Show that

$$\left(\frac{dp}{dV}\right)_{\text{adiabatic}} = \gamma \left(\frac{dp}{dV}\right)_{\text{isothermal}}$$

where $\gamma = C_p/C_V$.

Proof:

$$\left(\frac{dp}{dV}\right)_{\text{adiabatic}} = \frac{Q \wedge dp}{Q \wedge dV} = \frac{(\Lambda_p dp + C_p dT) \wedge dp}{(\Lambda_V dV + C_V dT) \wedge dV} = \frac{C_p \; dT \wedge dp}{C_V \; dT \wedge dV} = \frac{C_p}{C_V}\left(\frac{dp}{dV}\right)_{\text{isothermal}}$$

■

(e) An *ideal gas* is one that, in equillibrium, obeys the constraints

$$pV = nT \quad \text{and} \quad \gamma = \text{constant}$$

where n are the moles of gas. Use part (d) to show that the adiabatic curves for an ideal gas are given by

$$pV^\gamma = \text{constant.}$$

Proof: Using the coordinates (V, T) for M, for an isothermal vector τ,

$$dp(\tau) = \frac{\partial p}{\partial V} dV(\tau) + \frac{\partial p}{\partial T} dT(\tau) = \frac{\partial p}{\partial V} dV(\tau).$$

Because pressure has the formula $p(V, T) = \frac{nT}{V}$,

$$\left(\frac{dp}{dV}\right)_{\text{isothermal}} = \frac{\partial p}{\partial V} = -\frac{nT}{V^2} \quad \Longrightarrow \quad \left(\frac{dp}{dV}\right)_{\text{adiabatic}} = -\gamma \frac{nT}{V^2}$$

Let ξ be an adiabatic vector in $T_x M$.

$$\begin{aligned} d(pV^\gamma)(\xi) &= V^\gamma dp(\xi) + p\gamma V^{\gamma - 1} \, dV(\xi) = \left(-V^\gamma \gamma \frac{nT}{V^2} + p\gamma V^{\gamma-1}\right) dV(\xi) \\ &= \left(-V^\gamma \gamma \frac{nT}{V^2} + \frac{nT}{V}\gamma V^{\gamma - 1}\right) dV(\xi) = 0. \end{aligned}$$

So infinitesimal adiabatic vectors are in the kernel of $d(pV^\gamma)$. Thus as long as $Q \neq 0$, Q must be proportional to $d(pV^\gamma)$, which means that adiabatic curves must be level sets of pV^γ. ■

MAT561 Homework 1

Problem 6

Part a

Let y^i be another basis of linear coordinates on V. Then, there exists $A = (A^i{}_j) \in GL(\mathbb{R}^n)$ such that $y^i = A^i{}_j x^j$. We have

$$\begin{aligned} dy^i &= A^i{}_j dx^j \\ \frac{\partial}{\partial y^i} &= A_i{}^j \frac{\partial}{\partial x^j} \end{aligned}$$

where $(A_i{}^j) = A^{-1}$.
Using above, we get:

$$\begin{aligned} \mathrm{Hess} f &= \frac{\partial^2 f}{\partial y^i \partial y^j} dy^i \otimes dy^j \\ &= A_i{}^k \frac{\partial}{\partial x^k}\left(A_j{}^l \frac{\partial f}{\partial x^l}\right) A^i{}_m A^j{}_n dx^m \otimes dx^n \\ &= A_i{}^k A^i{}_m A_j{}^l A^j{}_n \frac{\partial}{\partial x^k}\left(\frac{\partial f}{\partial x^l}\right) dx^m \otimes dx^n \\ &= \delta^k{}_m \delta^l{}_n \frac{\partial}{\partial x^k}\left(\frac{\partial f}{\partial x^l}\right) dx^m \otimes dx^n \\ &= \frac{\partial^2 f}{\partial x^m \partial x^n} dx^m \otimes dx^n \end{aligned}$$

The fact that Hessian is symmetric follows from the commutativity of partial differentiations.

Part b

Let $f \colon V \to \mathbb{R}$. Then, we have

$$\begin{aligned} df \colon TV &\to T\mathbb{R} \\ (p, \dot{p}) &\mapsto (f(p), (df_p)(\dot{p})) \end{aligned}$$

So $\phi_f(p) = df_p = \frac{\partial f}{\partial x^i}(p) dx^i$. Now to find $d\phi_f$:

$$\begin{aligned} d\phi_f \colon TV &\to TV^* \simeq V^* \times V^* \\ (p, \dot{p}) &\mapsto (df_p, \frac{\partial^2 f}{\partial x^j \partial x^i}(p) \dot{p}^i dx^j) \end{aligned}$$

$d\phi_f = \text{Hess}f$ as a map from V to $V^* \otimes V^*$. So ϕ_f is a local homeomorphism at p if and only if the Hessian is non-singular at p.

Part c

From the computation in part b), we have:

$d\phi_{f,p} = H_{ij}(p)dx^i \otimes dx^j$, where H_{ij} denotes the Hessian of f.

By the chain rule, it follows that:

$$\begin{array}{rcl} d\phi_f^{-1} : TV^* & \to & TV \\ (\alpha, \dot{\alpha}) & \mapsto & \left(\phi_f^{-1}(\alpha), H^{ij}(\phi_f^{-1}(\alpha))\dot{\alpha}_i \dfrac{\partial}{\partial x^j}\right) \end{array}$$

Using this, we have:

$$\begin{array}{rcl} \mathcal{L}f(\alpha) & = & < id(\alpha), \phi_f^{-1}(\alpha) > -f \circ \phi_f^{-1}(\alpha) \\ \phi_{\mathcal{L}f}(\alpha) = d(\mathcal{L}f_\alpha) & = & < d(id)_\alpha, \phi_f^{-1}(\alpha) > + < id(\alpha), df_{\phi_f^{-1}}(\alpha) > -df_{\phi^{-1}(\alpha)} \circ d\phi_{f,\alpha}^{-1} \\ & = & < id, \phi_f^{-1}(\alpha) > +(\alpha - df_{\phi_f^{-1}(\alpha)})H^{-1}(\phi_f^{-1}(\alpha)) \\ & = & \phi_f^{-1}(\alpha) + (\alpha - df_{\phi_f^{-1}(\alpha)})H^{-1}(\phi_f^{-1}(\alpha)) \\ & = & \phi_f^{-1}(\alpha) + (\alpha - \phi_f(\phi_f^{-1}(\alpha)))H^{-1} = \phi_f^{-1}(\alpha) + (\alpha - \alpha)H^{-1} \\ \phi_{\mathcal{L}f}(\alpha) & = & \phi_f^{-1}(\alpha) \end{array}$$

where $H^{-1} = H^{ij}\frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$.

Part d

From the previous part, we have

$$\alpha = \phi_{\mathcal{L}f}^{-1}(x) = \phi_f(x)$$

for any x in V.

$$\begin{array}{rcl} \mathcal{L}\mathcal{L}f(x) = \mathcal{L}(\mathcal{L}f)(x) & = & < x, \phi_{\mathcal{L}f}^{-1}(x) > -\mathcal{L}f(\phi_{\mathcal{L}f}^{-1}(x)) \\ & = & < x, \alpha > -\mathcal{L}f(\alpha) \\ & = & < x, \alpha > -(< x, \alpha > -f \circ \phi_f^{-1}(\alpha)) \\ & = & f(\phi_f^{-1}(\phi_f(x))) = f(x) \end{array}$$

So it follows that $\mathcal{L}\mathcal{L} = id$.

Part e

One can generalize the notion of the Legendre transform to vector bundles in the following way:

Given $f : M \times V \to \mathbb{R}$, we define $\phi_f : M \times V \to M \times V^*$ by

$$\phi_f(m, v) = (m, d_V f_{(m,v)}),$$

where $d_V f$ is the "partial external differential" of f along V. (This is equivalent to fixing a point in M, and then taking the differential of f as a function on V.)

Since f is smooth on $M \times V$, so is ϕ_f. The local coordinates on the fibre are linear, so the "Hessian along V" is well-defined on each point of M, and it is precisely equal to $d_V \phi_f$.

ϕ_f is a local homeomorphism along the fibres, if for each point p on M, the Hessian along V in some local coordinates have none-zero determinant. Note that this doesn't necessarily give a local homeomorphism on the bundle.

Assuming that ϕ_f is a diffeomorphism along the fibre at each base point allows one to take the Legendre transform of f as before, only this time one has to mind the base point. This gives us a notion of Legendre transform on trivial vector bundles, and one can see that transforming a smooth function on $M \times V$ will yield a smooth function on $M \times V^*$ since it is the sum of compositions of smooth maps.

Finally, for a general vector bundle, one may perform Legendre transformation via the local trivializations.

Thermodynamics Homework
MAT 561

Christopher Bay

March 6, 2007

Let $\mathbb{V}$ be the 4-dimensional vector space with linear coordinates T, S, p, V, i.e.

$$\mathbb{V} = \mathbb{R}_T \oplus \mathbb{R}_S \oplus \mathbb{R}_p \oplus \mathbb{R}_V$$

where each coordinate has the appropriate physical units. Endow $\mathbb{V}$ with the symplectic form

$$\omega = dT \wedge dS - dp \wedge dV.$$

1. We use ω to identify $\mathbb{R}_p = (\mathbb{R}_V)^*$ via the linear isomorphism

$$\omega_p : \mathbb{R}_p \to (\mathbb{R}_V)^*, \ \partial_p \mapsto \iota(\partial_p)\omega = -dV = -V. \tag{1}$$

Let $U \in C^\infty(\mathbb{R}_S \oplus \mathbb{R}_V)$ be the internal energy. Our goal is to compute the Legendre transform $\mathcal{L}_V U : \mathbb{R}_S \oplus (\mathbb{R}_V)^* \to \mathbb{R}$ of U along the V direction and reinterpret it as a map $\mathbb{R}_S \oplus \mathbb{R}_p \to \mathbb{R}$. Let $s \in \mathbb{R}_S$ and $\alpha \in (\mathbb{R}_V)^*$, and let $U_s : \mathbb{R}_V \to \mathbb{R}$ be the function $U_s(v) = U(s, v)$. Interpret $dU_s : \mathbb{R}_V \to \mathbb{R}_V \times (\mathbb{R}_V)^*$ as a map from $\mathbb{R}_V$ to $(\mathbb{R}_V)^*$ and suppose that this map is a diffeomorphism for each $s \in S$. Then

$$\begin{aligned} \mathcal{L}_V U(s, \alpha) &= \mathcal{L} U_s(\alpha) && (2)\\ &= \langle \alpha, (dU_s)^{-1}(\alpha)\rangle - U(s, (dU_s)^{-1}(\alpha)). && (3) \end{aligned}$$

Here $\langle\ ,\ \rangle : (\mathbb{R}_V)^* \otimes \mathbb{R}_V \to \mathbb{R}$ denotes the canonical pairing. Consider the bilinear map $p \cdot V : \mathbb{R}_p \oplus \mathbb{R}_V \to \mathbb{R}$. Via the diffeomorphism $(dU_s)^{-1} : (\mathbb{R}_V)^* \to \mathbb{R}_V$ we can interpret $p \cdot V$ as a map on $\mathbb{R}_p \oplus (\mathbb{R}_V)^*$ (which is no longer linear, and depends on the choice of $s \in S$). Furthermore, the map ω_p allows us to consider this map as a function on $\mathbb{R}_p$ alone. Making the dependance on s explicit, we have a map $p \cdot V : \mathbb{R}_S \oplus \mathbb{R}_p \to \mathbb{R}$. For $s \in S$ and $\rho \in \mathbb{R}_p$ we have

$$\begin{aligned} p \cdot V(s, \rho) &= p(\rho) \cdot V((dU_s)^{-1}(\omega_p(\rho)))\\ &= (p(\rho)V)((dU_s)^{-1}(\omega_p(\rho)))\\ &= \langle p(\rho)V, (dU_s)^{-1}(\omega_p(\rho))\rangle\\ &= -\langle \alpha, (dU_s)^{-1}(\alpha)\rangle, \end{aligned}$$

where $\alpha := \omega_p(\rho) \in (\mathbb{R}_V)^*$, and the minus sign is due to the sign in (**??**). Now, we may also interpret the term on the far right of (**??**) as a function $U : \mathbb{R}_S \oplus \mathbb{R}_p \to \mathbb{R}$ via ω_p:

$$U(s, \rho) = U(s, (dU_s)^{-1}(\omega_p(\rho))) = U(s, (dU_s)^{-1}(\alpha)).$$

Therefore,

$$-\mathcal{L}_V U = pV + U = H \in C^\infty(\mathbb{R}_S \oplus \mathbb{R}_p),$$

where H is the enthalpy.

For a path γ, the change in enthalpy along γ is

$$\int_\gamma dH = \int_\gamma dU + \int_\gamma (pdV + Vdp).$$

According to the First Law of Thermodynamics,

$$\int_\gamma dU = W(\gamma) + Q(\gamma) = -\int_\gamma pdV + Q(\gamma).$$

Thus, the change in enthalpy is $Q(\gamma)+\int_\gamma Vdp$ and the change in enthalpy is equal to the heat added precisely when $\int_\gamma Vdp = 0$. In particular, this occurs for isobaric (constant pressure) processes.

2. The form ω also induces an isomorphism

$$\omega_T : \mathbb{R}_T \to (\mathbb{R}_S)^*,\ \ \partial_T \mapsto \iota(\partial_T)\omega = dS = S. \tag{4}$$

We now derive the relationship

$$-\mathcal{L}_S U = U - TS = F_{Helm} \in C^\infty(\mathbb{R}_T \oplus \mathbb{R}_V)$$

where the $U - TS$ term will be properly interpreted.

We assume that for each $v \in V$ the map dU_v is a diffeomorphism. By definition, for $\sigma \in (\mathbb{R}_S)^*$ and $v \in V$,

$$\begin{aligned} \mathcal{L}_S U(\sigma, v) &= \mathcal{L} U_v(\sigma) && (5)\\ &= \langle \sigma, (dU_v)^{-1}(\sigma)\rangle - U((dU_v)^{-1}(\sigma), v). && (6) \end{aligned}$$

The bilinear map $TS : \mathbb{R}_T \oplus \mathbb{R}_S \to \mathbb{R}$ can be considered as a map on $\mathbb{R}_T \oplus (\mathbb{R}_S)^*$ (which also depends on v) by precomposing with $(dU_v)^{-1} : (\mathbb{R}_S)^* \to \mathbb{R}_S$ in the second argument. It can then be considered as a map on $\mathbb{R}_T$ by precomposing with ω_T in the second argument. We then have a map $TS : \mathbb{R}_T \oplus \mathbb{R}_V \to \mathbb{R}$, given by

$$\begin{aligned} TS(t, v) &= T(t)S((dU_v)^{-1}(\omega_T(t)))\\ &= \langle T(t)S, (dU_v)^{-1}(\omega_T(t))\rangle\\ &= \langle \sigma, (dU_v)^{-1}(\sigma)\rangle, \end{aligned}$$

where $\sigma := \omega_T(t) \in (\mathbb{R}_S)^*$. The U term in (**??**) may be interpreted as a map $U \in C^\infty(\mathbb{R}_T \oplus \mathbb{R}_V)$ as follows:

$$U(t, v) = U((dU_v)^{-1}(\omega_T(t)), v) = U((dU_v)^{-1}(\sigma), v).$$

Therefore,

$$-\mathcal{L}_S U = U - TS = F_{Helm} \in C^\infty(\mathbb{R}_T \oplus \mathbb{R}_V).$$

For a path γ the change in F_{Helm} along γ is

$$\int_\gamma dF_{Helm} = \int_\gamma dU - \int_\gamma (SdT + TdS).$$

By the First Law of Thermodynamics,

$$\int_\gamma dU = W(\gamma) + Q(\gamma) = W(\gamma) + \int_\gamma TdS.$$

So the change in F_{Helm} is $W(\gamma) - \int_\gamma SdT$ and the change in the Helmholtz free energy is equal to the work done by the system precisely when $\int_\gamma SdT = 0$. In particular, this occurs for isothermal (constant temperature) processes.

3. Now use ω to identify $\mathbb{R}_p \oplus \mathbb{R}_T = (\mathbb{R}_S \oplus \mathbb{R}_V)^*$ via $\omega_p \oplus \omega_T$. Suppose $dU : \mathbb{R}_S \oplus \mathbb{R}_V \to (\mathbb{R}_S \oplus \mathbb{R}_V)^*$ is a diffeomorphism. Then for $\sigma \in (\mathbb{R}_S)^*$ and $\nu \in (\mathbb{R}_V)^*$,

$$\mathcal{L}U(\sigma, \nu) = \langle (\sigma, \nu), (dU)^{-1}(\sigma, \nu)\rangle - U((dU)^{-1}(\sigma, \nu)).$$

In the notation used above,

$$dU(s, v) = \frac{\partial U}{\partial S} \mid_{(s,v)} dS + \frac{\partial U}{\partial V} \mid_{(s,v)} dV = dU_v(s) + dU_s(v),$$

so that $dU(s, v) = (dU_v(s), dU_s(v))$. So $dU(s, v) = (\sigma, \nu)$ implies $(dU)^{-1}(\sigma, \nu) = ((dU_v)^{-1}(\sigma), (dU_s)^{-1}(\nu))$ where the pair (s, v) which depends on the pair (σ, ν). Then for some (s, v) we have

$$\mathcal{L}U(\sigma, \nu) = \langle \sigma, (dU_v)^{-1}(\sigma)\rangle + \langle \nu, (dU_s)^{-1}(\nu)\rangle - U((dU_v)^{-1}(\sigma), (dU_s)^{-1}(\nu)).$$

Since (s, v) depends on $(\sigma, \nu) \in (\mathbb{R}_S \oplus \mathbb{R}_V)^*$ the interpretations made above are still valid, but now the first term is $TS \in C^\infty((\mathbb{R}_S \oplus \mathbb{R}_V)^*)$ and the second term is $-pV \in C^\infty((\mathbb{R}_S \oplus \mathbb{R}_V)^*)$. For example, if $t = \omega_T^{-1}(\sigma)$ then

$$\begin{aligned} \langle \sigma, (dU_v)^{-1}(\sigma)\rangle &= \langle \omega_T(t), (dU_v)^{-1}(\sigma)\rangle \\ &= \langle T(t)S, (dU_v)^{-1}(\sigma)\rangle \\ &= T(\omega_T^{-1}(\sigma))S((dU_v)^{-1}(\sigma)) \end{aligned}$$

and similarly for $-pV$. By precomposing with $(\omega_p \oplus \omega_T)^{-1}$ we have $TS, -pV, U \in C^\infty(\mathbb{R}_p \oplus \mathbb{R}_T)$ and

$$-\mathcal{L}U = U - TS + pV = F_{Gibbs} \in C^\infty(\mathbb{R}_p \oplus \mathbb{R}_T).$$

For the most part, the lectures will be heavily influenced by the first two references. We only mention the other references for those who would like to revisit some of the prerequisite topics or look into some of the more advanced topics.

Overview of selected topics in theoretical physics

Wm D Linch III and Jerome A. Jenquin

September 14, 2006

Part I

Classical Theory

Chapter 1

Introduction, notation, and preliminaries

1.1 Our guiding philosophy

This course is meant to be an introduction to the topics usually taught to undergraduate physics major. These are classical mechanics in both the Newtonian and Relativistic setting; classical electromagnetism; thermodynamics and statistical mechanics; and quantum mechanics. We have two audiences in mind: former physics majors who have seen the general content before and pure math majors with an interest is studying physical topics.

For the first audience, we present the physics in a way that emphasizes some of the overall mathematical structure. This approach can be of great pedagogical benefit by shining new light on old topics and preparing one for further study in field and string theory, to which many of the mathematical ideas we discuss apply.

For the second audience, the mathematical structure is there for psychological reasons, as well as pedagogical ones, softening the culture shock and yet introducing math that is interesting in its own right. We also present specific examples and solutions to get a hands-on feel for the physical ideas that they display.

In some cases, particularly when we cover quantum mechanics, some may find our mathematical approach to be vague and hand-waving at best. While this is somewhat regrettable, we will not apologize for it. One of the goals of this course is to offer the students a sense (perhaps even an intuition) for how physicists achieve progress, not in spite of eschewing mathematical rigor, but sometimes *because* of it.

1.2 Physical and mathematical preliminaries

Dimensional analysis and "naturalness"
Vector fields, differential forms, and calculus

Chapter 2

Classical mechanics

2.1 Newtonian mechanics on Euclidean space

2.1.1 Space, time and particle

Space We will generally refer to *space* meaning the 3-dimensional real vector space $\mathbb{R}^3 = \{\mathbf{x} = (x, y, z) \in \mathbb{R}\times\mathbb{R}\times\mathbb{R}\}$ with the *right-handed orientation* and the Euclidean inner product $\langle\cdot\,,\cdot\rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, to which we will refer as *dot product*. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, $\langle\mathbf{x}, \mathbf{y}\rangle = \mathbf{x} \cdot \mathbf{y}$. We describe all events by coordinate expressions as that is the language most closely related to the realization of the event. We will also switch freely between various confusing but conventional notations to describe the coordinates. For example, it is common to write $\mathbf{x} \in \mathbb{R}^3$ variously as $\mathbf{x}^i$ for $i = 1, 2, 3$ or just x^i and also $\mathbf{x} = (x, y, z)$. Note that by convention $x^1 = x$ denotes the "x-coordinate", $x^2 = y$ denotes the "y-coordinate", and $x^3 = z$ denotes the "z-coordinate". In this language, $\mathbf{x} \cdot \mathbf{y} = \sum_{i,j=1}^3 \delta_{ij} x^i y^j$ where δ_{ij} is the *Kronicker-delta*, equal to $+1$ when $i = j$ and 0 otherwise. We will use the *Einstein summation convention* meaning that when covariant and contravariant indices are repeated, a summation over the full range of the indices is implied, that is, for a vector x^i and covector p_i, $p_i x^i = \sum_{i=1}^3 p_i x^i$.

For any vector $\mathbf{x}$ we define the unit vector $\hat{\mathbf{x}} = |\mathbf{x}|^{-1}\mathbf{x}$ where $|\mathbf{x}| \equiv \sqrt{\mathbf{x} \cdot \mathbf{x}} \equiv r$. The unit vectors in the x-, y-, and z-directions are denoted $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$. The orientation on space defines a *cross-product* $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$. The right-handed orientation is the one given by the *right-hand rule* $\hat{\mathbf{x}} \times \hat{\mathbf{y}} = +\hat{\mathbf{z}}$. This can be expressed using the totally anti-symmetric tensor ϵ_{ijk} normalized to $\epsilon_{123} = +1$, that is $\epsilon_{ijk}(\hat{\mathbf{x}})^i(\hat{\mathbf{y}})^j(\hat{\mathbf{z}})^k = 1$, in terms of which for any two vectors $\mathbf{a}$, $\mathbf{b}$, $(\mathbf{a}\times\mathbf{b})^i = \delta^{ii'}\epsilon_{i'jk}a^j b^k$.

Time and Particle In the Newtonian picture of nature, there is a universal clock defining time for all observers. We will denote this universal time by $t \in \mathbf{R}$. In general, particle motion is described, by definition, by a time-dependent vector $\mathbf{x}(t)$.[1] The *velocity* $\mathbf{v}$ of a particle is the derivative with respect to time of its position $\mathbf{v}(t) = \dot{\mathbf{x}}(t) \equiv \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t)$. Its acceleration $\mathbf{a}$ is the derivative of its velocity, or the second derivative of its position $a(t) = \ddot{\mathbf{x}}(t)$. We will often drop the argument of these physical quantities, leaving their time-dependence implicit. It is also common to denote the constant values of these quantities with a 'naught', e.g. $\mathbf{x}_0$ for constant position vector. We define the linear momentum $\mathbf{p}(t)$ of a particle as the product of its mass m and velocity $\dot{\mathbf{x}}$, $\mathbf{p} = m\dot{\mathbf{x}}$.

Symmetries The space symmetry group for Newtonian mechanics is given by the Euclidean group $\mathrm{SO}(3) \ltimes \mathbb{R}^3$ where the compact factor acts on the coordinates by rotations $x^i \mapsto \Lambda^i{}_j x^j : \Lambda^i{}_j \delta_{ik} \Lambda^k{}_l = \delta_{jl}$ and the non-compact factor acts by translations $x^i \mapsto x^i + a^i$. In Newtonian mechanics the time variable does not mix with the spacial coordinates. We therefore have a separate symmetry factor $\mathbb{R}$ of translations in time $t \mapsto t + c$. The physical interpretation of these space-time symmetries is that in writing equations, the origin and orientation of the coordinate system are conventions and in particular are not physical. That is, only the relative coordinates of space-time events are physical. In general, physical quantities are invariant under the space-time symmetry group. In practice we will always fix this ambiguity by specifying the coordinate system.

Note that when physical quantities are expressed in linear-algebraic language, the transformation laws are simple, that is, linear. When a physical formalism is expressed in the way, we say that the formalism is *covariant* – in this case with respect to the space symmetry group $\mathrm{SO}(3) \ltimes \mathbb{R}^3 \times \mathbb{R}$ – and that the space-time symmetry is *manifest*. It is always the case that a covariant formalism is expressed in terms of unphysical quantities because covariance means that the symmetries are manifest which means that they are realized linearly on the variables which, in turn, means that the variables are not invariant under the symmetries and hence not physical.

2.1.2 Newton's Laws

Newton I: An object of mass m in rectilinear motion $\mathbf{x}(t) = \mathbf{v}_0 t + \mathbf{x}_0$ will stay in rectilinear motion unless acted on by a force.

[1] This is the meaning of *particle* as opposed to an extended object for which we have to specify a distribution of positions as a function of time.

This statement defines the concept of *kinematics* or geodesic motion. It is equivalent to the statement that free particle trajectories satisfy the equation

$$\ddot{\mathbf{x}} = 0. \tag{2.1}$$

Which, in turn, is equivalent to the statement that, in the absence of force, momentum is conserved

$$\dot{\mathbf{p}} = 0. \tag{2.2}$$

Note that the equation is $\mathrm{SO}(3) \ltimes \mathbb{R}^3 \times \mathbb{R}$ covariant. In a local lagrangian system, the existence of a global symmetry implies, via Noether's theorem, the existence of a conserved current (c.f. section 2.1.5). Suffice it here to say that the current associated to the translations is the momentum $\mathbf{p}$. The kinematic equation (2.2) expresses that it is *conserved*, that is, constant in time. Similarly, there is a current associated to the rotational invariance – the *angular momentum* $\mathbf{L} = \mathbf{x} \times \mathbf{p}$. Noting that $\mathbf{p} \parallel \dot{\mathbf{x}}$ and using the kinematical equation, we see that the angular momentum is conserved $\dot{\mathbf{L}} = 0$. Finally, the current associated to a shift in the time variable is $T = \frac{1}{2}m\dot{\mathbf{x}}^2 = \frac{1}{2m}\mathbf{p}^2$ and is called the (*kinetic*) *energy* (the normalization is conventional). Again, the kinematic equations imply that it is conserved.

Newton II: An object of mass m, when acted on by a force $\mathbf{F}$ will deviate from rectilinear motion with an acceleration $\mathbf{a} = \ddot{\mathbf{x}}$ according to the relation

$$\mathbf{F}(\mathbf{x}, t) = m\mathbf{a}(t) \tag{2.3}$$

or, equivalently,

$$\mathbf{F}(\mathbf{x}, t) = \dot{\mathbf{p}}(t). \tag{2.4}$$

This is the statement of *dynamics* or the deviation from geodesic motion due to an external influence. An equivalent way to express this is that the second law defines the *source* $\frac{1}{m}\mathbf{F}(\mathbf{x}, t)$ for the kinematic (read "source-less") equation $\ddot{\mathbf{x}} = 0$ or $\dot{\mathbf{p}}$ of the first law. In this sense, it defines what is meant by a force.

An important point to note is that the second law is *linear* in the force. This implies that we have the

Principle of superposition: If there are 2 forces $\mathbf{F}_1$ and $\mathbf{F}_2$ acting on the same particle, the effective force $\mathbf{F}_{\text{total}}$ the particle experiences is the vector sum of the individual forces $\mathbf{F}_{\text{total}} = \mathbf{F}_1 + \mathbf{F}_2$. In particular, two opposing forces of equal magnitude and opposite direction applied to the same particle produce no net dynamics.

A second important point is that the second law can be interpreted as defining the mass of an object to be the ratio of a stimulus $|\mathbf{F}|$ to the response $|\mathbf{a}|$ in its motion by $m = \frac{|\mathbf{F}|}{|\mathbf{a}|}$. In this sense, we see that m refers to an *inertial mass*, that is, a property describing its resistance to a change in motion.

Newton III: An object, when acted on by some agent with a force $\mathbf{F}_{\text{action}}$ will exert a force $\mathbf{F}_{\text{reaction}}$ on the agent of equal magnitude and opposite direction, *id est*,

$$\mathbf{F}_{\text{action}} = -\mathbf{F}_{\text{reaction}}. \tag{2.5}$$

This is a statement of linear momentum conservation during a collision. Intuitively, when pressing on an object with some force, the object presses back (otherwise, we wouldn't be able to feel it). The third law is the statement that the reaction force is of precisely the same magnitude as the applied force.[2]

Newton's law of universal gravitation Consider two objects, one of mass m_1 and the other of mass m_2. They will exert a gravitational force on one another given by

$$\mathbf{F}_{\text{gravitation}} = -G\frac{m_1 m_2}{r^2}\hat{\mathbf{r}} \tag{2.6}$$

where $G \equiv \frac{1}{4\pi\kappa} \approx 6.67259 \times 10^{-11}\mathrm{Nm}^2\mathrm{kg}^{-2}$ is the least precisely known fundamental constant of nature.

This formula is fundamentally different from the second law. Firstly, it introduces a constant G which is claimed to be fundamental in the sense that it is the same number no matter what material form the masses take.

Secondly and related to this, the masses $m_{1,2}$ entering it could be called *gravitational masses* since they describe a property of an object we are calling gravitation and should probably have been called gravitational charge. A priori, this is a different type of mass than the inertial mass entering the dynamical second law. Therefore, Newton's law of universal gravitation is making the bold assertion that gravitational mass and inertial mass are equivalent.

Finally, we note that setting $\mathbf{a}_2 = -\frac{Gm_2}{r^2}\hat{\mathbf{r}}$ to be the acceleration due to the gravity of the mass m_2 at a distance r from its position, we find the form $\mathbf{F}_{\text{gravitation}} = m_1\mathbf{a}_2$. Taking $m = m_{\text{♁}}$ to be the mass of the earth and $r = r_{\text{♁}}$ its radius, we find the famous acceleration due to gravity $g = |\mathbf{a}_{\text{♁}}| \approx 9.8ms^{-2}$.

[2] This law causes some confusion when used in conjunction with the second to the effect that if the object pushes back with exactly the same force, the forces should cancel and there should be no resulting dynamics. Indeed, there is no relative dynamics between the hand and the object, rather, the object will accelerate relative to the ground against which we are also pushing when we try to accelerate the object.

2.1.3 Potentials

There are various drawbacks to the vector space formulation of Newtonian mechanics, not the least of which is that all defining equations are vector equations. In most cases of physical interest, drastic simplifications are made possible by switching to a description in terms of energy. Suppose the force is *holonomic* $\nabla \times \mathbf{F} = 0$. Then we can define the *potential energy function* $U(\mathbf{x}, t)$ s.t. $\mathbf{F} = -\nabla U$. The sign comes from the observation that a force acts so as to decrease the potential energy. The total energy $E = T + U$ is the sum of the kinetic and potential energy. Just as the kinetic energy was conserved in the absence of external forces, the total energy of a system is conserved when the force is the gradient of the potential energy and the latter does not depend explicitly on time: $\dot{E} = m\dot{\mathbf{x}} \cdot \ddot{\mathbf{x}} + \nabla U \cdot \dot{\mathbf{x}} + \frac{\partial U}{\partial t} = 0$ by Newton II. This is the famous *principle of the conservation of energy*. It is very powerful because it is, in the cases in which it applies, equivalent to the second law but it is a scalar equation, making it much easier to use.

2.1.4 Hamiltonian

Very closely related to the energy formulation of Newtonian mechanics is the *Hamiltonian formalism*. In this formulation, the fundamental variables are the position x^i and the momentum p_i vs. the position and the velocity (c.f. section 2.1.5). A physical trajectory is a graph in the *phase space* $\{x^i, p_i\}_{i=1}^3$.[3] Note that the momentum is treated as a 1-form in this formulation (which, as we will soon learn, is the proper interpretation of this quantity). The dynamics is encoded in the *Hamiltonian* $H(x, p)$ which, when evaluated on a point in the phase space, is equal to the energy E introduced in section 2.1.3. In particular, it is the sum of the kinetic energy function $T(p)$ which we take to be a function only of the momentum (usually $T = \frac{1}{2m}p^2$)[4] and the potential energy function $U(x)$ with we take to depend only on the position. We can now easily show that the definition of momentum and the second law imply

Hamilton's Equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}$$

[3] In general, the space parameterized by x may be any $\mathcal{C}^2$ 3-manifold M. Then the phase space is defined to be the co-tangent bundle T^*M. From this point of view, it is easy to see the symplectic structure.

[4] On a more general space the kinetic energy function will depend on x through the metric: $T = \frac{1}{2m} g^{ij}(x) p_i p_j$.

$$\dot{p}_i = -\frac{\partial H}{\partial x^i} \tag{2.7}$$

The form of these equations[5] displays an important aspect of the phase space, namely, its *symplectic structure*: The phase space comes equipped with its Poincaré 1-form $p_i \mathrm{d}x^i$ and therefore the symplectic 2-form $\mathrm{d}p_i \wedge \mathrm{d}x^i$. This statement is often implicit in a discussion of Hamilton's equations in which one considers transformations of the variables (x, p) which preserve the 'form' of Hamilton's equations. These *canonical transformations* are the symplectomophisms – smooth transformations on the phase space coordinates which preserve the symplectic structure.

From the Poincaré 1-form and a phase space trajectory γ (a path in phase space) we can construct the *action* (*functional*)

$$S[\gamma] = \int_\gamma p_i \mathrm{d}x^i. \tag{2.8}$$

A useful generalization of the phase space includes the time coordinate as an additional variable. This 7-dimensional space is called the *extended phase space*. Similarly to the action functional (2.8) on the un-extended phase space, from the Poincaré 1-form and the Hamiltonian function we can construct the action functional[6]

$$S[\gamma] = \int_\gamma \left[p_i \mathrm{d}x^i - H(x, p)\mathrm{d}t \right]. \tag{2.9}$$

It is important to remember that $(x^i(t), p_i(t))$ are functions of the time parameter t. As such, we are allowed to "vary" them. That is, we consider an infinitesimal deformation of the trajectory $\gamma \to \gamma' = \gamma + \delta\gamma$. The *variational* or *functional derivative* of the action functional is defined to be the linear part of $S[\gamma']$, that is

$$\frac{\delta S}{\delta \gamma} \equiv \lim_{\delta\gamma \to 0} S[\gamma + \delta\gamma]. \tag{2.10}$$

This notation δ for ∂ for the functional derivative is customary in the calculus of variations.

The path has two linearly independent variations in the x-direction and the p-direction. It is therefore possible to define the partial variations in these directions. The following notation is customary (and, hopefully, self-explanatory)

$$\delta S = \delta x^i \frac{\delta S}{\delta x^i} + \delta p_i \frac{\delta S}{\delta p_i}. \tag{2.11}$$

---

[5]Note that the Hamilton equations (2.7) are coupled ordinary differential equations of the *first* order which contain the same information as Newton's second law which is second order.

[6]Note that this is (the negative of) an integrated Legendre transformation of $H(x, p)$.

The action is called *stationary* when $\delta S = 0$. Since the x- and p-variations are independent, stationary action implies

$$\begin{aligned} 0 &= \frac{\delta S}{\delta x^i} = -\dot{p}_i - \frac{\partial H}{\partial x^i} \\ 0 &= \frac{\delta S}{\delta p_i} = \dot{x}^i - \frac{\partial H}{\partial p_i} \end{aligned} \tag{2.12}$$

and we recover Hamilton's equations (2.7).[7] This is the *principle of stationary action*; the physical trajectories in phase space are those which extremize (usually minimize) the action.

This point of view has many advantages. Firstly, it generalizes the intuitive idea that physical processes are such that they minimize the energy. Secondly, a modification of this formalism (c.f. section 2.1.5) us a very powerful tool to solve complicated concrete problems in analytical dynamics especially dynamical systems defined in terms of constrained degrees of freedom. Finally, the principle of stationary action will fit seamlessly into the description of quantum mechanical systems c.f. chapter **??**. There we will see that quantum mechanical corrections to classical mechanics have the interpretation of deviations $\delta\gamma$ of the phase space trajectories.

2.1.5 Lagrangian

A complementary formulation of Newtonian mechanics is the Lagrangian formulation. The Lagrangian formulation is a "Legendre transform of the Hamiltonian formulation". Indeed, the space replacing the phase space of Hamiltonian mechanics is the space parameterized by positions q^i and velocities $\dot{q}^i$.[8] The *Lagrangian function* $L(q^i, \dot{q}^i)$ is the Legendre transform of the Hamiltonian $H(x^i, p_i)$

$$L(q^i, \dot{q}^i) = p_i \dot{q}^i - H(x, p). \tag{2.13}$$

Plugging in the form $H = T + U$ and substituting $p_i = m\dot{q}_i$ we find that $L = T - U$. By the definition of the action (2.9), the Lagrangian function is the unintegrated action density

$$S[\gamma] = \int_\gamma L(q, \dot{q}) \tag{2.14}$$

[7] In the first equation, we have integrated the time derivative by parts. This is legal since the surface term is proportional to δx which vanishes when evaluated at the endpoints of the path (recall that we vary the path but keep the endpoints fixed).

[8] For a general space M the Lagrangian formalism is defined on the tangent bundle TM. It is conventional in this context to denote the positions by q^i instead of x^i.

where γ is re-interpreted as a section of the tangent bundle.[9] The stationary phase principle in this case implies the *Euler-Lagrange equation*

$$\frac{\partial L}{\partial q^i} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{q}^i} = 0. \tag{2.15}$$

The advantages of the Lagrangian formalism over the Hamiltonian one include the use of the stationary action principle to solve complicated problems in anaytical dynamics and the possibility to easily manifest Lorentz invariance in relativistic theory (c.f. **??**).

Noether's theorem Consider a time-independent infinitesimal transformation of coordinates $q^i \mapsto q'^i \approx q^i + \epsilon^i$ under which the action is invariant $S \mapsto S$, that is, a *symmetry* of the theory. Now promote the parameter $\epsilon^i \to \epsilon^i(t)$ to a function. The resulting change in the action must be proportional to $\dot{\epsilon}$ since, when ϵ is constant, the transformation is a symmetry. Given this, there must be a function $J_i(q, \dot{q}, t)$ such that $\delta S = \int \dot{\epsilon}^i J_i = \epsilon^i J_i| - \int \epsilon^i \dot{J}_i$. The first term is the "surface term" $\epsilon J|_{t_i}^{t_f}$ – the difference of the quantity ϵJ at the final time t_f and the initial time t_i. The second term vanishes by the equation of motion. (Prove it!) When ϵ is constant, we see that J is conserved $J(t_f) = J(t_i)$. Such conserved functions arising from symmetries of the theory are called *Noether currents*. In this case the symmetry is a translation and the current is $J_i = \partial L/\partial \dot{q}^i$, which is the definition of the momentum. In the absence of external forces, this is indeed conserved.

With an eye to the future we will refer to a time-independent symmetry as a *global symmetry*. Noether's theorem is the statement that for every global symmetry of the action, there is a conserved current and *vice versa*.

2.1.6 Examples

Gravitational potentials and solvable systems.
Potential theory and the need for fields.

[9] Usually this whole story is reversed: One defines the Lagrangian function as the difference between the potential and kinetic energy functions and develops the Lagrangian formalism and stationary action principle. Subsequently the Legendre transformation to the Hamiltonian function is performed. It is then proven that the resulting Hamiltonian is independent of $\dot{q}$ and the phase space picture is developed.

Index

Bibliography

2.2 Physics

[1] H. Georgi, "Waves,"

[2] Griffiths, "Quantum Mechanics,"

2.3 Math

[3] V.A. Arnold, "Classical Mechanics,"

[4] Bamberg and Sternberg, "vol. 2,"

Mathematical Physics
September 6, 2006

Gabriel C. Drummond-Cole

September 6, 2006

[Course overview.]

Let's get started with Newtonian mechanics. The prerequisite is that you know what a manifold is. I won't assume Riemannian geometry. I'm trying to keep things simple, so things will be on flat affine spaces. If you want to sup it up in your head as we go along, feel free.

So classical mechanics, we're talking about the physics (and math) of a single particle moving in some Euclidean space. So if you want to play along at home it could be a Riemannian manifold. The mathematical models are paths x, maps from time $M^1 \to X$ where X is a Euclidean target space. (or possibly a Riemannian manifold). Do I need to define a Euclidean space? Sorry, I guess that's a little bit insulting.

All right, so there are two spaces that are going to color our approach to this, M^1 and X. Let's look at the structure of both of these spaces.

Okay, so time has physical significance. We attribute certain mathematical structure to it to correspond to this. In particular,

1. It's affine, meaning that it's not necessarily a vector space. How do you add two instances in time? You can't. You can talk about how much time has passed, so it's an affine space over a one dimensional vector space $T \cong \mathbb{R}$.
2. There are units, like seconds or hours. What sort of mathematical structure would units be associated with? A norm, a metric. In particular we have a translation invariant metric on affine time. In other words, T has an inner product on it.
3. We could potentially also attribute to it an orientation, a differentiation between going forward and backward. We'll hold off on that for now.

To have a cogent discussion, we want to do math, so I want to fix some affine coordinate $t : M^1 \to \mathbb{R}$. We want to choose this so that $|dt| = 1$.

The structure, excluding the third, gives us a symmetry group that we will talk about again and again. The symmetry group is the Euclidean group for M^1, which includes translations and reflections. So there is a short exact sequence $1 \to \text{Translations} \to \text{Euc}(M^1) \to O(T) \to 1$. That's the structure we associate to the domain.

Now what structure do we have in the range?

1. X is an affine space over a vector space V.
2. There is a (translation invariant) metric so that V is an inner product space. This measures distance on $\mathbb{E}^d$.

If you're playing along with Riemannian geometry, these are the conditions that

1. X is a smooth manifold
2. X has a Riemannian metric
3. The metric is complete so we can work globally on X.

Let's go back to the Euclidean space. Again there is an associated symmetry group $\text{Euc}(\mathbb{E}^d)$, where you have translations and then reflections and rotations. Again you have a short exact sequence

$$1 \to \underbrace{V}_{\text{translations}} \to \text{Euc}(\mathbb{E}^d) \to O(V) \to 1.$$

For a general Riemannian manifold the group of isometries will be smaller, meaning lower dimensional. Sometimes this group will even be trivial.

There is one last piece of data that we need to define classical mechanics on X. We have the model of maps from affine Euclidean time into a Riemannian manifold. The last piece of data is

- a potential energy function $\mathscr{V} : X \to \mathbb{R}$.
- A mass $n > 0$ of the particle.

A quick note on units. In this course we'll come across a few fundamental units. In classical mechanics we'll only come across mass, length, and time. Energy will have units mL^2/t^2.

Actually, these are the only basic units. This is an empirical fact. We will see that extensions of classical mechanics will introduce new fundamental constants *such as the speed of light c with units L/T (special relativity), a fundamental angular momentum $\hbar$ with units ML^2/T (quantum mechanics), etc. but must reduce to classical mechanics in the appropriate limit, e.g. $c \to \infty$ and $\hbar \to 0$. Therefore, no new* fundamental units *are introduced even when more fundamental physics is uncovered.*

Now we can finally define what classical mechanics is, or at least the classical mechanics of the system. So all togethher, for any target space X, all possible particle motions are modeled on paths $P = \mathrm{Map}(M^1, X)$. We'll assume that the paths are smooth. Don't worry about putting a Frechet topology or whatever on this.

This is again actually an empirical fact. Discontinuous paths imply the disappearance and reappearance (at a later time) of particles while a kink in the path amounts to an instantaneous change in the particle's velocity. In order to change a Newtonian particle's velocity discontinuously one must apply an infinite force. Infinite forces are considered pathological and any use of such a thing should be considered only as an approximation.

Now given the potential energy function $\mathscr{V}$ the actual particle motions are paths x such that they satisfy Newton's second law

$$m\ddot{x}(t) = -\mathscr{V}'(x(t)) = -\nabla\mathscr{V}(x(t))$$

In the physics literature this equation is variously written as $m\ddot{\mathbf{x}} = -\nabla\mathscr{V}$ as a "vector" equation or as $m\ddot{x}^i = -\partial\mathscr{V}/\partial x_i$ in "components". Here $\mathbf{x}$ is a coordinate for the point x and x^i are its components (in some orthonormal coordinate system) where the indices $i = 1, 2, 3$ label the linearly independent directions. A common notation is $(x^1, x^2, x^3) = (x, y, z)$. The index on x^i is defined to be 'up' and is lowered with the metric or its inverse. In Euclidean space the placement of the indices (upper or lower) doesn't matter as the metric $g_{ij} = \delta_{ij}$ is just the Kronicker symbol but the more general case such as in Riemannian geometry or "curvilinear coordinates" (e.g. spherical coordinates) it, of course, matters a great deal.

We're going to look at the space of solutions to this equation $\mathscr{M}$, the space of states. A solution is a state. It's also called a phase space. Let me mention some properties right off the bat.

- It will be clear soon that $\mathscr{M}$ is a smooth manifold, so we can do calculus on it.
- The affine Euclidean group for time $\mathrm{Euc}(M^1)$ acts on $\mathscr{M}$ on the right so in particular time translation acts on it. So $T_s(x)(t) = x(t-s)$.
- One other thing, the potential was a function of X. It can also be a function of time, so that the symmetry is broken. So $\mathrm{Euc}(M^1)$ no longer preserves the space of solutions.

Let me argue that this is a smooth manifold. Let's see this by breaking some symmetry. Choose an instant t_0 in time and by picking this we break the affine symmetry. Then there's a natural map $\mathscr{M} \to T\mathbb{E}^d = V \times \mathbb{E}^d$ given by $x \mapsto (\dot{x}(t), x(t))$. This is a bijection and you can just transfer the differentiable structure across.

The physical intuition behind this diffeomorphism is the intuitively obvious fact that when you want to specify a particles trajectory, it suffices to give the initial position, its initial velocity, and the potential field in which the particle motion occurs (i.e. the force which acts on it). For example, the parabolic trajectory of a baseball depends on the gravitational potential field (in this case $\mathscr{V} = mgz$ with $g \approx 9.8ms^{-2}$) the hight of the ball upon release and its velocity *(speed and direction) upon release. Newton formulated his second law to be second order in time derivatives precisely to accommodate this empirical fact.*

Now we have a picture of what the space of solutions looks like. Let me give you some examples.

Example 1 *The free particle*
This is the case where it's moving in Euclidean space and $\mathscr{V}=0$. Then Newton's second law says $m\ddot{x}=0$. Then $\mathscr{M}=\{x(t)=p+vt|p\in\mathbb{E}^d, v\in V\}$. So given a t_0, the map takes $p+vt$ to (v,p). In this case the map doesn't depend on t.

Example 2 *This is a little less trivial but just as famous. It's the spring or harmonic oscillator.*
Implicitly you have to have a distinguished point where the spring is stable. I may as well take $X=\mathbb{R}^1$. Then the potential energy is $\mathscr{V}=\frac{1}{2}kx^2$, where the units of k are $\frac{M}{T^2}$.

Then Newton's second law says $m\ddot{x}(t)=-kx(t)$. Then $\mathscr{M}=\{p\cos(\sqrt{\frac{k}{m}}t)+\sqrt{\frac{m}{k}}v\sin(\sqrt{\frac{k}{m}}t)|p,v\in\mathbb{R}\}$. So then $p\cos(\sqrt{\frac{k}{m}}t)+\sqrt{\frac{m}{k}}v\sin(\sqrt{\frac{k}{m}}t)\overset{t=0}{\mapsto}(v,p)$.

There are other things I can point out about the space of solutions. We have the symmetry group of the domain that acts on the solutions. What about the symmetry group of the target? How does that naturally act on the space of solutions? it acts on the left, but only those isometries that preserve the potential. An isometry that changes the potential will not preserve the space of solutions. What's true about the two group actions? They commute. The time group will have to do with dynamics, the target group with kinematics. If $X=\mathbb{E}^d$, and $\mathscr{V}=0$, so that we're talking about the free particle, then the entire Euclidean group acts on the space of solutions, since everything preserves the 0 potential. In particular, if A is an affine Euclidean transformation and its derivative is in $O(V)$, then $p+vt\mapsto Ap+(dA\cdot v)t$.

Let me wrap up what I've said today, which isn't much. In summary, we've discerned that the space of solutions $\mathscr{M}$ to Newton's second law has the following structure:

- There's a right action by the Euclidean time group $\mathrm{Euc}(M^1)$ acting by right composition with the inverse.
- There's a left action by potential-preserving isometries of X, $\mathrm{Isom}(X,\mathscr{V})$.
- For $t_0\in M^1$ there's a natural diffeomorphism $\mathscr{M}\overset{t_0}{\cong}TX$.

Next time I hope to make this fit in with the idea of a symplectic structure.

Mathematical Physics
September 11, 2006

Gabriel C. Drummond-Cole

September 12, 2006

[Is the exact sequence from last time secretly Noether's theorem?]

We'll see Noether's theorem later. Let me recap what we've seen so far. So far we've seen particle motion, and the structure of the phase space $\mathscr{M}$ which are paths from the affine time to the target satisfying Newton's second law

$$\{x : M^1 \to X | \ddot{x} = -\mathscr{V}'(x(t))\}$$

This has a right action by $\mathrm{Euc}(M^1)$, a left action by $\mathrm{Isom}(X, \mathscr{V})$ and a for each $t_0 \in M^1$ a natural diffeomorphism $\mathscr{M} \overset{t_0}{\cong} TX$.

We saw something about symplectic geometry on a smooth manifold M^{2n}. This means there is a two form $\omega \in \Omega^2(M)$ such that ω^n is nowhere vanishing (nondegeneracy) and $d\omega = 0$ (closed).

The thing that will play a big role today is the symplectic gradient which takes smooth functions on a symplectic manifold into vector fields $C^\infty(M) \to \mathscr{X}(M)$ via $f \mapsto \xi_f$ characterised by $\iota(\xi_f) = df$. This gives us the Poisson bracket $\{\cdot, \cdot\}$ which makes $C^\infty(M)$ a Lie algebra. This is given by $\{f, g\} = \omega(\xi_f, \xi_g)$. Then this map ξ is a homomorphism of Lie algebras.

There is a typo in the equation above. The formula for ξ_f should be $\iota(\xi_f)\omega = \mathrm{d}f$, or in components (local coordinates) $\xi_f^i = \omega^{ij}\partial_j f$. The Poisson bracket is given in local coordinates z^i by

$$\{f, g\} = \frac{\partial f}{\partial z^i}\omega^{ij}\frac{\partial g}{\partial z^j}. \tag{1}$$

Note that $\xi_f = -\{f, \cdot\}$ which is handy to keep in mind.

The symplectic form on the cotangent bundle is compatible with the coordinates of the latter in the sense that it takes the form $\omega = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ in the basis where $(z^i) = (x^i)$ for $i = 1, \ldots, n$ and $(z^i) = (p_i)$ for $i = n+1, \ldots, 2n$. (Compare $\omega = \mathrm{d}p_i \wedge \mathrm{d}x^i$.) Then the Poisson bracket *takes the form often found in physics books*

$$\{f, g\} \quad = \quad \frac{\partial f}{\partial x^i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial x^i}\frac{\partial f}{\partial p_i}. \tag{2}$$

By the way, I should mention that the use of components such as ξ_f^i instead of the full vector ξ_f which in local coordinates x is given by $\xi_f^i(x)\partial/\partial x^i$ is actually not really coordinate dependent. As pointed out by Penrose, the expression ξ_f^i can be understood merely as a notation which keeps track of the tensorial nature of ξ_f which is that of a vector in this case. Of course this does not apply to the coordinates themselves which prompted Penrose to use a different label e.g. ξ^a vs. x^i. However, modulo such caveats the notation is a powerful way of keeping track of covariance of complicated expressions for objects composed of multiple tensor quantities with derivatives, etc. acting on them. Finally, note that the coordinate basis vector $\partial/\partial x^i \equiv \partial_i$ is written as ∂_a in Penrose's notation. Physicists generally have never hear of this Penrose notation and never distinguish a coordinate index i from an abstract index a.

The prime example of a symplectic manifold is when $M = T^*X$, the cotangent bundle. Then $\omega = d\theta$ where θ is the God-given one-form on T^*X.

Recall that in local coordinates on T^*X $\theta = p_i \mathrm{d}x^i$ where the x^i are local coordinates in the base and the momenta p_i are local coordinates on the fibre of T_x^*X.

Why is this interesting to us in the context of particle motion? These diffeomorphisms give us a relationship, but we want to get from the tangent to the cotangent bundles. So we use the Riemannian metric to get $\mathscr{M} \overset{t_0}{\cong} TX \cong T^*X$. So now the rest of this class will be spent investigating, take the natural structure on T^*X and pull it back by $\mathscr{M}$. So we have a symplectic structure for each t_0 and these could depend on the choice of t_0. So this is breaking the symmetry.

Recall that the $\mathscr{M} \cong TX$ isomorphism comes from the map $x^i(t) \mapsto (x^i(t_0), \dot{x}^i(t_0))$ which takes a solution to Newton II*, that is, a specific particle motion, and maps it to the particle coordinate at $t = t_0$ and its velocity at $t = t_0$. The second isomorphism takes $(x^i, \dot{x}^i) \mapsto (x^i, mg_{ij}(x)\dot{x}^j)$ and, in physics at least, depends explicitly on the mass parameter.*

This brings us to Hamiltonian mechanics. The goal of Hamiltonian mechanics is to encode the symmetries of our phase space into the Lie algebra of smooth functions with the Poisson bracket $(C^\infty(M), \{\cdot, \cdot\})$.

To point out, to be grandiose, where this fits in the grand scheme of physical systems, there's usually a phase space, and another space (of observables). There should be a duality $\sim$ between them, as observables are evaluated on states. In our particular situation in classical mechanics, our state space is our phase space. Our observables are the functions on our phase space. These would be things like momentum and energy that we can assign to a particular particle path.

Strictly speaking, and perhaps shockingly, these are not observables. The problem is that these quantities are coordinate dependent. The (potential) energy, for example is only defined up to *a constant since it appears with a derivative in Newton* II. *Therefore, only* energy differences *are physical. Similarly, the momentum is only defined up to a constant vector and only the relative momentum between us and the particle is physical. This is the reason the observables are required to be* scalar functions *on the phase space; they are not allowed to transform non-trivially under a change of coordinates.*

One would expect that the symmetries of the phase space should translate into symmetries of the symplectic structure. Let me talk about that, and symplectomorphisms. I'm never going to write out symplectomorphism again. I probably spelled it wrong in the first place. I'll call them whatever in the future, unless you want me to call them, like Bob. That might look bad in Gabe's notes.

Weinstein coined the term symplectic, from taking the Greek equivalent for the Latin word for complex. Before it was the Abelian linear group. It sounds like a Victorian word, like perambulator. It's the Greek root for intertwined.

Let (M, ω) be a symplectic manifold. Then $\varphi \in \mathrm{Diff}(M)$ is a symplectomorphism if and only if $\varphi^*\omega = \omega$. Let me give you some examples related to the cotangent space. Since this happened automatically, we might think that any diffeomorphism from a diffeomorphism of the underlying manifold would be a symplectomorphism. That is the case.

If $M = T^*X,\ \omega = d\theta$ and $\varphi \in \mathrm{Diff}(X)$, then

$$\varphi : T^*X \xrightarrow{\sim} T^*X$$

by $(x, p) \mapsto (\varphi(x), (d_x\varphi^{-1}) * p)$. So to check that this is a symplectomorphism, you just check that this preserves θ.

Let us denote the symplectomorphism on the coordinates z^i of M by $\varphi : z^i \mapsto \tilde{z}^i(z)$. In the case of the tangent bundle this gives $(x^i, p_j) \mapsto (\tilde{x}^i, \frac{\partial x^k}{\partial \tilde{x}^j} p_k)$ which is just the statement that p_i is a 1-form. It is then obvious that $\theta = p_i \mathrm{d}x^i$ is invariant.

Let's look at a subclass where $X = \mathbb{E}^d$, so $M = V^* \times \mathbb{E}^d$ and let $\varphi = A \in \mathrm{Euc}(\mathbb{E}^d)$. So for $x \in \mathbb{E}^d$ and $p \in V^*$ then $\Phi(x, p) = (Ax, (dA^{-1})^*p)$ is a symplectomorphism.

In the linear category last time this is analogous to the subgroup, we said $GL(L) \subset Sp(L \oplus L^*)$, and this is the general analogue of this linear statement.

The reason I harped on these examples is because when we talked about particles, there are transformations on the target space. The diffeomorphisms will give us special symplectomorphisms on the phase space.

Now I want to talk about infinitessimal symplectomorphisms. So $\xi \in \mathscr{X}(M)$ is an infinitessimal symplectomorphism if and only if $\mathrm{Lie}(\xi)\omega = 0$. This leads us to a special subset of vector fields $\mathscr{X}_\omega = \{\xi \in \mathscr{X}(M) | \mathrm{Lie}(\xi)\omega = 0\}$. This sits inside $\mathscr{X}(M)$ as a subalgebra, preserving the Lie bracket.

So as long as you stick with diffeomorphisms isotopic to the identity, these are the same requirements.

Maybe it is a good little exercise to show that the linear part of the finite symplectomorphism is the infinitesimal symplectomorphism. That is, check explicitly that $(\varphi^* - 1)\omega \approx \mathrm{Lie}(\xi)\omega$.

Okay, now the symplectic gradient. For any $f \in C^\infty(M)$ I get a vector field ξ_f. I claim that this lives in $\mathscr{X}_\omega(M)$. To see this note that

$$(Lie(\xi_f)\omega) = d \circ \iota(\xi_f)\omega + \iota(\xi_f)d\omega = 0,$$

because $\iota(\xi_f)\omega = df$ and $d\omega = 0$.

So what if I want to look at a particular symmetry. Can I find corresponding obervables? Does every infinitessimal symmetry have a corresponding observable? The answer will depend on H^1. The short answer is no. The long answer brings up the exact sequence

$$f \longrightarrow \xi_f$$

$$0 \longrightarrow H^0_{dR} \longrightarrow \Omega^0(M) \longrightarrow \mathscr{X}_\omega \longrightarrow H^1_{dR} \longrightarrow 0$$

$$\xi \longrightarrow [\iota(\xi)\omega]$$

So if $[\iota(\xi)\omega] \neq 0$ then ξ has no corresponding observables. If $\xi \in \mathscr{X}_\omega$ has an observable, it has many, only unique up to the constants.

This is not what a physicist would call many since it is as small as possible without being trivial. As we will probably see soon, the ambiguity inherent in some potentials can be hugh sometimes involving an infinite number of functions. These ambiguities called gauge invariances *have become one of the central themes in theoretical physics.*

Okay, now time translation. In classical mechanics, there is always a distinguished one-parametery group of time translations. Let's just assume for now that ξ_t is the corresponding infinitessimal generator of time translation, that is, $\iota(\xi)\omega$ is exact.

So we have a choice of corresponding observables. Pick one, up to a constant. Finally we meet the energy. This is the Hamiltonian, which in the sense of these infinitessimal symmetries, is negative the corresponding observable for time translation. In other words, energy, once you take the symplectic gradient, it generates motion which is the negative of time translation. So that is $\{x \mapsto x \circ T_s, s \in \mathbb{R}\}$.

This is an example of our previous observation that $\xi_f = -\{f, \cdot\}$. *The statement that there is a distinguished 1-parameter group of translations is equivalent to the statement that there is a distinguished observable* $f = H$, *the Hamiltonian. Time translation is usually written as* $\xi_t = \frac{\mathrm{d}}{\mathrm{d}t}$ *so that our formula becomes*

$$\frac{\mathrm{d}}{\mathrm{d}t} = -\{H, \cdot\}. \tag{3}$$

Exercise: Hamilton's equations *follow by plugging the coordinates* x^i *and* p_i *into the equation above. Write down Hamilton's equations. Given the relation* $p^i = m\dot{x}^i$*, show that Hamilton's equations are equivalent to Newton* II.

Mathematical Physics
September 13, 2006

Gabriel C. Drummond-Cole

September 13, 2006

Recall that for certain $\xi \in \mathscr{X}_\omega(\mathscr{M})$ (those for which the topological obstruction disappears) there exists a corresponding observable $\mathscr{O}_\xi \in C^\infty(\mathscr{M})$ such that $-d\mathscr{O}_\xi = \iota(\xi)_\omega$.

This describes the infinitessimal symmetry ξ via $\xi f = \{\mathscr{O}_\xi, f\}$ for any $f \in C^\infty(\mathscr{M})$.

We got as far as saying that there are a particular set of symmetries we're concerned with. There are the infinitessimal symmetries of time translation, $\zeta \in \mathscr{X}_\omega(\mathscr{M})$, and this has the observable $\mathscr{O}_\zeta$ where $-\mathscr{O}_\zeta$ is the energy or Hamiltonian. For a path x, $H(x) = \mathfrak{m}2|\dot{x}(t)|^2 + V(x(t))$. For $x \in \mathscr{M}$ this is independent of t.

[Is that obvious?]

Yes, I'll get to it in a second. That's where we left off last time. Any questions?

Before I finish off Hamiltonian dynamics, let me make some tangential but useful remarks about observables. Most of the observables we see in this class will be something like $\mathscr{O}_{(t,f)}$, defined for any time $t \in M^1$ and $f : X \to \mathbb{R}$. Then

$$\mathscr{O}_{(t,f)}(x) = f(x(t)).$$

Let me give another example, two examples.

1. If $X = \mathbb{E}^d$, then we can take $f = x^i$, and in this case $\mathscr{O}_{(t,f)}$ the x^i coordinate of the particle at time t.
2. The Hamiltonian, this is the energy of the particle at time t.

A jet of a function is essentially its Taylor series. The first type of observable depended on the 0-jet of the path; the Hamiltonian depends on the 1-jet. $\mathscr{O}_{(t,f)}$ is local in time, meaning it only depends on finitely many of these, only depends on a small neighborhood of a given time.

So what's the upshot? The structure on $\mathscr{M}$ is as follows. We have what is called a Hamiltonian system. That's the phase space with its symplectic structure and our function H, so the couple $(\mathscr{M}, H)$. So this is a symplectic manifold and a distinguished observable (energy) such that $-\xi_H$ is the infinitessimal time translation.

Let's look at the symmetries of this extended structure. Global symmetries are symplectomorphisms that preserve H. The infinitessimal symmetries are infintessimal symplectomorphisms $\xi \in \mathscr{X}(\mathscr{M})$ such that $\mathrm{Lie}(\xi)H = 0$. Now we're going to look at these symmetries in terms of observables.

Here "extended structure" should not be confused with the (closely related) "extended phase space" which is used in the construction of the Hamiltonian form of the variational principle. Extended phase space is just the Cartesian product of the original phase space (the cotangent bundle T^*X of the configuration space X) with the time $M^1 \cong \mathbb{R}$.

So if Q is an observable that corresponds to an infinitessimal symmetry, then we have the following relation: $\{H, Q\} = 0$. Now, for any observable, never mind that it's a symmetry of any system, time translation flow on phase space is induced by H. So we get that $\dot{\mathscr{O}} = \{H, \mathscr{O}\}$.

This equation deserves its own line:

$$\frac{\mathrm{d}\mathscr{O}}{\mathrm{d}t} = \{H, \mathscr{O}\} \ . \tag{1}$$

In the operator formalism of quantum mechanics this will be the equivalent of the Schrödinger equation.

Exercise: *Show that the classical evolution equation (1) is equivalent to Hamilton's equations.*

Recall that you already showed that Hamilton's equations (with the condition that the momentum is given in terms of the velocity by $p^i = m\dot{x}^i$ or, equivalently, that the Hamiltonian factorizes as $H = \frac{1}{2m}p^2 + V(x)$) are equivalent to Newton's second law. Therefore, once we fix the identification $TX \cong T^*X$, the evolution equation (1) is equivalent to Newton II.

Now we use the observable energy to tell us how things change with time. So now, thus, what we can conclude, assuming that the observable is a symmetry of the Hamiltonian system, for Q, if Q induces a symmetry of the Hamiltonian system, then we havea conservation law. We have that $\dot{Q} = \{H, Q\} = \mathrm{Lie}(-\xi_H)Q = \mathrm{Lie}(\xi_Q)H = 0$.

So look at $\dot{H}$. This is $\{H, H\}$ which is zero. So H is conserved. Such observables, here's more jargon, are called, and this is why I used Q, are called conserved charges.

So here's the $\boxed{\text{big idea}}$, big enough to put in a box. Symmetries imply conservation laws.

Exercise 1 *Compute these conserved charges. The physical situation is the free particle in Euclidean space. We have the huge symmetry group, which is the isometries of $\mathbb{E}^d$.*

Compute the conserved charges for translations and for rotations. These will be momentum and angular momentum.

The term linear momentum *is sometimes used to distinguish these two types of momenta.*

Okay, let's talk about Lagrangian mechanics. For particles we have solutions to Newton's second law, $\mathscr{M} \subset \mathscr{P} = \mathrm{Map}(M^1, X)$. The idea of Lagrangian mechanics is to describe $\mathscr{M}$ as the critical submanifold of a function $S : \mathscr{P} \to \mathbb{R}$.

This function S is called the action, and $\mathscr{M}$ would be paths x such that $\delta S(x) = d_{\mathscr{P}} S(x) = 0$. So δ is the exterior derivative on $\mathscr{P}$.

This is the variational principle: we want the action to be stationary with respect to variations $\delta x^i(t)$ (which form a basis of $H^*(\mathscr{P})$) in the path $x^i(t)$. This philosophy can be motivated in various ways with various degrees of rigor. One such (rigorless) way is the following. In Newtonian mechanics, particle motion tends to minimize the potential energy; if a ball is sitting on an inclined plane it will roll to the bottom. The action principle is the precise embodiment of this intuition.

[Is this why physicists want a path integral?]

That's for quantum mechanics.

The Path Integral and the Principle of Least Action: A preview *Quantum theory introduces a fundamental unit of action $\hbar$ called Planck's constant. The path integral Z is the probability amplitude for a particle at position $\mathbf{x}_i$ at time t_i to be found at a $\mathbf{x}_f$ at a later time t_f. It is given (schematically) by*

$$Z \quad = \quad \int_{\mathbf{x}(t_i)=\mathbf{x}_i}^{\mathbf{x}(t_f)=\mathbf{x}_f} [\mathrm{d}\mathbf{x}(t)] \exp\left(\frac{i}{\hbar} S[\mathbf{x}(t)]\right) \tag{2}$$

The boundary conditions on the path are indicated by the "limits of integration" and $[\mathrm{d}\mathbf{x}(t)]$ is a "measure" on the space of paths $\mathscr{P}$. This formula expresses the fact that the probability of finding a particle at $\mathbf{x}_f$ at time t_f given that it was at $\mathbf{x}_i$ at time t_i is given by a sum over all paths (a.k.a. "histories") with these boundary conditions weighted by a unimodular complex number whose phase is the action ($\hbar = 1$ in natural units). Now consider the classical limit $\hbar \to 0$. *When the action is away from its stationary point, any small deviation in the path causes wild fluctuations in the exponential with "frequency" $\frac{1}{\hbar} \to \infty$. The claim is that these fluctuations average to 0 so that the path integral has, in the classical limit, support only on those paths for which the action is extremal, that is $\delta S = 0$. The principle of least action therefore follows naturally from the quantum principle of "sum over histories".*

So these equations, call these x paths, they satisfy what are called Euler-Lagrange equations. We'll eventually see that these are just Newton's second law. Let me just continue with the philosophical baloney. This sort of variational principle is also found in geometry, where it used to obtain nice PDEs, like the harmonic PDE.

The terms "Newton's second law", "Euler-Lagrange equation" and "Hamilton's equations" are all examples of "equations of motion". The phrase "equation of motion" or EOM *is used interchangeably (and non-commitally) with any of these.*

The Lagrangian approach gives us back our phase space, but it gives us a lot more than that. The symplectic form was borrowed and depended on a time t. In the Lagrangian approach, we'll get, the information embedded in this Lagrangian mechanics, which are the Euler Lagrange equations and the submanifold $\mathscr{M}$, but also a family of one-forms on $\mathscr{M}$ parameterized by time. Finally, these one-forms will give us the symplectic structure naturally, and that won't depend on t.

I've kept you guys ten minutes long, I apologize. But in this sense, physicists equate "theory" with a particular Lagrangian, which has all of this information in it.

[I thought it was the action?]

That's the integral of the Lagrangian, which I think is more basic.

Questions for Thermodynamics

Disclaimer: Some of these problems are clearly more involved than others and some require an understanding of results from prior problems. You have two weeks to complete them.

1. The following questions regard the linear algebra and local "flavor" of closed, non-degenerate two forms.

   (a) Let V be a real $2n$-dimensional vector space and let $\omega \in \bigwedge^2 V^*$ be a non-degenerate two-form. Show that there exists a basis $e^1, \ldots, e^{2n}$ of V^* such that

   $$\omega = e^1 \wedge e^2 + e^3 \wedge e^4 + \cdots + e^{2n-1} \wedge e^{2n}.$$

   Hint: use induction on n.

   (b) Now let $\omega_0, \omega_1 \in \Omega^2(\mathbb{R}^{2n})$ be closed, non-degenerate two forms such that $\omega_0(0) = \omega_1(0)$. Construct a diffeomorphism f defined in some neighborhood of $0 \in \mathbb{R}^{2n}$ such that

   $$f^*\omega_1 = \omega_0.$$

   In particular, consider the one-parameter family of two-forms

   $$\omega_t = \omega_o + t\sigma \quad \text{where} \quad \sigma = \omega_1 - \omega_0.$$

   Show that there exists a one-form β defined in some neighborhood such that

   $$\sigma = d\beta \quad \text{and} \quad \beta(0) = 0.$$

   Now ω_t and β define a time-dependent vector field ξ_t such that $\iota(\xi_t)\omega_t = -\beta$. Show that ξ_t generates a one-parameter family of diffeomorphisms f_t such that $f_0 = id$ and $f_t(0) = 0$ for all t. Compute the derivative

   $$\frac{d}{dt}\left(f_t^*\omega_t\right)$$

   and based on your results show that $f_1^*\omega_1 - \omega_0 = 0$ so that f_1 is the desired diffeomorphism.

2. Assume the results of the first question.

   (a) Use both parts of problem (1) to show that for any closed, non-degenerate two-form $\omega \in \Omega^2(M)$ there exist local coordinates $x^1, \ldots, x^{2n}$ such that

   $$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2n-1} \wedge dx^{2n}.$$

   (b) A two-form $\omega \in \bigwedge^2 V^*$ is said to have *rank* k if, for $k \leq n$,

   $$\omega^k \neq 0 \quad \text{and} \quad \omega^{k+1} = 0.$$

   Use the results from part (a) to show that if $\omega \in \Omega^2(M)$ has constant rank k then there exist local coordinates such that

   $$\omega = dx^1 \wedge dx^2 + \ldots + dx^{2k-1} \wedge dx^{2k}.$$

   This is *Darboux's Theorem* for two-forms of constant rank.

3. Let $\alpha \in \Omega^1(M)$ have constant rank $2k+1$. Note that $d\alpha$ is a two-form of constant rank k. Use this fact and Darboux's theorem to show that there are local coordinates $z, x^1, \ldots, x^k, y^1, \ldots, y^k$ such that

$$\alpha = dz + \sum_{i=1}^{k} x^i dy^i.$$

4. Let $\alpha \in \Omega^1(M)$ have constant rank $2k$. We will show that there are local coordinates $x^1, \ldots, x^k, y^1, \ldots, y^k$ such that

$$\alpha = \sum_{i=1}^{k} x^i dy^i.$$

   We will break the proof into several steps.

   (a) Start with a one-form $\alpha \in \Omega^1(\mathbb{R}^{2k})$ of constant rank $2k$. Show that there exist local coordinates $x^1, \ldots, x^{2k}$ such that

   $$\alpha \wedge (d\alpha)^{k-1} = g dx^2 \wedge dx^3 \wedge \cdots \wedge dx^{2k}$$

   for some nowhere vanishing – in fact, positive – function g. Now define a function f so that $f^k = g$ and define a one-form σ so that $\alpha = f\sigma$. Show that σ has rank $2k-1$.

   (b) Prove that the above still holds when $n \geq k$ and $\alpha \in \Omega^1(\mathbb{R}^{2n})$ still has rank $2k$. Hint: reduce to the case $n = k$.

   (c) Use part (b) to show that if $\alpha \in \Omega^1(M)$ has constant rank $2k$ then we can find a positive function f and a one-form σ of rank $2k-1$ such that

   $$\alpha = f\sigma.$$

   Use this fact and question (3) to find coordinates that verify the claim.

5. Consider a thermodynamic system whose only configurational variable is volume V so that the equilibrium submanifold M is two-dimensional.

   (a) Let Q denote the heat one-form and let p denote pressure. Explain how we can interpret the ratio of two forms as a function

   $$f = \frac{Q \wedge dp}{Q \wedge dV}.$$

   (b) Fix a point $x \in M$ and any adiabatic vector $\xi \in T_x M$. Show that

   $$f(x) = \frac{dp(\xi)}{dV(\xi)} \quad \text{which justifies the expression} \quad f = \left(\frac{dp}{dV}\right)_{\text{adiabatic}}$$

   where the right hand side is just the ratio of one-forms evaluated on some adiabatic vector.

(c) Let T denote temperature. Based on the discussion above state what is meant by

$$\frac{dT \wedge dp}{dT \wedge dV} = \left(\frac{dp}{dV}\right)_{\text{isothermal}}.$$

(d) Since the differentials of V, T and p, T are linearly independent we may write

$$Q = \Lambda_V dV + C_V dT \quad \text{or} \quad Q = \Lambda_p dp + C_p dT$$

where the Λs and Cs are functions on M. (These are called "specific" and "latent" heats, respectively).Show that

$$\left(\frac{dp}{dV}\right)_{\text{adiabatic}} = \gamma \left(\frac{dp}{dV}\right)_{\text{isothermal}}$$

where $\gamma = C_p/C_V$.

(e) An *ideal gas* is one that, in equilibrium, obeys the constraints

$$pV = nT \quad \text{and} \quad \gamma = \text{constant}$$

where n are the *moles* of gas. (A mole is essentially a rescaled count of the number of molecules. It is has no physical units). Use part (d) to show that the adiabatic curves for an ideal gas are given by

$$pV^\gamma = \text{constant}.$$

6. The following questions cover *Legendre transforms*. They play a useful role in certain applications of thermodynamics as we shall see in subsequent questions.

(a) Let $f : V \to \mathbb{R}$ be a smooth function on a real vector space. Choose some basis $x^1, \ldots, x^n$ of linear coordinates and define the *Hessian of* L as

$$\text{Hess} f = \frac{\partial^2 f}{\partial x^i \partial x^j} dx^i \otimes dx^j.$$

Show that the Hessian is independent of the chosen basis and takes values in $Sym^2 V^*$.

(b) Work under the natural identification $T^*V = V \times V^*$ and interpret the differential of f as a map $\phi_f = df : V \to V^*$. What condition must $\text{Hess} f$ satisfy so that ϕ_f is a local diffeomorphism? Explain.

(c) Let $\langle \cdot, \cdot \rangle : V^* \otimes V \to \mathbb{R}$ denote the canonical pairing. Assume ϕ_f is a diffeomorphism. Then we can define the Legendre transform

$$\mathcal{L}f(\alpha) = \langle \alpha, \phi_f^{-1}(\alpha) \rangle - f \circ \phi_f^{-1}(\alpha) \quad \text{for} \quad \alpha \in V^*$$

so that $\mathcal{L}f$ is a function on V^*. Compute the differential of $\mathcal{L}f$ and – under the natural identification $(V^*)^* = V$ – compute the map $\phi_{\mathcal{L}f} = d(\mathcal{L}f) : V^* \to V$.

(d) Now compute $\mathcal{L}\mathcal{L}f : V \to \mathbb{R}$ and conclude that the Legendre transform is involutive, i.e. $\mathcal{L}^2 = id$.

(e) Let M be a smooth manifold and generalize the Legendre transform to functions $f : M \times V \to \mathbb{R}$ by "transforming" along the V directions. Thus you should have $\mathcal{L}f : M \times V^* \to \mathbb{R}$. Further generalize to functions $f : E \to \mathbb{R}$ where $\pi : E \to M$ is a vector bundle over M so that $\mathcal{L}f : E^* \to \mathbb{R}$.

7. Let $\mathbb{V}$ be the 4-dimensional vector space with linear coordinates T, S, p, V, i.e.

$$\mathbb{V} = \mathbb{R}_T \oplus \mathbb{R}_S \oplus \mathbb{R}_p \oplus \mathbb{R}_V$$

where each coordinate has the appropriate physical units. Endow $\mathbb{V}$ with the symplectic form $\omega = dT \wedge dS - dp \wedge dV$.

(a) Use ω to identify $\mathbb{R}_p = (\mathbb{R}_V)^*$. If the internal energy is a function $U \in C^\infty(\mathbb{R}_S \oplus \mathbb{R}_V)$ show that the Legendre transform along the V-direction is the enthalpy

$$\mathcal{L}_V U = U + pV = H \in C^\infty(\mathbb{R}_S \oplus \mathbb{R}_p).$$

The second expression must be properly interpreted. When is change in enthalpy equal to the heat added?

(b) Use ω to identify $\mathbb{R}_T = (\mathbb{R}_S)^*$. Show that

$$\mathcal{L}_S U = U - TS = F_{Helm} \in C^\infty(\mathbb{R}_T \oplus \mathbb{R}_V).$$

This is called the *Helmholtz Free Energy*. When is the change in F_{Helm} equal to the work done by the system?

(c) Now use ω to identify

$$\mathbb{R}_p \oplus \mathbb{R}_T = (\mathbb{R}_S \oplus \mathbb{R}_V)^*$$

and show that

$$\mathcal{L}U = U - TS + PV = F_{Gibbs} \in C^\infty(\mathbb{R}_T \oplus \mathbb{R}_p).$$

This is called the *Gibbs Free Energy*.

Homework set 2: Statistical Thermodynamics

1 Binary systems

We start with a system of N particles each of which can be in one of two states with equal probability. We will refer to the particles as *sites* and the two states as *spins* which can be *up* $\uparrow$ or *down* $\downarrow$. We define $N_\uparrow = \frac{N}{2} + s$ and $N_\downarrow = \frac{N}{2} - s$ to be the number of spins up and down, respectively. The *spin excess* is the quantity $N_\uparrow - N_\downarrow = 2s$. Define the function $g(N,s)$ to be the number of configurations of spin excess $2s$. Compute $g(N,s)$.
Using the Stirling approximation

$$N! = \sqrt{2\pi N} N^N \exp\left\{-N + \frac{1}{12N} + \dots\right\}, \tag{1}$$

show that

$$\log g \approx -\left(N_\uparrow + \frac{1}{2}\right)\log\left(\frac{N_\uparrow}{N}\right) - \left(N_\downarrow + \frac{1}{2}\right)\log\left(\frac{N_\downarrow}{N}\right) + \frac{1}{2}\log\left(\frac{1}{2\pi N}\right) \tag{2}$$

Using a further expansion of logarithms, show the Gaußian distribution

$$g(N,s) \approx g(N,0)\exp\left(-\frac{2s^2}{N}\right) \tag{3}$$

Recall that the entropy of the system in the state ρ is defined to be $S = -k_B\langle \log\rho\rangle_\rho$. Argue, based on the fundamental hypothesis of equal probability of accessible microstates, that if the system is in a state of spin excess $2s$, its entropy is given by $S = -k_B \log g(N,s)$:

$$-\frac{S}{k_B} \approx -\left(\frac{N}{2} + s\right)\log\left(\frac{1}{2} + \frac{s}{N}\right) - \left(\frac{N}{2} - s\right)\log\left(\frac{1}{2} - \frac{s}{N}\right) + \frac{1}{2}\log\left(\frac{1}{2\pi N}\right). \tag{4}$$

2 Paramagnetism

We return to the situation of problem **??**. Suppose each spin has a magnetic dipole moment $\mathbf{m}$ associated to it. Remind yourself from what we learned in electromagnetism that the energy of a magnetic dipole in an external magnetic field $\mathbf{B}$ is $U = -\mathbf{m} \cdot \mathbf{B}$. Let us suppose the magnetic field points "up" with magnitude B and that the total magnetic moment due to spin excess is $2sm$. Using the results of problem **??** argue for the following approximate form for the free energy:

$$\begin{aligned} F(T, s, B) &\approx -2smB + k_B T \left(\frac{N}{2} + s\right) \log\left(\frac{1}{2} + \frac{s}{N}\right) \\ &+ k_B T \left(\frac{N}{2} - s\right) \log\left(\frac{1}{2} - \frac{s}{N}\right) - \frac{1}{2} k_B T \log\left(\frac{1}{2\pi N}\right) \end{aligned} \tag{5}$$

Minimize (extremization will suffice) this free energy with respect to s to show that the expectation value of the spin excess in thermal equilibrium is given by

$$\langle 2s \rangle = N \tanh\left(\frac{mB}{k_B T}\right). \tag{6}$$

The *magnetization* M is defined as the magnetic field per unit volume $M = \langle 2s \rangle m/V$. Plot the magnetization as a function of external magnetic field and temperature. Interpret the results.

3 Ferromagnetism

Consider a system of magnetic moments as in problem **??** but without external magnetic field. In the *mean field approximation*, we assume that each magnetic moment experiences an effective magnetic field due to the other magnets surrounding it. In particular, we assume the field is proportional to the magnetization M (the magnetic moment per unit volume):

$$B_{\text{effective}} = cM \tag{7}$$

for some constant c. Substituting $B_{\text{effective}}$ into the result for the magnetization of problem **??** gives a transcendental equation for M. Note that $M = 0$ is a solution for any T. Investigate the qualitative behavior of the solutions to this equation as a you change T.[1] You should find that there is a special value T_c for which the number of solutions jumps from 2 to 1. What is this value in terms of the magnetic moment of a spin m, the number of spins per unit volume n, and the magnetic "permeability" c? The temperature T_c is called the *critical temperature*.

4 The ideal gas revisited

In class we derived various properties of an ideal gas using the Gibbs partition function. However, since the N particles were non-interacting, we expect that we obtain the same result by taking N copies of the single particle in a box. Recall that for the latter we derived in class the canonical partition function

$$Z_1 = \frac{n_Q}{n} \tag{8}$$

where

$$n_Q = \left(\frac{mk_BT}{2\pi\hbar^2}\right)^{\frac{3}{2}} \tag{9}$$

is the *quantum concentration*, $n = \frac{N}{V}$ is the concentration or particle density and V is the volume of the box. Argue that the partition function for the gas of N particles should be simply

$$Z_N = \frac{(Z_1)^N}{N!}. \tag{10}$$

Using this generating function, show that the internal energy of the gas is

$$U = \frac{3}{2}Nk_BT. \tag{11}$$

Recall from your first homework that the *(specific) heat capacity at constant volume* C_V is the amount of heat energy Q required to raise the temperature of a system by one degree when we keep the volume constant. Show that the first law then implies that $C_V = \frac{3}{2}N$ for an ideal gas.

[1] For example, solve this equation graphically for various values of T between 0 and ∞.

Using Stirling's formula (**??**), show that the Helmholtz free energy derived from the N-particle partition function (**??**) is approximately

$$F = Nk_BT\left\{\log\left(\frac{n}{n_Q}\right) - 1\right\}. \tag{12}$$

Use this to compute the pressure of the gas and show that it implies the equation of state (ideal gas law)

$$pV = Nk_BT. \tag{13}$$

5 Interacting gas: van der Waals' equation of state

The ideal gas is a very good approximation for dilute gasses. However, since the particles are point-like and do not interact, the behavior of the system is essentially the same for all densities n. Van der Waals modified the ideal gas law to include a long-range attractive inter-particle force and a short range repulsive force representing the finite size of the particle. In this problem we will include these effects by modifying the free energy of the ideal gas (**??**).

The short-range repulsive part of the force is modeled by a "hard core": The force is 0 outside a small volume b representing the volume of the particle and infinite inside. This subtracts an amount $-Nb$ from the total volume V, that is, we replace $V \to V - Nb$ everywhere in the expression for the free energy. For the attractive part, we invoke the *mean field approximation*: We consider the potential energy of a particle in the field of all the others in which we take the density n to be constant, that is, we are replacing the local number density $n(\mathbf{x})$ with its average value n. Then the potential energy for the first particle is proportional to the sum of the potential field φ of each of the other particles:

$$\int_b^{V-Nb} \varphi(\mathbf{x})\, n\, \mathrm{d}^3x \approx n\int_b^{V-Nb} \varphi(\mathbf{x})\, \mathrm{d}V = n(-2a) \tag{14}$$

where a is some constant with units of energy times volume. (The factor of 2 is there just for convenience and the sign is the assumption that the inter-particle force is attractive.) In all, we are saying that the contribution to the free energy per particle pair is $-2a$. Approximating $\binom{N}{2} \approx \frac{1}{2}N^2$ write down the Helmholtz free energy of the ideal gas with the two modifications above. You should find

$$F = -Nk_BT\left\{\log\frac{n_Q(V-Nb)}{N} + i\right\} - \frac{N^2a}{V}. \tag{15}$$

From this expression compute the pressure. This gives the *van der Waals equation of state*

$$\left(p + \frac{N^2a}{V^2}\right)(V-Nb) = Nk_bT. \tag{16}$$

Now set $p_c = a/27b^2$, $V_c = 3Nb$, and $T_c = 8a/27b$ and rewrite the v. d. Waals equation of state in terms of dimensionless variables $\hat{p} = p/p_c$, $\hat{V} = V/V_c$, and $\hat{T} = T/T_c$. This defines a family of *isothermal curves* on the $\hat{p}$-$\hat{V}$-plane depending on $\hat{T}$. Sketch these curves for various values of T. (You may also plot them using a graphing tool.) Show (analytically) that there is a unique curve with a horizontal inflection point which is given by $\hat{T} = 1$, $\hat{V} = 1$, and $\hat{p} = 1$.

6 Blackbody radiation

This problem concerns the spectrum of electromagnetic radiation in a cavity (box with conducting walls; like an oven) in thermal equilibrium at temperature T. Since the cavity is conducting, the electric field must satisfy Maxwell's equations as well as the boundary conditions we discussed last semester. Suppose the box is a cube with sides of length L. Then the eigenvalues of the wave operator $\Box$ are

$$c^2\pi^2\mathbf{n}\cdot\mathbf{n} - \omega^2L^2 \tag{17}$$

for $\mathbf{n} = (n_x, n_y, n_z) \in (\mathbb{Z}^+)^3$. Let $n = \sqrt{\mathbf{n}\cdot\mathbf{n}}$ denote the length of this vector. Then the solutions to the wave equation have frequency $\omega_n = n\pi c/L$. According to the rules of quantum mechanics, each such "mode" has energy $\varepsilon_n = \hbar\omega_n$. Using the Boltzmann distribution, write down the canonical partition function for all the modes in the cavity. Write a similar expression for the internal energy U of the system. The resulting expression is valid for a spin-less boson (boson with no internal degrees of freedom). Photons, on the other hand, have 2 physical polarizations (massless spin 1). Additionally, photons do not interact with themselves. Argue from this that we should therefore multiply our result for the energy by a factor of 2.

Now approximate the sum by an integral

$$\sum_{\mathbf{n}\in(\mathbb{Z}^+)^3} \to \frac{1}{8}4\pi \int_0^\infty \mathrm{d}n\, n^2 \tag{18}$$

where the factor of $\frac{1}{8}$ comes from restricting the integral to the positive octant. Compute the energy per unit volume per unit frequency or *spectral density* u_ω defined by the equation

$$\frac{U}{V} = \int_0^\infty u_\omega \,\mathrm{d}\omega \tag{19}$$

with $V = L^3$. You should find *Planck's law of radiation*:

$$u_\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\mathrm{e}^{\hbar\omega/k_B T} - 1}. \tag{20}$$

Plot this distribution as a function of $x = \hbar\omega/k_B T$. Find the (approximate/numerical) value of the frequency of the peak as a function of the temperature. Perform the integral to find the temperature-dependence of the total energy density. Note the characteristic T^4 temperature dependence. This is called the *Stephan-Boltzmann law*.

7 Bose-Einstein condensation

For a Bose-Einstein gas, any energy level can have any occupancy. Use this to show that the free energy is given by

$$F = \frac{1}{\beta} \sum_{i=0}^{\infty} \log\left(1 - \lambda \mathrm{e}^{-\beta\varepsilon_i}\right). \tag{21}$$

where $\lambda = \mathrm{e}^{\beta\mu}$ is the *absolute activity*. In the case of a dilute gas, we proceeded by replacing the sum with an integral over phase space. In doing so we have tacitly assumed that the lowest energy states do not contribute disproportionately to the sum. This is certainly true for temperatures not close to absolute zero, however, let us now revisit this assumption by considering the low-temperature limit.

Suppose there are N particles confined to a region of volume V in the system under consideration. Let us normalize the ground state energy $\varepsilon_0 = 0$ and assume that this state is non-degenerate, that is $\varepsilon_1 > 0$.[2] Taking the appropriate derivatives of the free energy, compute the expectation values of the total number of particles in the system $\langle N \rangle$ and the occupancy of the ground state $\langle N_0 \rangle$. From the expression for $\langle N \rangle$ show that as $T \to 0$ all particles have to be in the ground state $\langle N_0 \rangle \to \langle N \rangle$. Show that to lowest non-trivial order

$$\lambda \approx 1 - \frac{1}{\langle N_0 \rangle} \quad \text{or} \quad \mu \approx -\frac{k_B T}{\langle N_0 \rangle}. \tag{22}$$

(Notice that the chemical potential approaches 0 from below.) As μ is now very small, we can approximate $\varepsilon_i - \mu \approx \varepsilon_i$ for all excited states. Having separated out the lowest level as the only one contributing significantly to the sum, we can now compute $\langle N_{\text{ex}} \rangle := \langle N - N_0 \rangle$ by integration as before. Show that this gives

$$\langle N_{\text{ex}} \rangle = \mathrm{Li}_{3/2}(1) n_Q V \tag{23}$$

where n_Q is the quantum concentration (**??**) and the polylogarithm

$$\mathrm{Li}_s(\lambda) = \sum_{n=1}^{\infty} \frac{\lambda^n}{n^s}. \tag{24}$$

Numerically, $\mathrm{Li}_{3/2}(1) \approx 2.6$. Define the *critical temperature for Bose-Einstein condensation* as the temperature for which the quantum density $n_Q = \frac{n}{\mathrm{Li}_{3/2}(1)}$ (where $n = \frac{N}{V}$):

$$T_c = \frac{2\pi\hbar^2}{k_B m} \left(\frac{n}{\mathrm{Li}_{3/2}(1)} \right)^{\frac{2}{3}}. \tag{25}$$

Conclude that

$$\langle N_0 \rangle = \langle N \rangle \left[1 - \left(\frac{T}{T_c} \right)^{\frac{3}{2}} \right]. \tag{26}$$

Given the typical number of particles in a macroscopic system, what is the significance of this equation? In the case of ^{4}He, plugging the numbers into (**??**) gives $T_c \approx 3.1\,\mathrm{K}$. The actual number is $2.17\,\mathrm{K}$.

---

[2] Recall that the convergence of the geometric sum in the derivation of the Bose-Einstein distribution function requires that $\varepsilon_i - \mu > 0 \,\forall\, i$.

Questions for Quantum Mechanics

Choose one. Presentations will be on Monday, May 7th starting at 10 AM. TeX'd solutions are due Monday, May 14th. Do not hesitate to ask for help from me, William, or each other.

1. Let V be a real vector space with symplectic form $\omega \in \bigwedge^2 V^*$. We define the Heisenberg algebra to be the vector space $\mathfrak{heis}(\omega) = i\mathbb{R} \oplus V$ whose only non-trivial brackets are between elements $v, w \in V$:
$$[v, w] = i\omega(v, w) \in i\mathbb{R}.$$
   (a) Use ω to identify $V = V^*$. This identification gives $i\mathbb{R} \oplus V^*$ a Lie algebra structure. How is this structure related to the Poisson bracket?

   (b) Let $\mathrm{Heis}(\omega)$ denote the Heisenberg group. We can define it to be the exponentiation of the Heisenberg algebra. Thus for each $(i\theta, v) \in \mathfrak{heis}(\omega)$ we have
$$\exp{(i\theta, v)} \in \mathrm{Heis}(\omega).$$
   Use the Baker-Cambell-Hausdorf formula to find the group law. Show that as a set $\mathrm{Heis}(\omega) = \mathbb{T} \times V$ and that as a group it fits into the following short exact sequence
$$1 \to \mathbb{T} \to \mathrm{Heis}(\omega) \to V \to 1.$$
   (c) Let $Sp(\omega)$ be the symplectic group with Lie algebra $\mathfrak{sp}(\omega)$. Use ω to identify
$$\mathrm{Sym}^2 V = \mathfrak{sp}(\omega) = \mathrm{Sym}^2 V^*.$$
   The second identification gives the quadratic forms a Lie algebra structure. How is this structure related to the Poisson bracket?

   (d) Show that the automorphism group of $\mathfrak{heis}(\omega)$ is exactly $Sp(\omega)$. Based on your answers in parts (a) and (c) relate the action of $\mathfrak{sp}(\omega)$ on $\mathfrak{heis}(\omega)$ to the Poisson bracket.

2. Let $\mathcal{S}$ be a complex vector space that supports a representation
$$\rho : \mathfrak{heis}(\omega) \to \mathrm{End}(\mathcal{S})$$
such that $\rho(i\theta)$ is scalar multiplication by $i\theta \in i\mathbb{R}$.

   (a) Extend ρ to a representation of $\mathfrak{sp}(\omega)$: Let $A_{v\cdot w} \in \mathfrak{sp}(\omega)$ denote the element identified with the symmetric product $v \cdot w$ for $v, w \in V$. (See (1c) above). Then set
$$\rho\left(A_{v\cdot w}\right) = \frac{\rho(v)\rho(w) + \rho(w)\rho(v)}{2}.$$
   Show that this is indeed a Lie algebra homomorphism.

   (b) Show that $[\rho(A), \rho(v)] = \rho(Av)$ for all $A \in \mathfrak{sp}(\omega)$ and $v \in V$.

(c) Assume $\mathcal{S}$ has a hermitian inner product and that for each $v \in V$, $\rho(v)$ is a self-adjoint endomorphism. Explain why this is compatible with the requirement above for $\rho(i\theta)$. Show that $\rho(A)$ is also self-adjoint for $A \in \mathfrak{sp}(\omega)$.

3. Let $X \oplus X' = V$ be a real polarization. Take $\mathcal{S}$ to be the Schwartz space of half-forms on X with its natural hermitian inner product.

(a) For $v' \in X'$ we can interpret $\omega(v') = \omega(\cdot, v')$ as an element of X^*. For $\psi \in \mathcal{S}$ define

$$\rho(v)\psi = i\frac{\partial \psi}{\partial v} \quad \text{and} \quad \rho(v')\psi = \omega(v') \cdot \psi \quad \text{for} \quad v \in X, v' \in X'.$$

Show that this is a representation of $\mathfrak{heis}(\omega)$ and that $\rho(w)$ is self-adjoint for each $w \in V$.

(b) Show that we can "exponentiate" ρ to obtain a unitary representation of $\mathrm{Heis}(\omega)$. Write the precise action of $\exp(i\theta, v) \in \mathrm{Heis}(\omega)$ on $\psi \in \mathcal{S}$.

(c) With respect to the polarization we can write

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathfrak{sp}(\omega).$$

Show that $D = -A^*$ and that both C and B are symmetric with respect to ω. Use (2a) above to compute

$$\rho\begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}, \quad \rho\begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}, \quad \rho\begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix}.$$

4. As above, (V, ω) is a symplectic vector space and $\{\cdot, \cdot\}$ is the corresponding Poisson bracket. The Weyl algebra $W(\omega)$ is defined as the associative algebra over $\mathbb{C}$ generated by the real vector space V^* subject to the relations

$$ab - ba = i\{a, b\}, \quad a, b \in V^*.$$

(a) Show that a representation of $\mathfrak{heis}(\omega)$ on which $i\mathbb{R}$ acts by scalar multiplication extends to an representation of $W(\omega)$ on which $\mathbb{C}$ acts by scalar multiplication. Show that if $\mathfrak{heis}(\omega)$ acts irreducibly then so does $W(\omega)$.

(b) Given a real polarization $V = X \oplus X'$ show that $W(\omega)$ is isomorphic to the algebra of differential operators on X with polynomial coefficients.

(c) Let $\hbar \in \mathbb{R}^{>0}$ be the parameter that models the quantum scale. Show that $\mathfrak{heis}(\omega) \cong \mathfrak{heis}(\omega/\hbar)$ and that $Sp(\omega) = Sp(\omega/\hbar)$. Which associative algebra does $W(\omega/\hbar)$ approach in the classical limit?

5. Let $S^2_r \subset \mathbb{R}^3$ be the sphere of radius r. We would like to consider "quantizing" this compact manifold. Let x, y, z be the standard linear coordinates on $\mathbb{R}^3$. The algebra of polynomial functions on S^2_r is

$$P = \mathbb{C}[x, y, z]/(x^2 + y^2 + z^2 - r^2).$$

(a) Compute the standard symplectic form Ω on S^2_r – the one derived from the embedding in Euclidean $\mathbb{R}^3$ – in terms x, y, z.

(b) Let $\{\cdot,\cdot\}$ be the Poisson bracket for Ω. Show that $\{x, y\} = z$ with cyclic permutations holding as well. Conclude that the "linear" functions in P are isomorphic to $\mathfrak{so}_3$.

(c) Based on problem (4) it is natural think of the "Weyl algebra" $W(\Omega)$ as the associative algebra over $\mathbb{C}$ generated by x, y, z subject to the relations given by $xy-yx = iz$ with cyclic permutations and $x^2 + y^2 + z^2 = r^2$. Based on (b) show that

$$W(\Omega) \cong \mathcal{U}(\mathfrak{so}_3)/(X^2 + Y^2 + Z^2 = r^2)$$

where $\mathcal{U}(\mathfrak{so}_3)$ is the universal enveloping algebra for $\mathfrak{so}_3$ and $X, Y, Z \in \mathfrak{so}_3$ are the elements identified with x, y, z respectively.

(d) As in problem (4), when we "quantize" S^2_r we want our Hilbert space to be an irreducible representation of $W(\Omega)$ such that $\mathbb{C} \subset W(\Omega)$ acts by scalar multiplication. On the other hand, standard representation theory tells us that for irreducible representations of $\mathfrak{so}_3$ the $X^2+Y^2+Z^2$ must act as scalar multiplication by $n(n+2)/4$ for some integer n. Based on part (c) what restrictions do we have on the radius to obtain a non-trivial Hilbert space?

6. Consider $V = \mathbb{R}_x \oplus \mathbb{R}_p$ with symplectic form $\omega = dp \wedge dx$.

   (a) Show that the subspace $W = \mathbb{C} \cdot (e_x + ie_p)$ is a positive complex polarization.

   (b) Let $Mp(V)$ denote the metaplectic group. Describe the subgroup $\widetilde{U}(W) \subset Mp(V)$ that preserves W.

   (c) Recall the irreducible unitary representation of $\mathrm{Heis}(\omega)$ on $\mathcal{H} = L^2(\mathbb{R}_x)$ defined in class. Also recall that there is a natural unitary action of $Mp(V)$ on $\mathcal{H}$. Derive the Schrödinger equation

$$i\frac{\partial \psi}{\partial t} = \frac{1}{2}(-\frac{\partial^2}{\partial x^2} + x^2)\psi$$

   from the action of the one-paremter group $\exp(tJ/2) \in \widetilde{U}(W)$ on $\mathcal{H}$. Here

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{with respect to the basis} \quad e_x, e_p.$$

   (Hint: Use the natural isomorphism between representations of $\mathrm{Heis}(\omega)$

$$\mathcal{H} = \mathrm{Sym}^\bullet(W^*) \otimes \mathcal{H}^W$$

   and that $\mathcal{H}^W \otimes \mathcal{H}^W = W$).

7. Let M be a smooth manifold with a symplectic form ω. Assume that there is a one-form α such that $d\alpha = \omega$. Given $\xi \in \mathfrak{X}(M)$ we define a new operator on functions

$$\nabla_\xi = \xi + im_{\alpha(\xi)}$$

where m_f denotes multiplication by the function f.

(a) Show that for $\xi, \eta \in \mathfrak{X}(M)$ we have

$$\nabla_\xi \nabla_\eta - \nabla_\eta \nabla_\xi - \nabla_{[\xi,\eta]} = i\omega(\xi, \eta).$$

(b) Let ξ_g denote the symplectic gradient of $g \in C^\infty(M)$. Define an operator on functions

$$\hat{g} = i\nabla_{\xi_g} + m_g$$

Show that $[\hat{g}, \hat{f}] = i\widehat{\{f, g\}}$.

(c) Let $M = T^*X$ and let α be the canonical one-form and take $\omega = d\alpha$. Let $x^1, \ldots, x^n$ be local coordinates on X and let $p_1, \ldots, p_n$ be the corresponding linear coordinates along the fibers of T^*X. Write explicit formulas for $\widehat{x^i}$ and $\widehat{p_n}$.

(d) Restrict $\widehat{x^i}$ and $\widehat{p_n}$ to the subspace $C^\infty(X) \subset C^\infty(T^*X)$. Write explicit formulas for these restricted operators.

Homework set 4: Quantum Mechanics

Abstract

This homework assignment consists of 4 problems. Each student is required to complete all problems and submit his TEX-ed solutions by email or otherwise by Monday, May 14th. I wrote the problems myself so if you think there may be a typo or mistake in the problem, please let me know.

1 Harmonic Oscillator

This problem concerns the quantum analogue of a mass m on a spring of stiffness k in one dimension. Classically, the restoring force is proportional to the displacement x as $F = -kx$. Defining the angular frequency $\omega = \sqrt{k/m}$ (check the units), write down the Schrödinger equation. Define the annihilation operator

$$\hat{a} = \frac{1}{\sqrt{2m}}(\hat{p} - im\omega\hat{x}) \; . \tag{1}$$

It's Hermitian conjugate $\hat{a}^\dagger$ is called the creation operator. Compute the commutator $[a, a^\dagger]$. Define the Hermitian operator $\hat{n} = \hat{a}^\dagger\hat{a}$ and compute its commutator with the creation and annihilation operators. Write the Hamiltonian in terms of these operators. Let $|n\rangle$ denote the eigenvectors of $\hat{n}$ with eigenvalue n as usual. Show that $\hat{a}^\dagger|n\rangle$ is an eigenvector of $\hat{n}$ with eigenvalue $(n+1)$. Show the analogous statement for $\hat{a}|n\rangle$. Define the *Fock vacuum* $|0\rangle$ such that $a|0\rangle = 0$. This is the unique normalizable state of lowest energy. Check that it has energy $E_0 = \frac{1}{2}\hbar\omega$ called the *ground state energy*, *0-point energy*, or *vacuum energy*.

Solve the position-space Schrödinger equation for the ground state wave function $\psi_0(x) = \langle x|\, 0\rangle$ and normalize it. You should find

$$\psi_0(x) = \frac{1}{\pi^{1/4}\sqrt{x_0}} \exp\left(-\frac{x^2}{2x_0^2}\right) \tag{2}$$

where $x_0 = \sqrt{\hbar/m\omega}$ is the characteristic length scale of the harmonic oscillator with mass m and angular frequency ω.

The *Fock space* is spanned by the states $|n\rangle \propto (\hat{a}^\dagger)^n |0\rangle$. The wave functions can thus be found by repeated differentiation of the ground state wave function $\psi_0(x)$ to be

$$\psi_n(x) = \frac{1}{\sqrt{2^n n!}} H_n\left(\frac{x}{x_0}\right) \psi_0(x) \tag{3}$$

where

$$H_n(z) = (-1)^n \mathrm{e}^{z^2} \frac{\mathrm{d}^n}{\mathrm{d}z^n} \mathrm{e}^{-z^2} \tag{4}$$

are the *Hermite polynomials*.

2 Scattering off a localized source

Consider the following form for a potential in the 1-dimensional Schrödinger equation:

$$V(x) = -\alpha\delta(x) \ . \tag{5}$$

What are the units of α? For $\alpha > 0$ (< 0) this represents the idealized form of a localized impurity at $x = 0$ which attracts (repels) the otherwise free particle for which we are solving the Schrödinger equation. Write the Schrödinger equation in position space and integrate the equation from $[-\epsilon, \epsilon]$. Take the $\epsilon \to 0$ to show that the wave function experiences a discontinuity in it's first derivative across $x = 0$ proportional to its value at $x = 0$:

$$\Delta\psi'(0) = -\frac{2}{x_0}\psi(0) \tag{6}$$

where $x_0 = \hbar^2/m\alpha$ has the units of length.

Bound state Show that there is a unique energy $E < 0$ for which a normalizable solution to the Schrödinger equation exists. Show that the energy of this state is

$$E = -\frac{\alpha}{2x_0} \ . \tag{7}$$

Write down the wave function for this *bound state*.

Scattering states Let us now consider the positive energy solutions to the Schrödinger equation. Away from $x = 0$ the potential vanishes so that the solution takes the form

$$\psi(x) = \begin{cases} A_r \mathrm{e}^{ikx} + A_l \mathrm{e}^{-ikx} & x < 0 \\ B_r \mathrm{e}^{ikx} + B_l \mathrm{e}^{-ikx} & x > 0 \end{cases} \tag{8}$$

where the As and Bs are some complex coefficients and

$$\hbar k = \sqrt{2mE} \ . \tag{9}$$

Argue, by considering the time dependence of the complete wave function $\Psi(x,t)$ that the terms labeled with the subscripts r and l represent right- and left-moving waves respectively.[1] Scattering events are usually such that a wave comes in from the left- or the right-hand side but not both at once. Let us consider as the initial condition a right-moving wave coming in from the left – that is no left-moving incoming wave from the right. This means that we are setting $B_l = 0$. Find the relations between the remaining coefficients from the continuity of the wave function and the jump in the derivative.

We interpret the situation as follows: A right-moving wave comes in from the left with an "intensity" set by the amplitude A_r. It is reflected back to $x = -\infty$ with amplitude A_l and transmitted to $x = +\infty$ with amplitude B_r. Since the wave function is not normalizable, we cannot compute the absolute reflection and transmission probabilities. However, we can compute the *transmission* and *reflection coefficients*

$$T = \frac{|B_r|^2}{|A_r|^2} \ , \ R = \frac{|A_l|^2}{|A_r|^2} \ . \tag{10}$$

Show that

$$T = \frac{1}{1 + (m\alpha^2/2\hbar^2 E)} \ , \ R = \frac{1}{1 + (2\hbar^2 E/m\alpha^2)} \ . \tag{11}$$

Comment on the various limits of these equations. Note that contrary to the case of the bound state wave function, the scattering states do not care about the sign of α. Free particles will scatter identically off of a point-like impurity regardless of whether it is attractive or repulsive. However, the particle can get stuck on the impurity only if the force is attractive. Note that the coefficients are singular precisely at the bound state energy (7).

For a general localized potential the analogous story is of course much more complicated. However the basic lesson that we can infer from the calculation

[1]These should be called partial waves. Recall that the positive-energy solutions for definite k are not normalizable and are therefore not physical states. However, the physical wave functions are normalizable superpositions of these "partial wave functions" called wave packets.

above is that the Schrödinger equation gives us the coefficients A_l and B_r in terms of A_r and B_l. That is, there exists a unitary *scattering matrix* $\mathbf{S}$ such that

$$\begin{pmatrix} A_l \\ B_r \end{pmatrix} = \mathbf{S} \begin{pmatrix} A_r \\ B_l \end{pmatrix} \ . \tag{12}$$

In the example above, we also saw that the S-matrix had a pole at the bound state energy. This is a generic phenomenon implying that finding the analytic form of the S-matrix is the same thing as solving the Schrödinger equation.

3 Angular momentum and spin

Angular momentum in three dimensions is given by the *pseudo-vector*

$$\mathbf{L} = \mathbf{x} \times \mathbf{p} \ . \tag{13}$$

Write the corresponding operator in the position space representation and show that its components satisfy the commutation relations

$$[\hat{L}^i, \hat{L}^j] = i\hbar\, \epsilon^{ijk} \hat{L}^k \tag{14}$$

where repeated indices are summed. This indicates that the 3 operators $\{\hat{L}^i\}$ cannot be simultaneously diagonalized. We work in a basis in which the z-component of $\hat{\mathbf{L}}$ is diagonal. States can therefore be labeled by the eigenvalue of $\hat{L}_z$ but this does not suffice to identify the state.

Prove that on the other hand the square of the angular momentum generator (not to be confused with the y-component!) is a scalar, i.e. rotationally invariant:

$$[\hat{L}^2, \hat{L}_z] = 0 \ . \tag{15}$$

We will henceforth drop the $\hat{}$ notation for operators. Let $|\lambda, \mu\rangle$ denote a simultaneous eigenvector with L^2 eigenvalue $\lambda\hbar^2$ and L_z eigenvalue $\mu\hbar$:

$$L^2 \, |\lambda, \mu\rangle = \lambda\hbar^2 \, |\lambda, \mu\rangle \ \ , \ L_z \, |\lambda, \mu\rangle = \mu\hbar \, |\lambda, \mu\rangle \ \ . \tag{16}$$

Define the operators $L_\pm = L_x \pm i L_y$ and show that

$$[L_z, L_\pm] = \pm\hbar L_\pm \ . \tag{17}$$

Use this to show that the vector $L_\pm \, |\lambda, \mu\rangle$ has L_z eigenvalue $(\mu \pm 1)\hbar$.

Thus for every λ we get a tower of states with different L_z eigenvalues. Since the L_z component of the angular momentum is bounded by the total angular momentum of the state there is a lowest state $|\lambda, \mu_-\rangle$ and highest state $|\lambda, \mu_+\rangle$ such that

$$L_\pm |\lambda, \mu_\pm\rangle = 0 \ . \tag{18}$$

Show that

$$L^2 = L_\pm L_\mp + L_z^2 \mp \hbar L_z \ , \tag{19}$$

and use this result on the highest state to obtain $\lambda = \mu_+(\mu_+ + 1)$. Similarly, the lowest state gives $\lambda = \mu_-(\mu_- - 1)$. Use these results to argue that $\mu_- = -\mu_+$.

The L_z eigenvalue of a state with L^2 eigenvalue equal to λ goes under the action of L_+ from $-\mu_+$ to μ_+ in n integer steps for some n. Therefore, $\mu_+ = \frac{n}{2}$ is either integer or half-integer. Switching to more standard notation, we take $\mu_+ = \ell$ and label the L_z eigenvalues by m:

$$L^2 |\ell, m\rangle = \hbar^2 \ell(\ell+1) |\ell, m\rangle \ , \ L_z |\ell, m\rangle = \hbar m |\ell, m\rangle \ . \tag{20}$$

Normalization of the state gives

$$L_\pm |\ell, m\rangle = \hbar\sqrt{\ell(\ell+1) - m(m \pm 1)}\, |\ell, m \pm 1\rangle \ . \tag{21}$$

Note that this means that $L_\pm |\ell, \pm\ell\rangle = 0$.

The value ℓ of the state is called the *angular momentum* or *spin*. In the algebraic determination of the allowed values we find that 2ℓ is an integer. When realized on a function space, however, only the states with integer ℓ are recovered. This is the case of *orbital angular momentum* with generator (13). The algebraic realization implies that we can consider also *internal angular momentum* which is what is usually meant by the term *spin*. To distinguish these cases, we reserve the symbol $\mathbf{L}$ for the orbital part and write $\mathbf{S}$ for the internal part. We also change the notation $\ell \to s$ but keep m.

Show that a matrix realization of the angular momentum operators is $\hat{S}^i = \frac{\hbar}{2}\sigma^i$ where σ^i are the *Pauli spin matrices*

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ , \ \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ , \ \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{22}$$

Argue that the spin up and down states are represented by

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \ , \ |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ . \tag{23}$$

Show that

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ , \ S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \ , \tag{24}$$

are nilpotent of order 2 and map $|\downarrow\rangle \leftrightarrow |\uparrow\rangle$. Compute $\hat{S}^2$ on these states. What is their spin?

4 Identical Particles

Consider a system consisting of two particles 1 and 2 at $\mathbf{r}_1$ and $\mathbf{r}_2$. If the 1-particle states are denoted by $\psi_a(\mathbf{r}_1)$ and $\psi_b(\mathbf{r}_2)$, the 2-particle state for distinguishable particles would simply be the product

$$\psi_d(\mathbf{r}_1, \mathbf{r}_2) = \psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) \ . \tag{25}$$

Now consider the case when particle 1 is identical to particle 2, for example, both are electrons. In fundamental theory any electron is necessarily identical to any other electron: it is labeled by its mass m, spin s and that's it.[2] Contrary to classical particles which can be colored or continuously tracked, painting a quantum particle or observing it continuously would necessarily change the particle's state. Therefore, 1-particle quantum states are necessarily indistinguishable *in principle.* In quantum mechanics such states are constructed by not committing to which particle 1 or 2 is in state ψ_a or ψ_b:

$$\psi_\phi(\mathbf{r}_1, \mathbf{r}_2) = A\left[\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \mathrm{e}^{i\phi}\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2)\right] \ . \tag{26}$$

Claim: Of these only the symmetric and anti-symmetric states occur in nature. The symmetric states describe identical *bosons* and the anti-symmetric ones describe identical *fermions.*[3] It follows from anti-symmetry of the wave function that identical fermions obey the *Pauli exclusion principle*: The amplitude of finding two identical fermions in the same state $\psi_a = \psi_b$ is 0.

Let ψ_d denote the 2 distinguishable particle state (25). Denote by $\psi_\pm$ the bosonic and fermionic identical particle states (26) with $\phi = 0$ and $\phi = \pi$, respectively, normalized by the choice $A = 1/\sqrt{2}$. Denote the expectation value of the square of the separation $\Delta = \langle(\mathbf{r}_1 - \mathbf{r}_2)^2\rangle$ for the distinguishable (Δ_d), bosonic (Δ_+), and fermionic (Δ_-) cases. Write the result as

$$\Delta_\pm = \Delta_d \mp X \tag{27}$$

and give a formula for X. Show that if the single particle wave functions overlap, $X > 0$. This X term is called an *exchange force.* Although it is not really a force in any conventional sense, explain the name paying particular attention to the differing consequences for bosons and fermions.

[2]Actually, that is not it: It should also be labeled by lepton number, etc. Part of the job of a fundamental theory is to discover the exhaustive list of these "quantum numbers". This however, does not change the point that any two 1-particle states are identical.

[3]In relativistic quantum mechanics one can go on to prove that all particles with integer spin are bosons while particles with half-integer spin are fermions.

Hi guys,

Gabriel told me he won't TeX up his problem because he thinks there's something wrong with his face. So instead of a nicely typed solution you get this vaguely informative note from me. Have a nice summer.

Jerry

Problem 1. *Let $\mathcal{S}$ be a vector space with a representation*

$$\rho : \mathrm{heis}(\omega) \to \mathrm{End}(\mathcal{S})$$

such that $\rho(i\theta)$ acts by scalar multiplication.

(Note that by Problem 1, I mean problem 2.)

First we establish and identity we will need in the sequel.

$$\begin{aligned}
\rho(v)\rho(w) - \rho(w)\rho(v) &= [\rho(v), \rho(w)] \\
&= \rho([v, w]) \\
&= \rho(i\omega(v, w)) \\
&= i\omega(v, w) \cdot 1_{\mathcal{S}}
\end{aligned}$$

(a) Extend ρ to a represenation of $\mathrm{sp}(\omega)$ by

$$\rho(A_{v\cdot w}) = \frac{1}{2}(\rho(v)\rho(w) + \rho(w)\rho(v))$$

We check that the new ρ is indeed a Lie algebra homomorphism. Let $\{p_i, q_i\}$ be symplectic coordinates on V. It suffices to check that the Lie bracket is preserved for symplectic matrices corresponding to symmetric pairs of basis vectors. Two cases are easy:

$$\rho[A_{p_i\cdot p_j}, A_{p_k\cdot p_l}] = [\rho A_{p_i\cdot p_j}, \rho A_{p_k\cdot p_l}] = 0$$

and

$$\rho[A_{q_i\cdot q_j}, A_{q_k\cdot q_l}] = [\rho A_{q_i\cdot q_j}, \rho A_{q_k\cdot q_l}] = 0$$

The other cases are a real pain — I can't get them to come out right, so I'm skipping them.

(b) Show that $[\rho(A), \rho(v)] = \rho(Av)$ for $A \in \mathrm{sp}(\omega)$ and $v \in V$. First observe that $A = A_{x\cdot y}$ for some x and y in V. We compute:

$$\begin{aligned}
[\rho(A), \rho(v)] &= [\frac{\rho(x)\rho(y) + \rho(y)\rho(x)}{2}, \rho(v)] \\
&= \frac{1}{2}(\rho(x)\rho(y)\rho(v) + \rho(y)\rho(x)\rho(v) - \rho(v)\rho(x)\rho(y) - \rho(v)\rho(y)\rho(x)) \\
&= \frac{1}{2}(\rho(x)\rho(v)\rho(y) + i\omega(y, v)\rho(x) + \rho(y)\rho(v)\rho(x) + i\omega(x, v)\rho(y) \\
&\quad -\rho(x)\rho(v)\rho(y) + i\omega(x, v)\rho(y) - \rho(y)\rho(v)\rho(x) + i\omega(v, y)\rho(x)) \\
&= \rho(x)i\omega(y, v) + \rho(y)i\omega(x, v) \\
&= \rho(A_{x\cdot y}v)
\end{aligned}$$

(c) Let $\mathcal{S}$ have a Hermitian inner product such that $\rho(v)$ is self-adjoint for all v. Show that this is compatible with the condition that $i\theta$ acts by scalar multiplication and show that $\rho(A)$ is self-adjoint.

Since, $\rho([v, w]) = \rho(i\omega(v, w))$ is skew-adjoint for all v and w, we must check that $[\rho(v), \rho(w)]$ is skew-adjoint when $\rho(v)$ and $\rho(w)$ are self-adjoint. Indeed:

$$\begin{aligned}
[\rho(v), \rho(w)]^\dagger &= (\rho(v)\rho(w) - \rho(w)\rho(v))^\dagger \\
&= (\rho(v)\rho(w))^\dagger - (\rho(w)\rho(v))^\dagger \\
&= \rho(w)^\dagger \rho(v)^\dagger - \rho(v)^\dagger \rho(w)^\dagger \\
&= \rho(w)\rho(v) - \rho(v)\rho(w) \\
&= -[\rho(v), \rho(w)]
\end{aligned}$$

Yet another computation shows:

$$\begin{aligned}
\rho(A_{v\cdot w})^\dagger &= \frac{1}{2}(\rho(v)\rho(w) + \rho(w)\rho(v))^\dagger \\
&= \frac{1}{2}((\rho(v)\rho(w))^\dagger + (\rho(w)\rho(v))^\dagger) \\
&= \frac{1}{2}(\rho(w)^\dagger \rho(v)^\dagger + \rho(v)^\dagger \rho(w)^\dagger) \\
&= \frac{1}{2}(\rho(w)\rho(v) + \rho(v)\rho(w)) \\
&= \rho(A_{v\cdot w})
\end{aligned}$$

Hoo-ray.

Joseph Walsh
May 14, 2007
Jerry's QM Homework
Question 3:

Proof. (a) ρ is defined for all $w \in V$ if we extend by linearity: $w = x + x'$ for some $x \in X$, $x' \in X'$, and we define $\rho(w) = \rho(x) + \rho(x')$. To extend the domain of ρ to all of $\mathfrak{heis}(\omega)$, we define $\rho(i\theta)\psi = i\theta\psi$ for all $i\theta \in i\mathbb{R}$ and all $\psi \in \mathcal{S}$. To show that ρ is a Lie algebra homomorphism, we must show that it preserves brackets. Clearly $\rho(i\theta)$ commutes with $\rho(w)$ for all $w \in V$. It is also obvious that $[\rho(x_1), \rho(x_2)] = [\rho(x_1'), \rho(x_2')] = 0 = \rho(i\omega(x_1, x_2)) = \rho(i\omega(x_1', x_2'))$ for all $x_1, x_2 \in X$ and $x_1', x_2' \in X'$. For $x \in X$, $x' \in X'$,

$$\begin{aligned}
[\rho(x), \rho(x')]\psi &= i\frac{\partial}{\partial x}(\omega(x')\psi) - i\omega(x')\frac{\partial}{\partial x}\psi \\
&= i\frac{\partial}{\partial x}(\omega(\cdot, x'))\psi \\
&= i\omega(x, x')\psi \\
&= \rho(i\omega(x, x'))\psi
\end{aligned}$$

for all $\psi \in \mathcal{S}$. Here, we used the fact that $\omega(\cdot, x')$ is a linear functional on X, and so is its own derivative. Interpreting $\frac{\partial}{\partial x}$ as a directional derivative, we obtained $\frac{\partial}{\partial x}\omega(\cdot, x') = \omega(x, x')$. Thus, for all $w_1, w_2 \in V$, if $w_1 = x_1 + x_1'$, $w_2 = x_2 + x_2'$, then

$$\begin{aligned}
[\rho(w_1), \rho(w_2)] &= [\rho(x_1) + \rho(x_1'), \rho(x_2) + \rho(x_2')] \\
&= [\rho(x_1), \rho(x_2)] + [\rho(x_1), \rho(x_2')] + [\rho(x_1'), \rho(x_2)] + [\rho(x_1'), \rho(x_2')] \\
&= \rho(i\omega(x_1, x_2)) + \rho(i\omega(x_1, x_2')) + \rho(i\omega(x_1', x_2)) + \rho(i\omega(x_1', x_2')) \\
&= \rho(i\omega(x_1 + x_1', x_2 + x_2')) \\
&= \rho(i\omega(w_1, w_2))
\end{aligned}$$

So indeed, ρ preserves brackets.

The hermitian inner product on $\mathcal{S}$ is given by $\langle \psi, \phi \rangle = \int_X \psi\bar{\phi}$. So show that $\rho(w)$ is self-adjoint for all $w \in V$, it suffices to show that $\rho(x)$ and $\rho(x')$ are self-adjoint for all $x \in X$ and $x' \in X'$. Indeed, if we know this, then if we decompose $w = x + x'$,

$$\begin{aligned}
\langle \rho(w)\psi, \phi \rangle &= \langle (\rho(x) + \rho(x'))\psi, \phi \rangle = \langle \rho(x)\psi, \phi \rangle + \langle \rho(x')\psi, \phi \rangle \\
&= \langle \psi, \rho(x)\phi \rangle + \langle \psi, \rho(x')\phi \rangle = \langle \psi, (\rho(x) + \rho(x'))\phi \rangle = \langle \psi, \rho(w)\phi \rangle
\end{aligned}$$

As $\rho(x')$ is multiplication by a real-valued functional $\omega(\cdot, x')$ for all $x' \in X'$, $\rho(x')$ is obviously self-adjoint. Thus, it suffices to prove that $\rho(x)$ is self-adjoint for all $x \in X$. To do this, we show that $\langle \rho(x)\psi, \phi \rangle - \langle \psi, \rho(x)\phi \rangle = 0$.

$$\begin{aligned}
\langle \rho(x)\psi, \phi \rangle - \langle \psi, \rho(x)\phi \rangle &= \int_X i\frac{\partial\psi}{\partial x}\bar{\phi} - \psi\overline{i\frac{\partial\phi}{\partial x}} \\
&= i\int_X \frac{\partial\psi}{\partial x}\bar{\phi} + \psi\frac{\partial\bar{\phi}}{\partial x} \\
&= i\int_X \frac{\partial}{\partial x}(\psi\bar{\phi}) \\
&= i\int_X (d\iota_x + \iota_x d)(\psi\bar{\phi})
\end{aligned}$$

Here we interpreted $\frac{\partial}{\partial x}$ as a Lie derivative with respect to the constant vector field that has value x at every point and applied Cartan's formula. Since $\psi\bar{\phi}$ is a top form, we get $\langle \rho(x)\psi, \phi \rangle - \langle \psi, \rho(x)\phi \rangle = i\int_X d\iota_x(\psi\bar{\phi}) = 0$, as we are integrating an exact form on a boundaryless space.

(b) Exponentiating $\rho(i\theta)$ for $(i\theta, 0) \in \mathfrak{heis}(\omega)$ yields a unitary operator, namely multiplication by $e^{i\theta}$. To obtain a unitary operator from exponentiating an element $x \in X$ or $x' \in X'$, we first must multiply by i. Indeed, $\exp(i\rho(x'))$ is multiplication by the function $\exp(i\omega(\cdot, x'))$, which is a unitary operation since ω is real-valued. $\exp i\rho(x)$ acts on ψ by translation by $-x$, which again is clearly a unitary operation. For $w \in V$, let $x \in X$, $x' \in X'$ be such that $w = x + x'$. Since $i\theta$ commutes with w, $\exp(i\theta, w)$ acts on $\psi \in \mathcal{S}$ by $e^{i\theta}\exp(i(\rho(x) + \rho(x')))\psi$. By the Baker-Campbell-Hausdorff formula,

$$\begin{aligned}
\exp(i\rho(x'))\exp(i\rho(x)) &= \exp\left(i\rho(x') + i\rho(x) + \frac{1}{2}[i\rho(x'), i\rho(x)]\right) \\
&= \exp\left(i\rho(x') + i\rho(x) - \frac{1}{2}[\rho(x'), \rho(x)]\right) \\
&= \exp\left(i\rho(x') + i\rho(x) - \frac{i}{2}\omega(x', x)\right) \\
&= \exp\left(i\rho(x') + i\rho(x)\right) e^{-i\omega(x',x)/2}.
\end{aligned}$$

(All the higher terms in the sequence vanish since they involve commutators with $i\omega(x', x)$.) Therefore, $\exp(i\theta, x + x')$ acts on $\psi(t)$ by $e^{i\theta}\exp(i(\rho(x) + \rho(x')))\psi(t) = e^{i\theta}e^{i\omega(x',x)/2}\exp(i\omega(t, x'))\psi(t - x)$.

(c) Let $\{e_i\}_{i=1}^n$ be a basis for X, and let $\{f_i\}_{i=1}^n$ be the basis for X' corresponding via ω to the dual basis of $\{e_i\}_{i=1}^n$. With respect to this basis for V, ω has the block matrix form, $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Since $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ is in $\mathfrak{sp}(\omega)$, it is subject to the relation

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \left(\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right)^t.$$

Working out the multiplication yields

$$\begin{pmatrix} B & -A \\ D & -C \end{pmatrix} = \begin{pmatrix} B^t & D^t \\ -A^t & -C^t \end{pmatrix}.$$

Therefore, B and C are symmetric, and $D = -A^t = -A^*$ as desired. If $F \in \mathfrak{sp}(\omega)$, we associate the element $\frac{1}{2}\sum_{u_i}\sum_{v_j}\omega(u_i, Fv_j)u_i \otimes v_j \in Sym^2V$ with F, where u_i and v_j both run over the entire basis $\{e_k, f_k | 1 \le k \le n\}$ of V. If $F = \begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$, and $B = (b_{ij})$ with respect to the bases $\{f_j\}$ for X' and $\{e_i\}$ for X, then $\omega(f_i, Ff_j) = -b_{ij}$, and $\omega(\cdot, F\cdot) = 0$ for all other combinations of basis vectors of V. Thus, $\begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$ is associated to the polynomial $\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n -b_{ij}f_i \otimes f_j = \sum_{i=1}^n\sum_{j=i}^n -b_{ij}\frac{f_i\otimes f_j + f_j \otimes f_i}{2}$. Thus, using problem 2a and the fact that $\rho(f_i)$ and $\rho(f_j)$ commute yields $\rho\begin{pmatrix} 0 & B \\ 0 & 0 \end{pmatrix}$ is multiplication by $\sum_{i=1}^n\sum_{j=i}^n -b_{ij}\omega(\cdot, f_i)\omega(\cdot, f_j)$. Similarly, $\rho\begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix}$ is the application of the operator $\sum_{i=1}^n\sum_{j=i}^n -c_{ij}\frac{\partial}{\partial e_i}\frac{\partial}{\partial e_j}$.

Now if $F = \begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix}$, and $A = (a_{ij})$ with respect to the basis $\{e_i\}$ of X, then $\omega(e_i, Fe_j) = \omega(f_i, Ff_j) = 0$ for all i, j. Also, $\omega(e_i, Ff_j) = \omega(e_i, -a_{jk}f_k) = -a_{ji} = -\omega(f_j, a_{ki}e_k) = \omega(f_j, Fe_i)$. So the symmetric polynomial associated to F is $\sum_{i=1}^n \sum_{i=1}^n -a_{ji}\left(\frac{e_i \otimes f_j + f_j \otimes e_i}{2}\right)$. Therefore,

$$\begin{aligned}
\rho\begin{pmatrix} A & 0 \\ 0 & -A^* \end{pmatrix} &= \sum_{i=1}^n \sum_{j=1}^n \frac{-a_{ij}\sqrt{-1}}{2}\left(\frac{\partial}{\partial e_i}\omega(f_j) + \omega(f_j)\frac{\partial}{\partial e_i}\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \frac{-a_{ij}\sqrt{-1}}{2}\left(2\omega(f_j)\frac{\partial}{\partial e_i} + [\frac{\partial}{\partial e_i}, \omega(f_j)]\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \frac{-a_{ij}\sqrt{-1}}{2}\left(2\omega(f_j)\frac{\partial}{\partial e_i} + \omega(e_i, f_j)\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n \frac{-a_{ij}\sqrt{-1}}{2}\left(2\omega(f_j)\frac{\partial}{\partial e_i} + \delta_{ij}\right) \\
&= \sum_{i=1}^n \sum_{j=1}^n -a_{ij}\sqrt{-1}\left(\omega(f_j)\frac{\partial}{\partial e_i}\right) + \sum_{i=1}^n \frac{-a_{ii}\sqrt{-1}}{2} \\
&= \left(\sum_{i=1}^n \sum_{j=1}^n -a_{ij}\sqrt{-1}\omega(f_j)\frac{\partial}{\partial e_i}\right) - \frac{\sqrt{-1}}{2}tr(A)
\end{aligned}$$

□

4. (V,ω) is a symplectic vector space and $\{\cdot,\cdot\}$ is the corresponding Poisson bracket. The Weyl algebra $W(\omega)$ is defined as the associative algebra over $\mathbb{C}$ generated by the real vector space V^* subject to the relations
$$ab - ba = i\{a,b\}, \qquad a,b \in V^*.$$
so that
$$W(\omega) = T(\mathbb{C}\otimes V^*)/(ab - ba - i\{a,b\})$$

(a) Show that a representation of $\mathfrak{heis}(\omega)$ on which $i\mathbb{R}$ acts by scalar multiplication extends to a representation of $W(\omega)$ on which $\mathbb{C}$ acts by scalar multiplication. Show that if $\mathfrak{heis}(\omega)$ acts irreducibly then so does $W(\omega)$.

Consider the isomorphim $V \overset{\omega}{\simeq} V^*$, defined by $a = \omega(\cdot, v_a)$, for some $v_a \in V$. Under the isomorphism we can identify $\mathfrak{heis}(\omega)$ with $i\mathbb{R}\oplus V^*$, where the bracket is given by,
$$[i\theta_a + a, i\theta_b + b] = i\{a,b\} \tag{1}$$
where we identify $V^* \subset C^\infty(V)$ as linear functionals.
Then we consider a representation of $\mathfrak{heis}(\omega)$,
$$\rho : i\mathbb{R}\oplus V^* \to \mathrm{End}(S)$$
for some vector space S, assuming that $i\mathbb{R}$ acts by scalar multiplication. So, we have for any $(i\theta_a + a), (i\theta_b + b) \in \mathfrak{heis}(\omega)$,
$$\rho([i\theta_a + a, i\theta_b + b]) = [\rho(i\theta_a + a), \rho(i\theta_b + b)]$$
by property of ρ being a Lie-Algebra homomorphism.
Now we extend ρ to $W(\omega)$, thinking of $V^* \subset W(\omega)$. Given some basis of V^*, $\{a_i\}$, then $\{1\otimes a_i\}$ generates $T(\mathbb{C}\otimes V^*)$, so that it's sufficient to define a representation, $\tilde{\rho} : W(\omega) \to \mathrm{End}(S)$ on the generators defined by,
$$\tilde{\rho}(1\otimes a_i) = \rho(a_i) \tag{2}$$
and
$$\begin{aligned}\tilde{\rho}(a_{k_1}\cdots a_{k_n}) &= \tilde{\rho}(a_{k_1})\cdots\tilde{\rho}(a_{k_n}) \\ &= \rho(a_{k_1})\cdots\rho(a_{k_n})\end{aligned} \tag{3}$$
and then extend linearly, since ρ is an algebra homomorphism.
<u>Claim:</u> $\tilde{\rho}$ is a representation of $W(\omega)$.
Pf.: We must show that $\tilde{\rho}$ is well defined (since $W(\omega)$ consists of equivalence classes) and an associative Lie Algebra homomorphism.
By (**??**) and the linear extension, it is a Lie Algebra homomorphism and by (**??**) it preserves associativity, making $\tilde{\rho}$ an associative Lie Algebra homomorphism. Thus, we have left to show well-definedness.
It is sufficient to show that $\tilde{\rho}$ is well-defined on generators, so consider a,b generators of $W(\omega)$.
Need: $\tilde{\rho}(ab) \overset{?}{=} \tilde{\rho}(ba + i\{a,b\})$:
$$\begin{aligned}\tilde{\rho}(ab) &= \rho(a)\rho(b), \quad \text{(by (\textbf{??}))} \\ &= \rho(b)\rho(a) + \rho([a,b]), \quad \text{(since } \rho \text{ is a Lie Algebra homom.)} \\ &= \tilde{\rho}(ba) + i\rho(\{a,b\}), \quad \text{(by (\textbf{??}))} \\ &= \tilde{\rho}(ba + i\{a,b\}) \quad \text{(linearity).} \quad \square\end{aligned}$$

Claim: $\mathbb{C}$ acts by scalar multiplication **Pf.:** Consider any 0-tensor, $\lambda \otimes 1 \in W(\omega)$.
Then by property of tensor and linearity of rho_W,

$$\begin{aligned}\tilde{\rho}(\lambda \otimes 1) &= \lambda\tilde{\rho}(1 \otimes 1), \quad \text{(property of tensor and linearity of } \tilde{\rho}) \\ &= \lambda\rho(1), \quad (\mathbf{??}) \\ &= \lambda \cdot \mathrm{id}, \quad (\rho \text{ is homom.}) \quad \square\end{aligned}$$

Claim: $\tilde{\rho}$ acts irreducibly if ρ acts irreducibly
Pf.: Consider any $T \subseteq S$ and assume that it's invariant under action by $\tilde{\rho}$.
Want to show that $T = \phi$ or $T = S$.
But an isomorphic copy of $\mathbb{C} \otimes V^*$ is contained in $W(\omega)$ is the form of 1-tensors, so that

$$\tilde{\rho}|_{\{1-tensors\}} = \rho$$

Thus, irreducibility of ρ implies that $T = \phi$ or $T = S$. $\square$

(b) Given a real polarization $V = X \oplus X'$ show that $W(\omega)$ is isomorphic to the algebra of differential operators on X with polynomial coefficients, D.
Take linear coordinates $\{x^i, p_j\}_{0\leq i,j\leq n}$ on $V^* = X^* \oplus (X')^*$, then they generate $W(\omega)$ as an algebra.
D is generated by $\{x^i, \frac{\partial}{\partial x^j}\}_{0\leq i,j\leq n}$ as an algebra.
Define $\varphi : W(\omega) \to D$ on generators by,

$$\begin{aligned}\varphi(x^j) &= x^j \\ \varphi(p_j) &= -i\frac{\partial}{\partial x^j}\end{aligned} \tag{4}$$

Extending **??** with associative product and linearly and using multi-index notation,

$$\varphi(\lambda x^{\vec{k}} p_{\vec{l}}) = \lambda i^{|\vec{l}|} x^{\vec{k}} \left(\frac{\partial}{\partial x}\right)^{\vec{l}}$$

Thus, φ is a linear associative algebra homomorphism.
It is bijective, because it maps generators to generators surjectively and injectively.
It remains to show that φ preserves the Lie-bracket. It is sufficient to do so for generators.
Since V has a real polarization, we have in $W(\omega)$ that the only non-trivial bracket is,

$$[x^j, p_k] = -i\delta^j_k$$

Similarly, in D, both the x^i's and $\frac{\partial}{\partial x^j}$'s commute, respectively. Thus, the only non-trivial bracket is $[x^j, \frac{\partial}{\partial x^k}]$. For any $f \in C^\infty$,

$$\begin{aligned}[x^j, \frac{\partial}{\partial x^k}] &= (x^j, \frac{\partial}{\partial x^k})f \\ &= x^j\frac{\partial}{\partial x^k}f - \frac{\partial}{\partial x^k}(x^j f) \\ &= x^j\frac{\partial f}{\partial x^k} - \delta^j_k f - \frac{\partial f}{\partial x^k}x^j \\ &= -\delta^j_k f\end{aligned}$$

$$\varphi([x^j, \frac{\partial}{\partial x^k}]) = \varphi(i\delta^j_k) = i\delta^j_k$$

$$[\varphi(x^j), \varphi(p_k)] = -i[x^j, \frac{\partial}{\partial x^k}] = i\delta^j_k$$

Thus, φ preserves the Lie bracket. $\square$

(c) Let $\hbar \in \mathbb{R}^{>0}$ be the parameter that models the quantum scale. Show that $\mathfrak{heis}(\omega) \cong \mathfrak{heis}(\omega/\hbar)$ and that $Sp(\omega) = Sp(\omega/\hbar)$. Which associative algebra does $W(\omega/\hbar)$ approach in the classical limit?

Here we are using the representation of $\mathfrak{heis}(\omega) = i\mathbb{R} \oplus V$. Define $\varphi : \mathfrak{heis}(\omega/\hbar) \to \mathfrak{heis}(\omega)$ by,

$$\varphi(v) = \sqrt{\hbar}\, v$$

<u>Claim:</u> φ is Lie algebra isomorphism.

Pf.: Multiplication by a scalar is an isomorphism of algebras, so that it's sufficient to show that φ preserves the Lie bracket:

$$[\varphi(v), \varphi(w)] = \left[\frac{v}{\sqrt{\hbar}}, \frac{w}{\sqrt{\hbar}}\right] = i\omega\left(\frac{v}{\sqrt{\hbar}}, \frac{w}{\sqrt{\hbar}}\right) = \varphi\left(i\frac{\omega}{\hbar}(v,w)\right) = \varphi([v,w]). \quad \square$$

<u>Claim:</u> $Sp(\omega) = Sp(\omega/\hbar)$.

Pf.: We have that,

$$\begin{aligned} L \in Sp(\omega) &\Leftrightarrow \forall v, w \in V,\ \omega(Lv, Lw) = \omega(v, w) \\ &\Leftrightarrow \forall v, w \in V \frac{\omega}{\hbar}(Lv, Lw) = \frac{\omega}{\hbar}(v, w), \quad \text{(divided both sides by } \hbar) \\ &\Leftrightarrow L \in Sp(\omega/\hbar). \quad \square \end{aligned}$$

In the extension of $\mathfrak{heis}(\omega)$ to $W(\omega)$, the bracket must satisfy,

$$\frac{i}{\hbar}[a,b] = i\{a,b\}$$

so that in the classical limit,

$$\lim_{\hbar \to 0}[a,b] = \lim_{\hbar \to 0} i\hbar\{a,b\} = 0$$

Thus, in the limit,

$$W(\omega) = T(\mathbb{C} \otimes V^*)/(ab - ba - i\hbar\{a,b\}) \to T(\mathbb{C} \otimes V^*)/(ab - ba) = Sym(\mathbb{C} \otimes V^*)$$

Let $S^2_r \in \mathbb{R}^3$ be the sphere of radius r. We would like to consider "quantizing" this compact manifold. Let x, y, z be the standard linear coordinates on $\mathbb{R}^3$. The algebra of polynomial functions on S^2_r is

$$P = \mathbb{C}[x, y, z]/(x^2 + y^2 + z^2 - r^2).$$

(a) Compute the standard symplectic form Ω on S^2_r (the one derived from the embedding into Euclidean $\mathbb{R}^3$) in terms of x, y, z.

Computation: The standard symplectic form on S^2_r is the volume form, which is obtained by contracting the volume form on $\mathbb{R}^3$ with a radial vector field that has unit length on the sphere, and then pulling it back to S^2_r.

Let $\varphi : S^2_r \hookrightarrow \mathbb{R}^3$ be our embedding. Let $\partial_x = \frac{\partial}{\partial x}$ be a vector field on $\mathbb{R}^3$ and likewise for y and z. Let

$$\rho = \frac{1}{r^2}\left(x\partial_x + y\partial_y + z\partial_z\right).$$

Then

$$\omega = \iota_\rho(dx \wedge dy \wedge dz) = \frac{1}{r^2}\left(x\, dy \wedge dz \quad + \quad y\, dz \wedge dx \quad + \quad z\, dx \wedge dy\right)$$

as a 2-form on $\mathbb{R}^3$. If we let the $\mathsf{x}, \mathsf{y}, \mathsf{z}$ denote the pullback along φ of the respective coordinate functions, then

$$\Omega = \varphi^*\omega = \frac{1}{r^2}\left(\mathsf{x}\, d\mathsf{y} \wedge d\mathsf{z} \quad + \quad \mathsf{y}\, d\mathsf{z} \wedge d\mathsf{x} \quad + \quad \mathsf{z}\, d\mathsf{x} \wedge d\mathsf{y}\right)$$

as a 2-form on S^2_r. ■

(b) Let $\{\cdot, \cdot\}$ be the Poisson bracket for Ω. Show that $\{\mathsf{x}, \mathsf{y}\} = \mathsf{z}$ with cyclic permutations holding as well. Conclude that the "linear" functions in P are isomorphic to $\mathfrak{so}_3$.

Proof: The Poisson bracket with respect to Ω is defined by

$$\{f, g\} = dg(\mathrm{grad}_\Omega(f)) \qquad \text{where} \qquad \iota_{\mathrm{grad}_\Omega(f)}(\Omega) = -df.$$

Let $\tilde{\xi} = z\partial_y - y\partial_z$ be a vector field on $\mathbb{R}^3$. Note that on at a point in $\varphi(S^2_r) \subset \mathbb{R}^3$,

$$\iota_{\tilde{\xi}}(\omega) = \frac{1}{r^2}\left[x(zdz + ydy) - y^2dx - z^2dx\right] = \frac{x}{r^2}\left(zdz + ydy + xdx\right) - dx$$

Note that because $\left\langle \tilde{\xi}, \rho \right\rangle = 0$ with the standard metric on $\mathbb{R}^3$, $\tilde{\xi}$ is in the image of the injection $\varphi_* : TS^2_r \hookrightarrow T\mathbb{R}^3\big|_{\varphi(S^2_r)}$. Let $\varphi_*(\xi) = \tilde{\xi}\Big|_{\varphi(S^2_r)}$. Since $(xdx + ydy + zdz) \in \ker(\varphi^*)$, we have that

$$\iota_\xi(\Omega) = \varphi^*(\iota_{\tilde{\xi}}(\omega)) = \varphi^*\left(\frac{x}{r^2}\left(zdz + ydy + xdx\right) - dx\right) = -d(\varphi^* x)$$

Therefore, $\mathrm{grad}_\Omega(\mathsf{x}) = \xi$ and $\{\mathsf{x}, \mathsf{y}\} = d\mathsf{y}(\xi) = \mathsf{z}$. By cyclically permuting x, y, z in the above argument (note that ω is invariant under such permutations), we obtain that $\{\mathsf{y}, \mathsf{z}\} = \mathsf{x}$ and $\{\mathsf{z}, \mathsf{x}\} = \mathsf{y}$. Thus (the linear functions in P with the Poisson bracket) are isomorphic to $\mathfrak{so}_3$ as Lie algebras. ■

(c) Based on problem (4), it is natural to think of the "Weyl algebra" $W(\Omega)$ as the associative algebra over $\mathbb{C}$ generated by x, y, z subject to the relations given by $xy - yx = iz$ with cyclic permutations and $x^2 + y^2 + z^2 = r^2$. Based on (b), show that

$$W(\Omega) \cong \mathcal{U}(\mathfrak{g}) / \left(X^2 + Y^2 + Z^2 = -r^2\right)$$

where $\mathcal{U}(\mathfrak{g})$ is the universal enveloping algebra for $\mathfrak{g} := \mathbb{C} \otimes \mathfrak{so}_3$ and $X, Y, Z \in \mathfrak{so}_3$ are the elements identified with x, y, z respectively.

Proof: We define the universal enveloping algebra for $\mathfrak{g}$ as

$$\mathcal{U}(\mathfrak{g}) = T(\mathfrak{g}) / \langle R \rangle$$

where

$$T(\mathfrak{g}) = \bigoplus_{k=0}^{\infty} (\mathfrak{g})^{\otimes k}$$

is the tensor algebra of the vector space $\mathfrak{g}$,

$$R = \{ab - ba - [a, b] \mid a, b \in \mathfrak{so}_3\},$$

and $\langle R \rangle$ is the two-sided algebra ideal generated by R. Because $\{X, Y, Z\}$ are a basis for $\mathfrak{so}_3$, $\{X' = iX, Y' = iY, Z' = iZ\}$ are a basis for $\mathbb{C} \otimes \mathfrak{g}$, and $\langle R \rangle$ is generated by

$$R' = \left\{(X'Y' - Y'X') - iZ',\ \ (Y'Z' - Z'Y') - iX',\ \ (Z'X' - X'Z') - iY'\right\}.$$

Thus,

$$\mathcal{U}(\mathfrak{g}) / \left(X^2 + Y^2 + Z^2 = -r^2\right) \cong T(\mathfrak{g}) \,/ \left\langle \left\{X'^2 + Y'^2 + Z'^2 - r^2\right\} \cup R' \right\rangle,$$

which is precisely the definition (in terms of generators and relations) that we have for $W(\Omega)$. ■

(d) As in problem (4), when we "quantize" S_r^2, we want our Hilbert space to be an irreducible representation of $W(\Omega)$ such that $\mathbb{C} \subset W(\Omega)$ acts by scalar multiplication. On the other hand, standard representation theory tells us that for irreducible representations of $\mathfrak{so}_3$, the element $X^2 + Y^2 + Z^2$ must act as scalar multiplication by $n(n+2)/4$ for some integer n. Based on part (c), what restrictions do we have on the radius to obtain a non-trivial Hilbert space?

Answer: Complex representations of $\mathfrak{so}_3$ are in 1-1 correspondence with representations of $\mathcal{U}(\mathfrak{g})$ (i.e., $\mathcal{U}(\mathfrak{g})$-modules). Thus an irreducible representation of $W(\Omega)$ such that $\mathbb{C} \subset W(\Omega)$ acts by scalar multiplication (i.e., a simple $W(\Omega)$-module) corresponds to an irreducible representation of $\mathfrak{g}$ such that $X^2 + Y^2 + Z^2$ acts by scalar multiplication by $-r^2$. However, because of what was said above, for this to be a non-trivial vector space, $-r^2 = n(n+2)/4$ for some integer n, which implies that $n = -1 \implies r = 1/2$. ■

MAT 561 HW 3 Problem 6

Consider $V = \mathbb{R}_x + \mathbb{R}_p$ with symplectic form $\omega = dp \wedge dx$. Take $\mathbb{C} \otimes_{\mathbb{R}} V = W \oplus \overline{W}$, with $W = \mathbb{C}(e_x + ie_p)$ and $\overline{W} = \mathbb{C}(e_x - ie_p)$.

1 Part a

To see that W is a positive complex polarization, one has to check that $i\omega(\overline{w}, w) > 0$ for any $w \in W$. We only have to check for $e_x + ie_p$:

$$i\omega(e_x - ie_p, e_x + ie_p) = i(-i * 1 - 1 * i) = i(-2i) = 2 > 0$$

2 Part b

Take $v = v_x e_x + v_p e_p$. Note that $\omega(v, w) = \det\left(\begin{pmatrix} v_p & w_p \\ v_x & w_x \end{pmatrix}\right)$. For $A \in End(V)$, $\omega(Av, Aw) = \det\left(A * \begin{pmatrix} v_p & w_p \\ v_x & w_x \end{pmatrix}\right)$.

If $A \in Sp(V)$, then $\omega(Av, Aw) = \omega(v, w)$ so by above arguments $\det A = 1$. Therefore, $Sp(V) \approx SL_2(\mathbb{R})$.

$U(W)$ can be computed as follows:

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} &= \lambda \begin{pmatrix} 1 \\ i \end{pmatrix} \\ \begin{pmatrix} a + bi \\ c + di \end{pmatrix} &= \begin{pmatrix} \lambda \\ i\lambda \end{pmatrix} \\ a + bi &= d - ci \end{aligned}$$

Hence $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, with $ad - bc = a^2 + b^2 = 1$. It follows that $U(W) \approx \mathbb{T}$. Since the double cover of the circle is the circle, we have $\widetilde{U(W)} = \mathbb{T}$

3 Part c

In this problem, we'll use the fact that the representation of $Heis(\omega)$ is $\mathcal{H} = \overline{Sym^{\bullet}(W^*) \otimes \mathcal{H}^W}$ to compute the Schrödinger equation by looking at the action of the given evolution operator on a dense subset of $\mathcal{H}$.

The action of $\exp(tJ)$ on W is given by:

$$\begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = e^{-it} \begin{pmatrix} 1 \\ i \end{pmatrix}$$

So the action on W^* is multiplication by e^{it}.

Since $\mathcal{H}^{\otimes 2} = W^*$, we may choose the action of $\exp(tJ)$ to be $e^{it/2}$.

Let z be a generator of W^*. Then, $Sym^\bullet(W^*) = \mathbb{C}[z]$. One sees that

$$\frac{d}{dt}(e^{tJ}(z^n \otimes w))\Big|_{t=0} = i(n+\frac{1}{2})(z^n \otimes w)$$

where w is some element of $\mathcal{H}^W$.

(Note that either choice of the action on $\mathcal{H}^W$ would give us the same eigenvalues and eigenvectors.)

$z\frac{d}{dz}$ is an operator on $\mathbb{C}[z]$ that has eigenvalues n with corresponding eigenvectors z^n. One sees that these are the only eigenvalues and eigenvectors of this operator, and by a simple basis argument that this is the only such operator.

To get the desired operator on $\mathcal{H} = \mathbb{C}[z] \otimes \mathcal{H}^W$, we have to add $\frac{1}{2} * id$.

So we have $H = z\frac{d}{dz} + \frac{1}{2}$.

The representation of $Heis(\omega)$ is given by $\hat{z} = z$, and $\hat{\overline{z}} = \frac{d}{dz}$.

We have $\widehat{z\overline{z}} = \frac{\hat{z}\hat{\overline{z}} + \hat{\overline{z}}\hat{z}}{2}$, so one can rewrite:

$$\begin{aligned} H &= \hat{z}\hat{\overline{z}} + \frac{1}{2} \\ &= \frac{\hat{z}\hat{\overline{z}} + \hat{\overline{z}}\hat{z}}{2} + \frac{\hat{z}\hat{\overline{z}} - \hat{\overline{z}}\hat{z}}{2} + \frac{1}{2} \\ &= \widehat{z\overline{z}} + \frac{1}{2}([\hat{z}, \hat{\overline{z}}] + 1) \\ &= \widehat{z\overline{z}} \end{aligned}$$

The last line follows from $[\frac{d}{dz}, z] = id$.

To write H as an operator on $L^2(\mathbb{R}_x)$, we must convert everything with respect ot the real decomposition. So $z = \frac{1}{\sqrt{2}}(x+ip)$, $\overline{z} = \frac{1}{\sqrt{2}}(x-ip)$, where x and p are the coordinates with respect to the real polarization. By noting that $\hat{x} = x$ and $\hat{p} = -i\frac{\partial}{\partial x}$, we obtain the desired result:

$$H = \widehat{z\overline{z}} = \frac{\widehat{x^2+p^2}}{2} = \frac{\hat{x^2}}{2} + \frac{\hat{p^2}}{2} = \frac{1}{2}\left(-\frac{\partial^2}{\partial x^2} + x^2\right)$$

Jerry's Quantum Mechanics Homework
MAT 561

Christopher Bay

May 7, 2007

Let M be a smooth manifold with a symplectic form ω. Assume that there is a one-form α such that $d\alpha = \omega$. Give $\xi \in \mathcal{X}(M)$ we define a new operator on functions

$$\nabla_\xi = \xi + i m_{\alpha(\xi)}$$

where m_f denotes multiplication by the function f.

1. For $\xi, \eta \in \mathcal{X}(M)$ we have
$$\nabla_\xi \nabla_\eta - \nabla_\eta \nabla_\xi - \nabla_{[\xi,\eta]} = i\omega(\xi, \eta).$$

   Proof. Let $g \in C^\infty(M)$. Then,

$$\begin{aligned}
\{\nabla_\xi \nabla_\eta - \nabla_\eta \nabla_\xi - \nabla_{[\xi,\eta]}\}g &= \nabla_\xi(\nabla_\eta g) - \nabla_\eta(\nabla_\xi g) - \nabla_{[\xi,\eta]}g \\
&= \nabla_\xi(\eta g + i\alpha(\eta)g) - \nabla_\eta(\xi g + i\alpha(\xi)g) - ([\xi,\eta]g + i\alpha([\xi,\eta])g) \\
&= \xi(\eta g) + i\alpha(\xi)(\eta g) + i(\xi\alpha(\eta))g + i\alpha(\eta)(\xi g) + i\alpha(\xi)\alpha(\eta)g \\
&\quad -\{\eta(\xi g) + i\alpha(\eta)(\xi g) + i\eta(\alpha(\xi))g + i\alpha(\xi)(\eta g) + i\alpha(\xi)\alpha(\eta)g\} \\
&\quad -\{\xi(\eta g) - \eta(\xi g) + i\xi(\alpha(\eta))g - i\eta(\alpha(\xi))g - i\omega(\xi,\eta)g\}.
\end{aligned}$$

   In the last part of the last equality we have used the identity

$$\omega(\xi, \eta) = d\alpha(\xi, \eta) = \xi(\alpha(\eta)) - \eta(\alpha(\xi)) - \alpha([\xi, \eta]).$$

   All of the terms except the very last one cancel in pairs, proving the desired equality.

   □

2. Let ξ_g denote the symplectic gradient of $g \in C^\infty(M)$. Define an operator on functions
$$\hat{g} = i\nabla_{\xi_g} + m_g.$$

   Then $[\hat{f}, \hat{g}] = i\widehat{\{f, g\}}$.

Proof. For any $h \in C^\infty(M)$,

$$\begin{aligned}
[\hat{f}, \hat{g}]h &= \hat{f}(\hat{g}h) - \hat{g}(\hat{f}h) \\
&= \hat{f}(i\nabla_{\xi_g} h + gh) - \hat{g}(i\nabla_{\xi_f} h + fh) \\
&= -\nabla_{\xi_f}(\nabla_{\xi_g} h) + if(\nabla_{\xi_g} h) + i\nabla_{\xi_f}(gh) + fgh \\
&\quad +\nabla_{\xi_g}(\nabla_{\xi_f} h) - ig(\nabla_{\xi_f} h) - i\nabla_{\xi_g}(fh) - fgh.
\end{aligned}$$

Now use part 1 (for the terms with two ∇'s) and expand the other terms using the definitions.

$$\begin{aligned}
[\hat{f}, \hat{g}]h &= i\omega(\xi_g, \xi_f)h + \nabla_{[\xi_g,\xi_f]}h + if(\xi_g h) - \alpha(\xi_g)fh \\
&\quad +i\xi_f(gh) - \alpha(\xi_f)gh - ig(\xi_f h) + \alpha(\xi_f)gh - i\xi_g(fh) + \alpha(\xi_g)fh \\
&= i\{g, f\}h + \nabla_{[\xi_g,\xi_f]}h + if(\xi_g h) + i(\xi_f g)h + ig(\xi_f h) - ig(\xi_f h) - i(\xi_g f)h - if(\xi_g h) \\
&= i\{g, f\}h + \nabla_{[\xi_g,\xi_f]}h + i(\xi_f g)h - i(\xi_g f)h \\
&= i\{g, f\}h + \nabla_{[\xi_g,\xi_f]}h + i\{f, g\}h - i\{g, f\}h \\
&= -\nabla_{[\xi_f,\xi_g]}h + i\{f, g\}h \\
&= i(i\nabla_{\{f,g\}}h + \{f, g\}h) \\
&= i\widehat{\{f, g\}}h.
\end{aligned}$$

Here we have used that the Poisson bracket is defined by $\{f, g\} = \omega(\xi_f, \xi_g)$, as well as the fact that $f \mapsto \xi_f$ is a Lie algebra homomorphism.

□

3. Let $M = T^*X$ and let α be the canonical one-form and take $\omega = d\alpha$. Let $x^1, \ldots, x^n$ be local coordinates on X and let $p_1, \ldots, p_n$ be the corresponding linear coordinates along the fibers of T^*X. Then $\widehat{x^i}$ and $\widehat{p_i}$ are given by

$$\widehat{x^i} = -i\frac{\partial}{\partial p_i} + m_{x^i},$$

$$\widehat{p_i} = i\frac{\partial}{\partial x^i}.$$

Proof. The canonical one-form α is defined by $\alpha = \sum_i p_i dx^i$, so $\omega = \sum_i dp_i \wedge dx^i$. By definition, a symplectic gradient satisfies $\omega(\xi_f, \cdot) = -df$. Therefore $\xi_{x^i} = -\partial/\partial p_i$ and $\xi_{p_i} = \partial/\partial x^i$. One also checks that $\alpha(\xi_{x^i}) = -\alpha(\partial/\partial p_i) = 0$ and $\alpha(\xi_{p_i}) = \alpha(\partial/\partial x^i) = p_i$. So for any $f \in C^\infty(T^*X)$,

$$\begin{aligned}
\widehat{x^i}f &= i\nabla_{\xi_{x^i}} f + x^i f \\
&= i\xi_{x^i} f - \alpha(\xi_{x^i})f + x^i f \\
&= (-i\frac{\partial}{\partial p_i} + m_{x^i})f
\end{aligned}$$

and

$$\begin{aligned}
\widehat{p_i} f &= i\nabla_{\xi_{p_i}} f + p_i f \\
&= i\xi_{p_i} f - \alpha(\xi_{p_i}) f + p_i f \\
&= (i\frac{\partial}{\partial x^i}) f.
\end{aligned}$$

□

4. The restrictions of the operators $\widehat{x^i}$ and $\hat{p}_i$ to $C^\infty(X)$ are given by

$$\widehat{p_i} \mid_{C^\infty(X)} = i\frac{\partial}{\partial x^i},$$

$$\widehat{x^i} \mid_{C^\infty(X)} = m_{x^i}.$$

Proof. To properly define the restrictions, one should be slightly careful and compose with the maps $\pi : T^*X \to X$ (the bundle projection) and $\iota : X \to T^*X$ (inclusion as the zero-section) in the appropriate places. But doing this just verifies what is already clear. Since $\partial/\partial p_i \equiv 0$ and for any function of the form $f \circ \pi$, $f \in C^\infty(X)$, the identities follow immediately from 3. □

Jerry Jenquin

This semester (Spring 2007) I'm co-teaching one course. MAT 561 is a graduate course that covers the math and physics of thermodynamics, statistical and quantum mechanics. The goal is to relay the topics to a mathematical audience in a way that provides both a stepping stone to more advanced topics and some insight into physical intuition.

The various ways to contact me are listed in the contact link. The best way is via email, which I check habitually.

Jerry Jenquin

Nothing here, as of yet.