Lecture Notes of Quals Review for Physics PhD Candidates

Zhuoran He
Columbia University

February, 2019

Preface

The qualifying exams (Quals) in the department of physics of Columbia University are a required qualification before PhD students begin their research. As a former PhD student of Columbia University, I wrote these lecture notes when I was invited by the department of physics to give tutorial lectures to the Quals takers in 2014. The exams mainly focus on in-depth topics of undergrad physics, including mechanics, electrodynamics, relativity, quantum physics, thermodynamics and statistical physics. These lecture notes were recreated when I graduated from Columbia University with some new thoughts incorporated, in memory of the wonderful times during my PhD studies. The lecture notes are a good summary and revision of undergrad physics and can be useful for future Quals takers and for PhDs and postdocs as a quick mind refresher of the fundamentals.

The lecture notes contain 6 equal-length lectures, each can be covered in about 2 hours. The notes are typically a combination of formalism derivation and sample problems picked from past qualifying exams with solutions and explanations. The 6 lecture notes in this book correspond to the 6 qualifying exam sections that take place on 3 separate days with minor reorganizations. Among all 6 lectures, Lectures 4 and 5 on quantum mechanics have the richest and most diversified content. I can only cover some important selected topics and encourage the reader to also go over Griffith's *Introduction to Quantum Mechanics* for a more comprehensive revision. Other lectures tend to be more systematic and complete but no less interesting. There are also order-of-magnitude estimation, dimension analysis, and astrophysics problems that I did not have time to sufficiently talk about, but the techniques used in astrophysics, e.g., planetary motion, radiation, relativity, quantum physics, thermodynamics, fluid dynamics, etc., are covered in these lecture notes.

I acknowledge Columbia University for giving me the chance to teach the Quals and share my notes to future students. I wish all students to perform well in the Quals!

Zhuoran He
February, 2019

Contents

Lecture 1 Classical Mechanics and Lagrangian Mechanics

1 Derivation of the Lagrange equation

1.1 Virtual work and the d'Alembert principle

From a purely classical mechanical point of view, Lagrangian mechanics is introduced as a way to study particles moving under constraints. It is more convenient than Newtonian mechanics as it cancels the constraint forces and directly gives the motion of the particles given only the applied forces acting on them. Suppose we have N constrained interacting particles at positions given by

$$\vec{r}_i = \vec{r}_i(q,t), \quad i = 1, 2, \ldots, N, \tag{1}$$

where $q = (q_1, q_2, \ldots, q_s)$ is a set of independent real parameters called the *generalized coordinates*. The above N *mapping equations* map the set of generalized coordinates q into the N Cartesian coordinates of the particles. The particles will be subject to forces of two types: *applied forces* $\vec{F}_i^A$, which are active and drive the particles to move, and *constraint forces* $\vec{N}_i$, which are passive and keep the particles onto the constraint surface. The explicit time dependence t in the mapping equations allows for the possibility of a time-dependent constraint surface, which changes the mapping rules $\vec{r}_i(q,t)$ as time goes by. The Newtonian equations of the particles are given by

$$\vec{F}_i^A + \vec{N}_i = m_i \ddot{\vec{r}}_i, \quad i = 1, 2, \ldots, N. \tag{2}$$

In case of a *frictionless* constraint, the constraint forces $\vec{N}_i$ will be normal to the surface but can still do work if the surface moves. In order to cancel the constraint forces, we consider a *virtual displacement*

$$\delta\vec{r}_i = \sum_{\alpha=1}^{s} \frac{\partial \vec{r}_i}{\partial q_\alpha} \delta q_\alpha, \quad i = 1, 2, \ldots, N, \tag{3}$$

of the N particles with the constraint surface fixed at a certain time. This is like taking a *snapshot* of the constraint surface and imagining that the particles move on it. And then it is clear to say that frictionless constraint forces do no virtual work, i.e.,

$$\vec{N} \cdot \delta\vec{r}_i = 0, \quad i = 1, 2, \ldots, N. \tag{4}$$

We therefore arrive at the *d'Alembert principle*

$$\sum_{i=1}^{N} \vec{F}_i^A \cdot \delta\vec{r}_i = \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \delta\vec{r}_i. \tag{5}$$

The constraint forces $\vec{N}_i$ are now all canceled, with only the active forces $\vec{F}_i^A$ and particle accelerations left in the equation.

1.2 Generalized forces and the Lagrange equation

The left-hand side of Eq. (5) can be written as

$$\text{LHS} = \sum_{\alpha=1}^{s} \delta q_\alpha \sum_{i=1}^{N} \vec{F}_i^A \cdot \frac{\partial \vec{r}_i}{\partial q_\alpha} = \sum_{\alpha=1}^{s} Q_\alpha \delta q_\alpha, \tag{6}$$

where we have introduced the *generalized forces* defined as

$$Q_\alpha = \sum_{i=1}^{N} \vec{F}_i^A \cdot \frac{\partial \vec{r}_i}{\partial q_\alpha}, \quad \alpha = 1, 2, \ldots, s. \tag{7}$$

The generalized forces do not act on a single particle, but on the generalized coordinates q_α, which can describe collective motion. The right-hand side of Eq. (5) becomes

$$\text{RHS} = \sum_{\alpha=1}^{s} \delta q_\alpha \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_\alpha} = \sum_{\alpha=1}^{s} \delta q_\alpha \sum_{i=1}^{N} m_i \left[\frac{d}{dt}\left(\dot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_\alpha} \right) - \dot{\vec{r}}_i \cdot \frac{d}{dt}\left(\frac{\partial \vec{r}_i}{\partial q_\alpha} \right) \right]. \tag{8}$$

From the mapping equations $\vec{r}_i = \vec{r}_i(q, t)$, we have

$$\dot{\vec{r}}_i = \sum_{\alpha=1}^{s} \frac{\partial \vec{r}_i}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \vec{r}_i}{\partial t} = \dot{\vec{r}}_i(q, \dot{q}, t). \tag{9}$$

Therefore we can directly read from Eq. (9) that

$$\frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_\alpha} = \frac{\partial \vec{r}_i}{\partial q_\alpha}, \tag{10}$$

and by acting $\partial/\partial q_\alpha$ onto Eq. (9) obtain

$$\frac{\partial \dot{\vec{r}}_i}{\partial q_\alpha} = \sum_{\alpha=1}^{s} \dot{q}_\beta \frac{\partial}{\partial q_\beta}\left(\frac{\partial \vec{r}_i}{\partial q_\alpha} \right) + \frac{\partial}{\partial t}\left(\frac{\partial \vec{r}_i}{\partial q_\alpha} \right) = \frac{d}{dt}\left(\frac{\partial \vec{r}_i}{\partial q_\alpha} \right). \tag{11}$$

Using Eqs. (10) and (11), we rewrite the right-hand side of Eq. (5) into

$$\text{RHS} = \sum_{\alpha=1}^{s} \delta q_\alpha \sum_{i=1}^{N} m_i \left[\frac{d}{dt}\left(\dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_\alpha} \right) - \dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial q_\alpha} \right]. \tag{12}$$

Since the total kinetic energy of the mechanical system is given by

$$T = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\vec{r}}_i^2 = T(q, \dot{q}, t), \tag{13}$$

and the particle masses m_i remain constant over time, we have

$$\text{RHS} = \sum_{\alpha=1}^{s} \delta q_\alpha \left[\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} \right] = \sum_{\alpha=1}^{s} Q_\alpha \delta q_\alpha = \text{LHS}. \tag{14}$$

Finally, from our assumption that the generalized coordinates q_α are independent, we obtain

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_\alpha} \right) - \frac{\partial T}{\partial q_\alpha} = Q_\alpha = \sum_{i=1}^{N} \vec{F}_i^A \cdot \frac{\partial \vec{r}_i}{\partial q_\alpha}, \quad \alpha = 1, 2, \ldots, s. \tag{15}$$

These s equations are called the *Lagrange equations*. They govern the motion of the particles without involving the constraint forces.

One may wonder why we need to assume that the virtual variations δq_α of the generalized coordinates are independent. If they are not, why can't we just cancel some of the generalized coordinates q_α and keep only the rest? However, cancellation of generalized coordinates using the constraints works for *holonomic* constraints only, i.e., when the dependence of the variations δq_α is integrable. In case of a non-holonomic constraint, the δq_α's are inter-related but not integrable, so that there is still no one-to-one correspondence between the generalized coordinates q_α. The quantum-mechanical analogue of this situation is the Berry phase. We will not go into the details here.

1.3 Lagrange equation of a conservative system

In case that the applied forces $\vec{F}_i^A$ are conservative, we may introduce a potential function $V(\vec{r}_1, \ldots, \vec{r}_N, t)$, such that

$$\vec{F}_i^A = -\nabla_i V(\vec{r}_1, \ldots, \vec{r}_N, t). \tag{16}$$

In this case, the generalized forces can be found to be

$$Q_\alpha = \sum_{i=1}^{N} \vec{F}_i^A \cdot \frac{\partial \vec{r}_i}{\partial q_\alpha} = -\sum_{i=1}^{N} \frac{\partial \vec{r}_i}{\partial q_\alpha} \cdot \nabla_i V(\vec{r}_1(q,t), \ldots, \vec{r}_N(q,t), t) = -\frac{\partial V(q,t)}{\partial q_\alpha}, \tag{17}$$

where we have plugged in the mapping equations and used the chain rule. Since the potential $V(q,t)$ does not depend on the generalized velocities $\dot{q}$, we have $\partial V / \partial \dot{q}_\alpha = 0$. Therefore, the Lagrange equations become

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) = \frac{\partial L}{\partial q_\alpha}, \quad \alpha = 1, 2, \ldots, s, \tag{18}$$

where $L = T - V$ is introduces as the *Lagrangian* of the system. Eq. (18) is more frequently referred to as the *Lagrange equation* than Eq. (15).

In some systems, we may introduce a potential function $V = V(q, \dot{q}, t)$ that also depends on the generalized velocities $\dot{q}$. If the Lagrangian $L = T - V$ constructed yields the correct classical motion, we still accept the validity of the above Lagrangian equation, in a more generalized sense. For example, the Lagrangian of a non-relativistic particle moving in an electromagnetic field can be constructed as

$$L = \frac{1}{2} m \dot{\vec{r}}^2 - e \left[\phi(\vec{r}, t) - \dot{\vec{r}} \cdot \vec{A}(\vec{r}, t) \right]. \tag{19}$$

The vector potential $\vec{A}(\vec{r}, t)$ is coupled with the particle's velocity $\dot{\vec{r}}$ so as to give a velocity-dependent potential. Pluggin this Lagrangian into the Lagrange equation, we obtain

$$\begin{aligned} \frac{d}{dt} \left(\nabla_{\dot{\vec{r}}} L \right) - \nabla_{\vec{r}} L &= \frac{d}{dt} \left(m \dot{\vec{r}} + e \vec{A} \right) + e \nabla_{\vec{r}} \phi + e \nabla_{\vec{r}} (\dot{\vec{r}} \cdot \vec{A}) \\ = \frac{d}{dt} (m \dot{\vec{r}}) + e \left(\nabla \phi + \frac{\partial \vec{A}}{\partial t} \right) - e \dot{\vec{r}} \times (\nabla \times \vec{A}) &= \frac{d}{dt} (m \dot{\vec{r}}) - e (\vec{E} + \dot{\vec{r}} \times \vec{B}) = 0, \end{aligned} \tag{20}$$

which gives precisely the Lorentz force law.

2 Conservation laws of Lagrangian mechanics

2.1 Cyclic coordinates

In case that $\partial L / \partial q_\alpha = 0$, q_α is called a *cyclic coordinate* of the system. In such case, from the Lagrange equation (18), we have

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) = 0 \quad \Rightarrow \quad \frac{\partial L}{\partial \dot{q}_\alpha} = \text{const.} \tag{21}$$

Therefore, whenever the system is invariant under translation along q_α, the corresponding *conjugate momentum* $p_\alpha = \partial L / \partial \dot{q}_\alpha$ is conserved.

2.2 Energy conservation

If $\partial L / \partial t = 0$, the system exhibits time translation invariance. We can find another conserved quantity in this case. Multiplying $\dot{q}_\alpha$ onto the Lagrange equation (18), we obtain

$$\sum_{\alpha=1}^{s} \dot{q}_\alpha \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\alpha} \right) = \sum_{\alpha=1}^{s} \dot{q}_\alpha \frac{\partial L}{\partial q_\alpha} + \frac{\partial L}{\partial t}, \tag{22}$$

where the added term on the right-hand side $\partial L / \partial t = 0$ is zero by assumption. The left-hand

side of Eq. (22) can be written as

$$\sum_{\alpha=1}^{s} \dot{q}_\alpha \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\right) = \sum_{\alpha=1}^{s} \left[\frac{d}{dt}\left(\dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha}\right) - \ddot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha}\right]. \tag{23}$$

Equating this to the right-hand side of Eq. (22), we have

$$\frac{d}{dt}\left(\sum_{\alpha=1}^{s} \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L\right) = 0 \quad \Rightarrow \quad H = \sum_{\alpha=1}^{s} \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L = \text{const}, \tag{24}$$

where H is introduced as the *Hamiltonian* of the system. In case of a time-independent constraint surface, since

$$\dot{\vec{r}}_i = \sum_{\alpha=1}^{s} \frac{\partial \vec{r}_i}{\partial q_\alpha}\, \dot{q}_\alpha + \frac{\partial \vec{r}_i}{\partial t} = \sum_{\alpha=1}^{s} \frac{\partial \vec{r}_i}{\partial q_\alpha}\, \dot{q}_\alpha \tag{25}$$

is a linear homogeneous function of the generalized velocities $\dot{q}_\alpha$, the total kinetic energy T will be quadratic homogeneous while V is independent of $\dot{q}_\alpha$. In this case, from Euler's theorem of homogeneous functions, we have

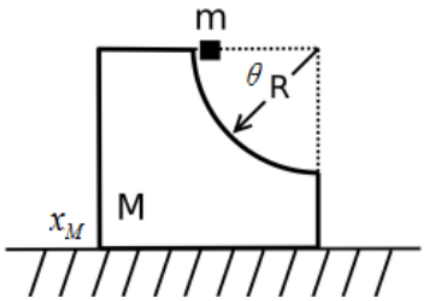
$$H = \sum_{\alpha=1}^{s} \dot{q}_\alpha \frac{\partial L}{\partial \dot{q}_\alpha} - L = 2T - (T - V) = T + V, \tag{26}$$

which is the total energy of the system. Therefore time translation invariance implies energy conservation. The situations discussed in § 2.1 and § 2.2 are two special cases of the more general Noether's theorem, which states that any differentiable symmetry of a system has a corresponding conservation law.

3 Sample problems

3.1 Problem 4 in Sec. 1, 2013

A mass m slides down a circularly curved surface on an object with mass M as shown in the diagram below. Mass M is free to slide on a frictionless surface. What are the final speeds of the two masses after m separates from M?



Answer: To find the Lagrangian, we have the kinetic and potential energies

$$T = \frac{1}{2}M\dot{x}_M^2 + \frac{1}{2}m\left(\dot{x}_M^2 + 2R\dot{\theta}\dot{x}_M\sin\theta + R^2\dot{\theta}^2\right), \quad V = -mgR\sin\theta. \tag{27}$$

Therefore, the Lagrangian is given by

$$L = T - V = \frac{1}{2}(M+m)\dot{x}_M^2 + mR\dot{\theta}\dot{x}_M\sin\theta + \frac{1}{2}mR^2\dot{\theta}^2 + mgR\sin\theta. \tag{28}$$

Since $\partial L/\partial x_M = 0$, we have

$$\frac{\partial L}{\partial \dot{x}_M} = (M+m)\dot{x}_M + mR\dot{\theta}\sin\theta = \text{const}, \tag{29}$$

which is the conservation of horizontal momentum. And since $\partial L/\partial t = 0$, we have

$$H = \dot{x}_M\frac{\partial L}{\partial \dot{x}_M} + \dot{\theta}\,\frac{\partial L}{\partial \dot{\theta}} - L = T + V \tag{30}$$

$$= \frac{1}{2}(M+m)\dot{x}_M^2 + mR\dot{\theta}\dot{x}_M\sin\theta + \frac{1}{2}mR^2\dot{\theta}^2 - mgR\sin\theta = \text{const}, \tag{31}$$

which is the conservation of energy. Plugging in the initial and final values, the two conserved quantities give us two equations

$$M\dot{x}_M + m(\dot{x}_M + R\dot{\theta}) = 0, \tag{32}$$

$$\frac{1}{2}M\dot{x}_M^2 + \frac{1}{2}m(\dot{x}_M + R\dot{\theta})^2 - mgR = 0, \tag{33}$$

which leads to the speeds

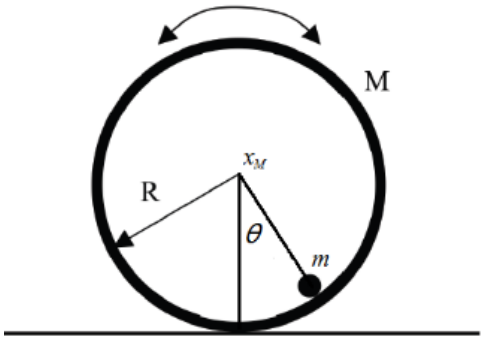
$$v_M = |\dot{x}_M| = \frac{m\sqrt{2gR}}{\sqrt{M(M+m)}}, \quad v_m = |\dot{x}_M + R\dot{\theta}| = \frac{M\sqrt{2gR}}{\sqrt{M(M+m)}}. \tag{34}$$

Lagrangian mechanics is an overskill for this simple problem, because we knew the energy and horizontal momentum were conserved. We don't need the Lagrangian to tell us that. But the next problem will be different.

3.2 Problem 5 in Sec. 1, 2013

A point particle of mass m slides without friction within a hoop of radius R and mass M. The hoop is free to roll without slipping along a horizontal surface. What is the frequency of small oscillations of the point mass, when it is close to the bottom of the hoop?

Answer: Since the hoop rolls on the horizontal surface instead of sliding frictionlessly, the horizontal momentum is not conserved. But the system definitely exhibits translational invariance along the horizontal direction. Now you see sometimes it's not so easy to find a



crucial conserved quantity without using Lagrangian mechanics. To find the Lagrangian, we have the kinetic energy

$$T = \frac{1}{2}M\dot{x}_M^2 + \frac{1}{2}I_M\omega_M^2 + \frac{1}{2}m\left(\dot{x}_M^2 + 2R\dot{\theta}\dot{x}_M\cos\theta + R^2\dot{\theta}^2\right), \tag{35}$$

and the potential energy $V = -mgR\cos\theta$. Since the hoop M rolls without slipping, we have $\dot{x}_M = \omega_M R$. The moment of inertia of a hoop is $I_M = MR^2$. Therefore, the Lagrangian

$$L = T - V = \frac{1}{2}(2M+m)\dot{x}_M^2 + mR\dot{\theta}\dot{x}_M\cos\theta + \frac{1}{2}mR^2\dot{\theta}^2 + mgR\cos\theta. \tag{36}$$

Since $\partial L/\partial x_M = 0$, we have

$$\frac{\partial L}{\partial \dot{x}_M} = (2M+m)\dot{x}_M + mR\dot{\theta}\cos\theta = \text{const}, \tag{37}$$

which is the conserved quantity we couldn't easily find using Newtonian mechanics. And since $\partial L/\partial t = 0$, we have again the conservation of total energy

$$H = \dot{x}_M\frac{\partial L}{\partial \dot{x}_M} + \dot{\theta}\,\frac{\partial L}{\partial \dot{\theta}} - L = T + V \tag{38}$$

$$= \frac{1}{2}(2M+m)\dot{x}_M^2 + mR\dot{\theta}\dot{x}_M\cos\theta + \frac{1}{2}mR^2\dot{\theta}^2 - mgR\cos\theta = \text{const}. \tag{39}$$

Since the system does an oscillation with no slippage in the horizontal direction, the conserved quantity $\partial L/\partial \dot{x}_M = 0$. This gives a constraint between $\dot{x}_M$ and $\dot{\theta}$ so we can cancel one of them, say $\dot{x}_M$. Then the Hamiltonian becomes

$$H = \frac{1}{2}\left(m - \frac{m^2\cos^2\theta}{2M+m}\right)R^2\dot{\theta}^2 - mgR\cos\theta. \tag{40}$$

Using the approximation $\theta \ll 1$ and keeping only quadratic terms of θ and $\dot{\theta}$, we obtain

$$H \approx \frac{1}{2}\left(\frac{2Mm}{2M+m}\right)R^2\dot{\theta}^2 + \frac{1}{2}mgR\theta^2 - mgR. \tag{41}$$

The last term is a constant and can be dropped without changing the physics. What we now have is the Hamiltonian of a simple harmonic oscillator. The angular frequency of oscillation can be found to be

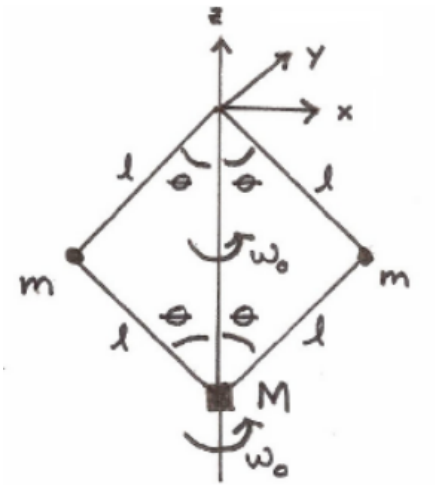
$$\omega = \sqrt{\left(1 + \frac{m}{2M}\right)\frac{g}{R}}. \tag{42}$$

3.3 Problem 2 in Sec. 1, 2011

A flyball governor consists of two masses m connected to arms of length l and a mass M as shown below. The assembly is constrained to rotate around a shaft on which the mass M can slide up and down without friction. Neglect the mass of the arms, air friction and assume that the diameter of mass M is small. Suppose that the shaft is constrained to rotate at angular velocity ω_0. Now answer the following questions

(a) Calculate the equilibrium height of the mass M.

(b) Calculate the frequency of small oscillations around this value.



Answer: This problem can be solved using Newtonian mechanics in the rotating frame and introducing the centrifugal (inertial) forces. And once the system oscillates, there will also be Coriolis forces. But these forces will be perpendicular to the velocities of the masses and will not affect their motion under the constraints. But still, using Lagrangian mechanics in the lab frame, we needn't worry about any of these inertial forces. All we need is write down the kinetic and potential energies to obtain the Lagrangian.

(a) For Part (a) we have the kinetic energy

$$T = \frac{1}{2}m\left[(\omega_0 l\sin\theta)^2 + (l\dot{\theta})^2\right] \times 2 + \frac{1}{2}M\left[\frac{d}{dt}(2l\cos\theta)\right]^2 \tag{43}$$

$$= (m + 2M\sin^2\theta)l^2\dot{\theta}^2 + m\omega_0^2 l^2\sin^2\theta, \tag{44}$$

and the potential energy

$$V = -2mgl\cos\theta - 2Mgl\cos\theta = -2(M+m)gl\cos\theta. \tag{45}$$

Therefore, the Lagrangian is given by

$$L = T - V = (m + 2M\sin^2\theta)l^2\dot{\theta}^2 + m\omega_0^2 l^2\sin^2\theta + 2(M+m)gl\cos\theta. \tag{46}$$

Since $\partial L/\partial t = 0$, we have the conserved quantity

$$H = \dot{\theta}\,\frac{\partial L}{\partial\dot{\theta}} - L = (m + 2M\sin^2\theta)l^2\dot{\theta}^2 - m\omega_0^2 l^2\sin^2\theta - 2(M+m)gl\cos\theta. \tag{47}$$

Notice that the system in this problem is subject to a time-dependent constraint (because the system is forced to rotate at a fixed angular frequency ω_0), so the Hamiltonian is not equal to $T+V$ any more. The term $-m\omega_0^2 l^2\sin^2\theta$ in H can be understood as the centrifugal potential energy in the rotating frame, and Coriolis force is not involved. To find the equilibrium angle θ_0, we set $\dot{\theta} = 0$ and minimize the total energy

$$H = V_{\text{eff}}(\theta) = -m\omega_0^2 l^2\sin^2\theta - 2(M+m)gl]\cos\theta, \tag{48}$$

by taking the derivative

$$\frac{dV_{\text{eff}}(\theta)}{d\theta} = -2m\omega_0^2 l^2\sin\theta\cos\theta + 2(M+m)gl\sin\theta = 0 \tag{49}$$

$$\Rightarrow \quad \sin\theta = 0, \quad \text{or} \quad \cos\theta = \frac{(M+m)g}{m\omega_0^2 l}. \tag{50}$$

If $\omega_0 < \sqrt{(M+m)g/ml} \equiv \omega_c$, we have $(M+m)g/m\omega_0^2 l > 1$, so the only solution is $\theta = 0$. The two arms will stay vertically down if the system rotates too slowly. Once $\omega_0 > \omega_c$, one may take the second-order derivative to find that the equilibrium position $\theta = 0$ becomes unstable, so the system will stay at the other equilibrium angle

$$\theta_0 = \cos^{-1}\left(\frac{\omega_c^2}{\omega_0^2}\right), \qquad \omega_0 > \omega_c = \sqrt{\frac{(M+m)g}{ml}}. \tag{51}$$

(b) Now we need to consider the motion of θ around θ_0. We need to approximate $V_{\text{eff}}(\theta)$ as a quadratic function by taking its second-order derivative at θ_0 to obtain

$$\left.\frac{d^2V_{\text{eff}}(\theta)}{d\theta^2}\right|_{\theta_0} = -2m\omega_0^2 l^2(2\cos^2\theta_0 - 1) + 2(M+m)gl\cos\theta_0. \tag{52}$$

Then using $\cos\theta_0 = (M+m)g/m\omega_0^2 l$, we have

$$\left.\frac{d^2V_{\text{eff}}(\theta)}{d\theta^2}\right|_{\theta_0} = 2\left[m\omega_0^2 l^2 - \frac{(M+m)^2 g^2}{m\omega_0^2}\right] = 2ml^2\left(\omega_0^2 - \frac{\omega_c^4}{\omega_0^2}\right). \tag{53}$$

Assuming $\omega_0 > \omega_c$, we have $V''_{\text{eff}}(\theta_0) \equiv d^2V_{\text{eff}}(\theta)/d\theta^2|_{\theta_0} > 0$. Therefore, the system oscillates around θ_0. The Hamiltonian under small angle approximation $|\theta - \theta_0| \ll 1$ is given by

$$H \approx (m + 2M\sin^2\theta_0)l^2\dot{\theta}^2 + V_{\text{eff}}(\theta_0) + \frac{1}{2}V''_{\text{eff}}(\theta_0)(\theta - \theta_0)^2 \tag{54}$$

$$= \left[m + 2M\left(1 - \frac{\omega_c^4}{\omega_0^4}\right)\right] l^2\dot{\theta}^2 + ml^2\left(\omega_0^2 - \frac{\omega_c^4}{\omega_0^2}\right)(\theta - \theta_0)^2 + \text{const.} \tag{55}$$

The angular frequency of small angle oscillation is therefore

$$\omega = \omega_0\sqrt{\frac{\left(1 - \frac{\omega_c^4}{\omega_0^4}\right)}{1 + \frac{2M}{m}\left(1 - \frac{\omega_c^4}{\omega_0^4}\right)}}, \qquad \omega_0 > \omega_c = \sqrt{\frac{(M+m)g}{ml}}. \tag{56}$$

This problem is nearly the hardest one involving small oscillations with only one degree of freedom. The Lagrangian method can also be used for oscillations with multiple degrees of freedom. The next problem gives an example.

3.4 Problem 5 in Sec. 1, 2011

A pendulum consisting of a massless rod of length L with a mass M at its end hangs from a fixed point. A second pendulum of the same construction hangs from the end of the first pendulum. The pendulums are constrained to move in the same plane. Find the frequencies of small oscillation and clearly describe the corresponding motions assuming that neither pendulum moves far from the vertical configuration.

Answer: Let the angle between the upper pendulum and the vertical direction be α, and the angle between the lower pendulum and the vertical direction be β. Given that both α and β are small, the total kinetic energy of the system can be found to be

$$T = \frac{1}{2}ML^2\left[\dot{\alpha}^2 + (\dot{\alpha} + \dot{\beta})^2\right]. \tag{57}$$

And the total potential energy is

$$V = -MgL\cos\alpha - MgL(\cos\alpha + \cos\beta) \tag{58}$$

$$\approx MgL\left(\alpha^2 + \frac{1}{2}\beta^2\right) + \text{const.} \tag{59}$$

Therefore, dropping the constant, the Lagrangian is given by

$$L = T - V = ML^2\left(\dot{\alpha}^2 + \dot{\alpha}\dot{\beta} + \frac{1}{2}\dot{\beta}^2\right) - MgL\left(\alpha^2 + \frac{1}{2}\beta^2\right). \tag{60}$$

We recognize that the system is a harmonic oscillator with two degrees of freedom. The

Lagrangian and the two Lagrange equations can be written in matrix form

$$L = \frac{1}{2}ML^2 \begin{pmatrix} \dot{\alpha} & \dot{\beta} \end{pmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \dot{\alpha} \\ \dot{\beta} \end{pmatrix} - \frac{1}{2}MgL \begin{pmatrix} \alpha & \beta \end{pmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{61}$$

The two Lagrange equations yield the matrix equation

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{pmatrix} \ddot{\alpha} \\ \ddot{\beta} \end{pmatrix} = -\frac{g}{L} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{62}$$

This is a generalized eigenvalue problem. The eigen-equation we get is given by

$$\det \left[\omega^2 \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} - \frac{g}{L} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \right] = \det \begin{bmatrix} 2(\omega^2 - \omega_0^2) & \omega^2 \\ \omega^2 & \omega^2 - \omega_0^2 \end{bmatrix} \tag{63}$$

$$= \omega^4 - 4\omega^2\omega_0^2 + 2\omega_0^4 = 0, \tag{64}$$

where $\omega_0 = \sqrt{g/L}$. The two eigen-frequencies are therefore

$$\omega = \sqrt{2 \pm \sqrt{2}}\,\omega_0 = \sqrt{(2 \pm \sqrt{2})\frac{g}{L}}. \tag{65}$$

The two corresponding eigenmodes satisfy $\beta = \mp\sqrt{2}\alpha$. The mode $\beta = -\sqrt{2}\alpha$ has a higher frequency $\sqrt{2+\sqrt{2}}\,\omega_0$ and the two pendulums oscillate in opposite directions. The mode $\beta = \sqrt{2}\alpha$ has a lower frequency $\sqrt{2-\sqrt{2}}\,\omega_0$ and the two pendulums oscillate in the same direction. The general small oscillation can be a superposition of the two eigenmodes. In a typical experiment, the high-frequency mode decays away faster due to air resistance and the oscillation decays towards the low-frequency mode.

The Lagrangian method is very useful and important in solving classical mechanical problems. Its advantage mainly reveals when the system is subject to complicated but frictionless constraints and the constraint forces aren't relevant to the problem. Following the Lagrangian method, the procedure is standard and the result is reliable. But one still needs to be aware of the limitations of the method. It is not applicable, for example, to dissipative systems (with friction) or open systems (with changing mass). These problems need to be solved using other methods.

3.5 Problem 5 in Sec. 1, 2012

Find the net acceleration of a raindrop falling through a cloud. Assume that the raindrop is a tiny sphere, and the cloud consists of similar small droplets that are uniformly suspended at rest; when the falling drop hits a droplet within the cloud, the two merge into a combined, still spherical, drop, which continues falling.

Answer: There are three parameters that fully describe the status of the raindrop: its radius r, its mass m and its falling speed v. Notice that this is a changing mass problem. Neither Newton's second law nor Lagrangian mechanics is applicable here. To establish equations for r, m and v, we start with the most obvious ones:

$$m = \frac{4}{3}\pi r^2 \rho_w, \qquad \frac{dm}{dt} = \pi r^2 v \rho, \tag{66}$$

where ρ_w is the density of water and ρ is the density of the cloud. For each tiny bit of mass dm added to the raindrop, we have from the conservation of vertical momentum

$$(m + dm)(v + dv) - mv - dm \cdot 0 = mg dt, \tag{67}$$

because the velocity of dm before hitting the raindrop is zero. Therefore we have

$$\frac{d(mv)}{dt} = mg. \tag{68}$$

You may feel that this equation is intuitively true. But it should not be taken for granted, because it depends on that the velocity of dm before hitting the raindrop be zero. Using $m = \frac{4}{3}\pi r^3 \rho_w$, we may cancel the mass m and only keep two dynamical variables r and v. We have from the other two equations

$$\dot{r} = \frac{\rho}{4\rho_w} v, \qquad \dot{v} = g - \frac{3v}{r}\dot{r}. \tag{69}$$

To solve the two coupled ODEs for r and v, we may plug the first ODE into the second one to cancel v and obtain

$$\ddot{r} = \frac{1}{2}\frac{d\dot{r}^2}{dr} = \frac{\rho g}{4\rho_w} - \frac{3}{r}\dot{r}^2. \tag{70}$$

Now we have a first-order linear ODE, the general form of which is given by

$$\frac{dy}{dx} = P(x)y + Q(x). \tag{71}$$

Using the similarity relationship of differential operators

$$\frac{d}{dx} - P(x) = \exp\left[\int_{x_0}^{x} P(x')dx'\right] \frac{d}{dx} \exp\left[-\int_{x_0}^{x} P(x')dx'\right], \tag{72}$$

we obtain the general solution formula

$$y(x) = \int_{x_0}^{x} Q(x')dx' \exp\left[\int_{x'}^{x} P(x'')dx''\right] + C \exp\left[\int_{x_0}^{x} P(x')dx'\right], \tag{73}$$

where $C = y(x_0)$ is an integration constant to be determined by the initial condition. Using this general solution formula, we obtain from Eq. (70) that

$$\dot{r}^2 = \frac{\rho g}{14\rho_w} r + \frac{C}{r^6}. \tag{74}$$

Given that initially the raindrop has $r = 0$ and $v = 0$ so that $\dot{r} = \frac{\rho}{4\rho_w} v = 0$, we get $C = 0$. Therefore, we have

$$\frac{dr}{dt} = \sqrt{\frac{\rho g}{14\rho_w} r} \quad \Rightarrow \quad r = \frac{\rho g}{56\rho_w} t^2, \tag{75}$$

given that initially $r = 0$ at $t = 0$. Therefore,

$$v = \frac{4\rho_w}{\rho} \dot{r} = \frac{1}{7} g t, \tag{76}$$

which means the acceleration of the raindrop is $\frac{1}{7} g$. Interestingly the result is independent of the cloud density ρ/ρ_w relative to the density of condensed raindrops. Compared with the standard Lagrangian procedure to follow for studying closed conservative systems, the dynamics of open systems can be tricky but interesting to solve.

Lecture 2 Electrostatics and Radiation

1 Poisson equation in electrostatics

The most fundamental equation in electrostatics is the Poisson equation. Given the total charge distribution $\rho(\vec{r})$, the electrostatic potential $\phi(\vec{r})$ satisfies

$$\nabla^2\phi(\vec{r}) = -\frac{\rho(\vec{r})}{\epsilon_0}. \tag{1}$$

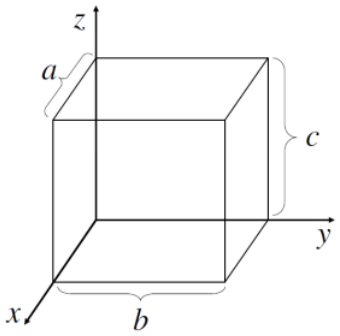
The solution can be obtained using Fourier transform:

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0}\int d^3r' \frac{\rho(\vec{r}')}{|\vec{r}-\vec{r}'|}, \tag{2}$$

in case that we know the charge distribution $\rho(\vec{r})$ in the whole 3D space. Unfortunately, in most electrostatic problems, we only know $\rho(\vec{r})$ in a finite region and some boundary conditions of ϕ. For general shapes of boundary, solution can be extremely hard. Analytically solvable cases are mostly 3 types: the box, the sphere, and the cylinder.

1.1 The box

1. (Problem 4 in Sec. 2, 2011) Consider a box with side lengths a, b, and c along the x, y, and z axes. Suppose there is no electric charge in the box and that $\phi = 0$ on the surface of the box except at $z = 0$ where $\phi = V_1 \sin\frac{\pi x}{a}\sin\frac{\pi y}{b}$ and at $z = c$ where $\phi = V_2 \sin\frac{2\pi x}{a}\sin\frac{2\pi y}{b}$. Find ϕ everywhere in the box.



Answer: Electrostatic problems in the box are comparatively easier than the sphere and the cylinder as they involve separation of variables along different axes and no special

functions (only trigonometry and exponentials). In this problem, since there is no charge inside the box, the inhomogeneous Poisson equation (1) becomes the homogeneous Laplace equation $\nabla^2\phi(\vec{r}) = 0$. In Cartesian coordinates, we have

$$\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2} = 0. \tag{3}$$

And the boundary conditions are given by

$$\phi(x = 0, y, z) = \phi(x = a, y, z) = 0, \tag{4}$$

$$\phi(x, y = 0, z) = \phi(x, y = b, z) = 0, \tag{5}$$

$$\phi(x, y, z = 0) = V_1 \sin\frac{\pi x}{a} \sin\frac{\pi y}{b}, \tag{6}$$

$$\phi(x, y, z = c) = V_2 \sin\frac{2\pi x}{a} \sin\frac{2\pi y}{b}. \tag{7}$$

The boundary conditions on the x and y faces are homogeneous while on the z face they are inhomogeneous. We may use the separation of variables

$$\phi(x, y, z) = X(x)Y(y)Z(z). \tag{8}$$

Then, from the Laplace equation and the boundary conditions, we have

$$X(x) = \sin\frac{m\pi x}{a}, \quad Y(y) = \sin\frac{n\pi y}{b}, \quad m, n = 1, 2, 3, \ldots, \tag{9}$$

and along the z direction we have

$$\frac{d^2Z}{dz^2} = \pi^2\left(\frac{m^2}{a^2} + \frac{n^2}{b^2}\right)Z, \tag{10}$$

which has the general solution

$$Z(z) = a_{mn}\exp\left(\frac{\pi z}{ab}\sqrt{n^2a^2 + m^2b^2}\right) + b_{mn}\exp\left(-\frac{\pi z}{ab}\sqrt{n^2a^2 + m^2b^2}\right). \tag{11}$$

Therefore, the general solution of $\phi(x, y, z)$ is

$$\phi(x, y, z) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left(a_{mn}e^{\frac{\pi z}{ab}\sqrt{n^2a^2+m^2b^2}} + b_{mn}e^{-\frac{\pi z}{ab}\sqrt{n^2a^2+m^2b^2}}\right)\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}. \tag{12}$$

Using Eqs. (6) & (7), we obtain for the a_{11} and b_{11} modes that

$$\begin{cases} a_{11} + b_{11} = V_1, & (13) \\ a_{11}e^{\frac{\pi\sqrt{a^2+b^2}}{ab}c} + b_{11}e^{-\frac{\pi\sqrt{a^2+b^2}}{ab}c} = 0, & (14) \end{cases}$$

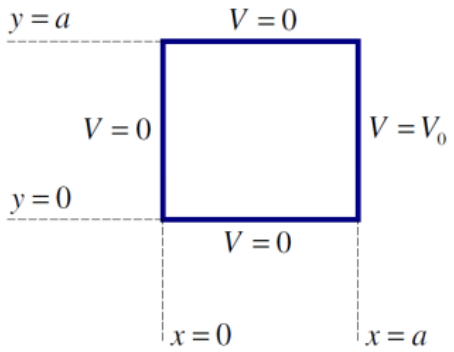
and for the a_{22} and b_{22} modes that

$$\begin{cases} a_{22} + b_{22} = 0, & (15) \\ a_{22} e^{\frac{2\pi\sqrt{a^2+b^2}}{ab}c} + b_{22} e^{-\frac{2\pi\sqrt{a^2+b^2}}{ab}c} = V_{22}. & (16) \end{cases}$$

From the orthogonality of different eigenfunctions $\sin\frac{m\pi x}{a}\sin\frac{n\pi y}{b}$, all other coefficients a_{mn} and b_{mn} are zero. We may now find the 4 nonzero coefficients $a_{11}, b_{11}, a_{22}, b_{22}$ to determine the specific solution:

$$\phi = \frac{V_1 \sinh\left[\frac{\pi\sqrt{a^2+b^2}}{ab}(c-z)\right]}{\sinh\left[\frac{\pi\sqrt{a^2+b^2}}{ab}c\right]} \sin\frac{\pi x}{a}\sin\frac{\pi y}{b} + \frac{V_2 \sinh\left[\frac{2\pi\sqrt{a^2+b^2}}{ab}z\right]}{\sinh\left[\frac{2\pi\sqrt{a^2+b^2}}{ab}c\right]} \sin\frac{2\pi x}{a}\sin\frac{2\pi y}{b}. \quad (17)$$

2. (Problem 5 in Sec. 2, 2012) Consider an infinite pipe with a square cross section, as is drawn in the figure, with three sides grounded and one side at potential V_0. Calculate the potential everywhere inside the pipe.



Answer: In this problem, $\phi(x, y, z) = \phi(x, y)$ is z-independent. The problem is therefore essentially 2D. Using the separation of variables $\phi(x, y) = X(x)Y(y)$, we obtain the general solution based on $\phi(x, y = 0 \text{ or } a) = 0$:

$$\phi(x, y) = \sum_{n=1}^{\infty} \left(a_n e^{\frac{n\pi x}{a}} + b_n e^{-\frac{n\pi x}{a}}\right) \sin\frac{n\pi y}{a}. \quad (18)$$

To satisfy $\phi(x = 0, y) = 0$, we require $a_n + b_n = 0$ for all n. This leads to

$$\phi(x, y) = \sum_{n=1}^{\infty} c_n \sinh\frac{n\pi x}{a}\sin\frac{n\pi y}{a}. \quad (19)$$

To determine the coefficients c_n, we need the fourth edge $x = a$. We have

$$\phi(x = a, y) = V_0 = \sum_{n=1}^{\infty} c_n \sinh(n\pi)\sin\frac{n\pi y}{a}. \quad (20)$$

Using the orthogonality of $\sin\frac{n\pi y}{a}$, we obtain

$$\int_0^a V_0 \sin\frac{n\pi y}{a} dy = c_n \sinh(n\pi) \int_0^a \sin^2\frac{n\pi y}{a} dy. \tag{21}$$

Therefore, the coefficients are found to be

$$c_n = \frac{2V_0[1-(-1)^n]}{n\pi \sinh(n\pi)}. \tag{22}$$

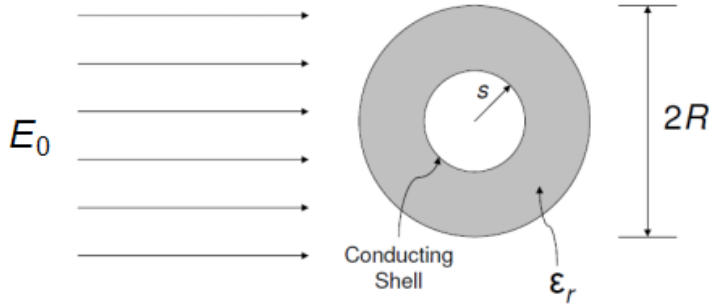
Since c_n is only nonzero for odd n, the final result can be simplified to

$$\phi = \frac{4V_0}{\pi} \sum_{n=0}^{\infty} \frac{\sinh\left(\frac{2n+1}{a}\pi x\right) \sin\left(\frac{2n+1}{a}\pi y\right)}{(2n+1)\sinh[(2n+1)\pi]}. \tag{23}$$

At $x = a$, the two sinh functions in the numerator and denominator cancel and you recognize the familiar trigonometric expansion of a square wave. So everything works. Later in the question, you are asked to find $\phi(x, y)$ when the boundary value $\phi(x, y = a) = V_0$ is also nonzero. This seems to have complicated the problem a lot, since we no longer have a pair of homogeneous boundary conditions in either x or y direction. But we can actually solve each nonzero edge separately and add up the two solutions by linearity.

1.2 The sphere

1. (Problem 2 in Sec. 2, 2012) A dielectric sphere of radius R is hollowed out in the region $0 \le r \le s$ and a thin, grounded, conducting shell is inserted at $r = s$. The sphere is placed in a uniform, external E-field $E = E_0\hat{z}$ along the z axis. The dielectric constant of the hollowed sphere is ϵ_r. Calculate the potential in the region $r \ge R$.



Answer: Unless they are spherically symmetric, electrostatic problems of the sphere typically involve the Legendre polynomials. The potential $\phi(r, \theta)$ in this problem has cylindrical symmetry around the z axis. Inside the grounded conducting shell, we have $\phi(r, \theta) = 0$ for $r \le s$. Outside the conducting shell, we have the Laplace equation $\nabla^2\phi = 0$ for both regions of $s < r < R$ and $r > R$. At $r = R$, due to the continuity condition of the normal

component of the electric displacement $\vec{D}$, we have

$$\epsilon_r \left.\frac{\partial \phi}{\partial r}\right|_{R-0} = \left.\frac{\partial \phi}{\partial r}\right|_{R+0}. \tag{24}$$

This will cause discontinuity of $\partial\phi/\partial r$ at $r = R$, while the continuity of ϕ at $r = R$ implies the continuity of the tangential component of the electric field $\vec{E}$. Now we need to solve the Laplace equation using the separation of variables in spherical coordinates. We have

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \phi}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \phi}{\partial \theta}\right) = 0. \tag{25}$$

Plugging in $\phi(r, \theta) = R(r)\,\Theta(\theta)$, we find that $\Theta(\theta) = P_l(\cos\theta)$ is a Legendre polynomial of $\cos\theta$ and the corresponding radial function $R(r)$ satisfies

$$r^2 \frac{d^2 R}{dr^2} + 2r\frac{dR}{dr} - l(l+1)R = 0. \tag{26}$$

This is an Euler differential equation. By doing a change of variable $r = e^u$, the varying coefficients become constant. We have

$$\frac{d^2 R}{du^2} + \frac{dR}{du} - l(l+1)R = 0. \tag{27}$$

There are two linearly independent solutions that form the general radial function

$$R = ae^{lu} + be^{-(l+1)u} = ar^l + \frac{b}{r^{l+1}}. \tag{28}$$

The general solution of $\phi(r, \theta)$ is therefore given by

$$\phi(r,\theta) = \begin{cases} \displaystyle\sum_{l=0}^{\infty} \left(a_l r^l + \frac{b_l}{r^{l+1}}\right) P_l(\cos\theta), & r \geq R, \\ \displaystyle\sum_{l=0}^{\infty} \left(c_l r^l + \frac{d_l}{r^{l+1}}\right) P_l(\cos\theta), & s \leq r \leq R. \end{cases} \tag{29}$$

At $r = s$, we require that $\phi(r, \theta) = 0$. Therefore, we have

$$c_l s^l + \frac{d_l}{s^{l+1}} = 0, \qquad l = 0, 1, 2, \ldots \tag{30}$$

At $r = R$, we require that ϕ is continuous. So we have

$$a_l R^l + \frac{b_l}{R^{l+1}} = c_l R^l + \frac{d_l}{R^{l+1}}, \qquad l = 0, 1, 2, \ldots \tag{31}$$

And the continuity condition in Eq. (24) gives

$$la_l R^{l-1} - \frac{(l+1)b_l}{R^{l+2}} = \epsilon_r \left[lc_l R^{l-1} - \frac{(l+1)d_l}{R^{l+2}} \right]. \tag{32}$$

Finally, by looking at the asymptotic behavior of

$$\phi(r,\theta) \approx -E_0 z = -E_0 r\cos\theta = -E_0 P_1(\cos\theta), \quad r \to \infty, \tag{33}$$

we have $a_l = -E_0\delta_{l1}$. Only the $l = 1$ component is nonzero. We now have four equations for a_1, b_1, c_1, d_1. We find that

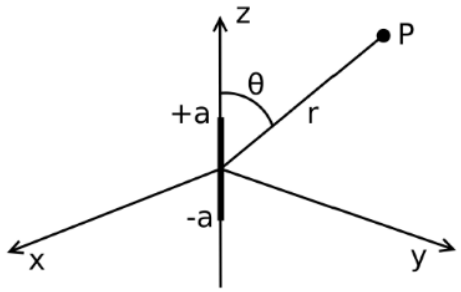
$$a_1 = -E_0, \quad b_1 = \frac{(\epsilon_r - 1)R^3 + (2\epsilon_r + 1)s^3}{(2+\epsilon_r)R^3 + 2(\epsilon_r - 1)s^3} E_0 R^3, \tag{34}$$

$$c_1 = -\frac{3E_0R^3}{(2+\epsilon_r)R^3 + 2(\epsilon_r - 1)s^3}, \quad d_1 = \frac{3E_0R^3s^3}{(2+\epsilon_r)R^3 + 2(\epsilon_r - 1)s^3}. \tag{35}$$

The potential ϕ outside the dielectric sphere is then given by

$$\phi = -E_0 r\cos\theta \left[1 - \frac{(\epsilon_r - 1)R^3 + (2\epsilon_r + 1)s^3}{(2+\epsilon_r)R^3 + 2(\epsilon_r - 1)s^3} \cdot \frac{R^3}{r^3} \right], \quad r \geq R. \tag{36}$$

2. (Problem 2 in Sec. 2, 2013) A linear uniform charge distribution of λ coulomb/meter extends along the z axis from $z = -a$ to $z = a$. Find a series expansion for $V(r,\theta)$ in terms of Legendre polynomials valid to all orders for $r > a$.



Answer: The solution is a straightforward integral of the Coulomb potential

$$\phi(r,\theta) = \frac{1}{4\pi\epsilon_0} \int_{-a}^{a} \frac{\lambda dz}{\sqrt{r^2 + z^2 - 2rz\cos\theta}}. \tag{37}$$

In the region $r > a$, we may Taylor expand the integrand using the generating function of the Legendre polynomials

$$\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n=0}^{\infty} P_n(x)t^n, \qquad |x| < 1, |t| < 1. \tag{38}$$

Physically this mathematical formula gives us the multipole expansion

$$\phi = \frac{\lambda}{4\pi\epsilon_0 r}\sum_{n=0}^{\infty} P_n(\cos\theta)\int_{-a}^{a}\left(\frac{z}{r}\right)^n dz = \frac{\lambda}{2\pi\epsilon_0}\sum_{n=0}^{\infty}\frac{P_{2n}(\cos\theta)}{2n+1}\left(\frac{a}{r}\right)^{2n+1}. \tag{39}$$

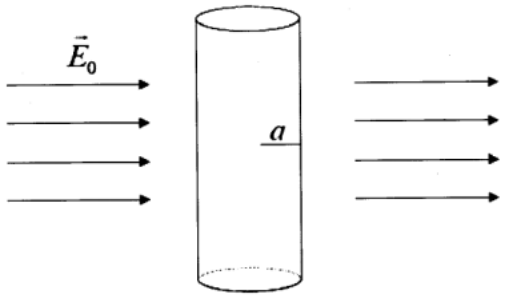
It is recommended to review these related topics to prepare for the Quals: the first few orders of Legendre polynomials, how they appear in the Poisson equation, their recursive relations, series expansions, Rodriguez formula, orthogonality relations, etc. This problem requires the generating function of Legendre polynomials.

1.3 The cylinder

1. (Problem 3 in Sec. 2, 2010) Consider an infinitely long, grounded conducting cylinder, of radius a, which is introduced into a uniform electric field $\vec{E}_0$. The axis of the cylinder is perpendicular to $\vec{E}_0$.

(a) Find an expression for the external potential after insertion of the cylinder;

(b) Find an expression for the surface charge induced on the cylinder.



Answer: (a) The problem involves solving the Laplace equation $\nabla^2\phi = 0$ in cylindrical coordinates, which is given by

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} = 0. \tag{40}$$

In the region $r > a$, using the separation of variables $\phi(r,\theta) = R(r)\,\Theta(\theta)$, we get $\Theta(\theta) = \sin n\theta$ or $\cos n\theta$ and the radial function $R(r)$ satisfies

$$r\frac{d}{dr}\left(r\frac{dR}{dr}\right) = n^2 R \quad\Rightarrow\quad R = ar^n + \frac{b}{r^n}, \qquad n \geq 1. \tag{41}$$

In case that $n = 0$, we have $R = a + b\ln r$. The general solution is therefore given by

$$\phi = a_0 + b_0\ln r + \sum_{n=1}^{\infty}\left[\left(a_n r^n + \frac{a_{-n}}{r^n}\right)\cos n\theta + \left(b_n r^n + \frac{b_{-n}}{r^n}\right)\sin n\theta\right]. \tag{42}$$

Defining $\theta = 0$ as the direction of $\vec{E}_0$, we have the asymptotic behavior

$$\phi(r, \theta) \approx -E_0 r \cos\theta, \quad r \to \infty. \tag{43}$$

Therefore $a_n = -E_0 \delta_{n1}$, so the only nonzero terms will be a_1 and a_{-1}. At $r = a$, we have $\phi = 0$. This gives us $a_{-1} = -a^2 a_1 = a^2 E_0$, so we have

$$\phi(r, \theta) = -E_0 \cos\theta \left(r - \frac{a^2}{r} \right). \tag{44}$$

(b) The surface charge induced can be found by Gauss's law:

$$\sigma(r = a, \theta) = -\epsilon_0 \left. \frac{\partial \phi}{\partial r} \right|_{a+0} = 2\epsilon_0 E_0 \cos\theta. \tag{45}$$

Positive charge gets repelled to the $\theta = 0$ side while negative charge gets pulled to the $\theta = \pi$ side. Not every cylinder involves Bessel functions. This problem is special because the potential ϕ has no z dependence. Bessel functions will typically not appear on Quals.

2 Radiation using multipole expansion

2.1 Fields in the radiation zone

In Gaussian units, Maxwell's equations in vacuum are given in Gaussian units by

$$\nabla \cdot \vec{E} = 4\pi\rho, \qquad \nabla \cdot \vec{B} = 0, \tag{46}$$

$$\nabla \times \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}, \qquad \nabla \times \vec{B} = \frac{4\pi}{c} \vec{j} + \frac{1}{c} \frac{\partial \vec{E}}{\partial t}. \tag{47}$$

After introducing the electromagnetic potentials $\vec{A}$ and ϕ, the fields are expressed asf

$$\vec{E} = -\nabla\phi - \frac{1}{c} \frac{\partial \vec{A}}{\partial t}, \qquad \vec{B} = \nabla \times \vec{A}. \tag{48}$$

Using the Lorentz gauge condition

$$\nabla \cdot \vec{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} = 0, \tag{49}$$

the potentials $\vec{A}$ and ϕ satisfy the d'Alembert wave equations

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = 4\pi\rho, \qquad \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2 \vec{A} = \frac{4\pi}{c} \vec{j}. \tag{50}$$

Radiation theory will focus on the vector potential $\vec{A}$. The retarded solution of the d'Alembert equation is given by

$$\vec{A}(\vec{x},t) = \frac{1}{c}\int \frac{d^3x'}{|\vec{x}-\vec{x}\,'|}\,\vec{j}\left(\vec{x}\,', t' = t - \frac{|\vec{x}-\vec{x}\,'|}{c}\right). \tag{51}$$

In the radiation zone, based on $|\vec{x}| \gg |\vec{x}\,'|$ we may do two approximations

$$t' = t - \frac{|\vec{x}-\vec{x}\,'|}{c} \approx t - \frac{|\vec{x}|}{c} + \left(\nabla\frac{|\vec{x}|}{c}\right)\cdot\vec{x}\,' = t - \frac{r}{c} + \hat{r}\cdot\frac{\vec{x}\,'}{c}, \tag{52}$$

where $r = |\vec{x}|$ and $\hat{r} = \vec{x}/|\vec{x}|$, and the other approximation being $\frac{1}{|\vec{x}-\vec{x}\,'|} \approx \frac{1}{|\vec{x}|}$, to obtain the vector potential in the radiation zone given by

$$\vec{A}(\vec{x},t) \approx \frac{1}{rc}\int \vec{j}\left(\vec{x}\,', t' = t - \frac{r}{c} + \frac{\hat{r}\cdot\vec{x}\,'}{c}\right) d^3x'. \tag{53}$$

Q Then the $\vec{E}$ and $\vec{B}$ fields and the energy flow are given by

$$\vec{B} = -\frac{1}{c}\,\hat{r}\times\frac{\partial\vec{A}}{\partial t}, \quad \vec{E} = \frac{1}{c}\,\hat{r}\times\left(\hat{r}\times\frac{\partial\vec{A}}{\partial t}\right), \quad \vec{S} = \frac{c}{4\pi}\,\vec{E}\times\vec{B} = \frac{\hat{r}}{4\pi c}\left|\hat{r}\times\frac{\partial\vec{A}}{\partial t}\right|^2. \tag{54}$$

We can now calculate the fields and the power distribution from these formulas.

2.2 Electric dipole radiation

If the typical length scale of the radiating source is much smaller than the wave length of the radiation, we can do a Taylor expansion of $\hat{r}\cdot\vec{x}\,'/c$ and obtain from Eq. (53) that

$$\vec{A}(\vec{x},t) = \frac{1}{rc}\int d^3x' \sum_{n=0}^{\infty}\frac{1}{n!}\left.\frac{\partial^n\vec{j}(\vec{x}\,',t')}{\partial t'^n}\right|_{t'=t-\frac{r}{c}}\left(\frac{\hat{r}\cdot\vec{x}\,'}{c}\right)^n \tag{55}$$

$$= \frac{1}{rc}\sum_{n=0}^{\infty}\frac{1}{n!}\left.\frac{\partial^n}{\partial t'^n}\right|_{t'=t-\frac{r}{c}}\int \vec{j}(\vec{x}\,',t')\left(\frac{\hat{r}\cdot\vec{x}\,'}{c}\right)^n d^3x'. \tag{56}$$

This simplifies the retardation to $t' = t - \frac{r}{c}$ and all we need is find the nth order moment of the current distribution $\vec{j}(\vec{x}\,',t')$. In most radiation problems in the uals, we only need to look at the leading order term in the above expansion, which is the $n = 0$ term

$$\vec{A}_0(\vec{x},t) = \frac{1}{rc}\int \vec{j}(\vec{x}\,',t')\,d^3x'. \tag{57}$$

This term is called the *electric dipole radiation*, because the integral of current density over the whole 3D space is equal to the change rate of the electric dipole moment, which can be

obtained from charge conservation. We have

$$\frac{d}{dt'}\int \vec{x}\,'\rho(\vec{x}\,',t')d^3x' = \int \vec{x}\,'\frac{\partial\rho(\vec{x}\,',t')}{\partial t'}d^3x' = -\int \vec{x}\,'\nabla'\cdot\vec{j}(\vec{x}\,',t')d^3x' \tag{58}$$

$$= -\int \nabla'\cdot[\vec{j}(\vec{x}\,',t')\vec{x}\,']d^3x' + \int \vec{j}(\vec{x}\,',t')\cdot(\nabla'\vec{x}\,')d^3x' = \int \vec{j}(\vec{x}\,',t')d^3x', \tag{59}$$

where we have used the formulas in vector calculus

$$\nabla\cdot(\vec{A}\vec{B}) = \partial_i(A_iB_j)\vec{e}_j = \partial_i(A_i)B_j\vec{e}_j + A_i\partial_i(B_j\vec{e}_j) = (\nabla\cdot\vec{A})\vec{B} + \vec{A}\cdot\nabla\vec{B}, \tag{60}$$

$$\nabla\vec{x} = \vec{e}_i\partial_i x_j\vec{e}_j = \vec{e}_i\delta_{ij}\vec{e}_j = \vec{e}_i\vec{e}_i = \overleftrightarrow{1}, \tag{61}$$

and that the total divergence term integrated over the whole 3D space is zero because

$$\int \nabla'\cdot[\vec{j}(\vec{x}\,',t')\vec{x}\,']d^3x' = \lim_{V\to\infty}\oint_{\partial V} d\vec{S}\cdot\vec{j}(\vec{x}\,',t')\vec{x}\,' = 0, \tag{62}$$

following from our locality assumption of the radiating source. We define the *electric dipole moment* of the source as

$$\vec{p}(t') = \int \vec{x}\,'\rho(\vec{x}\,',t')d^3x', \tag{63}$$

so that the leading-order term of the vector potential becomes

$$\vec{A}_0(\vec{x},t) = \frac{1}{rc}\int \vec{j}(\vec{x}\,',t')d^3x' = \frac{1}{rc}\frac{d\vec{p}(t')}{dt'}\bigg|_{t'=t-\frac{r}{c}}. \tag{64}$$

Since $t' = t - \frac{r}{c}$ only differs from t by a constant, we have $dt = dt'$. Therefore, the vector potential under the electric dipole approximation can be written as

$$\vec{A}(\vec{x},t) \approx \vec{A}_0(\vec{x},t) = \frac{1}{rc}\dot{\vec{p}}|_{t'=t-\frac{r}{c}}. \tag{65}$$

From this, the $\vec{E}$ and $\vec{B}$ fields are found to be

$$\vec{E}(\vec{x},t) = \frac{1}{c}\hat{r}\times\left(\hat{r}\times\frac{\partial\vec{A}}{\partial t}\right) = \frac{1}{rc^2}\hat{r}\times(\hat{r}\times\ddot{\vec{p}})|_{t'=t-\frac{r}{c}}, \tag{66}$$

$$\vec{B}(\vec{x},t) = -\frac{1}{c}\hat{r}\times\frac{\partial\vec{A}}{\partial t} = -\frac{1}{rc^2}(\hat{r}\times\ddot{\vec{p}})|_{t'=t-\frac{r}{c}}. \tag{67}$$

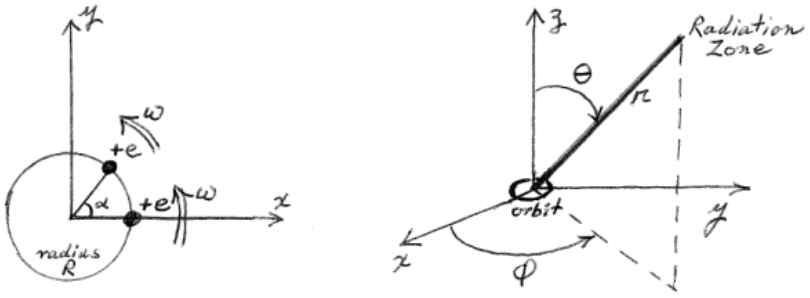
Finally, the angular distribution of radiation power is given by

$$\frac{dP}{d\Omega} = \frac{c}{4\pi}r^2(\vec{E}\times\vec{B}) = \frac{1}{4\pi c^3}|\hat{r}\times\ddot{\vec{p}}|^2\bigg|_{t'=t-\frac{r}{c}}. \tag{68}$$

The next order of expansion $n = 1$ gives us the magnetic dipole and electric quadruple terms, and the radiation power can have cross terms between different orders.

2.3 Sample problems

1. (Problem 1 in Sec. 2, 2012) Two point particles, each having the same electric charge $+e$, move in the xy plane around the circumference of a circle with radius R. Both charges travel at the same angular velocity ω but maintain a fixed angular separation α throughout the motion. Assume that the motion of the particles is non-relativistic ($\omega R \ll c$). The particles radiate energy at distance r far from the circular orbit.



Find the time-averaged power radiated per solid angle in the (θ, ϕ) direction in the radiation zone ($r \gg R$). Hint: use multipole expansion.

Answer: The condition $\omega R \ll c$ enables us to do a multipole expansion in the radiation zone $r \gg R$. So long as the angular separation $\alpha \neq \pi$, there's going to be change in the electric dipole moment of the system as the particles move. Therefore an electric dipole approximation will be enough to solve the problem. We have

$$|\vec{p}| = 2eR\cos\frac{\alpha}{2}, \tag{69}$$

and this dipole moment $\vec{p}$ rotates about the origin at angular frequency ω. Therefore,

$$|\ddot{\vec{p}}| = 2eR\omega^2\cos\frac{\alpha}{2}, \tag{70}$$

and $\ddot{\vec{p}}$ also rotates about the origin at angular frequency ω. The component of $\ddot{\vec{p}}$ that is parallel to $\hat{r}$ can be found to be

$$(\ddot{\vec{p}})_\parallel = \ddot{\vec{p}}\cdot\hat{r} = |\ddot{\vec{p}}|\cos\phi'\sin\theta, \tag{71}$$

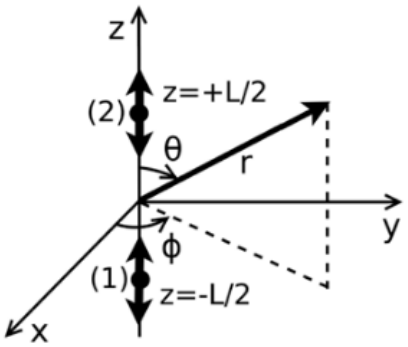
where ϕ' is the angle between $\ddot{\vec{p}}$ and the projection of $\hat{r}$ in the xy plane. This component generates no radiation power in the $\hat{r}$ direction. Therefore,

$$\langle(\ddot{\vec{p}})_\perp\rangle = \langle|\ddot{\vec{p}}\times\hat{r}|^2\rangle = |\ddot{\vec{p}}|^2 - \langle(\ddot{\vec{p}})_\parallel^2\rangle = |\ddot{\vec{p}}|^2\left(1-\frac{1}{2}\sin^2\theta\right), \tag{72}$$

using $\langle\cos^2\phi'\rangle = \frac{1}{2}$. Then from Eqs. (68), (70) & (72), we obtain the power distribution

$$\frac{dP}{d\Omega} = \frac{e^2R^2}{\pi c^3}\,\omega^4\cos^2\frac{\alpha}{2}\left(1-\frac{1}{2}\sin^2\theta\right). \tag{73}$$

2. (Problem 1 in Sec. 2, 2013) Consider two very tiny oscillating electric dipoles, which are located on the z axis, as is shown in the figure below.



The two dipoles is located at $z = \mp L/2$ and have time-varying electric dipole moments

$$\vec{D}_{1,2} = D_0 \hat{z} \cos\left(\omega t \mp \frac{\alpha}{2}\right) = \mathrm{Re}\left[D_0 \hat{z} e^{-i(\omega t \mp \frac{\alpha}{2})}\right]. \tag{74}$$

The two dipoles therefore oscillate in the same z direction with the same amplitude D_0 and the same angular frequency ω but are out of phase by phase angle α. Consider a field at point (r, θ, ϕ) in the radiation zone $r \to \infty$ and keep only the leading $1/r$ terms.

(a) Find the electric field $\vec{E}$ at (r, θ, ϕ) in the radiation zone.

(b) Find the time-averaged power detected per unit solid angle $dP/d\Omega$ as a function of the direction (θ, ϕ) in the radiation zone.

Answer: (a) From the electric dipole radiation formula in Eq. (65), the vector potential of this problem in the radiation zone is given by

$$\vec{A}(\vec{x}, t) = \frac{1}{rc}\dot{\vec{D}}_1\big|_{t'_1 = t - \frac{|\vec{x} - \vec{x}'_1|}{c}} + \frac{1}{rc}\dot{\vec{D}}_2\big|_{t'_2 = t - \frac{|\vec{x} - \vec{x}'_2|}{c}}, \tag{75}$$

where $\vec{x}'_1 = -\frac{L}{2}\hat{z}$ and $\vec{x}'_2 = \frac{L}{2}\hat{z}$ are the positions of the two dipoles. There are two retarded times t'_1 and t'_2 because the two dipoles are at different positions. We need to consider L because we are not told whether $\omega L \ll c$ or not. But we do know $r = |\vec{x}| \gg L$ since we are in the radiation zone. Therefore, the retarded times can be approximated as

$$t'_1 \approx t - \frac{r}{c} + \frac{\hat{r} \cdot \vec{x}'_1}{c} = t - \frac{r}{c} - \frac{L}{2c}\cos\theta. \tag{76}$$

$$t'_2 \approx t - \frac{r}{c} + \frac{\hat{r} \cdot \vec{x}'_2}{c} = t - \frac{r}{c} + \frac{L}{2c}\cos\theta. \tag{77}$$

Putting together Eqs. (74)–(77), we have the vector potential

$$\vec{A}(\vec{x},t) = \frac{1}{rc}\frac{d}{dt}\,\mathrm{Re}\Big[D_0\hat{z}e^{-i(\omega t_1' - \frac{\alpha}{2})} + D_0\hat{z}e^{-i(\omega t_2' + \frac{\alpha}{2})}\Big] \tag{78}$$

$$= \frac{2D_0\hat{z}}{rc}\frac{d}{dt}\,\mathrm{Re}\left[e^{-i\omega(t-\frac{r}{c})}\cos\left(\frac{\omega L}{2c}\cos\theta + \frac{\alpha}{2}\right)\right] \tag{79}$$

$$= -\frac{2\omega D_0\hat{z}}{rc}\sin\Big[\omega\left(t-\frac{r}{c}\right)\Big]\cos\left(\frac{\omega L}{2c}\cos\theta + \frac{\alpha}{2}\right). \tag{80}$$

The $\vec{E}$ and $\vec{B}$ fields are then obtained from Eqs. (66) & (67) to be

$$\vec{E}(\vec{x},t) = \frac{1}{c}\,\hat{r}\times\left(\hat{r}\times\frac{\partial\vec{A}}{\partial t}\right) = -\frac{2\omega^2 D_0\hat{\theta}}{rc^2}\cos\Big[\omega\left(t-\frac{r}{c}\right)\Big]\cos\left(\frac{\omega L}{2c}\cos\theta + \frac{\alpha}{2}\right)\sin\theta, \tag{81}$$

$$\vec{B}(\vec{x},t) = -\frac{1}{c}\,\hat{r}\times\frac{\partial\vec{A}}{\partial t} = -\frac{2\omega^2 D_0\hat{\phi}}{rc^2}\cos\Big[\omega\left(t-\frac{r}{c}\right)\Big]\cos\left(\frac{\omega L}{2c}\cos\theta + \frac{\alpha}{2}\right)\sin\theta, \tag{82}$$

where we have used $\hat{r}\times\hat{z} = -\hat{\phi}\sin\theta$ and $\hat{r}\times(\hat{r}\times\hat{z}) = \hat{\theta}\sin\theta$.

(b) The instantaneous radiation power is given by the Poynting vector

$$\vec{S} = \frac{c}{4\pi}\vec{E}\times\vec{B} = \frac{\omega^4 D_0^2}{\pi r^2 c^3}\hat{r}\cos^2\Big[\omega\left(t-\frac{r}{c}\right)\Big]\cos^2\left(\frac{\omega L}{2c}\cos\theta + \frac{\alpha}{2}\right)\sin^2\theta. \tag{83}$$

Therefore, the time-averaged radiation power is given by

$$\left\langle\frac{dP}{d\Omega}\right\rangle = r^2\langle\vec{S}\cdot\hat{r}\rangle = \frac{\omega^4 D_0^2}{2\pi c^3}\cos^2\left(\frac{\omega L}{2c}\cos\theta + \frac{\alpha}{2}\right)\sin^2\theta, \tag{84}$$

where $\langle\cos^2\big[\omega(t-\frac{r}{c})\big]\rangle = \frac{1}{2}$ when averaged over many periods. The above power distribution exhibits an interference pattern of two electric dipole radiators.

Lecture 3 Special Relativity and Covariant Electrodynamics

1 Special relativity theory

1.1 Lorentz transformation

Consider a 1D space, in which there are two inertial reference frames (x,t) and (x',t'). Suppose at $t=t'=0$, the origins of the two frames $x=0$ and $x'=0$ coincide. Due to the translational symmetry of spacetime, the spatiotemporal points of the two reference frames should be related by a linear transformation

$$x' = Ax + Bt, \qquad t' = Cx + Dt. \tag{1}$$

Suppose the origin $x'=0$ of frame (x',t') moves in the (x,t) frame at velocity v. Therefore $B=-vA$, so we have

$$x' = A(x - vt). \tag{2}$$

Then if the length and time scales of the two coordinate systems are set equal and the positive directions are the same, the origin $x=0$ of frame (x,t) should be moving in the (x',t') frame at velocity $-v$ in the low speed limit $v \ll c$. We can make an intuitive guess that this symmetry of speed measurement $v \leftrightarrow -v$ between two relatively moving frames still holds for high speed and therefore obtain $D=A$.

Here comes the crucial assumption: the *constancy* of speed of light. When the wavefront of a beam of light travels in frame (x,t) according to $x=ct$, its motion in the (x',t') frame will be given by $x'=ct'$, where the speed of light c is the same. This crucial assumption gives us $C=-vA/c^2$, so we have

$$t' = A\left(t - \frac{v}{c^2}x\right). \tag{3}$$

No speed measurements can determine the last parameter A, because it is a universal scaling factor that equally affects x' and t'. We can only determine A by some symmetry argument. We notice that when frame (x',t') moves relative to frame (x,t) at velocity v, so does frame $(-x,t)$ (a frame that remains stationary relative to frame (x,t) but has an x axis in the opposite direction) relative to frame $(-x',t')$. The transformation from $(-x',t')$ to $(-x,t)$ should therefore be the same as the transformation from (x,t) to (x',t'), i.e., the

transformation should be given by

$$-x = A(-x' - vt'), \qquad t = A\left(t' + \frac{vx'}{c^2}\right). \tag{4}$$

These equations are consistent with Eqs. (2) & (3) if the scaling factor A satisfies

$$A^2 \begin{bmatrix} 1 & -v \\ -\frac{v}{c^2} & 1 \end{bmatrix} \begin{bmatrix} 1 & v \\ \frac{v}{c^2} & 1 \end{bmatrix} = A^2 \begin{bmatrix} 1-\frac{v^2}{c^2} & 0 \\ 0 & 1-\frac{v^2}{c^2} \end{bmatrix} = I_{2\times 2}, \tag{5}$$

or equivalently,

$$A^2\left(1 - \frac{v^2}{c^2}\right) = 1. \tag{6}$$

The equation has no real solution if $|v| \geq c$, in which case the whole theory breaks down. To get out of trouble, we have to boldly assume $|v| < c$ for all moving objects. Then we can prove using the relativistic velocity-addition formula that all physically relevant reference frames, in which there are moving objects, also satisfy $|v| < c$. Remember relativity states that a reference frame with no moving objects in it is meaningless and metaphysical.

In the low speed limit $v \ll c$, we should return to the Galilean transformation. Hence $A(v \ll c) \approx 1$, which lets us keep the positive square root by continuity and obtain

$$x' = \gamma(x - vt), \qquad t = \gamma\left(t - \frac{vx}{c^2}\right), \tag{7}$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$ is the Lorentz factor. We have changed the symbol A to the more common symbol γ. Eq. (7) is called the Lorentz transformation. To generalize the results to 3D, we intuitively assume $y = y'$ and $z = z'$ and find that the constancy of speed of light is preserved, since we have

$$x^2 + y^2 + z^2 = c^2t^2 \quad \Leftrightarrow \quad x'^2 + y'^2 + z'^2 = c^2t'^2. \tag{8}$$

We have outlined a brief derivation, or more precisely, a construction of Einstein's theory of special relativity. Building physical theories is unlike proving mathematical theorems, in that we always have to make intuitive and reasonable assumptions, and extrapolations from what daily life experiences and experiments have confirmed at the time. Einstein's theory is of no exception. What makes him exceptional is his brilliance of guessing everything right in a beautifully consistent way.

1.2 Formulas in 1D relativistic kinematics

1.2.1 Length contraction

To measure the length of a moving ruler, we need to read its two endpoint coordinates x_1, x_2 at the same time t. According to the Lorentz transformation in Eq. (7), in the co-moving frame of the ruler that moves at velocity v relative to our lab frame, the rest length L_0 of

the ruler is longer than the measured moving length L. We have

$$L_0 = x_2' - x_1' = \gamma(x_2 - x_1) = \gamma L. \tag{9}$$

Therefore, the length of a moving ruler L is shorter than its rest length L_0 by a factor of $1/\gamma$. This effect is called *length contraction*.

1.2.2 Relativity of simultaneity

Two simultaneous events (x_1, t) and (x_2, t) can happen at different times

$$t_2' - t_1' = -\frac{\gamma v}{c^2}(x_2 - x_1), \tag{10}$$

in a different reference frame (x', t') moving at velocity v relative to frame (x, t). But this does not violate causality because

$$x_2' - x_1' = \gamma(x_2 - x_1), \tag{11}$$

so we always have

$$\left|\frac{x_2' - x_1'}{t_2' - t_1'}\right| = \frac{c^2}{v} > c \tag{12}$$

for any physically accessible reference frames ($v < c$). This means no causal link has enough time to get from (x_1', t_1') to (x_2', t_2'), or from (x_2', t_2') to (x_1', t_1'), whichever is earlier.

1.2.3 Time dilation

The time interval between two consecutive events (x, t_1) and (x, t_2) at the same position is viewed larger in a different reference frame. We have

$$\Delta t' = t_2' - t_1' = \gamma(t_2 - t_1) = \gamma \Delta t. \tag{13}$$

The interval Δ is called the *proper time*. It is often measured in the co-moving frame by a clock that travels with the moving object.

1.2.4 Velocity-addition formula

One may verify that two Lorentz transformations along x still give a Lorentz transformation. The resultant velocity is given by

$$v = \frac{v_1 + v_2}{1 + \frac{v_1 v_2}{c^2}}. \tag{14}$$

The derivation is a straightforward matrix multiplication. We have from multiplying two

Lorentz transformation matrices of Eq. (7) that

$$L(v_1)L(v_2) = \frac{1}{\sqrt{1-\frac{v_1^2}{c^2}}}\begin{pmatrix} 1 & -v_1 \\ -\frac{v_1}{c^2} & 1 \end{pmatrix} \frac{1}{\sqrt{1-\frac{v_2^2}{c^2}}}\begin{pmatrix} 1 & -v_2 \\ -\frac{v_2}{c^2} & 1 \end{pmatrix} \tag{15}$$

$$= \frac{1}{\sqrt{(1-\frac{v_1^2}{c^2})(1-\frac{v_2^2}{c^2})}}\begin{pmatrix} 1+\frac{v_1 v_2}{c^2} & -(v_1+v_2) \\ -\frac{v_1+v_2}{c^2} & 1+\frac{v_1 v_2}{c^2} \end{pmatrix} = L(v), \tag{16}$$

with the resultant velocity v given in Eq. (14). The result velocity v of two subluminal velocities $|v_1| < c$ and $|v_2| < c$ is therefore always less than c. This proves that if all moving objects are subluminal in all reference frames, then all reference frames that have moving objects in them are also subluminal to us. This rescues our theory for now. But we will definitely need a new theory if a superluminal object is found someday.

2 Covariant form

We have now established a new set of rules for spacetime. Galilean transformation is replaced by Lorentz transformation and the speed of light c in vacuum now becomes reference frame independent. Consequently, old physical laws like Newton's laws of motion that are covariant under Galilean transformation have to be modified into new laws that are Lorentz covariant. The correctness of the new laws will then have to be tested experimentally. Other physical laws like Maxwell's equations that are already Lorentz covariant will be kept. To make this sifting of physical laws easier, we introduce the covariant form of 4-vectors.

2.1 4-vectors

The spacetime vector $\mathrm{x} = (ct, \vec{x})$ is a 4-vector. It transforms under a Lorentz boost of velocity v in the x direction as

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}, \tag{17}$$

where $\beta = v/c$, $\gamma = 1/\sqrt{1-\beta^2}$. All 4-vectors transform exactly the same way. We may construct may other 4-vectors from the spacetime vector $\mathrm{x} = (ct, \vec{x})$. We have, for example, the 4-momentum of a particle

$$\mathrm{p} = m_0 \frac{d\mathrm{x}}{d\tau} = \gamma m_0 \frac{d(ct, \vec{x})}{dt} = (\gamma m_0 c, \gamma m_0 \vec{v}), \tag{18}$$

where the rest mass m_0 and the proper time $d\tau$ are Lorentz scalars, i.e., quantities that remain unchanged under Lorentz transformations. The proper time $d\tau$ is measured in the co-moving frame of the particle, and that $dt = \gamma d\tau$ is due to time dilation.

In Newtonian mechanics, the total $m_0\vec{v}$ of two colliding particles is conserved. But since $(m_0c, m_0\vec{v})$ without γ is not a 4-vector, the conservation of total $m_0\vec{v}$ and total m_0 is in general not true in a different reference frame. We therefore modify the law to the conservation of the total 4-momentum $\mathrm{p} = (\gamma m_0 c, \gamma m_0 \vec{v})$ and this new law is justified experimentally. The new law will mean that the mass of a moving particle is γm_0, greater than its rest mass m_0 by a factor of γ that depends on the particle's velocity v.

The 4-current (density) given by

$$\mathrm{j} = \rho_0 \frac{d\mathrm{x}}{d\tau} = \gamma\rho_0 \frac{d(ct, \vec{x})}{dt} = (\rho c, \vec{j}) \tag{19}$$

is a 4-vector, where ρ_0 is the proper charge density in the co-moving frame of the charge. Charge is scalar, but volume is subject to Lorentz contraction. Therefore $\rho = \gamma\rho_0$. The 4-vectors j and p are constructed in very similar ways.

Finally, in the next section, from Maxwell's equations, we will find that the 4-potential $\mathrm{A} = (\phi, \vec{A})$ in Gaussian units is a 4-vector. The derivation uses the pseudo inner product (length is not positive definite)

$$\mathrm{a} \cdot \mathrm{b} = a_0 b_0 - \vec{a} \cdot \vec{b} \tag{20}$$

of two 4-vectors $\mathrm{a} = (a_0, \vec{a})$ and $\mathrm{b} = (b_0, \vec{b})$. This inner product is useful because Lorentz transformation leaves the speed of light invariant, i.e.,

$$c^2t^2 - x^2 - y^2 - z^2 = 0 \quad \Leftrightarrow \quad c^2t'^2 - x'^2 - y'^2 - z'^2 = 0. \tag{21}$$

This condition can be written as $\mathrm{x} \cdot \mathrm{x} = \mathrm{x}' \cdot \mathrm{x}'$ using the inner product, which, to a linear transformation, will mean that it leaves the inner product

$$\mathrm{a} \cdot \mathrm{b} = \frac{1}{2}\left[(\mathrm{a} + \mathrm{b}) \cdot (\mathrm{a} + \mathrm{b}) - \mathrm{a} \cdot \mathrm{a} - \mathrm{b} \cdot \mathrm{b}\right] \tag{22}$$

of any two 4-vectors a and b invariant, so that $\mathrm{a} \cdot \mathrm{b}$ is Lorentz scalar. From this technique you can derive, for example, that the 4-wavenumber $\mathrm{k} = (\frac{\omega}{c}, \vec{k})$ of a beam of light is a 4-vector, because the phase $\mathrm{k} \cdot \mathrm{x} = \omega t - \vec{k} \cdot \vec{x}$ should be a Lorentz scalar.

2.2 Covariant electrodynamics

In Gaussian units, Maxwell's equations in vacuum are given by

$$\nabla \cdot \vec{E} = 4\pi\rho, \qquad \nabla \cdot \vec{B} = 0, \tag{23}$$

$$\nabla \times \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}, \qquad \nabla \times \vec{B} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial \vec{E}}{\partial t}. \tag{24}$$

We have solved Maxwell's equations in the radiation zone in Lecture 2. Let's briefly review

the procedure here and this time we want to rewrite the equations in covariant form. After introducing the electromagnetic potentials $\vec{A}$ and ϕ, the fields are expressed as

$$\vec{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}, \qquad \vec{B} = \nabla \times \vec{A}. \tag{25}$$

Using the Lorentz gauge condition

$$\nabla \cdot \vec{A} + \frac{1}{c}\frac{\partial \phi}{\partial t} = 0, \tag{26}$$

the potentials $\vec{A}$ and ϕ satisfy d'Alembert wave equations as given by

$$\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} - \nabla^2\phi = 4\pi\rho, \qquad \frac{1}{c^2}\frac{\partial^2 \vec{A}}{\partial t^2} - \nabla^2\vec{A} = \frac{4\pi}{c}\vec{j}. \tag{27}$$

We notice that if we define $\mathrm{A} = (\phi, \vec{A})$, from the 4-current $\mathrm{j} = (\rho c, \vec{j})$ we have

$$\Box \mathrm{A} = \frac{4\pi}{c}\mathrm{j}, \qquad \Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2, \tag{28}$$

where $\Box$ is called the *d'Alembertian* operator and can be proven to be a Lorentz scalar. The proof uses the 4-gradient

$$\partial = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right), \tag{29}$$

which can be shown to be a 4-vector because from Eq. (17) we have using the chain rule

$$\begin{pmatrix} \partial_{ct} \\ \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \partial_{ct'} \\ \partial_{x'} \\ \partial_{y'} \\ \partial_{z'} \end{pmatrix}. \tag{30}$$

The symmetric Lorentz matrix is unchanged after transposition. Then the sign of β changes twice via inversion and the negative sign of the spatial gradient in Eq. (29). So the 4-gradient ∂ is a 4-vector and the d'Alembertian $\Box = \partial \cdot \partial$ is a Lorentz scalar. This means if we have found A in one reference frame by solving the d'Alembert equation

$$\Box \mathrm{A} = \frac{4\pi}{c}\mathrm{j} \tag{31}$$

under the Lorentz gauge condition $\partial \cdot \mathrm{A} = 0$, we may transform A and j like 4-vectors we we go to a different frame and both the equation and the gauge condition will be satisfied, which uniquely determine A under certain boundary conditions. We do not need to solve the equation again. This proves that A is a 4-vector.

2.3 Transformation of $\vec{E}$ and $\vec{B}$ fields

The transformation of $\vec{E}$ and $\vec{B}$ fields can be derived using the transformation of the 4-potential $\mathrm{A} = (\phi, \vec{A})$ and the 4-gradient $\partial = (\partial_{ct}, -\nabla)$. They satisfy

$$\begin{pmatrix} \partial_{ct'} \\ \partial_{x'} \\ \partial_{y'} \\ \partial_{z'} \end{pmatrix} = \begin{bmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \partial_{ct} \\ \partial_{x} \\ \partial_{y} \\ \partial_{z} \end{pmatrix}, \qquad \begin{pmatrix} \phi' \\ A'_x \\ A'_y \\ A'_z \end{pmatrix} = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \phi \\ A_x \\ A_y \\ A_z \end{pmatrix}. \tag{32}$$

From Eq. (25), we obtain the $\vec{E}$ field in the new frame given by

$$E'_x = -\begin{pmatrix} \partial_{x'} & \partial_{ct'} \end{pmatrix} \begin{pmatrix} \phi' \\ A'_x \end{pmatrix} = -\begin{pmatrix} \partial_x & \partial_{ct} \end{pmatrix} \begin{bmatrix} \gamma & \beta\gamma \\ \beta\gamma & \gamma \end{bmatrix} \begin{bmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{bmatrix} \begin{pmatrix} \phi \\ A_x \end{pmatrix} = E_x. \tag{33}$$

$$E'_y = -\partial_{y'}\phi' - \partial_{ct'}A'_y = -\partial_y(\gamma\phi - \beta\gamma A_x) - (\gamma\partial_{ct} + \beta\gamma\partial_x)A_y \tag{34}$$

$$= \gamma\left[(-\partial_y\phi - \partial_{ct}A_y) - \beta(\partial_x A_y - \partial_y A_x)\right] = \gamma(E_y - \beta B_z), \tag{35}$$

$$E'_z = -\partial_{z'}\phi' - \partial_{ct'}A'_z = -\partial_z(\gamma\phi - \beta\gamma A_x) - (\gamma\partial_{ct} + \beta\gamma\partial_x)A_z \tag{36}$$

$$= \gamma\left[(-\partial_z\phi - \partial_{ct}A_z) + \beta(\partial_z A_x - \partial_x A_z)\right] = \gamma(E_z + \beta B_y), \tag{37}$$

and the $\vec{B}$ field in the new frame is given by

$$B'_x = \partial_{y'}A'_z - \partial_{z'}A'_y = \partial_y A_z - \partial_z A_y = B_x, \tag{38}$$

$$B'_y = \partial_{z'}A'_x - \partial_{x'}A'_z = \partial_z(-\beta\gamma\phi + \gamma A_x) - (\beta\gamma\partial_{ct} + \gamma\partial_x)A_z \tag{39}$$

$$= \gamma\left[(\partial_z A_x - \partial_x A_z) + \beta(-\partial_z\phi - \partial_{ct}A_z)\right] = \gamma(B_y + \beta E_z), \tag{40}$$

$$B'_z = \partial_{x'}A'_y - \partial_{y'}A'_x = (\beta\gamma\partial_{ct} + \gamma\partial_x)A_y - \partial_y(-\beta\gamma\phi + \gamma A_x) \tag{41}$$

$$= \gamma\left[(\partial_x A_y - \partial_y A_x) - \beta(-\partial_{ct}A_y - \partial_y\phi)\right] = \gamma(B_z - \beta E_y). \tag{42}$$

The equations tell us that even in case of a non-relativistic limit, the $\vec{E}$ and $\vec{B}$ fields are still reference frame dependent. This makes the electromagnetic field quite bizarre from the point of view of classical physics and also inspired Einstein to establish his great theory of relativity. The calculations of these formulas can be a bit lengthy to go through. But you are allowed to put them on your own formula sheet to prepare for the Quals.

3 Sample problems

3.1 Problem 1 in Sec. 4, 2013

Object A has a rest mass m and moves at constant velocity $v = 0.8c$. Object A inelastically collides with another object B, which is at rest and has the same mass m.

(a) Find their common velocity u after the collision.

(b) Suppose the heat released in this collision is radiated in photons distributed isotropically in the rest frame of $A + B$. How much energy is radiated in the lab frame?

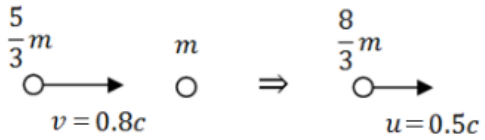
(c) Would the answer change if the emission were anisotropic?

Answer: This problem may be a bit confusing. In relativity, inelastic collisions, unlike what the name suggests, also satisfy energy conservation, and thus the total 4-momentum is conserved. In classical physics, we say that in an inelastic collision, part of the kinetic energy is dissipated into heat. In relativity, that heat corresponds to an increase in the rest mass (energy). An inelastic collision therefore converts part of the moving mass into rest mass, but the total mass (energy) is still conserved before the heat radiates away into photons, which can take away both energy and momentum.

(a) From the conservation of mass and momentum, we have

$$\left(\frac{m}{\sqrt{1-v^2/c^2}}+m\right)u=\frac{mv}{\sqrt{1-v^2/c^2}}. \tag{43}$$

Plugging in $v = 0.8c$, we find $u = 0.5c$. Results of the collision are plotted below:



Once the relativistic description of an inelastic collision is understood, the calculation of the common velocity u is not difficult.

(b) We may then find the rest mass M of the composite object $A + B$. It satisfies

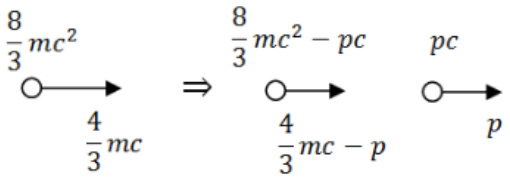
$$\frac{M}{\sqrt{1-0.5^2}}=\frac{8}{3}m \quad \Rightarrow \quad M=\frac{4}{\sqrt{3}}m=2.309m>2m. \tag{44}$$

In the rest frame of $A + B$, the emitted photons have zero total momentum and a total mass of $\Delta M = M - 2m = 0.309m$, as the composite object $A + B$ cools down and its mass returns to $2m$. We may transform the photon energy back to the lab frame to obtain the total energy emitted

$$\Delta E=\frac{\Delta Mc^2}{\sqrt{1-0.5^2}}=\left(\frac{8}{3}-\frac{4}{\sqrt{3}}\right)mc^2=0.357mc^2. \tag{45}$$

Recall that the energy formula is also applicable to multiple mass points (photons).

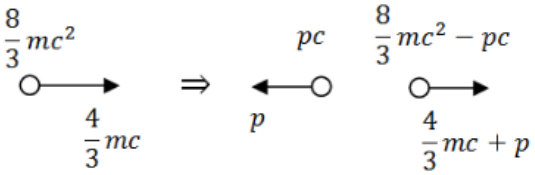
(c) If the photons are not emitted isotropically in the rest frame of $A+B$, there will be a recoil of $A+B$ in the opposite direction of the total momentum of the emitted photons because of momentum conservation. This will change the energy radiated. For a quantitative discussion, we consider two extreme situations. The most convenient calculates are done in the lab frame.



In case that all photons are emitted in the same direction of $u = 0.5c$, we have

$$\left(\frac{8}{3}mc - p\right)^2 = \left(\frac{4}{3}mc - p\right)^2 + (2mc)^2 \quad \Rightarrow \quad \Delta E = pc = 0.5mc^2 > 0.357mc^2, \tag{46}$$

where we require that the rest mass of $A+B$ returns to $2m$ when it is fully cooled down.



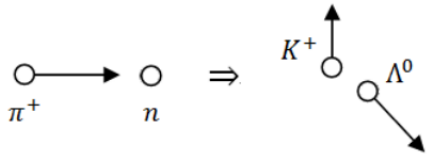
In case that all photons are emitted in the opposite direction of $u = 0.5c$, we have

$$\left(\frac{8}{3}mc - p\right)^2 = \left(\frac{4}{3}mc + p\right)^2 + (2mc)^2 \quad \Rightarrow \quad \Delta E = pc = \frac{1}{6}mc^2 < 0.357mc^2. \tag{47}$$

In general the emitted energy ΔE will be in between $\frac{1}{6}mc^2$ and $\frac{1}{2}mc^2$. The recoil affects the final speed of the $A+B$ composite. The three cases satisfy $\frac{5}{13}c < 0.5c < \frac{3}{5}c$. Thus the available energies of photon emission satisfy $\frac{1}{2}mc^2 > 0.357mc^2 > \frac{1}{6}mc^2$.

3.2 Problem 2 in Sec. 4, 2013

Consider the reaction $\pi^+ + n \to K^+ + \Lambda^0$. The rest masses of the particles are $m_\pi c^2 = 140$ MeV, $m_n c^2 = 940$ MeV, $m_K c^2 = 494$ MeV, $m_\Lambda c^2 = 1115$ MeV. What is the threshold kinetic energy of the π^+ to create a K^+ moving at an angle of 90° in the initial direction of the π^+ in the lab frame, in which the n is at rest?



Answer: This type of 2D scattering problems are the commonest relativistic collision problems in the Quals. The standard method is to use the conservation of 4-momentum

$$(\gamma_\pi m_\pi c, \vec{p}_\pi) + (m_n c, 0) = (\gamma_K m_K c, \vec{p}_K) + (\gamma_\Lambda m_\Lambda c, \vec{p}_\Lambda). \tag{48}$$

Since we know the angle between $\vec{p}_K$ and $\vec{p}_\pi$ is $90°$, and we don't know the direction of $\vec{p}_\Lambda$, we may use the inner product to cancel $\vec{p}_\Lambda$ to obtain

$$\begin{aligned} m_\Lambda^2 c^2 = (\gamma_\Lambda m_\Lambda c, \vec{p}_\Lambda) \cdot (\gamma_\Lambda m_\Lambda c, \vec{p}_\Lambda) &= \left((\gamma_\pi m_\pi + m_n - \gamma_K m_K) c, \vec{p}_\pi - \vec{p}_K\right)^2 && (49)\\ &= (\gamma_\pi m_\pi + m_n - \gamma_K m_K)^2 c^2 - (p_\pi^2 + p_K^2) && (50)\\ &= (\gamma_\pi^2 m_\pi^2 c^2 - p_\pi^2) + (\gamma_K^2 m_K^2 c^2 - p_K^2) - 2\gamma_\pi \gamma_K m_\pi m_K c^2 \\ &\quad + 2m_n(\gamma_\pi m_\pi - \gamma_K m_K) c^2 + m_n^2 c^2 && (51)\\ &= m_\pi^2 c^2 + m_K^2 c^2 + m_n^2 c^2 - 2\gamma_\pi \gamma_K m_\pi m_K c^2 + 2m_n(\gamma_\pi m_\pi - \gamma_K m_K) c^2. && (52) \end{aligned}$$

Therefore, we have

$$\gamma_\pi m_\pi = \frac{m_\Lambda^2 - m_\pi^2 - m_n^2 - m_K^2 + 2\gamma_K m_K m_n}{2(m_n - \gamma_K m_K)}. \tag{53}$$

The right-hand side monotonically increases with γ_K. The minimum is found by plugging in $\gamma_K = 1$, which means we generate a K^+ at rest. Generating a moving K^+ will definitely cost more energy. Therefore, we get the threshold kinetic energy

$$E_\pi = (\gamma_\pi - 1) m_\pi \geq \frac{m_\Lambda^2 - m_\pi^2}{2(m_n - m_K)} - \frac{1}{2}(m_n - m_K) - m_\pi = 1009\,\mathrm{MeV}. \tag{54}$$

3.3 Problem 3 in Sec. 2, 2013

Suppose we have a point charge q in uniform motion with relativistic velocity v in the x direction. What is the electric field configuration sourced by this charge? To be concrete, suppose the charge position as a function of time t is $\vec{x}_q(t) = vt\hat{x}$, and we are interested in the electric field at some arbitrary location and time.

Answer: This problem can be solved using either the retarded potential (no radiation zone approximation) or the relativistic transformation of the $\vec{E}$ and $\vec{B}$ fields. The second method will be simpler if you remember the formulas. In the co-moving frame, the $\vec{E}$-field is just the Coulomb field of a charge q (Lorentz scalar) at the origin $\vec{x}_q'(t') = 0$ and there is no $\vec{B}$-field. The fields are thus given by

$$\vec{E}\,'(\vec{x}', t') = \frac{q\vec{x}'}{|\vec{x}'|^3}, \qquad \vec{B}\,'(\vec{x}', t') = 0. \tag{55}$$

We may then get the components of the $\vec{E}$ and $\vec{B}$ fields in the lab frame one by one, using the formulas we derived for the relativistic transformation of $\vec{E}$ and $\vec{B}$ fields. We get

$$E_x(\vec{x},t) = E'_x(\vec{x}\,',t') = \frac{qx'}{(x'^2+y'^2+z'^2)^{3/2}} = \frac{q\gamma(x-vt)}{[\gamma^2(x-vt)^2+y^2+z^2]^{3/2}}, \tag{56}$$

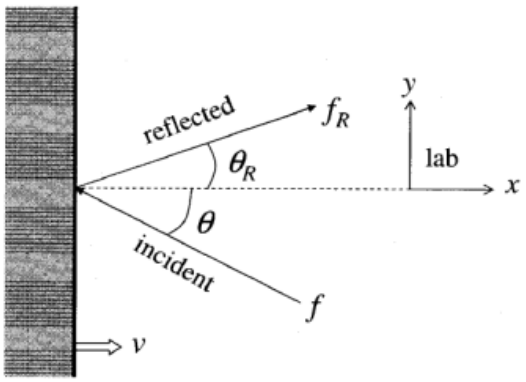
$$E_y(\vec{x},t) = \gamma E'_y(\vec{x}\,',t') = \frac{q\gamma y'}{(x'^2+y'^2+z'^2)^{3/2}} = \frac{q\gamma y}{[\gamma^2(x-vt)^2+y^2+z^2]^{3/2}}, \tag{57}$$

$$E_z(\vec{x},t) = \gamma E'_z(\vec{x}\,',t') = \frac{q\gamma z'}{(x'^2+y'^2+z'^2)^{3/2}} = \frac{q\gamma z}{[\gamma^2(x-vt)^2+y^2+z^2]^{3/2}}, \tag{58}$$

because $\vec{B}\,'(\vec{x}\,',t') = 0$. However, $\vec{B}(\vec{x},t)$ is nonzero, which can be understood as being generated by the moving charge in the lab frame. Electricity and magnetism are closely related in relativity.

3.4 Problem 3 in Sec. 4, 2009

A monochromatic beam of light is incident onto a flat mirror. In the lab frame, the mirror is traveling at the speed v in the $+x$ direction. The plane of the mirror is perpendicular to the x axis. The incident light beam has frequency f and is propagating at angle θ from the x axis. Find the frequency f_R and the angle θ_R of the reflected light beam as measured in the lab frame.



Answer: The problem can be solved using the Lorentz transformation of either the 4-wavevector $\mathrm{k} = (\omega/c, \vec{k})$ if we think of light as a wave or the 4-momentum $\mathrm{p} = (E/c, \vec{p})$ if we think of light as photons. Since both k and p are 4-vectors, they transform the same way and give the same answer. The 4-wavevectors of the incident and reflected beams are

$$\mathrm{k} = (k, -k\cos\theta, k\sin\theta, 0)^T, \qquad k = \frac{2\pi f}{c}, \tag{59}$$

$$\mathrm{k}_R = (k_R, k_R\cos\theta_R, k_R\sin\theta_R, 0)^T, \qquad k_R = \frac{2\pi f_R}{c}. \tag{60}$$

In the co-moving frame of the mirror, the 4-wavevector of incident beam becomes

$$\mathrm{k}' = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} k \\ -k\cos\theta \\ k\sin\theta \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k(1+\beta\cos\theta) \\ -\gamma k(\beta+\cos\theta) \\ k\sin\theta \\ 0 \end{pmatrix}, \tag{61}$$

and the 4-wavevector of reflected beam becomes

$$\mathrm{k}'_R = \begin{bmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} k_R \\ k_R\cos\theta_R \\ k_R\sin\theta_R \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma k_R(1-\beta\cos\theta_R) \\ \gamma k_R(\cos\theta_R-\beta) \\ k_R\sin\theta_R \\ 0 \end{pmatrix}, \tag{62}$$

In the coming frame, k' and k'_R are symmetric about the normal direction (x axis) and equal in frequency. So we have

$$\begin{cases} \gamma k(1+\beta\cos\theta) = \gamma k_R(1-\beta\cos\theta_R), & (63) \\ -\gamma(\beta+\cos\theta) + \gamma k_R(\cos\theta_R-\beta) = 0, & (64) \\ k\sin\theta = k_R\sin\theta_R. & (65) \end{cases}$$

Since $\mathrm{k}'\cdot\mathrm{k}' = \mathrm{k}'_R\cdot\mathrm{k}'_R = 0$, only two of the above three equations are independent, which is consistent with the two unknowns k_R and θ_R to be solved for. Nevertheless, we can still use all three equations to double check. The first two equations are simultaneous linear equations of k_R and $k_R\cos\theta_R$, which give us

$$\begin{cases} k_R = \gamma^2 k(1+\beta^2+2\beta\cos\theta), & (66) \\ k_R\cos\theta_R = \gamma^2 k\left[(1+\beta^2)\cos\theta + 2\beta\right]. & (67) \end{cases}$$

One may get from these equations $k_R^2\sin^2\theta_R = k^2\sin^2\theta$, which is consistent with Eq. (65). We may therefore write our answer as

$$\frac{2\pi f_R}{c} = \gamma^2\frac{2\pi f}{c}(1+\beta^2+2\beta\cos\theta) \quad\Rightarrow\quad f_R = \gamma^2 f(1+\beta^2+2\beta\cos\theta), \tag{68}$$

$$\sin\theta_R = \frac{k}{k_R}\sin\theta = \frac{\sin\theta}{\gamma^2(1+\beta^2+2\beta\cos\theta)} = \frac{(1-\beta^2)\sin\theta}{1+\beta^2+2\beta\cos\theta}. \tag{69}$$

The law of reflection $\theta_R = \theta$ is recovered if $\beta = 0$. Relativity generalizes physical laws to high speed scenarios, enriches our understanding of mass and energy and reveals the essence of electricity, magnetism, light and spacetime.

Lecture 4 Quantum Mechanics - Theory

There are various topics that can appear in Secs. 3 (only QM) and 4 (with relativity) of the Quals. A thorough revision of the topics in Griffith's *Introduction to Quantum Mechanics* can be very helpful and necessary. This lecture only covers some important selected topics.

1 Simple harmonic oscillator

The time-independent Shcrödinger equation of a 1D simple harmonic oscillator is given by

$$\hat{H}\psi(x) = \left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2\right)\psi(x) = E\psi(x). \tag{1}$$

We define the *lowering* and *raising* operators as

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} + \frac{i\hat{p}}{\sqrt{2m\hbar\omega}} = \sqrt{\frac{m\omega}{2\hbar}}\,x + \sqrt{\frac{\hbar}{2m\omega}}\,\frac{d}{dx}, \tag{2}$$

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\,\hat{x} - \frac{i\hat{p}}{\sqrt{2m\hbar\omega}} = \sqrt{\frac{m\omega}{2\hbar}}\,x - \sqrt{\frac{\hbar}{2m\omega}}\,\frac{d}{dx}, \tag{3}$$

so that in terms of them, the Hamiltonian is written as

$$\hat{H} = \hbar\omega\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right), \tag{4}$$

and we can prove the commutation relations

$$[\hat{a}, \hat{a}^\dagger] = 1, \quad [\hat{H}, \hat{a}^\dagger] = \hbar\omega\hat{a}^\dagger, \quad [\hat{H}, \hat{a}] = -\hbar\omega\hat{a}. \tag{5}$$

From these commutation relations, together with the fact that the Hamiltonian $\hat{H}$ is bounded below, we deduce that there is a ground state $\psi_0(x)$ such that

$$\hat{a}\psi_0(x) = \sqrt{\frac{m\omega}{2\hbar}}\,x\psi_0(x) + \sqrt{\frac{\hbar}{2m\omega}}\,\frac{d\psi_0(x)}{dx} = 0 \quad \Rightarrow \quad \psi_0(x) = Ce^{-\frac{m\omega}{2\hbar}x^2}, \tag{6}$$

where the arbitrary coefficient C can be determined by the normalization condition

$$\int_{-\infty}^{+\infty} |\psi_0(x)|^2 = 1 \quad \Rightarrow \quad C = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}, \tag{7}$$

up to an arbitrary phase. Once the ground state $\psi_0(x)$ is found, we can use the raising operator $\hat{a}^\dagger$ to find the excited state $\psi_n(x)$ by acting $\hat{a}^\dagger$ onto the ground state n times, i.e., $\psi_n(x) \propto (\hat{a}^\dagger)^n \psi_0(x)$. Since we have

$$\|\hat{a}|n\rangle\|^2 = \langle n|\hat{a}^\dagger \hat{a}|n\rangle = n\langle n|n\rangle = n, \tag{8}$$

$$\|\hat{a}^\dagger|n\rangle\|^2 = \langle n|\hat{a}\hat{a}^\dagger|n\rangle = \langle n|\hat{a}^\dagger\hat{a} + 1|n\rangle = n + 1, \tag{9}$$

where $|n\rangle$ is the Dirac ket of the nth normalized eigenstate $\psi_n(x)$, we obtain

$$\hat{a}|n\rangle = \sqrt{n}\,|n-1\rangle, \qquad \hat{a}^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle. \tag{10}$$

Therefore, the wave function of the nth eigenstate can be found to be

$$\psi_n(x) = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}\psi_0(x) = \frac{1}{\sqrt{2^n n!}}\left(\sqrt{\frac{m\omega}{\hbar}}x - \sqrt{\frac{\hbar}{m\omega}}\frac{d}{dx}\right)^n \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2}. \tag{11}$$

2 Hydrogen atom

2.1 General central force field

The Schrödinger equation in a general central potential $V(r)$ in 3D, where $V(r)$ is only a function of the distance r from the origin, is given in spherical coordinates by

$$-\frac{\hbar^2}{2m_e}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi + V(r)\psi = E\psi. \tag{12}$$

For separable solutions $\psi(r,\theta,\phi) = R(r)Y(\theta,\phi)$, we have

$$\frac{2m_e}{\hbar^2}[E - V(r)]r^2 + \frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) = -\frac{1}{Y\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial Y}{\partial\theta}\right) - \frac{1}{Y\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}. \tag{13}$$

The left-hand side is a function of r only, and the right-hand side is a function of θ and ϕ. Both of them have to equal a constant. Let's call that constant $l(l+1)$, so we have

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\,\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2} + l(l+1)Y = 0. \tag{14}$$

The solutions to this equation under proper boundary conditions are the spherical harmonics $Y_{lm}(\theta,\phi)$, where $l = 0, 1, 2, \ldots$ and $m = 0, \pm1, \pm2, \ldots, \pm l$. The angular part of the wave function ψ is always $Y_{lm}(\theta,\phi)$. Only the radial equation

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = \frac{1}{r}\frac{d^2}{dr^2}(rR) = \left(\frac{l(l+1)}{r^2} + \frac{2m_e}{\hbar^2}[V(r) - E]\right)R \tag{15}$$

may depend on the specific form of $V(r)$. Let $u(r) = rR(r)$ be the effective radial wave function, so we have an effective 1D Schrödinger equation

$$-\frac{\hbar^2}{2m_e}\frac{d^2u}{dr^2} + \left[V(r) + \frac{\hbar^2}{2m_e}\frac{l(l+1)}{r^2}\right]u = Eu. \tag{16}$$

Apart from the central potential $V(r)$, the electron is also subject to a centrifugal potential which propels it away from the origin. Bound states can form if $V(r)$ is attractive and its attractive part at $r \to 0$ is less singular than $1/r^2$. The Coulomb potential satisfies these conditions, so the electron can form bound states in the hydrogen atom.

2.2 Coulomb potential

In the hydrogen atom, an electron orbits around a proton due to the Coulomb attraction between them. Assuming the more massive proton is approximately at rest when the electron moves around it, the electron is subject to a central Coulomb potential given by

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r} \propto -\frac{1}{r}. \tag{17}$$

Plugging this into Eq. (16) and by introducing dimensionless variables

$$\rho = \frac{\sqrt{-2m_eE}}{\hbar}r, \qquad \rho_0 = \frac{e^2}{4\pi\epsilon_0\hbar}\sqrt{\frac{2m_e}{-E}}, \tag{18}$$

the effective radial Schrödinger equation (16) becomes

$$\frac{d^2u}{d\rho^2} = \left[1 - \frac{\rho_0}{\rho} + \frac{l(l+1)}{\rho^2}\right]u. \tag{19}$$

A full solution of this equation can be found in § 4.2.1 of Griffith's book, which is a bit lengthy to go through and will not be duplicated here. We will only give an estimate of the energy and length scales of the hydrogen atom using dimension analysis. Given an integer quantum number l, there should be a set of discrete values that the dimensionless variable ρ_0 can take in order to make $u(\rho \to \infty)$ converge. Since ρ_0 is dimensionless, the ground state energy can be estimated by setting $\rho_0 \sim 1$, so that

$$E \approx -2m_e\left(\frac{e^2}{4\pi\epsilon_0\hbar}\right)^2 = -2m_ec^2\alpha^2, \tag{20}$$

which deviates from the exact ground state energy $-\frac{1}{2}m_ec^2\alpha^2 = -13.6$ eV of the hydrogen atom by a positive dimensionless factor 4. In expressing the results, we have used the fine structure constant $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137}$. The length scale, or atomic size, of the hydrogen atom

can be obtained by setting $\rho \sim 1$, which leads to

$$r \approx \frac{\hbar}{\sqrt{-2m_e E}} \approx \frac{\hbar}{2m_e c\alpha} = \frac{1}{2}a_B, \tag{21}$$

where $a_B = \frac{\hbar}{m_e c\alpha} = 0.529\text{Å}$ is the Bohr radius. Dimension analysis is an important technique tested in the Quals. Though inaccurate, it gives the correct energy and length scales of the hydrogen atom and is applicable to other problems that are not analytically solvable.

3 Angular momentum

In quantum mechanics, the three components $\hat{L}_x$, $\hat{L}_y$, $\hat{L}_z$ of an angular momentum $\hat{\vec{L}}$ do not commute. The commutation relations are given by

$$[\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x, \quad [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y, \quad [\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z. \tag{22}$$

Defining $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$, we obtain

$$[\hat{L}^2, \hat{L}_x] = [\hat{L}^2, \hat{L}_y] = [\hat{L}^2, \hat{L}_z] = 0. \tag{23}$$

We conventionally use the eigenvalues $\hat{L}^2 = l(l+1)\hbar^2$ and $\hat{L}_z = m\hbar$ to completely describe a basis state $|l, m\rangle$ with a definite angular momentum. We may further define raising and lowering operators $\hat{L}_\pm = \hat{L}_x \pm i\hat{L}_y$ in the z direction. They satisfy

$$[\hat{L}^2, \hat{L}_\pm] = 0, \qquad [\hat{L}_z, \hat{L}_\pm] = \pm\hbar\hat{L}_\pm. \tag{24}$$

Since we have for the raising operator $\hat{L}_+$ that

$$\|\hat{L}_+|l, m\rangle\|^2 = \langle l, m|\hat{L}_-\hat{L}_+|l, m\rangle = \langle l, m|\hat{L}_x^2 + \hat{L}_y^2 + i[\hat{L}_x, \hat{L}_y]|l, m\rangle \tag{25}$$

$$= \langle l, m|\hat{L}^2 - \hat{L}_z^2 - \hbar\hat{L}_z|l, m\rangle = (l-m)(l+m+1)\hbar^2, \tag{26}$$

and for the lowering operator $\hat{L}_-$ we have

$$\|\hat{L}_-|l, m\rangle\|^2 = \langle l, m|\hat{L}_+\hat{L}_-|l, m\rangle = \langle l, m|\hat{L}_x^2 + \hat{L}_y^2 - i[\hat{L}_x, \hat{L}_y]|l, m\rangle \tag{27}$$

$$= \langle l, m|\hat{L}^2 - \hat{L}_z^2 + \hbar\hat{L}_z|l, m\rangle = (l+m)(l-m+1)\hbar^2, \tag{28}$$

the normalization constants of $\hat{L}_\pm|l, m\rangle$ are determined and we obtain

$$\hat{L}_\pm|l, m\rangle = \sqrt{(l \mp m)(l \pm m + 1)}\,\hbar\,|l, m \pm 1\rangle. \tag{29}$$

These normalization constants are useful for calculating the Clebsch-Gordan coefficients by hand. The formalism applies to both orbital and spin angular momenta.

4 Electron in an electromagnetic field

4.1 Spinless theory

Recall that in Lecture 1, when a non-relativistic electron (with charge $-e$ and mass m_e) moves in an electromagnetic field, the classical Lagrangian is given by

$$L = \frac{1}{2}m_e v^2 + e\Big[\phi(\vec{r},t) - \vec{v}\cdot\vec{A}(\vec{r},t)\Big]. \tag{30}$$

The conjugate momentum of $\vec{r}$ is

$$\vec{p} = \nabla_{\vec{v}} L = m_e\vec{v} - e\vec{A}(\vec{r},t). \tag{31}$$

Therefore the classical Hamiltonian is given by

$$H = \vec{p}\cdot\vec{v} - L = \frac{1}{2}m_e v^2 - e\phi(\vec{r},t) = \frac{1}{2m_e}(\vec{p}+e\vec{A})^2 - e\phi. \tag{32}$$

The quantized Hamiltonian operator is therefore given by

$$\hat{H} = \frac{1}{2m_e}(-i\hbar\nabla + e\vec{A})^2 - e\phi. \tag{33}$$

The time-dependent Schrödinger equation is then

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = \left[\frac{1}{2m_e}(-i\hbar\nabla + e\vec{A})^2 - e\phi\right]\psi(\vec{r},t). \tag{34}$$

4.2 Probability flux

From the above Schrödinger equation, we have

$$i\hbar\psi^*\frac{\partial\psi}{\partial t} = \psi^*\left[\frac{1}{2m_e}(-i\hbar\nabla + e\vec{A})^2 - e\phi\right]\psi, \tag{35}$$

$$-i\hbar\psi\frac{\partial\psi^*}{\partial t} = \psi\left[\frac{1}{2m_e}(i\hbar\nabla + e\vec{A})^2 - e\phi\right]\psi^*. \tag{36}$$

Taking the difference of the two equations, we therefore obtain

$$i\hbar\frac{\partial}{\partial t}(\psi^*\psi) = \frac{1}{2m_e}\left[\psi^*\left(-i\hbar\nabla + e\vec{A}\right)^2\psi - \psi\left(i\hbar\nabla + e\vec{A}\right)^2\psi^*\right] \tag{37}$$

$$= -\frac{\hbar^2}{2m_e}(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*) - \frac{i\hbar}{m_e}\nabla\cdot(e\vec{A}\,\psi^*\psi) \tag{38}$$

$$= -\frac{i\hbar}{m_e}\nabla\cdot\Big[\hbar\,\mathrm{Im}(\psi^*\nabla\psi) + e\vec{A}\,\psi^*\psi\Big]. \tag{39}$$

Since we know $\psi^*\psi$ is the probability density, the *probability flux* is given by

$$\vec{J}_p = \frac{\hbar}{m_e}\,\mathrm{Im}(\psi^*\nabla\psi) + \frac{e\vec{A}}{m_e}\psi^*\psi. \tag{40}$$

The extra term due to $\vec{A}$ in $\vec{J}_p$ can be understood from the classical conjugate momentum $\vec{p} = m_e\vec{v} - e\vec{A}$, from which we obtain $\vec{v} = (\vec{p} + e\vec{A})/m_e$. The probability flux $\vec{J}_p$ and the electric current $\vec{j} = -e\vec{J}_p$ are therefore still determined by $\vec{v}$ rather than $\vec{p}$.

4.3 Spin-1/2 theory

In quantum mechanics, electrons are spin-1/2 particles. This leads to an extra term in $\hat{H}$ due to the Zeeman coupling of the spin magnetic moment with $\vec{B} = \nabla \times \vec{A}$. We have

$$\hat{H} = \frac{1}{2m_e}(-i\hbar\nabla + e\vec{A})^2 - e\phi + \frac{g_s e}{2m_e}\frac{\hbar}{2}\,\hat{\vec{\sigma}}\cdot\vec{B}, \tag{41}$$

where $g_s \approx 2$ is the electron spin g-factor, and $\hat{\vec{\sigma}} = \vec{e}_x\hat{\sigma}_x + \vec{e}_y\hat{\sigma}_y + \vec{e}_z\hat{\sigma}_z$ is the Pauli vector with its three components being the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{42}$$

In Pauli representation, the time-evolution of a spin wave function is given by

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\psi_\uparrow \\ \psi_\downarrow\end{pmatrix} = \left[\frac{1}{2m_e}(-i\hbar\nabla + e\vec{A})^2 - e\phi + \frac{g_s e\hbar}{4m_e}\begin{pmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z\end{pmatrix}\right]\begin{pmatrix}\psi_\uparrow \\ \psi_\downarrow\end{pmatrix}. \tag{43}$$

The total probability density $n(\vec{r}) = |\psi_\uparrow(\vec{r})|^2 + |\psi_\downarrow(\vec{r})|^2$ in real space is independent of the spin representation (try σ_x-representation, for example, the total density $n(\vec{r})$ is the same). One may verify that the probability flux $\vec{J}_p$ corresponding to $n(\vec{r})$ still satisfies Eq. (40) because the magnetic field $\vec{B}(\vec{r})$ only flips the spins locally. So the spatial probability flux still agrees with the spinless theory.

5 Identical particles

(Problem 4 in Sec. 3, 2012) Consider two identical electrons with spin-1/2 and mass m that are confined to a 1D box with length L. The potential energy is 0 inside the box and infinity outside. The interaction between two particles can be expressed by the potential energy $V(x_1, x_2)$, where x_1 and x_2 are the coordinates of the two particles.

(a) We first consider noninteracting particles, i.e., $V(x_1, x_2) = 0$. What are the ground state energy and wave function, considering spatial and spin state symmetry?

(b) What are the energy and degeneracy of the first excited state? Write down explicit wave function considering spatial and spin state symmetry.

(c) Now consider a simple form of electron-electron interaction $V(x_1, x_2) = V_0\,\delta(x_1 - x_2)$. Assuming V_0 is small, find the correction for the ground state energy in (a) due to this e-e interaction.

(d) Describe how degeneracy in the excited state discussed in (b) would be lifted.

Answer: We use a sample problem to help review our knowledge of identical fermions. Assuming you are familiar with single-electron eigenstates of an infinite square well, we directly begin with placing two electrons into two spin states to satisfy the Pauli exclusion principle. Then we require some perturbation theory for parts (c) and (d).

(a) The ground state is reached by putting both electrons in the ground state $n = 1$ of the infinite square well. Since electrons are fermions, the spins of the two electrons will have to be opposite. This gives us a singlet spinor times a symmetric spatial wave function as the two-body ground-state spin wave function

$$\Psi_0(x_1, x_2) = \psi_1(x_1)\psi_1(x_2) \times \frac{\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2}{\sqrt{2}}, \tag{44}$$

where $\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$, $n = 1, 2, 3, \ldots$ are the single-electron eigenstates of an infinite square well. The ground state energy is given by

$$E_0 = \frac{1}{2m_e}\left(\frac{\pi\hbar}{L}\right)^2 \times 2 = \frac{\pi^2\hbar^2}{m_e L^2}. \tag{45}$$

(b) In the first excited state, one of the electrons gets excited to $n = 2$, while the other stays in $n = 1$. When the two-body spinor is a singlet, the spatial part of the two-body wave function should be symmetric. When the spinor is a triplet, the spatial part should be antisymmetric. Therefore, we have a four-fold degeneracy given by

$$\Psi_{\text{singlet}}(x_1, x_2) = \frac{\psi_1(x_1)\psi_2(x_2) + \psi_2(x_1)\psi_1(x_2)}{\sqrt{2}} \times \frac{\uparrow_1\downarrow_2 - \downarrow_1\uparrow_2}{\sqrt{2}}, \tag{46}$$

$$\Psi_{\text{triplet}}(x_1, x_2) = \frac{\psi_1(x_1)\psi_2(x_2) - \psi_2(x_1)\psi_1(x_2)}{\sqrt{2}} \times \begin{cases} \uparrow_1\uparrow_2 \\ \dfrac{\uparrow_1\downarrow_2 + \downarrow_1\uparrow_2}{\sqrt{2}} \\ \downarrow_1\downarrow_2 \end{cases}. \tag{47}$$

All four states have the same energy

$$E_1 = \frac{1}{2m_e}\left(\frac{\pi\hbar}{L}\right)^2 + \frac{1}{2m_e}\left(\frac{2\pi\hbar}{L}\right)^2 = \frac{5\pi^2\hbar^2}{2m_e L^2}. \tag{48}$$

(c) Since the interaction strength V_0 is small and the singlet ground state Ψ_0 is non-degenerate, we can use the first-order non-degenerate perturbation theory to obtain the energy correction of E_0, which is given by

$$\Delta E_0 = \langle \Psi_0 | V | \Psi_0 \rangle = V_0 \iint dx_1 dx_2 \, |\psi_1(x_1)|^2 |\psi_1(x_2)|^2 \, \delta(x_1 - x_2) \tag{49}$$

$$= V_0 \int dx \, |\psi_1(x)|^4 = \frac{4V_0}{L^2} \int_0^L \sin^4 \frac{\pi x}{L} dx = \frac{3V_0}{2L}. \tag{50}$$

(d) In the triplet states, the two electrons never meet, because the spatial wave function $\Psi_{\text{singlet}}(x_1, x_2)$ is antisymmetric so that $\Psi_{\text{triplet}}(x, x) = 0$. Therefore, the matrix elements of the perturbation $V(x_1, x_2) = V_0 \delta(x_1 - x_2)$ between any triplets or a triplet and a singlet are zero. The perturbation matrix is thus block-diagonal and only the energy of the singlet state gets shifted up by

$$\Delta E_{1s} = V_0 \iint dx_1 dx_2 \left| \frac{\psi_1(x_1)\psi_2(x_2) + \psi_2(x_1)\psi_1(x_2)}{\sqrt{2}} \right|^2 \delta(x_1 - x_2) \tag{51}$$

$$= 2V_0 \int_0^L dx \, |\psi_1(x)\psi_2(x)|^2 = \frac{8V_0}{L^2} \int_0^L \sin^2 \frac{\pi x}{L} \sin^2 \frac{2\pi x}{L} dx = \frac{2V_0}{L}. \tag{52}$$

The four-fold degeneracy of the first excited states are therefore split into $3 + 1$ due to the electron-electron repulsion $V_0 \, \delta(x_1 - x_2)$. The singlet state becomes higher in energy than the triplet states. When the repulsion strength V_0 is large enough to make the singlet ground state energy E_0 higher than the triplet excited states, the triplet states become the new ground states, which leads to a ferromagnetic phase transition when there are many identical copies of the system. This problem tells us that magnetism can arise from spin-independent electron-electron repulsion once the repulsion is strong enough.

6 Perturbation theory

The problem in the previous section uses perturbation theory to deal with electron-electron interactions. Here we give a revision of the general formalism of both time-independent and time-dependent perturbation theories and include some sample problems to apply these theories to. Perturbation theory is a powerful and important technique.

6.1 Time-independent perturbation theory

For simplicity we consider the non-degenerate case up to first and second order. Consider the Hamiltonian

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1, \tag{53}$$

where $\hat{H}_0$ is the unperturbed system, $\hat{H}_1$ is the perturbation applied, and λ is a tuning parameter that controls the strength of perturbation. Suppose the energy of the eigenstate

$|\psi_n^{(0)}\rangle$ of $\hat{H}_0$ is $E_n^{(0)}$. When $\lambda = 1$, the perturbation is fully turned on, $|\psi_n^{(0)}\rangle$ becomes $|\psi_n\rangle$ and $E_n^{(0)}$ becomes E_n. We expand $|\psi_n\rangle$ and E_n to all orders of λ as follows:

$$|\psi_n\rangle = |\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots , \tag{54}$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots . \tag{55}$$

We plug the two series into the eigen-equation $\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle$ and use the normalization condition $\langle\psi_n|\psi_n\rangle = 1$ and the phase condition $\mathrm{Im}\,\langle\psi_n^{(0)}|\psi_n\rangle = 0$ for all $\lambda \in [0, 1]$ to uniquely fix the eigenstate $|\psi_n\rangle$. We have

$$(\hat{H}_0 + \lambda\hat{H}_1)(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \cdots) = (E_n^{(0)} + \lambda E_n^{(1)} + \cdots)(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \cdots), \tag{56}$$

$$(\langle\psi_n^{(0)}| + \lambda\langle\psi_n^{(1)}| + \lambda^2\langle\psi_n^{(2)}| + \cdots)(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \lambda^2|\psi_n^{(2)}\rangle + \cdots) = 1, \tag{57}$$

$$\langle\psi_n^{(0)}|(|\psi_n^{(0)}\rangle + \lambda|\psi_n^{(1)}\rangle + \cdots) = (\langle\psi_n^{(0)}| + \lambda\langle\psi_n^{(1)}| + \cdots)|\psi_n^{(0)}\rangle. \tag{58}$$

Comparing the terms on both sides of Eq. (56) in each order of λ, we obtain

$$\hat{H}_0\,|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(0)}\rangle, \tag{59}$$

$$\hat{H}_0\,|\psi_n^{(1)}\rangle + \hat{H}_1|\psi_n^{(0)}\rangle = E_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}|\psi_n^{(0)}\rangle, \tag{60}$$

$$\hat{H}_0\,|\psi_n^{(2)}\rangle + \hat{H}_1|\psi_n^{(1)}\rangle = E_n^{(0)}|\psi_n^{(2)}\rangle + E_n^{(1)}|\psi_n^{(1)}\rangle + E_n^{(2)}|\psi_n^{(0)}\rangle. \tag{61}$$

Applying the same technique to Eqs. (57) & (58), we have

$$\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle = 1, \quad \langle\psi_n^{(0)}|\psi_n^{(1)}\rangle = 0, \quad \langle\psi_n^{(0)}|\psi_n^{(2)}\rangle = -\frac{1}{2}\langle\psi_n^{(1)}|\psi_n^{(1)}\rangle. \tag{62}$$

One can keep going to higher orders. Projecting Eqs. (60) & (61) onto $\langle\psi_n^{(0)}|$ and acting $\hat{H}_0$ to the left to cancel the $E_n^{(0)}$ term, we obtain

$$\langle\psi_n^{(0)}|\hat{H}_0|\psi_n^{(1)}\rangle + \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(0)}\rangle = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(0)}\rangle \tag{63}$$

$$\Rightarrow \quad E_n^{(1)} = \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(0)}\rangle, \tag{64}$$

$$\langle\psi_n^{(0)}|\hat{H}_0|\psi_n^{(2)}\rangle + \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(1)}\rangle = E_n^{(0)}\langle\psi_n^{(0)}|\psi_n^{(2)}\rangle + E_n^{(1)}\langle\psi_n^{(0)}|\psi_n^{(1)}\rangle + E_n^{(2)} \tag{65}$$

$$\Rightarrow \quad E_n^{(2)} = \langle\psi_n^{(0)}|\hat{H}_1|\psi_n^{(1)}\rangle. \tag{66}$$

The first-order energy correction $E_n^{(1)}$ in Eq. (64) is easy to calculate. The perturbation $\hat{H}_1$ averaged over the unperturbed state $|\psi_n^{(0)}\rangle$ gives the energy correction $E_n^{(1)}$ to first order. The second-order energy $E_n^{(2)}$ requires the first-order correction $\psi_n^{(1)}$ of the wave function, which we now compute. Eq. (60) can be rearranged to become

$$(E_n^{(0)} - \hat{H}_0)|\psi_n^{(1)}\rangle = (\hat{H}_1 - E_n^{(1)})|\psi_n^{(0)}\rangle. \tag{67}$$

From Eq. (62), we know $|\psi_n^{(1)}\rangle$ is orthogonal to $|\psi_n^{(0)}\rangle$, which in the non-degenerate case means that $|\psi_n^{(1)}\rangle$ is in the invertible eigenspace of $E_n^{(0)} - \hat{H}_0$. We may then expand both sides in the eigenstate basis of $\hat{H}_0$ and obtain

$$|\psi_n^{(1)}\rangle = \sum_m{}' \frac{|\psi_m^{(0)}\rangle\langle\psi_m^{(0)}|\hat{H}_1|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}, \tag{68}$$

where the prime $'$ means to exclude the $m = n$ term. Plugging the expansion of $|\psi_n^{(1)}\rangle$ into Eq. (66), we have the second-order energy correction

$$E_n^{(2)} = \sum_m{}' \frac{\langle\psi_n^{(0)}|\hat{H}_1|\psi_m^{(0)}\rangle\langle\psi_m^{(0)}|\hat{H}_1|\psi_n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}} = \sum_m{}' \frac{|\langle\psi_m^{(0)}|\hat{H}_1|\psi_n^{(0)}\rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{69}$$

Therefore $E_n^{(2)}$ is always negative if $\psi_n^{(0)}$ is the ground state of $\hat{H}_0$. Aside from eigenstate expansion, Eq. (67) can also sometimes be solved as a differential equation using the Green's function method as in scattering theory. Generally speaking, the expansion is only recommended if it truncates, otherwise the infinite series can be hard to sum.

6.2 Sample problem

(Problem 3 in Sec. 3, 2011) Consider a 1D harmonic oscillator. The Hamiltonian is

$$\hat{H}_0 = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}m\omega_0^2 x^2. \tag{70}$$

We now consider the effect of a perturbation Hamiltonian

$$\hat{H}' = \lambda x f(t), \tag{71}$$

where λ is a constant, and $f(t)$ is a time-dependent dimensionless function.

(a) Assume that the perturbation Hamiltonian is time-independent, i.e., $f(t) = 1$ at all time t. Find the energy shifts of the ground state to the first nonzero order in f.

Answer: Since the ground state $\psi_0(x) = \langle x|\psi_0\rangle$ is an even function of x, while $\hat{H}'$ is odd in x, the mean value $\langle\psi_0|\hat{H}'|\psi_0\rangle = 0$. So there is no first-order correction to the ground state energy due to $\hat{H}'$. We therefore have to go to the second order. From Eq. (69), the second-order correction of ground-state energy is given by

$$\Delta E_0^{(2)} = \sum_{n=1}^{\infty} \frac{|\langle\psi_n^{(0)}|\hat{H}_1|\psi_0^{(0)}\rangle|^2}{E_0^{(0)} - E_n^{(0)}} = -\lambda^2 \sum_{n=1}^{\infty} \frac{|\langle n|\hat{x}|0\rangle|^2}{n\hbar\omega_0}, \tag{72}$$

where $|n\rangle = |\psi_n^{(0)}\rangle$ is the nth eigenstate of the unperturbed simple harmonic oscillator $\hat{H}_0$. The matrix element can be evaluated using the raising and lowering operators. We have

from Eqs. (2) & (3) in this lecture that

$$\hat{x} = \sqrt{\frac{\hbar}{2m\omega_0}}(\hat{a} + \hat{a}^\dagger). \tag{73}$$

Therefore, the matrix elements can be evaluated using

$$\langle n|\hat{a} + \hat{a}^\dagger|0\rangle = \langle n|\hat{a}^\dagger|0\rangle = \langle n|1\rangle = \delta_{n1}, \tag{74}$$

not to forget checking the normalization constants from Eq. (10). Therefore,

$$\Delta E_0^{(2)} = -\frac{\lambda^2}{2m\omega_0}\sum_{n=1}^{\infty}\frac{\delta_{n1}}{n\omega_0} = -\frac{\lambda^2}{2m\omega_0^2}. \tag{75}$$

We will do a degenerate perturbation problem in solid-state physics in Lecture 5. The part (b) of this problem involves the time-dependent perturbation theory. Instead of going through the general formalisms of Fermi's golden rule, Rabi oscillations and adiabatic theorem, etc., we only briefly review this part by working out the sample problem.

6.3 Time-dependent perturbation theory

(Problem 3 in Sec. 3, 2012) (Continued)

(b) Now assume the perturbation is time-dependent, such that the power spectral density of the time-stationary signal $f(t)$ is given by

$$\tilde{\rho}(\omega) \equiv \lim_{T\to+\infty}\frac{1}{2\pi T}\left|\int_0^T f(t)e^{i\omega t}dt\right|^2 = \rho(\omega) + \rho(-\omega), \tag{76}$$

where the line shape is a Gaussian function

$$\rho(\omega) = \sqrt{\frac{1}{\pi\omega_0^2}}e^{-(\omega-\omega_0)^2/\omega_0^2}, \tag{77}$$

with the constant ω_0 given in the harmonic oscillator. You can verify that

$$\int_{-\infty}^{\infty}\tilde{\rho}(\omega)d\omega = \lim_{T\to+\infty}\frac{1}{T}\int_0^T |f(t)|^2 dt, \tag{78}$$

which gives the total time-averaged power. Assume that the oscillator is in its ground state at $t = 0$. What are the excited states of this oscillator that the ground state can make a transition into? Find the transition rate into these states using Fermi's golden rule.

Answer: In this problem the total Hamiltonian $\hat{H}(t) = \hat{H}_0 + \lambda\hat{x}f(t)$ is time-dependent. We are first asked to derive the selection rule and then determine the transition rates. The

time-independent part $\hat{H}_0$ is a 1D simple-harmonic oscillator. At any time t, we may expand the state $|\psi(t)\rangle$ in terms of the eigenstates of $\hat{H}_0$ as

$$|\psi(t)\rangle = \sum_k c_k(t)|k\rangle e^{-iE_k t/\hbar}. \tag{79}$$

Therefore, from the Schrödinger equation of the state $|\psi(t)\rangle$, we obtain

$$i\hbar \sum_k \dot{c}_k(t)|k\rangle e^{-iE_k t/\hbar} = \lambda f(t) \sum_k c_k(t)\,\hat{x}\,|k\rangle e^{-iE_k t/\hbar}. \tag{80}$$

Projecting both sides onto a harmonic eigenstate $\langle j|$, we have

$$\frac{dc_j(t)}{dt} = -\frac{i\lambda}{\hbar} f(t) \sum_k c_k(t)\langle j|\hat{x}|k\rangle e^{i(E_j - E_k)t/\hbar}. \tag{81}$$

Initially the system is in the ground state $|0\rangle$, so we have $c_k(t=0) = \delta_{k0}$. In case λ is very small, the coefficients will evolve very slowly. We may therefore solve this problem iteratively. We start from a zeroth order approximation $c_k^{(0)}(t) = \delta_{k0}$, which means the coefficients never evolve. We plug this into the right-hand side and get a first-order corrected set of coefficients $c_j^{(1)}(t)$ for all $j \neq 0$ (the initial state) as given by

$$c_j^{(1)}(t) = -\frac{i\lambda}{\hbar} \int_0^t dt' f(t') \sum_k \delta_{k0}\langle j|\hat{x}|k\rangle e^{i(E_j - E_k)t'/\hbar} \tag{82}$$

$$= -\frac{i\lambda}{\hbar}\langle j|\hat{x}|0\rangle \int_0^t dt' f(t') e^{ij\omega_0 t'}. \tag{83}$$

From the results in Part (a), we know $\langle j|\hat{x}|0\rangle$ is nonzero for $j = 1$ only, which tells us that to a first-order approximation, the system can only transition to the first excited state $n = 1$. This is the so-called *selection rule*. Transitions to other states are called *forbidden transitions* (although higher-order terms may allow the system to get to those states eventually). The probability of transitioning to $n = 1$ is therefore given by

$$P_{0\to 1}^{(1)}(t) = |c_1^{(1)}(t)|^2 = \frac{\lambda^2}{\hbar^2}|\langle 1|\hat{x}|0\rangle|^2 \left| \int_0^t dt' f(t') e^{i\omega_0 t'} \right|^2. \tag{84}$$

In case that λ is small, the only way of getting a large transition probability $P_{0\to 1}^{(1)}(t)$ is to wait for a long time. Since we have the power spectral density $\tilde{\rho}(\omega)$ defined in Eq. (76), the asymptotic behavior of $P_{0\to 1}^{(1)}(t)$ as $t \to +\infty$ is given by

$$P_{0\to 1}^{(1)}(t) \approx \frac{2\pi\lambda^2}{\hbar^2}|\langle 1|\hat{x}|0\rangle|^2 \tilde{\rho}(\omega_0) t = \frac{\pi\lambda^2}{m\hbar\omega_0}[\rho(\omega_0) + \rho(-\omega_0)]t, \quad t \to +\infty. \tag{85}$$

This is Fermi's golden rule. Perturbation theory breaks down if t is so large that $P_{0\to 1}^{(1)}(t)$ is no longer small. Nevertheless, it will give us the right transition rate on an intermediate

timescale, which is much longer than $1/\omega_0$ but much shorter than the time required for $P^{(1)}_{0\to 1}(t)$ to get large (which means comparable to 1). Such a timescale definitely exists if λ is small. On the intermediate timescale, we have

$$P^{(1)}_{0\to 1}(t) = \frac{dP^{(1)}_{0\to 1}(t)}{dt} \approx \frac{\pi\lambda^2}{m\hbar\omega_0}[\rho(\omega_0) + \rho(-\omega_0)] = \frac{\sqrt{\pi}\lambda^2}{m\hbar\omega_0^2}\left(1 + \frac{1}{e^4}\right), \tag{86}$$

which is the transition rate from the ground state $n = 0$ to the first excited state $n = 1$. The $1/e^4 = 1.83\%$ term can be neglected as an approximation.

Note: The original problem was formulated in a less rigorous way. It said

$$f(t) = \int_{-\infty}^{\infty} \rho(\omega)(e^{i\omega t} + e^{-i\omega t})d\omega, \tag{87}$$

where the line shape function is the same as Eq. (77). If the problem is formulated this way, then $f(t)$ would be a Gaussian pulse (or wavelet) in the time domain

$$f(t) = \sqrt{\frac{1}{\pi\omega_0^2}} \cdot 2\mathrm{Re}\int_{-\infty}^{\infty} \exp\left[-\frac{(\omega-\omega_0)^2}{\omega_0^2} + i\omega t\right] d\omega = 2e^{-\frac{1}{4}\omega_0^2 t^2}\cos\omega_0 t, \tag{88}$$

so we would not get Fermi's golden rule but some steady-state probability $P^{(1)}_{0\to 1}(t \to +\infty)$ instead. This would make the problem very different.

Lecture 5 Quantum Mechanics - Applications

In this lecture, we talk about some more advanced topics of quantum mechanics that are not covered in Griffith's book. There will be more sample problems than derivations of general theories in this lecture. We try to be broad to cover various topics from high energy physics, particle physics, condensed matter physics, etc.

1 Galilean transformation of QM

(Problem 1 in Sec. 3, 2010) Two observers in different inertial frames will need different wave functions to describe the same physical system. To make things simple, consider how it works in the non-relativistic case. The first observer uses coordinates $(\vec{x}, t)$ and a wave function $\psi(\vec{x}, t)$, while the second observer uses $(\vec{x}\,', t)$ and $\tilde{\psi}(\vec{x}\,', t)$. From Galilean transformation, $\vec{x}\,' = \vec{x} - \vec{v}t$, where $\vec{v}$ is a constant velocity. The wave functions for the two observers are said to be related as follows:

$$\tilde{\psi}(\vec{x}\,', t) = \psi(\vec{x}, t) \exp\left(-\frac{i}{\hbar}\left[m\vec{v}\cdot\vec{x} - \frac{mv^2}{2}t\right]\right). \tag{1}$$

Despite the innocuous look (it's just a phase!), this transformation has interesting effects.

(a) Let us first verify that it makes sense. Suppose $\psi(\vec{x}, t)$ is the wave function of a free particle of momentum $\vec{p} = (p_x, p_y, p_z)$. Show that $\tilde{\psi}(\vec{x}\,', t)$ indeed describes a free particle with the proper momentum.

Answer: Since we are already given the transformation from $\psi(\vec{x}, t)$ to $\tilde{\psi}(\vec{x}\,', t)$, which can be derived from the Galilean covariance of the Schrödinger equation, solving the problem now becomes some straightforward math work. The wave function of a free particle in the $\vec{x}$ frame with momentum $\vec{p}$ is given by

$$\psi(\vec{x}, t) = \exp\left(\frac{i}{\hbar}\left[\vec{p}\cdot\vec{x} - \frac{p^2}{2m}t\right]\right), \tag{2}$$

up to an arbitrary normalization constant with an arbitrary phase. In the new frame of $\vec{x}\,' = \vec{x} - \vec{v}t$, from Eq. (1), the wave function becomes

$$\tilde{\psi}(\vec{x}\,', t) = \exp\left(\frac{i}{\hbar}\left[(\vec{p} - m\vec{v})\cdot(\vec{x}\,' + \vec{v}t) - \frac{p^2}{2m}t + \frac{mv^2}{2}t\right]\right) \tag{3}$$

$$= \exp\left(\frac{i}{\hbar}\left[(\vec{p}-m\vec{v})\cdot\vec{x}\,' - \frac{(\vec{p}-m\vec{v})^2}{2m}t\right]\right), \tag{4}$$

which is the wave function of a free particle with momentum $\vec{p}\,' = \vec{p} - m\vec{v}$.

(b) Now let's put this to work. Suppose we have a hydrogen atom, which at $t < 0$ was at rest with the electron in the ground state $|1s\rangle$ described by the wave function

$$\psi(\vec{x}) = \psi_{1s}(\vec{x}) = \frac{1}{\sqrt{\pi a_B^3}}\exp\left(-\frac{r}{a_B}\right), \quad r = |\vec{x}|, \tag{5}$$

where a_B is the Bohr radius. Suppose at $t = 0$ the proton suddenly starts to move (e.g. due to a collision with a neutron) in the z direction with the velocity v. Let the change in the velocity be so abrupt that the electron wave function remains the same. What is the probability at $t > 0$ to find the moving hydrogen atom with the electron to be still in the ground state $\psi_{1s}(\vec{x})$?

Answer: After $t = 0$, the proton moves at a constant velocity v relative to the lab frame. Therefore, in its co-moving inertial frame, we still have a time-independent Coulomb potential, in which the probabilities of finding the electrons in different $|nlm\rangle$ states are constant. To go into that frame, we transform the wave function $\psi(\vec{x}, t = 0)$ to

$$\tilde{\psi}(\vec{x}\,', t = 0) = \frac{1}{\sqrt{\pi a_B^3}}\exp\left(-\frac{r}{a_B}\right)\exp\left(-\frac{i}{\hbar}m\vec{v}\cdot\vec{x}\right). \tag{6}$$

Since the velocity $\vec{v}$ is in the z direction, we have at $t = 0$ that

$$\vec{v}\cdot\vec{x} = \vec{v}\cdot\vec{x}\,' = vr'\cos\theta', \tag{7}$$

so that the wave function is written as

$$\tilde{\psi}(\vec{x}\,', t = 0) = \frac{1}{\sqrt{\pi a_B^3}}\exp\left(-\frac{r'}{a_B}\right)\exp\left(-\frac{i}{\hbar}mvr'\cos\theta'\right). \tag{8}$$

Suppose now we measure the $|nlm\rangle$ state of $\tilde{\psi}(\vec{x}\,', t = 0)$. The probability amplitude of finding the electron in the ground state $\tilde{\psi}_{1s}(\vec{x}\,')$ is given by

$$\int d^3x'\tilde{\psi}^*_{1s}(\vec{x}\,')\tilde{\psi}(\vec{x}\,', t = 0) = \int \frac{d^3x'}{\pi a_B^3}\exp\left(-\frac{2r'}{a_B}\right)\exp\left(-\frac{i}{\hbar}mvr'\cos\theta'\right) \tag{9}$$

$$= 2\pi\int_0^\infty \frac{r'^2dr'}{\pi a_B^3}\exp\left(-\frac{2r'}{a_B}\right)\int_0^\pi \sin\theta' d\theta'\exp\left(-\frac{i}{\hbar}mvr'\cos\theta'\right) \tag{10}$$

$$= \frac{4\hbar}{mva_B^3}\int_0^\infty r'dr'\exp\left(-\frac{2r'}{a_B}\right)\sin\frac{mvr'}{\hbar} = \left(\frac{4\hbar^2}{4\hbar^2 + m^2v^2a_B^2}\right)^2. \tag{11}$$

The above integral is most conveniently calculated by taking the imaginary part of a gamma integral using Euler's formula $e^{i\theta} = \cos\theta + i\sin\theta$. The ground state probability is therefore the modulus squared of the probability amplitude given by

$$P = \left(\frac{4\hbar^2}{4\hbar^2 + m^2v^2a_B^2}\right)^4. \tag{12}$$

The probability $P \to 1$ if $mva_B \ll \hbar$, and $P \to 0$ if $mva_B \gg \hbar$.

2 Dirac equation (Relativistic QM)

(Problem 1 in Sec. 3, 2011) The Dirac equation can be written in terms of two Pauli-type 2-component spinors, ϕ and χ. For a free particle of mass m and momentum $\vec{p}$, these are plane waves whose spinor coefficients satisfy

$$\begin{cases} m\phi + \vec{p}\cdot\vec{\sigma}\chi = E\phi, & (13) \\ \vec{p}\cdot\vec{\sigma}\phi - m\chi = E\chi, & (14) \end{cases}$$

where E is the energy eigenvalue and $\vec{\sigma} = (\sigma^1, \sigma^2, \sigma^3)$ are the Pauli matrices. We work in natural units $c = \hbar = 1$.

(a) Solve the two equations for E and give a physical interpretation of every solution for a given momentum $\vec{p}$.

Answer: The Dirac equation is the relativistic version of the Schrödinger equation of a spin-1/2 particle. In this problem, we solve for free-particle energy eigenstates of a given momentum $\vec{p}$. The two coupled matrix equations can be rewritten as

$$(E - m)\phi = \vec{p}\cdot\vec{\sigma}\chi, \qquad (E + m)\chi = \vec{p}\cdot\vec{\sigma}\phi. \tag{15}$$

Since $E \pm m$ are c-numbers that commute with any 2×2 matrix $\vec{p}\cdot\vec{\sigma}$, we have

$$(E^2 - m^2)\phi = (\vec{p}\cdot\vec{\sigma})^2\phi = \begin{pmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{pmatrix}^2 \phi = p^2\phi, \tag{16}$$

where $p = |\vec{p}| = \sqrt{p_x^2 + p_y^2 + p_z^2}$ is the magnitude of the momentum $\vec{p}$. The same eigen-equation holds for χ. To get a nonzero ϕ (or χ), we have

$$E = \pm\sqrt{m^2 + p^2}. \tag{17}$$

Clearly, the $E = \sqrt{m^2 + p^2}$ solution gives the relativistic dispersion relation of a free particle with mass m and momentum $\vec{p}$. But what about $E = -\sqrt{m^2 + p^2}$? Do the negative-energy states really exist? If so, why would particles occupy the positive-energy states instead of the negative ones that are lower in energy?

This also puzzled Dirac. To solve this problem, he made a wildly brilliant assumption based on the the Pauli exclusion principle. He assumed that the negative-energy states have already been occupied, forming the so-called *Dirac sea*, so that a positive-energy particle cannot drop into one of these states by Pauli exclusion. We can't feel the existence of the Dirac sea because it is everywhere and we are used to it (like the atmospheric pressure). Dirac also deduced that if a negative-energy particle in the Dirac sea gets excited into a positive-energy state, it will leave behind a hole in the Dirac sea. This hole corresponds to a "deficit" of an electron with momentum $\vec{p}$ and energy $-\sqrt{m^2+p^2}$. The hole as an excitation would thus appear to be an anti-particle with energy $\sqrt{m^2+p^2}$ and momentum $-\vec{p}$. The antiparticle of an electron is a *positron*. Dirac predicted its existence in 1928, and four years later, in 1932, the positron was observed experimentally.

(b) The scattering by a spin-independent central potential V between the initial and final momenta $\vec{p}$ and $\vec{p}\,'$ is described in the Born approximation by the matrix element

$$V(\vec{p}\,',\vec{p})\left(\phi^\dagger_{\vec{p}\,'}\phi_{\vec{p}}+\chi^\dagger_{\vec{p}\,'}\chi_{\vec{p}}\right),$$

where $V(\vec{p}\,',\vec{p})$ is the spatial matrix element of V between momentum eigenstates. Using your solution to part (a), show that

$$\chi^\dagger_{\vec{p}\,'}\chi_{\vec{p}}=\phi^\dagger_{\vec{p}\,'}\Sigma(\vec{p}\,',\vec{p})\phi_{\vec{p}}, \tag{18}$$

where $\Sigma(\vec{p}\,',\vec{p})=F_1(\vec{p}\,',\vec{p})+\vec{\sigma}\cdot\vec{F}_2(\vec{p}\,',\vec{p})$. Determine F_1 and $\vec{F}_2$.

Answer: This problem wants us to find the relativistic correction to a spin-independent scattering problem. In non-relativistic quantum mechanics, we only have one 2-component spinor ϕ, instead of ϕ and χ. From Eq. (15), we see that ϕ and χ are related by

$$\chi^\dagger_{\vec{p}\,'}\chi_{\vec{p}}=\phi^\dagger_{\vec{p}\,'}\frac{(\vec{p}\,'\cdot\vec{\sigma})(\vec{p}\cdot\vec{\sigma})}{(E_{\vec{p}}+m)(E_{\vec{p}\,'}+m)}\phi_{\vec{p}}=\phi^\dagger_{\vec{p}\,'}\Sigma(\vec{p}\,',\vec{p})\phi_{\vec{p}}. \tag{19}$$

One may verify the formula $(\vec{A}\cdot\vec{\sigma})(\vec{B}\cdot\vec{\sigma})=(\vec{A}\cdot\vec{B})I_{2\times2}+i(\vec{A}\times\vec{B})\cdot\vec{\sigma}$. Therefore,

$$\Sigma(\vec{p}\,',\vec{p})=\frac{(\vec{p}\cdot\vec{p}\,')I_{2\times2}+i(\vec{p}\,'\times\vec{p})\cdot\vec{\sigma}}{(E_{\vec{p}}+m)(E_{\vec{p}\,'}+m)}. \tag{20}$$

Coefficients F_1 and $\vec{F}_2$ can be read from this expression. The relativistically corrected matrix element of the scattering potential is therefore given by

$$V(\vec{p}\,',\vec{p})\left(\phi^\dagger_{\vec{p}\,'}\phi_{\vec{p}}+\chi^\dagger_{\vec{p}\,'}\chi_{\vec{p}}\right)=V(\vec{p}\,',\vec{p})\,\phi^\dagger_{\vec{p}\,'}\left[I_{2\times2}+\frac{(\vec{p}\cdot\vec{p}\,')I_{2\times2}+i(\vec{p}\,'\times\vec{p})\cdot\vec{\sigma}}{(E_{\vec{p}}+m)(E_{\vec{p}\,'}+m)}\right]\phi_{\vec{p}}. \tag{21}$$

(c) Suppose that the electron is initially moving in the x_3 direction, with its spin aligned also in the x_3 direction, and that it is scattered by a very weak potential V into the x_2

direction. Show that, in the non-relativistic limit, the probability of *spin flip* in the course of such a scattering event is small. Quantify what is meant by "small" here.

Answer: In relativity, spin and spatial motion are related and mutually convertible. The initial state of the electron is given by $\phi = (1,0)^T$, i.e., spin up in x_3 direction. Given that its initial momentum $\vec{p} = p\hat{z}$ is also in the x_3 direction, we obtain

$$\chi = \frac{\vec{p}\cdot\vec{\sigma}}{E+m}\phi = \frac{p}{E+m}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \frac{p}{E+m}\begin{pmatrix}1\\0\end{pmatrix}. \tag{22}$$

We are interested in the transition probability into the final state with $\phi' = (0,1)^T$, i.e., spin down in x_3 direction, and $\vec{p}\,' = p\hat{y}$ in the x_2 direction. We are assuming an elastic scattering $E_{\vec{p}} = E_{\vec{p}\,'} = E$. The spinor χ' corresponding to ϕ' with momentum $\vec{p}\,'$ is given by

$$\chi' = \frac{\vec{p}\,'\cdot\vec{\sigma}}{E+m}\phi' = \frac{p}{E+m}\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \frac{p}{E+m}\begin{pmatrix}-i\\0\end{pmatrix}. \tag{23}$$

The relativistically corrected matrix element in Part (b) is therefore

$$\langle\vec{p}\,'\downarrow|V|\vec{p}\uparrow\rangle = V(\vec{p}\,',\vec{p})\left(\phi'^{\dagger}\phi + \chi'^{\dagger}\chi\right) = \frac{ip^2}{(E+m)^2}\,V(\vec{p}\,',\vec{p}). \tag{24}$$

One can obtain the same result using Eq. (21). The transition rate as in Fermi's golden rule from $|\vec{p}\uparrow\rangle$ to $|\vec{p}\,'\downarrow\rangle$ is therefore proportional to

$$|\langle\vec{p}\,'\downarrow|V|\vec{p}\uparrow\rangle|^2 = \frac{p^4}{(E+m)^4}|V(\vec{p}\,',\vec{p})|^2. \tag{25}$$

In order for this to be small, we require $p \ll E+m = \sqrt{p^2+m^2}+m$, or equivalently $p \ll 2mc$ in SI units. This problem tells us that the spin cannot be flipped by a spin-independent potential in non-relativistic quantum mechanics, but can be flipped when p is comparable to mc, where m is the rest mass.

3 Spin-orbit interaction

(Problem 2 in Sec. 4, 2010) Consider a hydrogen atom. The spin-orbit interaction

$$H_{\rm so} = A(r)\vec{S}\cdot\vec{L}, \tag{26}$$

where $\vec{S}$ and $\vec{L}$ are the spin and orbital angular momenta of the electron. In Gaussian units,

$$A(r) = \frac{e^2}{2m_e^2c^2r^3}. \tag{27}$$

(a) Describe in words the origin of the spin-orbit interaction.

Answer: In Griffith's book §6.3.2, the spin-orbit interaction is explained using the classical picture of an electron orbiting a proton. In the co-moving frame of the electron, we see the proton orbiting the electron, forming a circular current that generates a magnetic field along $\vec{L}$ at the electron. The magnetic dipole moment of the electron spin is along $-\vec{S}$, so the coupling term in the Hamiltonian is proportional to $\vec{S}\cdot\vec{L}$. The co-moving frame of the electron is a non-inertial frame. When transforming back to the lab frame, the coefficient of $\vec{S}\cdot\vec{L}$ gets corrected by a factor of $\frac{1}{2}$ due to the Thomas precession.

Griffith's explanation implies that the type of force that binds the electron in the atom matters. The spin-orbit coupling constant would be different if the electrostatic potential were replaced by, e.g., a gravitational potential of the same value. This is because the spin-orbit coupling is a relativistic effect. Even though in the lab frame a gravitational potential can be made equal to an electrostatic potential, the difference emerges in the co-moving frame of the electron, because the electromagnetic potentials $(\phi, \vec{A})$ and the gravitational curvature tensor of spacetime would Lorentz transform differently.

(b) Construct the basis wave functions that diagonalize $H_{\rm so}$.

Answer: Without $H_{\rm so}$, electronic states are labeled by $|nlmm_s\rangle$. Now that $\vec{S}$ and $\vec{L}$ are coupled, we transform from the uncoupled basis to the coupled basis $|nljm_j\rangle$, where j and m_j are the quantum numbers of the total angular momentum $\vec{J}=\vec{L}+\vec{S}$. The two bases are related by the Clebsch-Gordan coefficients. In the coupled basis, the spin-orbit interaction $H_{\rm so}$ in Eq. (26) is angular diagonal. We have

$$\langle n'l'j'm'_j|H_{\rm so}|nljm_j\rangle = \langle n'l'|A(r)|nl\rangle\langle l'j'm'_j|\vec{S}\cdot\vec{L}|ljm_j\rangle. \tag{28}$$

Since $\vec{S}\cdot\vec{L}=\frac{1}{2}(\vec{J}^2-\vec{L}^2-\vec{S}^2)$ takes eigenvalues in the coupled basis, we have

$$\langle l'j'm'_j|\vec{S}\cdot\vec{L}|ljm_j\rangle = \frac{\hbar^2}{2}\left[j(j+1)-l(l+1)-\frac{3}{4}\right]\delta_{ll'}\delta_{jj'}\delta_{m_jm'_j}. \tag{29}$$

The radial integral $\langle n'l'|A(r)|nl\rangle$ need not be orthogonal with respect to n and n'. But the mixing of $n\neq n'$ states is only a small relativistic correction compared with the energy splitting due to the principal quantum numbers n, n' before considering $H_{\rm so}$. So we normally neglect this effect and restrict ourselves to the same electron shell $n=n'$.

(c) Obtain the spin-orbit interaction energies for hydrogen in the state with principal quantum number $n=2$.

Answer: For $n=2$, we have the $l=0$ and $l=1$ states ($2s$ and $2p$). From Eq. (29), there is no spin-orbit coupling for $l=0$, because $j=1/2$ and the matrix element of $\vec{S}\cdot\vec{L}$ is zero. We are left with the $l=1$ and $j=1/2$ or $3/2$ states.

The rest of this answer is not required for the exam problem. Since the radial wave function $R_{21}(r)$ of hydrogen is not given, you are not required to evaluate the radial integral $\langle 21|A(r)|21\rangle$ explicitly but only need to give a formal answer. From

$$R_{21}(r) = \frac{r}{2\sqrt{6}a_B^{5/2}}\, e^{-r/2a_B}, \tag{30}$$

we can evaluate the radial integral

$$\langle 21|A(r)|21\rangle = \int_0^\infty 4\pi r^2 dr \left(\frac{r}{2\sqrt{6}a_B^{5/2}}\, e^{-r/2a_B}\right)^2 \frac{e^2}{2m_e^2c^2r^3} \tag{31}$$

$$= \frac{\pi e^2}{12m_e^2c^2a_B^3} = \frac{\pi\alpha^4 m_e c^2}{12\hbar^2}, \tag{32}$$

where $\alpha = e^2/\hbar c = 1/137$ is the fine structure constant in Gaussian units, $a_B = \hbar/\alpha m_e c = 0.529$ Å is the Bohr radius. From Eq. (29), we have for $l = 1$ that

$$\langle ljm_j|\vec{S}\cdot\vec{L}|ljm_j\rangle = \frac{\hbar^2}{2} \text{ or } -\hbar^2 \tag{33}$$

which corresponds to $j = 3/2$ or $1/2$, respectively. From Eq. (28), the spin-orbit coupling energy is thus correspondingly given by

$$\langle nljm_j|H_{\rm so}|nljm_j\rangle = \frac{\pi\alpha^4 m_e c^2}{24} \text{ or } -\frac{\pi\alpha^4 m_e c^2}{12}. \tag{34}$$

The energy splitting due to spin-orbit coupling is $\mathcal{O}(\alpha^4 m_e c^2)$, smaller than the ground-state energy $\mathcal{O}(\alpha^2 m_e c^2)$ by a factor of α^2. Numerically we have $m_e c^2 = 0.511$ MeV, $\alpha^2 m_e c^2 = 27.2$ eV $= 1$ Hartree, and $\alpha^4 m_e c^2 = 1.45$ meV is the energy scale of spin-orbit effects.

4 Solid-state physics

(Problem 4 in Sec. 4, 2013) Consider a two-dimensional square lattice in which noninteracting electrons move. We will treat the motion of the electrons in the nearly free motion approximation. We consider the effects of the lattice using a two-dimensional potential. Two possibilities to describe the potential are:

Choice A,

$$V(x,y) = V_0 \cos\frac{2\pi x}{a}\cos\frac{2\pi y}{a}. \tag{35}$$

Choice B,

$$V(x,y) = V_0\left(\cos\frac{2\pi x}{a} + \cos\frac{2\pi y}{a}\right). \tag{36}$$

Here a is the lattice constant and V_0 is the strength of the potential. Let us ask whether these potentials will open a band gap in the electronic states at the point in momentum space located at $(\pi/a, \pi/a)$.

(a) In the limit where V_0 goes to zero, what is the degeneracy of the lowest-energy state at the point $(k_x, k_y) = (\pi/a, \pi/a)$?

Answer: This problem requires some understanding of solid-state physics and band theory. In the nearly free limit, the problem is also a good testing ground for degenerate perturbation theory. In a crystal with many noninteracting electrons, the single-particle Schrödinger equation is given by

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + V(\vec{x})\right]\psi(\vec{x}) = E\psi(\vec{x}), \tag{37}$$

where $V(\vec{x})$ is the periodic lattice potential that can be expanded into a Fourier series

$$V(\vec{x}) = \sum_{\vec{G}} V_{\vec{G}} e^{i\vec{G}\cdot\vec{x}}, \tag{38}$$

where $\vec{G}$ sums over the whole reciprocal lattice. In the momentum space, the Schrödinger equation (37) becomes

$$\frac{p^2}{2m_e}\psi(\vec{p}) + \sum_{\vec{G}} V_{\vec{G}}\,\psi(\vec{p} - \hbar\vec{G}) = E\psi(\vec{p}). \tag{39}$$

As we can see from Eq. (39), $\psi(\vec{p})$ only gets related with $\psi(\vec{p} - \hbar\vec{G})$ in the eigen-equation, where $\vec{G}$ is any lattice wave vector. In the limit of $V_0 \to 0$, we have a free electron with energy $E = \hbar^2\pi^2/ma^2$ when it is in momentum space located at $\vec{k}_1 = (\pi/a, \pi/a)$. There are 3 other momenta in the first Brillouin zone $\vec{k}_2 = (-\pi/a, \pi/a)$, $\vec{k}_3 = (-\pi/a, -\pi/a)$, and $\vec{k}_4 = (\pi/a, -\pi/a)$, which have the same energy and differ from $\vec{k}_1$ by a lattice wave vector. Therefore, we need to solve a 4-fold degenerate perturbation problem.

(b) In the presence of a nonzero value of V_0, which one of the two potentials will open a band gap at the point $(k_x, k_y) = (\pi/a, \pi/a)$? Justify.

Answer: For choices A and B, we have

$$V_A(x, y) = \frac{V_0}{4}\left[e^{i\frac{2\pi}{a}(x+y)} + e^{-i\frac{2\pi}{a}(x+y)} + e^{i\frac{2\pi}{a}(x-y)} + e^{-i\frac{2\pi}{a}(x-y)}\right]. \tag{40}$$

$$V_B(x, y) = \frac{V_0}{2}\left(e^{i\frac{2\pi}{a}x} + e^{-i\frac{2\pi}{a}x} + e^{i\frac{2\pi}{a}y} + e^{-i\frac{2\pi}{a}y}\right). \tag{41}$$

The matrices of $V_A(x,y)$ and $V_B(x,y)$ under plane wave basis $\vec{k}_1$, $\vec{k}_3$, $\vec{k}_2$, $\vec{k}_4$ are given by

$$V_A = \frac{V_0}{4}\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad V_B = \frac{V_0}{2}\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}. \tag{42}$$

We changed the order of the 4 wave vectors to give V_A and V_B better block structures. For small V_0, we restrict ourselves to the 4-dimensional subspace spanned by $\vec{k}_1, \vec{k}_2, \vec{k}_3, \vec{k}_4$, because other momenta that have a different energy will contribute very little to the wave function. This is the idea of degenerate perturbation theory. We now find the eigenvalues of V_A and V_B to get the energy corrections to the plane waves. We have

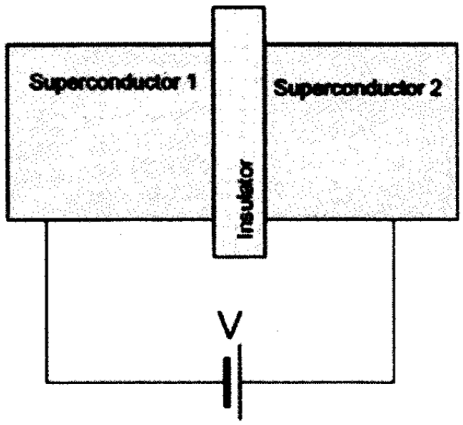
$$\det\left(EI_{4\times 4} - V_A\right) = \left[E^2 - \left(\frac{V_0}{4}\right)^2\right]^2 = 0 \quad \Rightarrow \quad E = \pm\frac{V_0}{4}, \tag{43}$$

$$\det\left(EI_{4\times 4} - V_B\right) = E^2(E^2 - V_0^2) = 0 \quad \Rightarrow \quad E = 0, \pm V_0. \tag{44}$$

Therefore, a band gap is opened at $(\pi/a, \pi/a)$ by V_A but not by V_B, because there is an $E = 0$ solution that is unaffected by V_B to first order. The 4-fold degeneracy split into $2 + 2$ for both choices A and B.

5 Josephson junction

(Problem 2 in Sec. 4, 2010) Two superconductors separated by a thin insulator (called a Josephson junction) are connected to a DC voltage V, as shown in the figure below.



Let ψ_1 be the wave function of the Bose-condensed superconducting electron pairs on one side of the superconductor and ψ_2 be the wave function on the other side. The two wave

functions are related by the time-dependent Schrödinger equations

$$\begin{cases} i\hbar\dfrac{d\psi_1}{dt} = eV\psi_1 + K\psi_2, & (45) \\ i\hbar\dfrac{d\psi_2}{dt} = -eV\psi_2 + K\psi_1, & (46) \end{cases}$$

where the constant K is a real parameter of the junction related to the tunneling process of the electron pairs across the insulator, and V is the DC voltage applied. In this problem, the wave functions are expressed as $\psi_1 = \sqrt{n_1}\, e^{i\theta_1}$ and $\psi_2 = \sqrt{n_2}\, e^{i\theta_2}$, where n_1 and n_2 are the condensation densities, and θ_1 and θ_2 are the phases of the condensate wave functions of superconductors 1 and 2, respectively.

(a) Show that the current density of this junction is given by

$$J = \frac{dn_1}{dt} = -\frac{dn_2}{dt} = J_0 \sin\delta, \tag{47}$$

where $\delta = \theta_2 - \theta_1$. Find the expression for J_0 in terms of K, n_1 and n_2.

Answer: From the Schrödinger equations for ψ_1 and ψ_2, we can derive

$$\begin{cases} i\hbar\dfrac{d}{dt}(\psi_1^*\psi_1) = K(\psi_1^*\psi_2 - \psi_2^*\psi_1), & (48) \\ i\hbar\dfrac{d}{dt}(\psi_2^*\psi_2) = K(\psi_2^*\psi_1 - \psi_1^*\psi_2). & (49) \end{cases}$$

Using $\psi_1 = \sqrt{n_1}\, e^{i\theta_1}$ and $\psi_2 = \sqrt{n_2}\, e^{i\theta_2}$, we obtain

$$J = \frac{dn_1}{dt} = -\frac{dn_2}{dt} = \frac{2K}{\hbar}\,\mathrm{Im}\,(\psi_1^*\psi_2) = \frac{2K}{\hbar}\sqrt{n_1 n_2}\,\sin(\theta_2 - \theta_1). \tag{50}$$

Since $\delta = \theta_2 - \theta_1$, our result agrees with Eq. (47), with $J_0 = \frac{2K}{\hbar}\sqrt{n_1 n_2}$.

(b) Assume that initially at $t = 0$ the condensation densities n_1 and n_2 are equal and large, and the tunneling probability K is small so we have $n_1(t) \approx n_2(t)$ afterwards. Show that the current density J derived in Part (a) oscillates periodically over time. Find the frequency of oscillation in terms of the applied DC voltage V.

Answer: From the Schrödinger equations of ψ_1 and ψ_2 again, we can derive

$$\begin{cases} i\hbar\dfrac{d}{dt}(\psi_2^*\psi_1) = 2eV\psi_2^*\psi_1 + K(n_2 - n_1), & (51) \\ i\hbar\dfrac{d}{dt}(\psi_1^*\psi_2) = -2eV\psi_1^*\psi_2 + K(n_1 - n_2). & (52) \end{cases}$$

Adding these two equations, we cancel the $K(n_2 - n_1)$ term and obtain

$$\frac{d}{dt}(\sqrt{n_1 n_2}\cos\delta) = -\frac{2eV}{\hbar}\sqrt{n_1 n_2}\sin\delta. \tag{53}$$

Since the densities $n_1(t) \approx n_2(t)$ are approximately equal and $n_1(t) + n_2(t) = \text{const}$, the amplitude $J_0 \propto \sqrt{n_1 n_2}$ takes its extremum and its oscillation is negligible. We can then assume $\sqrt{n_1 n_2} \approx \text{const}$ and obtain

$$\frac{d}{dt}\cos\delta = -\sin\delta\,\frac{d\delta}{dt} = -\frac{2eV}{\hbar} \quad \Rightarrow \quad \delta = \frac{2eV}{\hbar}t + \delta_0, \tag{54}$$

where δ_0 is the initial phase difference of ψ_1 and ψ_2. So we have from Eq. (50) that

$$J = \frac{2K}{\hbar}\sqrt{n_1 n_2}\,\sin\left(\frac{2eV}{\hbar}t + \delta_0\right), \tag{55}$$

which gives us the oscillating frequency $\omega = 2eV/\hbar$. Interestingly, we don't get an equation of motion for J that resembles the simple-harmonic oscillator, but get a linear equation for δ which produces the oscillation. This is because the frequency is determined by the applied voltage V rather than some intrinsic parameters.

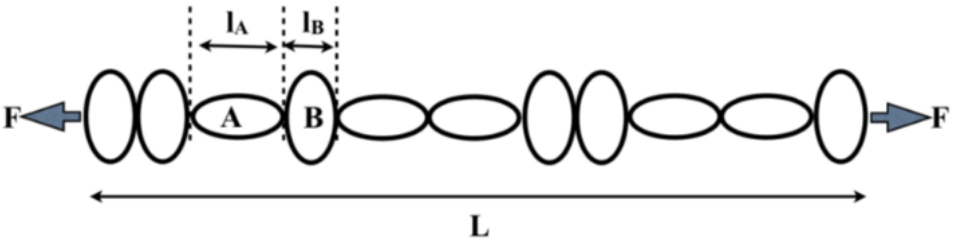
Lecture 6 Thermodynamics and Other Topics

Thermodynamics and lots of other topics such as statistical physics of noninteracting particles, dynamics of ideal fluids, geometric optics and many more can appear in Secs. 5 – 6 of the Quals. Aside from exact calculations, order of magnitude estimation and dimension analysis can be useful as well.

1 Statistical physics

1.1 Boltzmann distribution

(Problem 5 in Sec. 5, 2013) Consider a one-dimensional chain consisting of N molecules. Each molecule can exist in one of two configurations, A and B, with energies E_A and E_B and lengths l_A and l_B, respectively.



(a) Write down the partition function for the system at temperature T, assuming the external force applied $F = 0$.

Answer: Let's denote the state of the ith molecule as n_i, which can be either A or B. The molecules are assumed distinguishable so that it is valid to label them, and sequenced so that they never swap positions. The partition function is the sum of the Boltzmann factor over all microstates

$$Z = \sum_{\{n_i\}} \exp\left(-\beta \sum_{i=1}^{N} E_{n_i}\right), \tag{1}$$

where $\beta = 1/k_B T$ and $\{n_i\} = (n_1, n_2, \ldots, n_N)$. Some manipulation of Z yields that

$$Z = \sum_{\{n_i\}} \prod_{i=1}^{N} \exp(-\beta E_{n_i}) = \prod_{i=1}^{N} \sum_{n_i = A,B} \exp(-\beta E_{n_i}) = Z_1^N, \tag{2}$$

where $Z_1 = e^{-\beta E_A} + e^{-\beta E_B}$ is the partition function of a single molecule. For noninteracting and distinguishable particles, the partition function of the canonical system they form is simply the product of the individual partition functions, i.e., $Z = Z_1^N$.

(b) Assume that $E_A > E_B$ and $l_A > l_B$ (still with $F = 0$). Give an expression for the average length L of the molecule chain as a function of temperature at fixed large N. Sketch the result, give the high and low-T limits and indicate the characteristic temperature at which the length changes from the high-T to the low-T limits.

Answer: The mean value of the total length $L = \sum_j l_{n_j}$ is given by

$$\langle L \rangle = \frac{1}{Z} \sum_{\{n_i\}} \left[\left(\sum_{j=1}^{N} l_{n_j} \right) \prod_{i=1}^{N} \exp(-\beta E_{n_i}) \right] = \frac{1}{Z_1^N} \sum_{j=1}^{N} \sum_{\{n_i\}} l_{n_j} \prod_{i=1}^{N} \exp(-\beta E_{n_i}) \tag{3}$$

$$= \frac{1}{Z_1^N} \sum_{j=1}^{N} \sum_{n_j = A,B} l_{n_j} \exp(-\beta E_{n_j}) \sum_{\{n_i\} \backslash n_j} \prod_{i \neq j} \exp(-\beta E_{n_i}) \tag{4}$$

$$= \frac{1}{Z_1} \sum_{j=1}^{N} \sum_{n_j = A,B} l_{n_j} \exp(-\beta E_{n_j}) = N \cdot \frac{l_A \exp(-\beta E_A) + l_B \exp(-\beta E_B)}{\exp(-\beta E_A) + \exp(-\beta E_B)}. \tag{5}$$

The mean total length is the sum of the mean lengths of each molecule. In the low-T limit, $\beta = 1/k_B T \to +\infty$, we have $\exp(-\beta E_A) \ll \exp(-\beta E_B)$ given $E_A > E_B$. All molecules want to stay in the lower-energy state B. The length is then $\langle L \rangle \approx N l_B$. In the high-$T$ limit, $\beta = 1/k_B T \to 0$, we have $\exp(-\beta E_A) \approx \exp(-\beta E_B) \approx 1$. This makes the A and B states nearly equally likely as their energy difference $E_A - E_B$ is much smaller than the thermal fluctuation $k_B T$. The length is then $\langle L \rangle \approx N(l_A + l_B)/2$. The transition temperature will be of the scale $T \sim (E_A - E_B)/k_B$.

(c) Calculate the linear response function $\chi = \partial \langle L \rangle / \partial F$ showing how the length changes if a small tensile force is applied to the two ends of the chain as shown in the figure.

Answer: Since the force applied is a constant, we may introduce its potential energy and include it into the total energy of the system. The partition function becomes

$$Z = \sum_{\{n_i\}} \exp \left[-\beta \left(\sum_{i=1}^{N} E_{n_i} - F \sum_{i=1}^{N} l_{n_i} \right) \right] = \prod_{i=1}^{N} \sum_{n_i = A,B} \exp \left[-\beta (E_{n_i} - F l_{n_i}) \right] = Z_1^N, \tag{6}$$

where the partition function of a single molecule now becomes

$$Z_1(\beta, F) = \exp \left[-\beta (E_A - F l_A) \right] + \exp \left[-\beta (E_B - F l_B) \right]. \tag{7}$$

Repeating the same derivation as in Part (b), we obtain

$$\langle L\rangle = N\cdot\frac{l_A\exp\left[-\beta(E_A-Fl_A)\right]+l_B\exp\left[-\beta(E_B-Fl_B)\right]}{\exp\left[-\beta(E_A-Fl_A)\right]+\exp\left[-\beta(E_B-Fl_B)\right]}. \tag{8}$$

Let's define $\Delta(F)=E_A-E_B-F(l_A-l_B)$, so the isothermal linear response function

$$\chi_T=\left(\frac{\partial\langle L\rangle}{\partial F}\right)_T=N\frac{\partial}{\partial F}\left(\frac{l_A+l_B\,e^{\beta\Delta(F)}}{1+e^{\beta\Delta(F)}}\right)_\beta=N\frac{\partial}{\partial F}\left(l_B+\frac{l_A-l_B}{1+e^{\beta\Delta(F)}}\right)_\beta \tag{9}$$

$$=-\frac{N(l_A-l_B)}{\left[1+e^{\beta\Delta(F)}\right]^2}\,e^{\beta\Delta(F)}\,\beta\frac{d\Delta(F)}{dF}=\frac{N(l_A-l_B)^2}{4k_BT}\,\mathrm{sech}^2\left[\frac{\Delta(F)}{2k_BT}\right]. \tag{10}$$

The elasticity function $\chi_T=\chi_T(T,F)$ is the reciprocal of the spring constant (or stiffness). In case $E_A>E_B$ and $l_A>l_B$, the elasticity χ_T increases as F increases if $F<F_c=(E_A-E_B)/(l_A-l_B)$. Then χ_T decreases as F increases in the $F>F_c$ regime. If E_A-E_B and l_A-l_B have opposite signs, the elasticity χ_T decreases monotonically. This problem gives a model for the non-constant elasticity of rubber bands.

1.2 Bose-Einstein distribution

(Problem 6 in Sec. 5, 2013) A large isolated box of volume V is filled with black body radiation in thermal equilibrium. The radiation has an initial total energy U_0 (relative to vacuum in the same volume). A cyclic engine pulls heat from the boxed radiation, which is subsequently dumped into a large thermal reservoir whose temperature remains fixed at T.

(a) What is the minimum value of work W that must be expended on the cyclic engine to bring the radiation in the box arbitrarily close to being a vacuum if the boxed radiation temperature was initially T?

Answer: The partition function of the boxed black body radiation is given by

$$Z=\prod_{\vec{p},\zeta}\sum_{n=0}^{\infty}e^{-\beta n\epsilon_{\vec{p}}}=\prod_{\vec{p},\zeta}\frac{1}{1-e^{-\beta cp}}, \tag{11}$$

where $\vec{p}$ sums over all single-photon momenta allowed by the box V and ζ sums over the two possible polarizations. The dispersion relation is $\epsilon_{\vec{p}}=cp$ for photons. The Helmholtz free energy $F=-k_BT\ln Z$ of the boxed black body radiation (photon gas) is given by

$$F=k_BT\sum_{\vec{p},\zeta}\ln(1-e^{-\beta cp})=k_BT\,\frac{2V}{(2\pi\hbar)^3}\int_0^\infty 4\pi p^2\ln(1-e^{-\beta cp})dp \tag{12}$$

$$=\frac{k_B^4T^4V}{\pi^2\hbar^3c^3}\int_0^\infty u^2\ln(1-e^{-u})du=-\frac{\pi^2k_B^4T^4V}{45\hbar^3c^3}. \tag{13}$$

Recall that the integral can be evaluated using the Taylor expansion of logarithm, i.e.,

$$\int_0^\infty u^2 \ln(1-e^{-u})du = -\sum_{n=1}^\infty \frac{1}{n}\int_0^\infty u^2 e^{-nu}du = -\sum_{n=0}^\infty \frac{2}{n^4} = -2\zeta(4), \tag{14}$$

where $\zeta(4) = \pi^4/90$. The Riemann zeta function $\zeta(2n)$ at even integers can be computed analytically by applying Parseval's identity to a constructed Fourier series. We will not reproduce the result here. The numbers $\zeta(2) = \pi^2/6$, $\zeta(3) = 1.202$, $\zeta(4) = \pi^4/90$ can be put on your formula sheet if necessary. Now the total internal energy of the photon gas

$$U_0 = \frac{1}{Z}\prod_{\vec{p},\zeta}\sum_{n=0}^\infty n\epsilon_{\vec{p}} e^{-\beta n\epsilon_{\vec{p}}} = -\frac{1}{Z}\left(\frac{\partial Z}{\partial \beta}\right)_V = -T^2\frac{\partial}{\partial T}\left(\frac{F}{T}\right)_V = \frac{\pi^2 k_B^4 T^4 V}{15\hbar^3 c^3}. \tag{15}$$

Therefore, if we want the heat engine to dump all the energy U_0 of the photon gas into the heat reservoir at temperature T, the minimum amount of heat we must dump into the reservoir, in order not to violate the increase of entropy, is TS, where S is the entropy of the photon gas. The minimum work that has to be done by the heat engine is therefore

$$W = TS - U_0 = -F = \frac{\pi^2 k_B^4 T^4 V}{45\hbar^3 c^3} = \frac{U_0}{3}. \tag{16}$$

(b) What is the average energy in the lowest-frequency mode $\omega_0 = 2\pi f_0$ at arbitrary temperature T?

Answer: The average energy of a single mode is equal to the energy $\epsilon_0 = \hbar\omega_0$ of that mode times the average number of photons, which is given by Bose-Einstein distribution

$$\langle n_0\rangle = \sum_{n=0}^\infty n e^{-\beta n\epsilon_0} \Big/ \sum_{n=0}^\infty e^{-\beta n\epsilon_0} = \frac{1}{e^{\beta\epsilon_0}-1}. \tag{17}$$

Therefore, the average energy is given by

$$\langle E_0\rangle = \epsilon_0\langle n_0\rangle = \frac{\hbar\omega_0}{e^{\beta\hbar\omega_0}-1}. \tag{18}$$

1.3 Fermi-Dirac distribution

(Problem 5 in Sec. 5, 2019) Electron-positron equilibrium

(a) A box is filled with Planckian radiation of temperature $k_B T \gg m_e c^2$, where m_e is the electron mass. A population of $e^\pm$ pairs is maintained in equilibrium with radiation through the reaction $e^+ + e^- \leftrightarrow \gamma + \gamma$. Find the number density of positrons in the box.

Answer: In this problem, the box is a fixed-volume system with variable number of particles of 3 types. And it is assumed to be electrically neutral. Therefore, at equilibrium,

all three species have zero chemical potential. The number of positrons (or electrons) in the box volume V is therefore given by

$$\langle N_e \rangle = \sum_{\vec{p},s} \frac{1}{e^{\beta\epsilon_{\vec{p}}}+1} = \frac{2V}{(2\pi\hbar)^3}\int_0^\infty \frac{4\pi p^2 dp}{e^{\beta\sqrt{m_e^2c^4+p^2c^2}}+1}, \tag{19}$$

where $\vec{p}$ sums over all single-positron (electron) momenta allowed by the boundary conditions of the box V and s sums over the 2 possible spins. Let $\epsilon = \sqrt{m_e^2c^4+p^2c^2}$ be the new integration variable, so that the density

$$\langle n_e \rangle = \frac{\langle N_e \rangle}{V} = \frac{1}{\pi^2\hbar^3c^3}\int_{\epsilon_0}^\infty \frac{\sqrt{\epsilon^2-\epsilon_0^2}}{e^{\beta\epsilon}+1}\,\epsilon\, d\epsilon, \tag{20}$$

where $\epsilon_0 = m_ec^2$ is the rest-energy of a positron (or electron). In case that $k_BT \gg m_ec^2$, we have $\beta\epsilon_0 \to 0$. Therefore we set $\epsilon_0 \approx 0$ as an approximation to obtain

$$\langle n_e \rangle \approx \frac{1}{\pi^2\hbar^3c^3}\int_0^\infty \frac{\epsilon^2 d\epsilon}{e^{\beta\epsilon}+1} = \frac{3\zeta(3)}{2\pi^2}\frac{k_B^3T^3}{\hbar^3c^3} = 0.183\left(\frac{k_BT}{\hbar c}\right)^3, \tag{21}$$

where we have used the formula

$$\int_0^\infty \frac{x^n dx}{e^x+1} = \left(1-\frac{1}{2^n}\right) n!\,\zeta(n+1), \tag{22}$$

and $\zeta(3) = 1.202$, both of which were given in the original exam problem.

(b) Consider the same problem but now assume that the box is also filled with neutral electron-proton matter at the equilibrium temperature T. The proton number density n_p is given. Find the electron chemical potential μ assuming $\mu \ll k_BT$.

Answer: At temperature $k_BT \gg m_ec^2$, the neutral electron-proton matter quickly dissociates into free electrons and protons. We therefore have 4 noninteracting species: protons, positrons, electrons and photons. The charge neutrality condition $n_e^- = n_e^+ + n_p$ requires that there are more electrons than positrons. This makes the chemical potential μ of the electrons higher than the chemical potential μ' of the positrons. Another condition to satisfy is that the reversible reaction

$$e^+ + e^- \leftrightarrow \gamma + \gamma \tag{23}$$

is in phase equilibrium, so we have $\mu + \mu' = 0$. Therefore, $\mu' = -\mu$ and from Eq. (20), the positron and electron densities are given by

$$\langle n_e^\pm \rangle = \frac{1}{\pi^2\hbar^3c^3}\int_{\epsilon_0}^\infty \frac{\sqrt{\epsilon^2-\epsilon_0^2}}{e^{\beta(\epsilon\pm\mu)}+1}\,\epsilon\, d\epsilon. \tag{24}$$

Assuming $\mu \ll k_B T$, we have $\beta\mu \ll 1$ and can then use the Taylor expansion

$$\frac{1}{e^{\beta(\epsilon\pm\mu)}+1} = \frac{1}{e^{\beta\epsilon}+1} \mp \frac{\beta\mu e^{\beta\epsilon}}{(e^{\beta\epsilon}+1)^2} + \cdots \tag{25}$$

in terms of $\beta\mu$ to obtain

$$n_p = \langle n_e^- \rangle - \langle n_e^+ \rangle \approx \frac{2}{\pi^2\hbar^3 c^3} \int_{\epsilon_0}^{\infty} \frac{\beta\mu e^{\beta\epsilon}}{(e^{\beta\epsilon}+1)^2} \sqrt{\epsilon^2 - \epsilon_0^2}\, \epsilon\, d\epsilon. \tag{26}$$

From the condition $k_B T \gg m_e c^2$ given in Part (a), we have $\beta\epsilon_0 \ll 1$ and therefore

$$n_p \approx \frac{2}{\pi^2\hbar^3 c^3} \int_0^{\infty} \frac{\beta\mu e^{\beta\epsilon}}{(e^{\beta\epsilon}+1)^2}\, \epsilon^2\, d\epsilon = -\frac{2\beta\mu}{\pi^2\hbar^3 c^3} \frac{\partial}{\partial\beta} \int_0^{\infty} \frac{\epsilon\, d\epsilon}{e^{\beta\epsilon}+1} \tag{27}$$

$$= \frac{4k_B^2 T^2 \mu}{\pi^2\hbar^3 c^3} \int_0^{\infty} \frac{x\, dx}{e^x+1} = \frac{2k_B^2 T^2 \mu}{\pi^2\hbar^3 c^3} \zeta(2) = \frac{k_B^2 T^2 \mu}{3\hbar^3 c^3}, \tag{28}$$

using Eq. (22) and $\zeta(2) = \pi^2/6$. Therefore the chemical potential of electrons

$$\mu = \frac{3\hbar^3 c^3}{k_B^2 T^2}\, n_p. \tag{29}$$

This formula is for small n_p only such that $\mu \ll k_B T$ is satisfied.

2 Thermodynamics

2.1 Equation of state

(Problem 6 in Sec. 5, 2011) Suppose you have discovered a substance whose equation of state is given by $p = 5T^3/V$. Note that the units for the constant "5" must be $\mathrm{Nm/K^3}$. Your colleagues have carried out another experiment where they have found that the internal energy is a function of V and T given by $U(V,T) = BT^x \ln(V/V_0) + f(T)$, where B, x and V_0 are constants, and $f(T)$ is an unmeasured function that only depends on T. What are the values of B and x? (In terms of the constants you already know).

Answer: From the TdS equations of thermodynamics for a closed system, we have

$$dU = TdS - pdV = T\left[\left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV\right] - pdV. \tag{30}$$

We can derive Maxwell's relations using the Helmholtz free energy $F = U - TS$, which gives

$$dF = d(U-TS) = -SdT - pdV \quad \Rightarrow \quad \left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V. \tag{31}$$

Therefore, we can determine the internal energy from the equation of state. We have

$$dU = T\left(\frac{\partial S}{\partial T}\right)_V dT + \left[T\left(\frac{\partial p}{\partial T}\right)_V - p\right] dV = c_V dT + \frac{10T^3}{V} dV, \tag{32}$$

where we have plugged in the specific heat at constant volume $c_V = T(\partial S/\partial T)_V$ and the equation of state $p = 5T^3/V$. Now from the colleague's expression for the internal energy $U(V,T) = BT^x \ln(V/V_0) + f(T)$, we have

$$dU = \left[xBT^{x-1} \ln\left(\frac{V}{V_0}\right) + f'(T)\right] dT + \frac{BT^x}{V} dV. \tag{33}$$

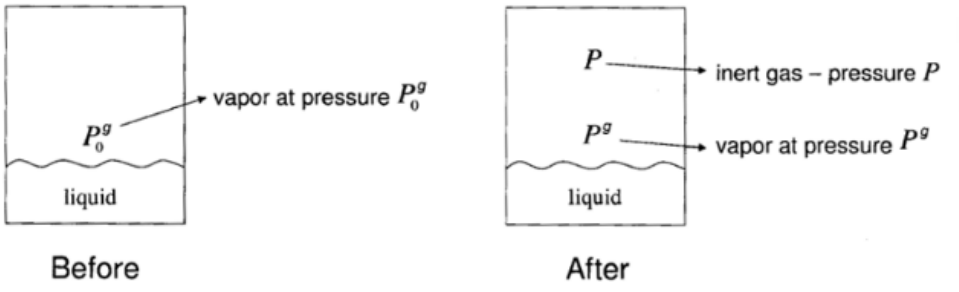
Comparing Eqs. (32) & (33), we obtain

$$\frac{BT^x}{V} = \frac{10T^3}{V} \quad \Rightarrow \quad B = 10\,\mathrm{Nm/K^3},\ x = 3, \tag{34}$$

so that the internal energy $U(V,T) = 10T^3 \ln(V/V_0) + f(T)$ is determined. This problem gives an example that in thermodynamics, we can use the equation of state to determine the internal energy from Maxwell's relations.

2.2 Liquid-gas phase transition

(Problem 6 in Sec. 5, 2009) Consider a liquid and its vapor, both at the same temperature in equilibrium with each other. Let the equilibrium vapor pressure be P_0^g. An inert, insoluble gas is added to the closed container holding the liquid. Find an expression for the new equilibrium partial vapor pressure P^g of the liquid in terms of V^l_{mol}, the molar volume of the liquid, P, the partial pressure of the inert gas, P_0^g, the vapor pressure without the inert gas, and T, the temperature of the system. Assume the gas phase of the liquid obeys the ideal gas law, and the liquid is incompressible.



Answer: The TdS equation for the Gibbs free energy of an open system is given by

$$dG = -SdT + Vdp + \mu dn, \tag{35}$$

where $n = N/N_A$ is the amount of substance. And from $G = \mu n$, we have

$$d\mu = -\frac{S}{n}dT + \frac{V}{n}dp = -S_{\text{mol}}dT + V_{\text{mol}}dp. \tag{36}$$

This is a thermodynamic identity that applies to both the liquid and its vapor. At a constant temperature, we have $dT = 0$, so we only need to look at the second term. The change of the chemical potential μ^l of the liquid is given by

$$\Delta\mu^l = V^l_{\text{mol}}\Delta p^l = V^l_{\text{mol}}(P + P^g - P^g_0), \tag{37}$$

assuming V^l_{mol} is incompressible. The change of the chemical potential μ^g of the vapor is

$$\Delta\mu^g = \int_{P^g_0}^{P^g} V^g_{\text{mol}}(p)dp = \int_{P^g_0}^{P^g} \frac{RT}{p}dp = RT\ln\frac{P^g}{P^g_0}, \tag{38}$$

where $R = k_B N_A = 8.314$ J/(mol $\cdot$ K) is the ideal gas constant. Therefore, by requiring $\Delta\mu^l = \Delta\mu^g$, we obtain

$$V^l_{\text{mol}}(P + P^g - P^g_0) = RT\ln\frac{P^g}{P^g_0} \quad \Rightarrow \quad P^g = P^g_0\,\exp\left[\frac{V^l_{\text{mol}}}{RT}(P + P^g - P^g_0)\right]. \tag{39}$$

This is a transcendental equation for P^g, which may be solved iteratively using P^g_0 as an initial guess. The final pressure $P^g > P^g_0$ when some inert gas with partial pressure P is added to the system. This result may be a bit counter-intuitive. When we add pressure, we normally expect vapor to liquefy. But in this problem, it's the opposite that happens.

The reason lies in the difference between applying pressure to the whole liquid-vapor system and introducing inert gas that only increases the pressure of the liquid while leaving the partial pressure of the vapor unchanged. The first situation is what we normally call "adding pressure". In this case, both μ^l and μ^g increase, but since $V^l_{\text{mol}} \ll V^g_{\text{mol}}$, we have $\mu^g > \mu^l$ so that some vapor has to liquefy to balance the two phases. In the second situation, only μ^l increases because of the extra partial pressure P from the inert gas while μ^g stays the same, so that $\mu^l > \mu^g$. Therefore, the inert gas will cause some liquid to vaporize.

Based on the analysis above, V^l_{mol}/RT should be small because liquid is dense compared with gas. This means good results of P^g can be found by just a few iterations starting from the initial guess P^g_0. Therefore, we obtain the approximate answer

$$P^g \approx P^g_0\,\exp\left(\frac{V^l_{\text{mol}}}{RT}P\right), \tag{40}$$

by doing only one such iteration. You may plug in data of water and find the approximate formula quite accurate compared with the iteration fixed point of Eq. (39) within a much wider range of inert gas pressure P than linear response.

3 Fluid dynamics

3.1 Variables and equations

A classical fluid is described by its local density $\rho(\vec{r},t)$, velocity field $\vec{v}(\vec{r},t)$ and pressure distribution $p(\vec{r},t)$. Fluid dynamics is often studied without considering its coupling with heat conduction. As a result, there is a simple *constitutive relation* of pressure and density

$$p = p(\rho). \tag{41}$$

The mass of the fluid is conserved in a flow. Therefore, we have the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0. \tag{42}$$

Newton's laws of motion applied to a fluid is called *Euler's equation*, which is given by

$$\rho \frac{d\vec{v}}{dt} \equiv \rho \left[\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} \right] = -\nabla p - \rho \nabla \phi, \tag{43}$$

where the total time derivative d/dt is introduced in the sense that we follow a moving mass point in the fluid and consider the changing rate of the velocity $\vec{v}$ of that mass point, i.e., its acceleration. The quantity $\phi = \phi(\vec{r})$ is the potential energy per unit mass of the fluid. A typical example will be the gravitational potential $\phi = gz$. No fluid resistance is included in the equation, which implies a *perfect fluid* (frictionless).

3.2 Bernoulli's equation

In the *steady flow* of a perfect (frictionless) *incompressible* fluid under gravity, at any point along the same streamline, we have *Bernoulli's equation* (1752)

$$p + \rho g z + \frac{1}{2}\rho v^2 = \text{const}, \tag{44}$$

where the mass density $\rho = \text{const}$ no longer depends on the pressure p in an incompressible fluid. We can derive Bernoulli's equation from Euler's equation and the continuity equation. Since $\rho = \text{const}$, the continuity equation (42) simplifies to $\nabla \cdot \vec{v} = 0$. From the steady flow condition, we have $\partial p/\partial t = 0$. And in a time-independent potential field $\phi(\vec{r})$, we have $\partial \phi/\partial t = 0$. Therefore, from Euler's equation (43), we have

$$\rho \vec{v} \cdot \frac{d\vec{v}}{dt} = -\vec{v} \cdot \nabla p - \rho \vec{v} \cdot \nabla \phi = -\frac{d}{dt}(p + \rho\phi), \tag{45}$$

$$\Rightarrow \quad \frac{d}{dt}\left(p + \rho\phi + \frac{1}{2}\rho v^2 \right) = 0. \tag{46}$$

From our definition of d/dt in Eq. (43), the quantity $p + \rho\phi + \frac{1}{2}\rho v^2$ is a constant along any streamline. Plugging in $\phi = gz$, we obtain Bernoulli's equation.

3.3 Sound waves in air

As sound waves travel in air, the uniform pressure p_0 and density ρ_0 (neglecting gravity) get slightly perturbed. Since the air is a poor conductor of heat and sound waves oscillate comparatively very quickly, the constitutive equation can be written as

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0}\right)^\gamma, \tag{47}$$

assuming adiabatic compression of ideal gas, where $\gamma = 7/5$ is the ratio of specific heats of diatomic molecules because the essential constituents N_2 and O_2 are both diatomic. Neglecting gravity, we may simplify Euler's equation (43) into

$$\rho\left[\frac{\partial\vec{v}}{\partial t} + (\vec{v}\cdot\nabla)\vec{v}\right] = -\nabla p. \tag{48}$$

And we also have the continuity equation (42). Now we linearize all three equations in the vicinity of $\rho = \rho_0$, $p = p_0$ and $\vec{v} = 0$. The advective acceleration $(\vec{v}\cdot\nabla)\vec{v}$ is second-order in $\vec{v}$ and thus neglected. By keeping only first-order terms, we have

$$\frac{p_0+\delta p}{p_0} = \left(\frac{\rho_0+\delta\rho}{\rho_0}\right)^\gamma \quad\Rightarrow\quad \delta p = \gamma\,\frac{p_0}{\rho_0}\,\delta\rho, \tag{49}$$

$$(\rho_0+\delta\rho)\,\frac{d\vec{v}}{dt} = -\nabla(p_0+\delta p) \quad\Rightarrow\quad \rho_0\frac{\partial\vec{v}}{\partial t} = -\nabla\delta p, \tag{50}$$

$$\frac{\partial(\rho_0+\delta\rho)}{\partial t} + \nabla\cdot[(\rho_0+\delta\rho)\,\vec{v}] = 0 \quad\Rightarrow\quad \frac{\partial\delta\rho}{\partial t} + \rho_0\nabla\cdot\vec{v} = 0, \tag{51}$$

Plugging Eq. (49) into Eq. (50), applying $\partial/\partial t$ to Eq. (51) and the divergence to Eq. (50), we obtain a linear wave equation for the density fluctuation $\delta\rho$ given by

$$\frac{\partial^2\delta\rho}{\partial t^2} = \gamma\,\frac{p_0}{\rho_0}\,\nabla^2\delta\rho. \tag{52}$$

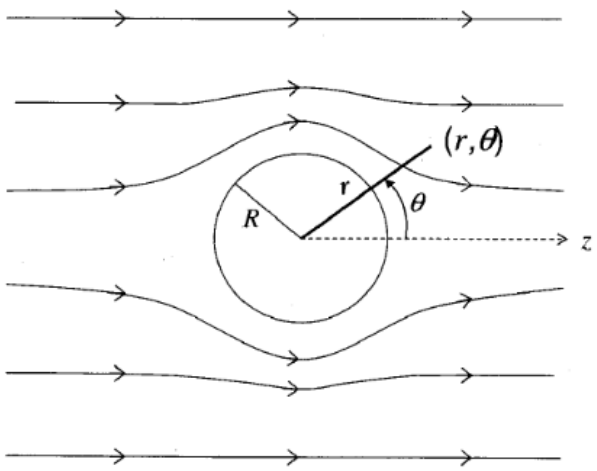
Therefore, the speed of sound in air is found to be

$$c = \sqrt{\gamma\,\frac{p_0}{\rho_0}} = \sqrt{\frac{\gamma RT}{M_{\text{air}}^{\text{mol}}}}, \tag{53}$$

where $R = 8.314$ J/(mol · K) is the ideal gas constant and $M_{\text{air}}^{\text{mol}} = 29$ g/mol is the molar mass of air. At 15 °C (59 °F), $T = 288$ K, the speed of sound in air is $u = 340$ m/s. Eq. (53) is called the *Newton-Laplace equation*, which was first published by Newton (1686) and then corrected by Laplace (1816) to include the γ factor by considering the compression of air due to sound waves to be adiabatic rather than isothermal.

3.4 Velocity potential

(Problem 3 in Sec. 5, 2009) The steady-state flow of a fluid is specified by the fluid velocity $\vec{v}(\vec{r})$ at position $\vec{r}$. Consider an incompressible fluid ($\nabla \cdot \vec{v} = 0$) undergoing irrotational flow ($\nabla \times \vec{v} = 0$). An impenetrable solid sphere of radius R is fixed in position while the fluid flows around it. Far from the sphere, the fluid flows uniformly in the z direction ($\vec{v} = v_0 \hat{z}$ for $r \to +\infty$). For this azimuthally symmetric problem, find the $\hat{r}$ and $\hat{\theta}$ components of the fluid velocity $\vec{v}(r, \theta)$ for all $r > R$.



Answer: Since the velocity field of the flow satisfies $\nabla \times \vec{v} = 0$, we may introduce a scalar potential ϕ such that $\vec{v} = -\nabla\phi$. Then from $\nabla \cdot \vec{v} = 0$, we have

$$\nabla^2 \phi = 0. \tag{54}$$

Therefore, we get essentially an electrostatic problem. The Laplace equation (54) can be solved by separation of variables in spherical coordinates. Since we already had a lecture on electrostatics (see Lecture 2), we will directly give the general solution here:

$$\phi(r, \theta) = \sum_{l=0}^{\infty} \left(a_l r^l + \frac{b_l}{r^{l+1}} \right) P_l(\cos\theta), \quad r > R, \tag{55}$$

where the $P_l(\cdot)$'s are the Legendre polynomials. However, the boundary conditions here are slightly different. Since the sphere is impenetrable, we have the normal component of velocity equals zero, i.e.,

$$\left. \frac{\partial \phi}{\partial r} \right|_{r=R+0} = 0. \tag{56}$$

At $r \to +\infty$, we require that

$$\phi \approx -v_0 r \cos\theta. \tag{57}$$

The first boundary condition Eq. (56) gives us

$$la_l R^{l-1} - \frac{(l+1)b_l}{R^{l+2}} = 0, \quad , l = 1, 2, 3, \ldots \tag{58}$$

The second boundary condition Eq. (57) gives us

$$a_l = -v_0\, \delta_{l1}. \tag{59}$$

Therefore, we obtain the specific solution

$$\phi(r, \theta) = -v_0 \left(r + \frac{R^3}{2r^2} \right) \cos\theta, \quad r > R. \tag{60}$$

The velocity field for $r > R$ is then given by

$$\begin{cases} v_r = -\dfrac{\partial \phi}{\partial r} = v_0 \left(1 - \dfrac{R^3}{r^3} \right) \cos\theta, & (61) \\ v_\theta = -\dfrac{1}{r} \dfrac{\partial \phi}{\partial \theta} = -v_0 \left(1 + \dfrac{R^3}{2r^3} \right) \sin\theta. & (62) \end{cases}$$

A scalar field is mathematically easier to solve than a vector field. An irrotational vector field is equivalent to the gradient of a scalar potential. This technique is used extensively in electrostatics (electrostatic potential), magnetostatics (magnetic scalar potential), and fluid dynamics (velocity potential), etc.

All these contemporary studies in various fields of 18th–19th-century physics were amazingly related by the same set of PDEs we reviewed in Lecture 2. Modern fluid dynamics involves a lot more richness and complexity, e.g., viscosity, turbulence, Reynolds number, nonlinearity, chaos, etc. These topics are beyond the scope of this book and the Quals, but they do open a new era of complexity science and emergent phenomena.

Congratulations! You have finished my lecture notes of Quals review. May you pass the exams smoothly and be a successful researcher!