

Chapter 1

Historical overview of the developments of quantum mechanics

1.1 Quantum Ideas Course Overview

Course synopsis: The overall purpose of this course is to introduce you all to the core concepts that underlie quantum physics, the key experimental and theoretical developments in the advent of quantum mechanics, the basic mathematical formalism, and give you a flavour of current research in quantum physics. Many of you will have likely been introduced to one or more of the topics that we will be covering, but perhaps not in the same way or in such detail. The aim is to ensure that everyone has the same introduction to this exciting area of physics that underpins nearly all physical phenomena that are observed.

Syllabus and resources: I have put together a course syllabus and website located at (<http://www.physics.ox.ac.uk/users/smithb/qid.html>). There will be 12 lectures over four weeks (three lectures per week). The lectures will take place in the Martin Wood Lecture Theatre in the Clarendon laboratory from 9:00-10:00am on Wednesday, Thursday and Friday of Trinity Term weeks 1, 2, 4, and 5.

Books: There are four recommended books and six supplementary books for your reading pleasure.

Recommended Books:

- *QED: The strange Theory of Light and Matter*, R. P. Feynman (Princeton University Press, 2006)
- *Quantum Theory: A Very Short Introduction*, J. C. Polkinghorne (Ox-

ford University Press, 2002)

- *The New Quantum Universe*, T. Hey and P. Walters (Cambridge University Press, 2003)
- *The Strange World of Quantum Mechanics*, D. F. Styer (Cambridge University Press, 2000)

Supplementary Reading:

- *Thirty Years That Shook Physics: Story of Quantum Theory*, G. Gamow (Dover, 1985)
- *Beyond Measure: Modern Physics, Philosophy and the Meaning of Quantum Theory*, J. Baggott (Oxford University Press, 2003)
- *Theoretical Concepts in Physics*, 2nd ed., M. Longair (Cambridge University Press, 2003)
- *Modern Physics*, K. Krane (Wiley, 2012)
- *Quantum Generations: A History of Physics in the Twentieth Century*, H. Kragh (Princeton University Press, 2002)
- *Feynman Lectures on Physics*, vol. 3, R. P. Feynman, R. B. Leighton, and M. Sands (Addison Wesley, 1971)

Personally, I find Polkinghorne's book to be useful for the core concepts in that it is quite compact, but covers the topics. However, for calculations, I suggested Krane's book, which covers much more than the course.

Lecture notes: The lectures will follow closely Dr. Axel Kuhn's lecture notes, available at

<http://webnix.physics.ox.ac.uk/atomphoton/index.php?dish=lectures>

and I will be writing additional notes as I work through the course. My notes will be posted on the course website as they are completed. Be sure you refresh your web browser when looking for new notes!

Topics covered: We will break down the material into four different topics:

Particle-like properties of light: The success of classical physics, measurements in classical physics. The nature of light, the ultraviolet catastrophe, the photoelectric effect and the quantization of radiation.

Wave-like properties of matter: Interference of atomic beams, discussion of two-slit interference, Bragg diffraction of atoms, quantum eraser experiments. Interferometry with atoms and large molecules (de Broglie wavelength). The Uncertainty principle by considering a

microscope and the momentum of photons, zero point energy, stability and size of atoms. Atomic spectral lines and the discrete energy levels of electrons in atoms, the Frank-Hertz experiment and the Bohr model of an atom.

Introductory mathematical formalism: Schrödinger equation and boundary conditions. Free-space, potential barrier, and tunneling solutions. Solution for a particle in an infinite potential well, to obtain discrete energy levels and wave functions. Superposition of energy-eigenstates and time evolution. Measurements in quantum physics: Expectation values. Dirac notation.

Modern quantum physics: Measurements in quantum physics. Schrödinger's cat and the many-worlds interpretation of quantum mechanics. Magnetic dipoles in homogeneous and inhomogeneous magnetic fields and the Stern-Gerlach experiment showing the quantization of the magnetic moment. The impossibility of measuring two orthogonal components of magnetic moments. A glimpse of quantum engineering and quantum computing: The EPR paradox, entanglement, hidden variables, non-locality and Aspect experiment, quantum cryptography and the BB84 protocol.

Lecture Plan: See the course syllabus for the lecture plan.

1.2 Historical development of quantum physics

1.2.1 Introduction

Scientific enquiry has a long and interesting history. Being a scientist, I cannot help but find the story of how this discipline of human activity has come to be. However, aside from being of purely personal interest I find that it can often be extremely insightful and helpful to understand how certain concepts were developed by the pioneers of science. By understanding the approaches that were taken to conceive, refine, and develop a theory, I believe we can better our own approach to perform research today. For example, knowledge of how quantum physics was developed at the beginning of the 1900s, including the dead ends that did not lead to success, has helped me formulate the theory of photon wave mechanics [?]. Now just because I find the history of a subject to be interesting and useful for research or insight does not imply that a historical perspective is the most suitable for teaching and learning a subject for the first time. This is, in my opinion, the case for quantum mechanics. Thus, I will only outline the key historical events that led to the development of quantum physics. The goal of this is to understand how this subject arose from the interplay of increasingly

accurate experimental observation and theoretical developments to explain newly observed phenomena.

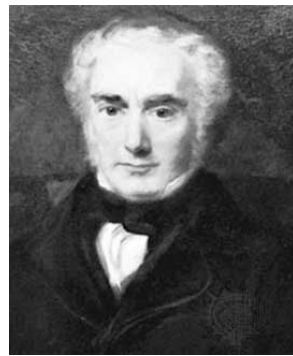
1.2.2 Summary and success of classical physics

The three pillars of classical physics: mechanics, thermodynamics, and electromagnetism, were well understood at the end of the 1800s. The backbone of classical physics can be summarized by the following basic sub-disciplines

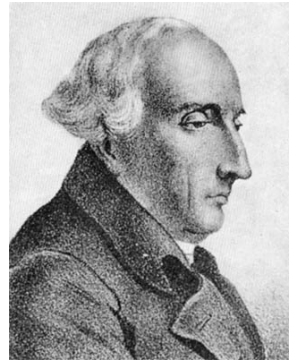
- Mechanics: Based upon Newton's laws, generalized by Lagrange and Hamilton: $\sum \mathbf{F} = m\mathbf{a}$ (Works well enough to send man to the moon and describe micron size objects move).



(a) Isaac Newton



(b) William R. Hamilton



(c) Joseph Louis Lagrange

- Electromagnetism: A crowning unification of three previously independent concepts (electricity, magnetism and optics) largely due to the work of Oersted, Faraday and Maxwell. Light is an electromagnetic wave.



(a) Hans Christian Ørsted



(b) Michael Faraday

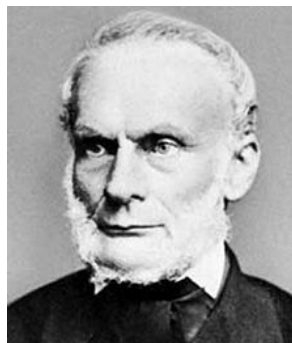


(c) James Clerk Maxwell

- Thermodynamics: Energy conservation and equilibration as set out by Carnot, Clausius, and Lord Kelvin.



(a) Nicolas Léonard Sadi Carnot



(b) Rudolf Clausius



(c) William Thomson

Classical physics has two key foundational differences from quantum theory (as we will see):

Measurement without disturbance: In classical physics a core assumption is that a measurement on a system could reveal information without any disturbance to the system. All that is required to do this is to turn down the interaction strength of the measurement. For example, if we wanted to determine the position and velocity of a particle all we must do is use a weak light source turned on and off quickly to take two snapshots of the particle. The first snapshot gives the initial position \mathbf{x}_i , and the second, taken at a time Δt later, gives its final position \mathbf{x}_f . We can thus determine its velocity as $\mathbf{v} = (\mathbf{x}_f - \mathbf{x}_i)/\Delta t$. According to classical physics one can measure precisely both the position and velocity to arbitrary accuracy. As we will see this contradicts what is observed and predicted for quantum systems. Another perhaps philosophical aside associated with this classical assumption is that the position and velocity are well defined for a system, simply awaiting our observation. In quantum theory, these interrelated ideas do not hold.

Determinism: Classical physics is based upon deterministic differential equations. If one knows precisely the initial values of the position and velocity of all particles in a system at one moment in time, then all future behaviour of the system can be predicted. The universe was likened to a clock, which when set about running deterministically evolves for all time. This concept is nicely summarized by Laplace's demon

“We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this

intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.”

- Pierre Simon Laplace, *A Philosophical Essay on Probabilities*, (1814)



Figure 1.1: Pierre Simon Laplace

We see that uncertainty in classical physics was only due to our own ignorance of the position and velocity of all particles at a given instant in time. The universe was simply one massive initial-value problem and if we knew the initial conditions at one instant in time, we could know the outcome of everything in the future, as well as look back in the past and know all that has been!

1.2.3 Cracks in the foundations

The foundations of classical physics started to show cracks well before the advent of quantum theory. In the 1880s, Michelson and Morley performed a series of experiments that demonstrated the earth does not travel through the luminiferous ether that was thought to permeate the universe. Recall that if light were a wave, then it must be an excitation of some medium and as the earth travelled through the ether, it should have a different velocity depending on the relative velocity of the earth and ether. No difference was observed, and the resulting theory that describes this is none other than Einstein’s special theory of relativity.

It is a common misconception that physicists of the 1890s believed that the fundamental laws of nature had been discovered and all that was left to do was work out the consequences of these. In fact, this was a tumultuous period in which several fundamental unresolved problems occupied some of

the greatest minds of the time. For example, the kinetic theory of gases and the equipartition theorem put forth by Clausius, Maxwell, and Boltzmann was not widely accepted. The atomic and molecular theory of the structure of matter came under attack, as the origin of the “resonances” in molecules, which were assumed to be the origin of spectral lines, was unknown. You must recall that at that time there was no direct evidence for fundamental particles until 1897, when J. J. Thomson discovered the electron.

These unexplained phenomena came about because of increasingly precise experimental results, which had slipped past previous generations because the observations of earlier times were not as precise. This does not limit the applicability of classical physics, which works quite well in its realm of applicability. For example, the new physics that arose at the outset of the 1900s (relativity and quantum physics) are not required for a transatlantic flight or phone call (at least a phone call when I was born).¹ There were several observations that led to the development of quantum theory, which I briefly describe below.

Spectral lines: Starting with a Swiss schoolmaster by the name of Johann Balmer in 1885, who was trying to understand the spectral lines observed in emission from hydrogen.



Figure 1.2: Johann Jakob Balmer

He noticed that there were regularities in the wavelengths of the emitted

¹It is useful to note here the relationship between increased measurement precision, technological advance, and discovery of new physical phenomena. When an experimentalist develops a method to measure some quantity with significantly increased precision beyond what was capable before, this new capability enables one to measure and manipulate systems on a more precise scale. This allows for the development of smaller, faster and more complex technologies that we use in our daily lives. Further more, these technologies can be used to probe new areas of physics previously unexplored. New discoveries from these experiments can then feed back into the development of new measurement techniques and technologies.

lines and found that he could determine the wavelengths with the following formula

$$\lambda = \lambda_0 \left(\frac{1}{4} - \frac{1}{n^2} \right)^{-1}, \quad (1.1)$$

where n is an integer greater than two, and λ_0 is a constant length of 364.56 nm.

In 1900, Swedish physicist Johannes Rydberg refined the expression for hydrogen emission to describe all observed lines, which gave

$$\lambda = \left(\frac{R}{hc} \right)^{-1} \left(\frac{1}{m} - \frac{1}{n^2} \right)^{-1}, \quad (1.2)$$



Figure 1.3: Johannes Robert Rydberg

where m and n are integers ($m < n$), R is known as the Rydberg constant ($R = 13.6$ eV), h is Planck's constant (6.626×10^{-34} Js) and c is the speed of light in vacuum. Although a concise formula for predicting the emission wavelengths for hydrogen were known, there was no physical description for the origin of these discrete lines. The leading theory of the day was that atoms and molecules had certain resonance frequencies at which they would emit, but there was no satisfactory description of the physical origins of these resonances. Furthermore, there were no other closed formulae to predict the emission spectral lines of other, more complex, materials.

Photoelectric effect: When a metal surface is illuminated by light, electrons can be emitted from the surface. This phenomenon is known as the photoelectric effect, and was first discovered by Heinrich Rudolph Hertz in 1887 while investigating electromagnetic radiation. In 1902 Philipp Lénárd observed that the maximum photoelectron kinetic energies are independent of intensity but depend on frequency, which could not be explained by a wave theory of light. Einstein extends the work of Planck and applies it to describe the photoelectric effect.

Heat capacity of solids: Another of the minor details that needed fixing, according to the wisdom of classical physics, was the theory of heat capacity for solids. The specific heat capacity of a material system C , is defined as the amount of heat Q , required to raise the temperature of a set amount of material (typically one mole) by a given amount ΔT

$$C = Q/\Delta T. \quad (1.3)$$

The SI unit of specific heat capacity is a J/mol-K. Recall that according to the equipartition theorem of classical statistical mechanics, the total energy contained in an assembly of a large number of individual particles exchanging energy amongst themselves through mutual interactions is shared equally on average by all the particles. In other words, at temperature T , each atom has an energy of $k_B T/2$ per degree of freedom for both the kinetic and potential energies. For an atom in a crystalline solid, there are three degrees of freedom (associated with the three directions they can wiggle about their equilibrium positions), and thus they have kinetic energy $K = 3/2 k_B T$, and potential energy $U = 3/2 k_B T$, giving total thermal energy stored in the system of $E = 3k_B T$. Thus the amount of heat required to increase the temperature a one mole of atoms by ΔT is given by the difference between the final and initial energies $Q = 3k_B(T + \Delta T) - 3k_B T = 3k_B \Delta T$. Thus the heat capacity is given by $C = 3k_B$, which is independent of temperature. This is known as the Dulong-Petit law after its French discoverers (1819), and derived theoretically by Boltzmann in 1876.

This prediction of classical physics agrees fairly well with experimental observations for most materials near room temperature. However, this temperature independent behaviour was not observed at low temperatures for certain materials, particularly diamond (carbon), boron and silicon. The carbon anomaly had been known since 1841. In experiments published in 1905 it was shown by the Scottish chemist James Dewar that the heat capacity of diamond essentially vanishes near 20 K, and as the temperature of any material approaches absolute zero, the heat capacity should approach zero as well. The solution to this problem was partially solved by Einstein in 1907 by extending the ideas of Max Planck, by assuming that the atoms are constrained to oscillate about their equilibrium positions in a lattice at frequency ν , can oscillate with only discrete energies given by integer multiples of $h\nu$, where h is Planck's constant. Einstein's theory was further refined and gives excellent agreement with experiments by the Dutch physicist P. Debye in 1912. The results are shown in Fig. 1.4.

Blackbody radiation and the ultraviolet catastrophe: The previously mentioned phenomena all played a role in the development of the quantum theory we know today. However, there was one key unresolved problem that led to the discovery of quantum physics and contributed to the resolution of many of the other problems of the time: the origin of the spectrum of

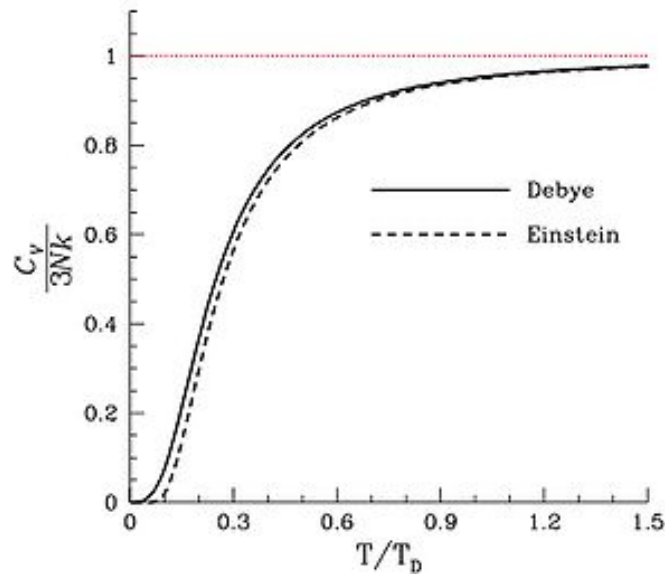


Figure 1.4: Classical, Einstein, and Debye treatment of heat capacity of solids.

blackbody radiation. We will go through this in detail shortly so I delay our discussion until then.

The beginnings of quantum physics: 1900-1932

1900 Blackbody radiation: Quantum physics was introduced to the world by Max Planck, speaking at the meeting of the German Physical Society on 14 December 1900, when he put forth his derivation of the blackbody spectrum. The key assumption was that the energy absorbed and emitted by the oscillators that make up the blackbody is proportional to the frequency of oscillator.

1905 Photoelectric effect: The success of Planck's theory to describe the observed spectral emission for blackbody radiation by constraining the absorption and emission of radiation to discrete energy values prompted a young patent clerk working in Bern Switzerland to apply this idea to an outstanding problem of the time, namely the photoelectric effect. Albert Einstein, one of the few scientists to take Planck's ideas seriously, proposed a quantum of light (the photon), which behaves like a particle.

1907 Heat capacity of solids: Einstein extended the ideas of Max Planck, by assuming that the atoms in a solid are which are constrained to oscillated about their equilibrium positions in a lattice at frequency ν ,



Figure 1.5: Peter Joseph William Debye

can only oscillate with only discrete energies given by integer multiples of $h\nu$, where h is Planck's constant.

- 1911 Nuclear model of atom: Ernest Rutherford infers the nucleus as the result of the alpha-scattering experiment performed by Hans Geiger and Ernest Marsden and proposes the nuclear model of atom, superseding Thomson's "plumb-pudding" model.
- 1913 Bohr's atom: Niels Bohr succeeds in constructing a theory of atomic structure based on Rutherford's nuclear planetary model of the atom and the quantum ideas of Planck and Einstein. The key insight was that there were only discrete energies that the system could have. The electrons were said to occupy stationary states at these energies, which do not radiate electromagnetic energy.
- 1914 Franck-Hertz experiment: James Franck and Gustav Hertz confirm the existence of stationary states through an electron-scattering experiment.
- 1923 X-ray - electron scattering: Arthur Compton discovers the quantum (particle) nature of x rays, thus confirming photons as particles.
- 1924 de Broglie waves: Louis de Broglie proposes that matter has wave properties.
- 1924 Bosons: Satyendra Nath Bose and Albert Einstein find a new way to count quantum particles, later called Bose-Einstein statistics, and they predict that extremely cold atoms should condense into a single quantum state, later known as a Bose-Einstein condensate.
- 1925 Matrix mechanics: Werner Heisenberg, Max Born, and Pascual Jordan develop matrix mechanics, the first complete version of quantum

mechanics, and make an initial step toward quantum field theory.

- 1925 Exclusion principle: Wolfgang Pauli formulates the exclusion principle for electrons in an atom.
- 1926 Wave mechanics: Erwin Schrödinger develops wave mechanics by trying to determine the equations of motion that describe de Broglie's waves. Max Born gives a probability interpretation of quantum mechanics. G.N. Lewis proposes the name "photon" for a light quantum.
- 1926 Fermions: Enrico Fermi and Paul A.M. Dirac find that quantum mechanics requires a second way to count particles, Fermi-Dirac statistics, opening the way to solid-state physics.
- 1926 Quantum theory of light: Dirac publishes seminal paper on the quantization of electromagnetism and quantum field theory is born.
- 1927 Heisenberg Uncertainty Principle: Werner Heisenberg formulates the uncertainty principle: the more you know about a particle's energy, the less you know about the time of the energy (and vice versa.) The same uncertainty applies to momenta and coordinates.
- 1928 Paul Dirac combines quantum mechanics and special relativity to describe the electron.
- 1931 Paul Dirac realizes that the positively-charged particles required by his equation are new objects (he calls them "positrons" which he mistakenly believes is the proton). They are exactly like electrons, but positively charged. This is the first example of antiparticles.
- 1932 Carl David Anderson discovers antimatter, an antielectron called the positron.

Further developments: 1932-1995 Nuclear physics, quantum field theory, superconductivity, and spooky action at a distance

- 1934 Enrico Fermi puts forth a theory of beta decay that introduces the weak interaction. This is the first theory to explicitly use neutrinos and particle flavor changes.
- 1934 Hideki Yukawa combines relativity and quantum theory to describe nuclear interactions by an exchange of new particles (mesons called "pions") between protons and neutrons. From the size of the nucleus, Yukawa concludes that the mass of the conjectured particles (mesons) is about 200 electron masses. This is the beginning of the meson theory of nuclear forces.

- 1935 Albert Einstein, Boris Podolsky, and Nathan Rosen raise concerns about the consequences of quantum theory for correlated quantum systems and put forth the EPR paradox.
- 1942 Richard Feynman puts forth his path integral formulation of quantum mechanics in his PhD thesis.
- 1946-48 Experiments by Isidor Rabi, Willis Lamb, and Polykarp Kusch reveal discrepancies in the Dirac theory of hydrogen.
- 1947 Richard Feynman, Julian Schwinger, and Sin-Itiro Tomonaga develop the first complete theory of the interaction of photons and electrons, quantum electrodynamics, which accounts for the discrepancies in the Dirac theory, giving procedures to calculate electromagnetic properties of electrons, positrons, and photons. Introduction of Feynman diagrams.
- 1951 David Bohm introduced a simplified version of the EPR paradox consisting of correlated spins.
- 1957 Julian Schwinger writes a paper proposing unification of weak and electromagnetic interactions.
- 1957 John Bardeen, Leon Cooper, and J. Robert Schrieffer show that electrons can form pairs whose quantum properties allow them to travel without resistance, providing an explanation for the zero electrical resistance of superconductors. This theory was later termed the BCS theory (after the surname initials of the three physicists).
- 1959 Yakir Aharonov and David Bohm predict that a magnetic field affects the quantum properties of an electron in a way that is forbidden by classical physics. The Aharonov-Bohm effect is observed in 1960 and hints at a wealth of unexpected macroscopic effects.
- 1964 John S. Bell proposes an experimental test, “Bell’s inequalities,” of whether quantum mechanics provides the most complete possible description of a system.
- 1982 Alain Aspect carries out an experimental test of Bell’s inequalities and confirms the completeness of quantum mechanics.
- 1995 Eric Cornell, Carl Wieman, and Wolfgang Ketterle trap clouds of metallic atoms cooled to less than a millionth of a degree above absolute zero, producing Bose-Einstein condensates, which were first predicted 70 years earlier. This accomplishment leads to the creation of the atom laser and superfluid gases.



Figure 1.6: George Gamow

Further reading:

For a well-written, entertaining, and lucid historical perspective on the development of quantum physics, I highly recommend Gamow's book, *Thirty Years That Shook Physics: Story of Quantum Theory*. Not only did he experience this transitional period firsthand, but he also contributed significantly to broad developments in physics from nuclear physics to cosmology. Another interesting view can be found in Longair's, *Theoretical Concepts in Physics*, which gives a useful description of not only the development of quantum physics, but also the other foundations of modern physics. Not only does it present the historical background, but also the physics. Finally, for those interested in the historical development of physics during the last century, Kragh's book, *Quantum Generations: A History of Physics in the Twentieth Century*, is an invaluable resource.

Chapter 2

Particle-like behavior of light

2.1 Photoelectric effect



(a) Heinrich Rudolf Hertz



(b) Albert Einstein



(c) Robert Andrews Millikan

When a metal surface is illuminated by light, electrons can be emitted from the surface. This phenomenon is known as the photoelectric effect, and was first discovered by Heinrich Rudolf Hertz in 1887 while investigating electromagnetic radiation.

The emitted electrons are known as photoelectrons. An example experimental setup to observe this phenomenon is shown in Fig. 2.1. Light incident on the metal surface (the emitter) can release electrons (e) from the surface. These photoelectrons can travel towards a second metal plate (the collector). A variable external voltage (V_{ext}) is applied between the emitter and collector, measured by a voltmeter (V), while an ammeter (A) registers the current i flowing between the emitter and collector. The current measured reflects the number of photoelectrons emitted. The experiment is performed in a vacuum tube so that the photoelectrons do not

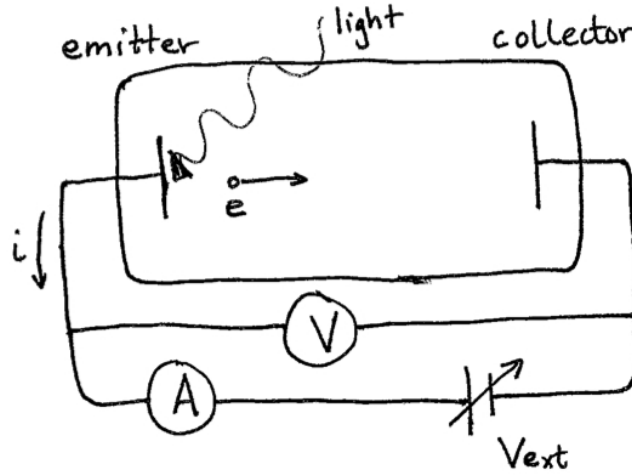


Figure 2.1: Schematic of experimental setup to observe the photoelectric effect. Light is incident on a metal surface (emitter) which can release an electron. The released photoelectron can travel toward a collector, held at potential V_{ext} above the emitter. If they electrons have sufficient energy, they reach the collector and contribute to the current i measured by the ammeter (A). A voltmeter (V) monitors the potential difference between the emitter and collector.

lose any energy from collisions with molecules of the air. This setup allows the rate of photoelectron emission to be measured, which is proportional to the current i , as well as the maximum kinetic energy of the photoelectrons, K_{max} . This is determined by applying a negative potential to the collector that is just sufficient to repel the most energetic photoelectrons, which do not have sufficient energy to climb the potential barrier. This potential, known as the stopping potential V_s , is determined by increasing the magnitude of the voltage until the current measured drops to zero. At this point the maximum kinetic energy is given by $K_{\text{max}} = eV_s$, where e is the magnitude of the electron charge. Typical values of V_s are on the order of 1 V.

2.1.1 Classical analysis

In the classical treatment, the metal surface is illuminated with light of intensity I (W/m^2). An electron on the surface absorbs energy continuously from the wave until the binding energy of the electron to the metal is exceeded, and then the electron is released. The minimum value of energy required to free an electron from the surface of a metal is called the work function or binding energy E_{bind} . Typical binding energies are on the order of a few electron volts (eV). Classical theory predicts

1. The maximum kinetic energy of the photoelectrons should be proportional to the intensity of incident radiation. As the brightness of the light is increased, more energy is delivered to the surface and the emitted photoelectrons should be released with more energy. Another way to think about this is the incident electric field is larger ($I = \epsilon_0 c |\mathbf{E}|^2 / 2$), which implies an increased force on the electron $\mathbf{F} = -e\mathbf{E}$, and thus increased energy when it is emitted from the surface.
2. The photoelectric effect should occur for light of any frequency. According to the wave theory of light, as long as the intensity of the incident radiation is large enough photoelectrons should be able to be emitted.
3. The first photoelectrons should be emitted a finite amount of time after the light hits the surface (on the order of seconds). To estimate the time required to emit a photoelectron, assume that light of intensity $I = 120 \text{ W/m}^2$ (about that of a helium-neon laser) is incident on aluminium ($E_{\text{bind}} = 4 \text{ eV}$). Furthermore, assume the atoms to occupy an area of radius equal to that of an aluminium atom ($r = 0.1 \text{ nm}$). The time taken for the electrons in the atoms to absorb at least the binding energy, and thus be freed, is then

$$\Delta t = \frac{E_{\text{bind}}}{I\pi r^2}, \quad (2.1)$$

which gives approximately 0.2 seconds.

2.1.2 Experimental observations

The experimental characteristics of the photoelectric effect were well known by 1902. However, these differed significantly from the classical predictions.

1. The maximum kinetic energy as determined by the stopping potential is completely independent of the intensity of the incident radiation. Figure 2.2 shows the observed behaviour of the measured current. Doubling the intensity does not change the stopping potential, indicating no change in the kinetic energy of the electrons. The experimental result conflicts with the predictions of wave theory, which predicts the maximum kinetic energy should depend on the intensity of light.
2. Photoelectrons are not emitted for light of any frequency. The light must have a minimum frequency for the photoelectric effect to be observed. The value of this frequency, called the cutoff frequency ν_c , is a characteristic of the material. For light with frequency above ν_c , any intensity will produce photoelectrons. However, for light with frequency below ν_c , no amount of intensity will enable photoelectrons to

be emitted. This experimental result again disagrees with the classical theory.

3. The first photoelectrons are emitted nearly instantaneously (within a few nanoseconds) of the light hitting the surface. Again, the experiment does not fit with classical predictions.

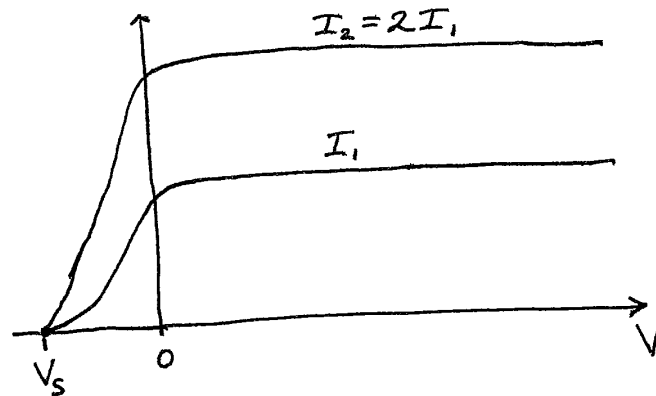


Figure 2.2: Photoelectric current as a function of measured potential difference between emitter and collector for two different values of incident radiation intensity. When the intensity is doubled from I_1 to I_2 , the current is doubled (twice as many photoelectrons emitted), but the stopping potential, V_s remains constant.

2.1.3 Einsteins treatment based upon Plancks idea of light quanta

In 1905, a satisfactory description of the photoelectric effect was given by Einstein.¹ Einstein based his theory on the work of Max Planck, who five years earlier developed a theory to explain the spectral distribution of light emitted by hot, glowing objects (known as blackbody radiation). We will discuss blackbody radiation and Planck's theory in the next lecture. Einstein proposed that the energy of the light wave was not continuously distributed across the wavefront, but instead concentrated in localized bundles, (light quanta), which were later called photons. The energy of a photon associated

¹Although it is often said that the photoelectric effect requires the electromagnetic field to be quantized, i.e. comprised of photons, this is not true. A so-called "semi-classical approach in which the electromagnetic field is treated as a classical wave, while the atoms that make up the material are treated quantum mechanically. The quantized atoms, which can only absorb and emit light at discrete energies, is sufficient to describe the photoelectric effect. This fact should not detract from Einsteins insight at the time, since the quantum theory of atomic systems was still eight years away.

with an electromagnetic wave of frequency ν is

$$E = h\nu, \quad (2.2)$$

where h is known as Planck's constant. Note this can also be written in terms of the wavelength

$$E = \frac{hc}{\lambda}. \quad (2.3)$$

In Einstein's treatment, a photoelectron is released as a result of the absorption of a single photon. The entire energy of the photon is instantaneously transferred to a single photoelectron. If the photon energy $h\nu$ is greater than the work function (binding energy) E_{bind} , of the material then the photoelectron will be released. However, if the photon energy is less than the binding energy, then no photoelectron can be emitted. This accounts for two of the failures of the classical theory: the existence of the cutoff frequency and lack of any measurable time delay in emission.

If the photon energy exceeds the work function, the excess is taken up in the electron kinetic energy

$$K_{\text{max}} = h\nu - E_{\text{bind}}. \quad (2.4)$$

The intensity of light does not appear here. Doubling the intensity of light means that twice as many photons are incident on the surface and thus twice as many photoelectrons should be emitted, but they have the same maximum kinetic energy.

The photon theory given by Einstein gave a satisfactory description of all the observed features of the photoelectric effect. The most detailed experimental test of the theory was performed by the American physicist Robert Millikan in 1915.

He measured the maximum kinetic energy (the stopping potential times the electron charge: $K_{\text{max}} = eV_s$), for different frequencies of light. Plotting V_s versus frequency, we expect

$$V_s = h\nu/e - E_{\text{bind}}/e. \quad (2.5)$$

The slope of the line should give h/e , as shown in Fig. 3. The intercept should give the cutoff frequency and thus the binding energy. Millikan had obtained a value for the electron charge ($e = 1.592 \times 10^{-19}$ C, which is within 1% of the currently accepted value) from his famous "oil drop" experiments, which he used to determine Planck's constant. Millikan obtained a value of $h = 6.57 \times 10^{-34}$ Js, which is quite close to the presently accepted value of approximately 6.626×10^{-34} Js. In part for his careful experiments on the photoelectric effect, Millikan was awarded the 1923 Nobel prize in physics. Einstein was awarded the Nobel prize in physics for his theory of the photoelectric effect.

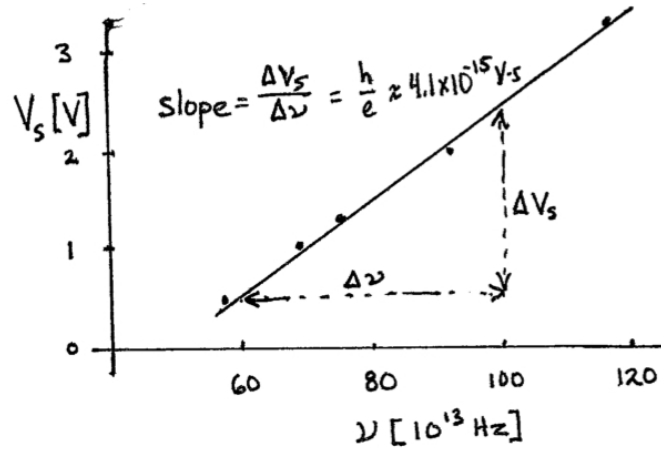


Figure 2.3: Experimental results of photoelectric effect. Stopping voltage V_s as a function of light frequency ν is linear. The slope of the line should be h/e . With previous knowledge of the electron charge from his earlier experiments, Millikan was able to measure the slope of this line and experimentally determine a value for Planck's constant.

2.2 Blackbody radiation

At the turn of the 20th century the tip of the quantum iceberg was revealed by the seemingly everyday phenomenon of thermal radiation, that is the spectral intensity distribution of electromagnetic radiation emitted by an object because of its temperature.

2.2.1 Introduction: Blackbody radiation

We see ordinary objects because light is reflected from their surface from another source and is reflected to our eyes. At room temperature, most thermal radiation is emitted in the infrared region of the spectrum ($\lambda_{\max} \approx 10 \mu\text{m}$), where our eye is not sensitive. However as the temperature of an object is increased, for instance by passing a current through a metal as in an electric filament, the material begins to emit visible light. Unfortunately, the amount of radiation emitted by everyday materials depends not only on the temperature, but also other properties of the material such as the reflectivity of the surface. To simplify these difficulties in calculating the spectrum of heated objects, we consider an idealized object called a blackbody, whose surface is completely absorbing. A blackbody absorbs all incident radiation regardless of frequency or angle of incidence.

A blackbody can be modeled by considering a large cubic cavity, which consists of perfectly reflecting mirrors for the internal walls, with a small hole

cut in one wall as depicted in Fig. 2.4. The hole acts as a blackbody, in that any radiation from outside the cavity that is incident on the hole, will become trapped inside the cavity, bouncing around indefinitely. Thus the hole effectively has perfect absorption, regardless of incidence angle or frequency. Since radiation emerging from the hole is just a sample of the radiation inside the cavity, understanding the radiation inside the cavity allows us to determine the emitted spectra from the hole and thus the blackbody.

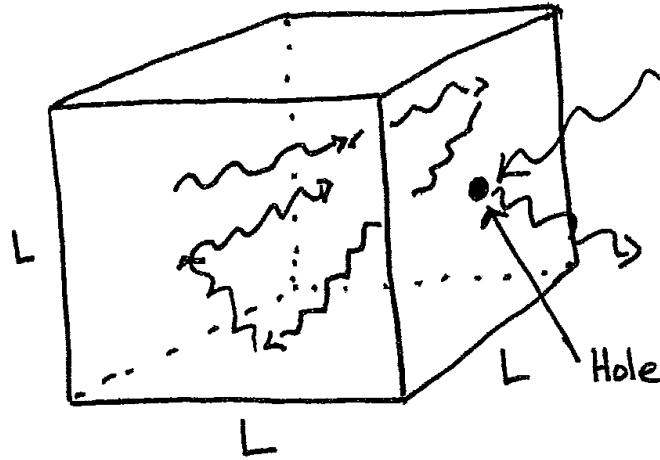


Figure 2.4: An optical cavity with perfect internal mirrors acts as a simplified blackbody.

We must now define what we mean by blackbody spectrum. For our purposes, we are concerned with the energy of radiation per unit frequency per unit volume inside the blackbody cavity with walls at temperature T , called the spectral energy density, $\rho(\nu, T)$. This represents the energy per unit frequency per unit volume inside our blackbody cavity and has units of $\text{J Hz}^{-1} \text{m}^{-3}$. The interpretation of $\rho(\nu, T)d\nu$ is that the amount of electromagnetic energy inside the box, per unit volume in the frequency interval $(\nu, \nu + d\nu)$, is $\rho(\nu, T)d\nu$. The radiation emitted from the hole in the cavity is proportional to this energy density, and thus $\rho(\nu, T)$ determines the spectrum measured by an observer outside the cavity.

The calculation to determine the spectral energy density inside the blackbody cavity can be broken into two parts. First, due to the fact that the radiation inside the blackbody cavity is confined to a finite region of space by the walls, electromagnetic boundary conditions will restrict the number of electromagnetic standing waves (called electromagnetic modes) with frequencies between ν and $\nu + d\nu$. The number of modes inside the cavity, per unit volume, per unit frequency is denoted by $g(\nu)$ and known as the

density of states. A state here refers to a mode of the electromagnetic field, which is determined by the polarization and allowed wavelength, as set by the boundary conditions imposed by the cavity. The total number of modes per unit volume inside the cavity is thus $g(\nu)d\nu$. Each of these modes is then assumed to be in thermal equilibrium with the cavity walls, which are held at temperature T .

The second part of the calculation involves the determination of the average energy of the electromagnetic field at a given frequency of oscillation, $\langle E(\nu, T) \rangle$, in thermal equilibrium at temperature T . The average energy at a given frequency is

$$\langle E(\nu, T) \rangle = \int E P(E, T) dE, \quad (2.6)$$

where $P(E, T)$ is the probability density of energy at temperature T . To determine this quantity, we must rely upon statistical mechanics, which we will discuss shortly.

Thus to calculate the spectral energy density, we multiply the density of states $g(\nu)$ by the average energy of the electromagnetic field at a given frequency $\langle E(\nu, T) \rangle$, giving

$$\rho(\nu, T) = g(\nu) \langle E(\nu, T) \rangle. \quad (2.7)$$

2.2.2 Density of electromagnetic states for a cavity

1D cavity

To calculate the density of states for a cubic cavity of side L , we begin by considering a 1D cavity and then generalize this result. If we consider electromagnetic waves confined to oscillate between two perfectly reflecting mirrors positioned at $x = 0$ and $x = L$, the boundary conditions imply that the field must go to zero at the mirrors. This leads to standing wave solutions such that the an integer number of wavelengths should equal the twice the cavity length

$$n\lambda = 2L. \quad (2.8)$$

The factor of 2 comes from the fact that you need only half integer values of wavelength to satisfy the boundary conditions, as depicted in Fig. ??.

This equation thus determines the allowed wavelengths of radiation in the cavity $\lambda_n = 2L/n$, where n is a mode label, running from 0 to ∞ . The frequency associated with a wavelength λ is simply given by $\nu = c/\lambda$, which combined with the constraints on wavelength imposed by the cavity gives the allowed frequencies of the cavity

$$\nu_n = \frac{c}{2L} n. \quad (2.9)$$

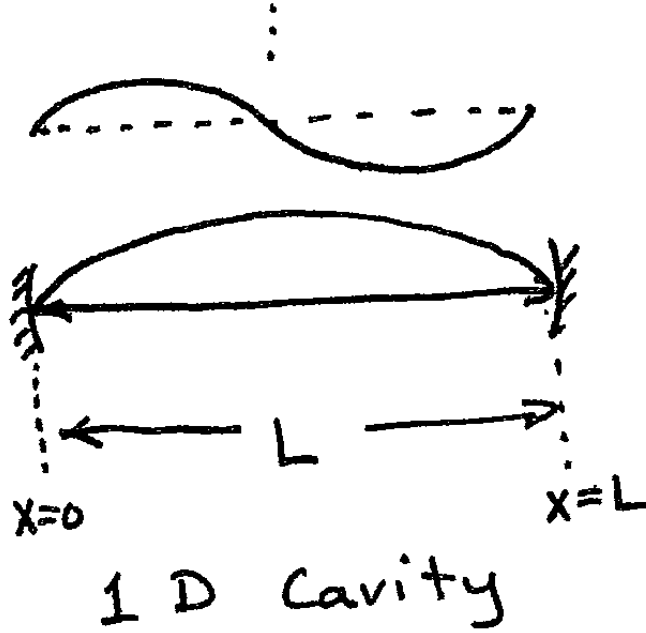


Figure 2.5: One-dimensional cavity with the first two allowed states.

Thus, the electric field amplitude can be written as

$$E(x, t) = \sum_n A_n \sin(k_n x - \omega_n t), \quad (2.10)$$

where the sum is taken over mode labels n . Here A_n is the amplitude, $k_n = 2\pi/\lambda_n$ is the wave vector, and $\omega_n = 2\pi c/\lambda_n = \pi$ is the angular frequency of each mode labeled by n .

The number of modes with frequency in the range $(\nu, \nu + d\nu)$ can be calculated by counting the number of mode indices that give frequencies in this range (and multiplying by 2 for the two possible polarizations associated with each wavelength). The number of mode indices from 0 to ν is given by $\nu/\delta\nu$ where $\delta\nu = c/2L$ is the frequency spacing between subsequent modes. Similarly, the number of mode indices from 0 to $\nu + d\nu$ is $(\nu + d\nu)/\delta\nu$. Thus, the number of mode indices between ν and $\nu + d\nu$ is given by the difference

$$\frac{\nu + d\nu - \nu}{\delta\nu} = \frac{2L}{c} d\nu. \quad (2.11)$$

Thus, the density of states for a 1D cavity is given by the number of modes divided by the cavity “volume”, which for a 1D cavity is just L

$$g_{1D}(\nu) d\nu = 2 \frac{1}{L} \frac{2L}{c} d\nu = \frac{4}{c} d\nu, \quad (2.12)$$

where the factor of 2 out front arises from the two possible polarizations for each mode.

Another way to think of this is in terms of the wave vector $k = 2\pi/\lambda = 2\pi\nu/c$. The allowed wave vectors, $k_n = 2\pi/\lambda_n = n\pi/L$ are on a lattice of equal spacing $\delta k = \pi/L$. The number of modes with frequency in the range $(\nu, \nu + d\nu)$ is equal to the number of modes with wave vectors that have length between k and $k + dk$, where $k = 2\pi\nu/c$ and $dk = 2\pi d\nu/c$. The number of modes is given by the “volume” of wave vector space in the shell with wave vector magnitude in the range $(k, k + dk)$ divided by the “volume spacing” between allowed wave vectors $\delta k = \pi/L$. The volume of the shell in 1D k -space with wave vector magnitude in the range $(k, k + dk)$ for a 1D cavity is given by

$$k + dk - k = dk = \frac{2\pi}{c}d\nu. \quad (2.13)$$

The spacing between allowed wave vectors is just $\delta k = n\pi/L$, so that the number of modes with wave vector in the range $(k, k + dk)$ is

$$\frac{dk}{\delta k} = \frac{dk}{\pi/L} = \frac{2L}{c}d\nu, \quad (2.14)$$

which agrees with Eq. (2.11) above.

2D cavity

This can be generalized to a 2D square cavity of sides L in a similar manner. The allowed wave vectors now have two components, an x - and y -component, each with a mode label n_x and n_y , respectively. The allowed wave vectors are given by

$$\mathbf{k}_{n_x, n_y} = \left(\frac{\pi}{L}n_x, \frac{\pi}{L}n_y \right), \quad (2.15)$$

where we assume the cavity lengths are the same in both directions. The allowed modes lie on a lattice in k -space, with each lattice point labeled by (n_x, n_y) , as depicted in Fig. 2.6 below.

The number of modes with frequencies between ν and $\nu + d\nu$ can be calculated in a manner similar to that for the 1D case, by considering the “volume” of a shell in the 2D k -space. The radius of the shell should be between k and $k + dk$, where $k = 2\pi\nu/c$ and $dk = 2\pi d\nu/c$. The number of modes in this shell is given by the “volume” of the shell divided by the volume associated with each mode, which is equivalent to the mode spacing. The volume of the shell in the 2D k -space is the difference the areas of the circles of radii $k + dk$ and k , divided by 4

$$\frac{1}{4} (\pi(k + dk)^2 - \pi k^2) \approx \frac{\pi k dk}{2}, \quad (2.16)$$

where we have only kept terms to leading order in dk . The factor of $1/4$ arises from the fact that we are the allowed values of n_x and n_y are positive,

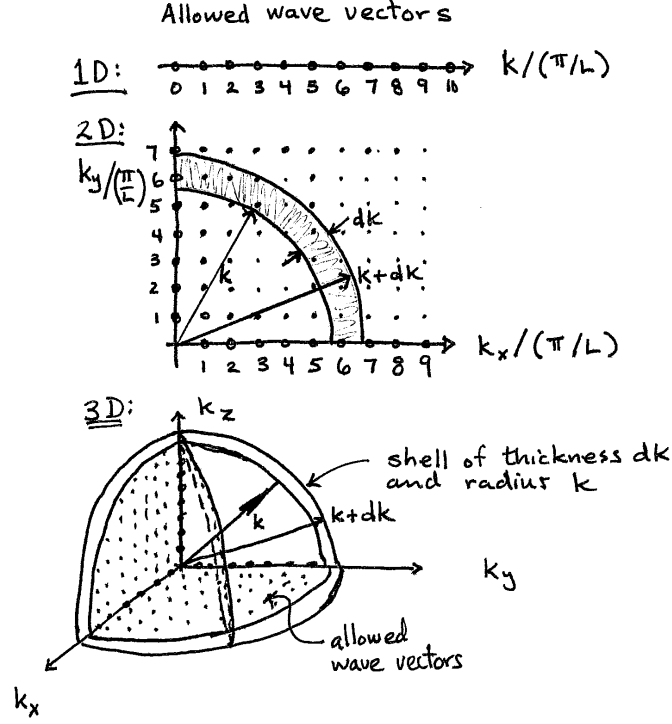


Figure 2.6: Geometry of wave vector space to calculate the density of optical states for a 1D, 2D and 3D cavity with sides of equal length L . The boundary conditions on the electromagnetic field impose constraints on the allowed modes as depicted by the various lattices.

and thus we must only consider the upper right quadrant of the 2D k -space. This gives the 2D volume of the allowed shell. To determine the number of allowed standing-wave modes, we take this volume and divide by the lattice spacing, which is given by $\delta k^2 = (\pi/L)^2$ for our 2D lattice with equal spacing, and multiply by 2 for each polarization

$$2 \frac{\pi k dk / 2}{(\pi/L)^2} = \frac{L^2}{\pi} k dk. \quad (2.17)$$

This can be written in terms of frequency by recalling that $k = 2\pi\nu/c$ and $dk = 2\pi d\nu/c$, giving the number of allowed modes

$$\frac{L^2}{\pi} k dk = \frac{4\pi L^2}{c^2} \nu d\nu. \quad (2.18)$$

Thus, the density of states is given by the number of allowed modes in our cavity divided by its volume, which for the 2D cavity of sides L is L^2 , giving

$$g_{2D}(\nu) d\nu = \frac{4\pi}{c^2} \nu d\nu. \quad (2.19)$$

3D cavity

Finally, to determine the density of states for a 3D cavity of sides L , we can generalize the above analysis by considering the volume of a shell in 3D k -space with radii between k and $k + dk$ as depicted in Fig. ???. The volume of this shell is given by the difference in volumes of the two spheres of radii $k + dk$ and k

$$\frac{4}{3}\pi(k + dk)^3 - \frac{4}{3}\pi k^3 \approx 4\pi k^2 dk, \quad (2.20)$$

where we again only keep the leading order in dk . Note we have used the expansion

$$(x + y)^3 = x^3 + 3x^2y + 3xy^2 + y^3,$$

to simplify the expression. The number of allowed modes with wave vector magnitudes in the range $(k, k + dk)$, is then given by the volume of this shell in 3D k -space given in Eq. (2.20), divided by the mode volume $\delta k^3 = (\pi/L)^3$, divided by 8 since the allowed values of mode labels (n_x, n_y, n_z) are positive only, and thus we only take the positive octant of the spherical shell, and multiplied by 2 for the possible polarizations associated with each wave vector

$$2 \left(\frac{1}{8} \right) \frac{4\pi k^2 dk}{\delta k^3} = \frac{L^3}{\pi^2} k^2 dk. \quad (2.21)$$

This can also be expressed in terms of the frequency using $k = 2\pi\nu/c$ and $dk = 2\pi d\nu/c$, giving

$$\frac{L^3}{\pi^2} k^2 dk = \frac{8\pi L^3}{c^3} \nu^2 d\nu. \quad (2.22)$$

Thus the density of states for the 3D cavity is given by the number of modes in Eq. (2.22) divided by its volume L^3

$$g_{3D}(\nu) d\nu = \frac{8\pi}{c^3} \nu^2 d\nu. \quad (2.23)$$

2.2.3 Average energy for an electromagnetic oscillator

We now have half the information we need to determine the spectral energy density in Eq. (2.7). All that is needed is to determine $\langle E(\nu, T) \rangle$, the average energy of the electromagnetic field at a given frequency when in thermal equilibrium with the cavity walls held at temperature T . This can be accomplished by using a statistical mechanics approach. We will discuss the two key methods that were taken in 1900 to explain the blackbody spectrum in the following two sections.

Classical Rayleigh-Jeans treatment and the ultraviolet catastrophe

In 1900, Lord Rayleigh published a paper that attempted to describe the spectrum of blackbody radiation using a classical theory of waves. He followed the same line of reasoning that led to Eq. (2.7) above. He obtained



(a) Lord Rayleigh (John William Strutt)



(b) James Hopwood Jeans

the same expression for the density of states as we have in Eq. (2.23) above. To determine the average energy per frequency mode of the electromagnetic field, Rayleigh relied upon the results of Boltzmann and Maxwell, known as the equipartition theorem. The equipartition theorem essentially says that for a system of particles in thermal equilibrium at temperature T each have mean energy of $k_B T/2$ per degree of freedom, i.e. per quadratic term in the Hamiltonian, where

$$k_B = 1.381 \times 10^{-23} \text{ J/K},$$

is the Boltzmann constant. For the radiation in the blackbody cavity, each mode is in thermal equilibrium with the atoms in the walls that oscillate at the radiation frequency ν . At temperature T each atom has average kinetic energy $K = k_B T/2$ and average potential energy $U = k_B T/2$ so that the total thermal energy associated with a frequency mode is simply

$$\langle E(\nu, T) \rangle = k_B T. \quad (2.24)$$

Note that this is completely independent of frequency. Each frequency mode of oscillation, of which there are an infinite number, has the same amount of thermal energy. This is a key issue with the present treatment.

An equivalent way to determine the average energy per frequency mode is by using Boltzmann's theory of statistical mechanics, in which the probability for a system to have energy between E and $E+dE$ is given by the Boltzmann probability density

$$P_B(E, T) = \frac{1}{k_B T} \exp\left(\frac{-E}{k_B T}\right), \quad (2.25)$$

as shown in Fig. pe.

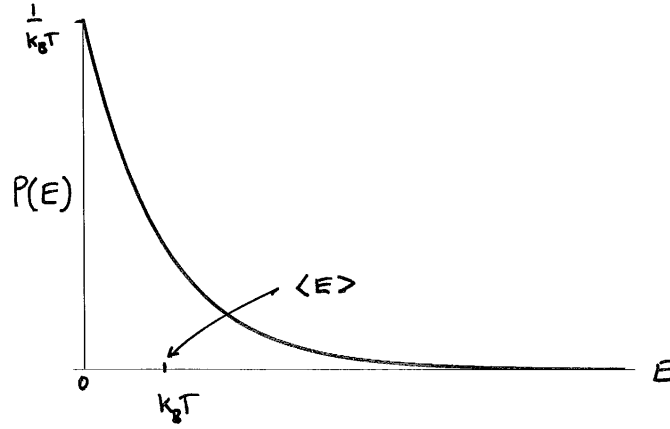


Figure 2.7: Boltzmann probability distribution.

The average energy is thus given by the integral in Eq. (2.6), with the Boltzmann probability density. This integral is equal to the area under the curve $EP_B(E, T)$, shown in Fig. 2.8.

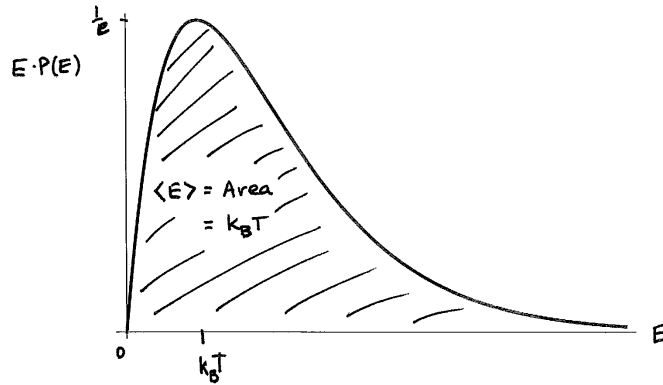


Figure 2.8: Energy distribution $EP_B(E, T)$ for classical continuous energy.

$$\langle E(\nu, T) \rangle = \int_0^\infty E \frac{1}{k_B T} \exp\left(\frac{-E}{k_B T}\right) dE = k_B T, \quad (2.26)$$

where I have used the following integral

$$\int_0^\infty x \exp\left(\frac{-x}{a}\right) dx = a^2.$$

Combining the expressions for the density of states and average energy per frequency mode, we obtain the spectral energy density predicted by Rayleigh

$$\rho_{RJ}(\nu, T) = \frac{8\pi}{c^3} \nu^2 k_B T. \quad (2.27)$$

Figures 2.9 and 2.10 show the behaviour of the Rayleigh-Jeans theory along with observed spectral energy density for $T = 6000$ K. At short wavelengths in the infrared and into the visible region of the spectrum the theory works well. However, as the frequency is increased (thus decreasing wavelength), the theoretical predictions deviate from the observations. Upon inspection of this result, you should note that as the frequency increases towards the ultraviolet part of the spectrum (away from the visible), so does the spectral energy density of the blackbody radiation according to this model. However, this is not what is observed as depicted in Figs. 2.9 and 2.10. In fact, Eq. (2.27) predicts that the spectral energy density diverges to infinity as the frequency increases implying that we should observe an ever-increasing amount of energy at higher frequencies. The failure of the Rayleigh-Jeans formula at high frequencies is known as the *ultraviolet catastrophe*.

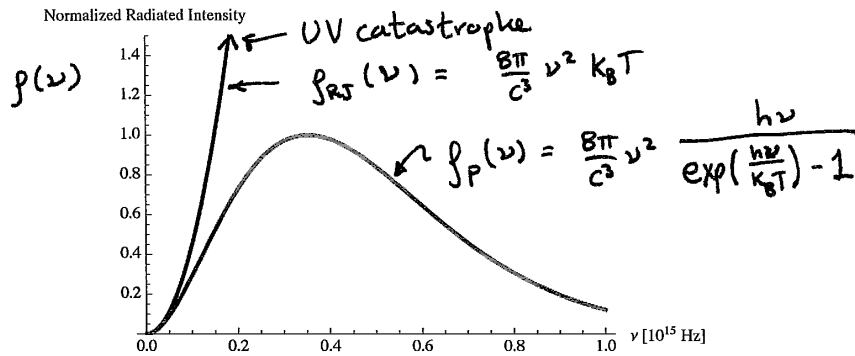


Figure 2.9: Normalized frequency distribution at 6000 K.

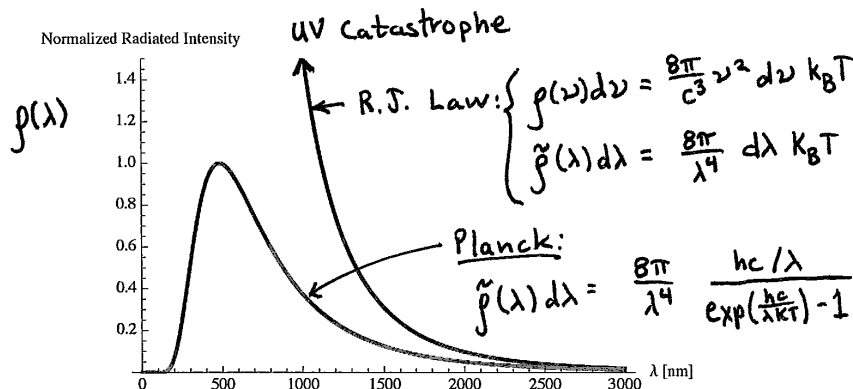


Figure 2.10: Normalized wavelength distribution at 6000 K.



Figure 2.11: Max Karl Ernst Ludwig Planck

Planck treatment and the introduction of light quanta

In the closing months of 1900, a successful theory to describe the spectral energy density of blackbody radiation was developed by the German theorist, Max Planck. The method by which Planck arrived at the equation that now bears his name, the Planck distribution, is a complex story and for a detailed exposition I suggest the text by M. Longair. In the treatment below, we will not follow Planck's original formulation, but rather a more modern description.

According to Eq. (2.27), the ultraviolet catastrophe results from too much energy being distributed at high frequencies. What is needed to fix this is a way to make $\rho(\nu, T) \rightarrow 0$ as $\nu \rightarrow \infty$. In Planck's analysis, he assumed that the reflections at the walls of the blackbody cavity resulted from radiation being absorbed and then quickly reemitted by the atoms of the wall. During this absorption-reemission time the atoms would oscillate at the frequency of the radiation. Since the energy of an oscillating system depends on the frequency,² Planck tried to find a way to reduce the number of high-frequency oscillators in the cavity walls, since the radiation field was in thermal equilibrium with the walls of the cavity. He did this by suggesting that an oscillating atom can absorb or reemit energy only in discrete bundles, or *quanta*, later called photons. If the energy of the quanta were proportional to the frequency of the radiation, then as the frequencies became large the energy also became increasingly large. In his theory, the energy absorbed or

²Recall that the total energy for a 1D harmonic oscillator is

$$E = \frac{p_x^2 + \omega^2 x^2}{2m},$$

where $\omega = 2\pi\nu$ is the resonant frequency of the oscillator.

reemitted by an atom oscillating with frequency ν is an integer multiple of the energy of a single quantum

$$E(\nu, n) = nh\nu, \quad (2.28)$$

where the constant of proportionality,

$$h = 6.6260755 \times 10^{-34} \text{ J s}, \quad (2.29)$$

is now known as the Planck constant and has units of J s or action. **The key difference between the classical and quantum treatment is that in the quantum treatment any physical quantity that oscillates with frequency ν in time, has total energy E satisfying Eq. (2.28).** Whereas classically, any continuum of values for the total energy could be obtained, as depicted in Fig. 2.12

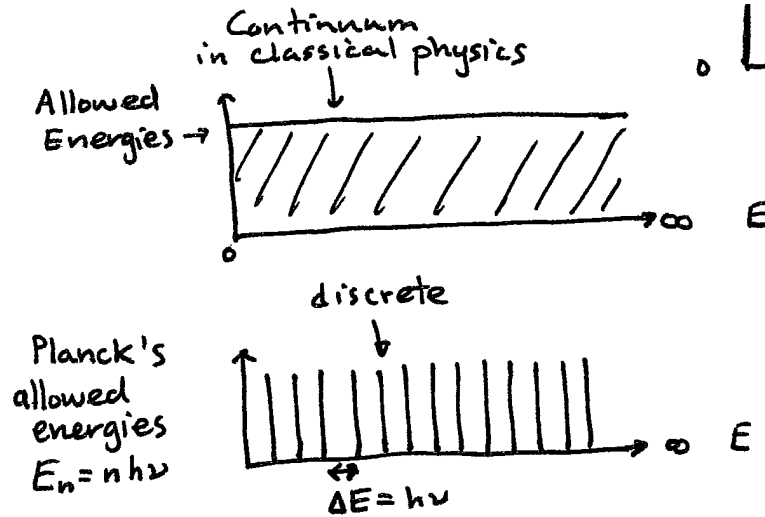


Figure 2.12: Allowed classical and quantized energy levels.

To show how this assumption fixes the ultraviolet catastrophe, let us calculate the average energy for electromagnetic modes of frequency ν . The atoms of the wall that oscillate at frequency ν are in thermal equilibrium with the corresponding electromagnetic frequency modes. Thus, we use the Boltzmann density of Eq. (2.25) to describe the probability of having energy between E and $E + dE$ at thermal equilibrium. However, instead of integrating over a continuum of allowed energies, as in Eq. (2.26), we have a discrete sum over the allowed energies in Eq. (2.28). This gives the following expression for the average energy at frequency ν

$$\langle E(\nu, T) \rangle = \frac{\sum_{n=0}^{\infty} E(\nu, n) P_B(E(\nu, n), T)}{\sum_{n=0}^{\infty} P_B(E(\nu, n), T)}. \quad (2.30)$$

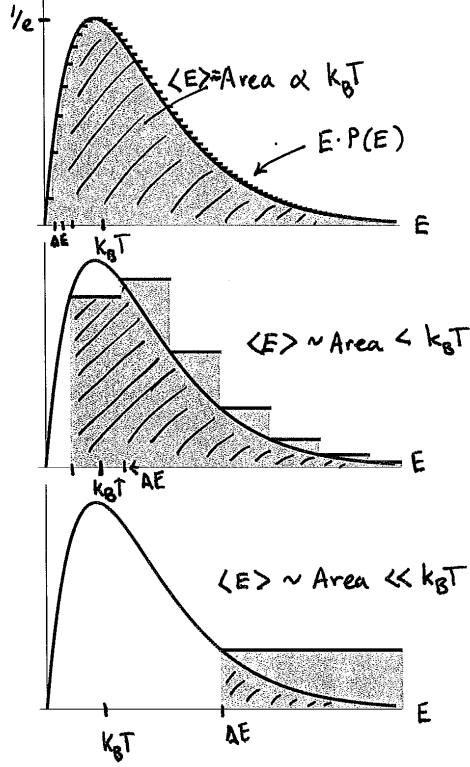


Figure 2.13: Energy distribution $E_n P_B(E_n, T)$ for quantized energy values. Note that for small quantization energies, the quantized sum is nearly the same as the continuous classical case in Fig. 2.8. However, the larger the quantization energy, the larger the deviation from the classical continuous situation.

This argument of this sum $E_n P(E_n, T)$ is show in Fig. 2.13.

To evaluate the sums we first simplify the notation, but introducing the dimensionless quantity $\beta = h\nu/(k_B T)$, which allows us to write the numerator and denominator as

$$\sum_{n=0}^{\infty} E(\nu, n) P_B(E(\nu, n), T) = k_B T \sum_{n=0}^{\infty} n \beta e^{-n\beta} \quad (2.31)$$

and

$$Z = \sum_{n=0}^{\infty} P_B(E(\nu, n), T) = \sum_{n=0}^{\infty} e^{-n\beta} = \sum_{n=0}^{\infty} (e^{-\beta})^n, \quad (2.32)$$

respectively. Here we have used the Boltzmann factor for the probability

$$P_B(E(\nu, n), T) = \exp\left(-\frac{E(\nu, n)}{k_B T}\right). \quad (2.33)$$

Note that the denominator can be expressed as a geometric series

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad (2.34)$$

where we identify

$$x = e^{-\beta},$$

which gives the denominator

$$Z = \frac{1}{1 - e^{-\beta}} = \frac{1}{1 - \exp\left(-\frac{h\nu}{k_B T}\right)}. \quad (2.35)$$

To determine the numerator, we note that it is proportional to a derivative of the denominator, which can be seen by noting that

$$\frac{dZ}{d\beta} = \sum_{n=0}^{\infty} -n(e^{-n\beta}). \quad (2.36)$$

From this we see that the numerator can be expressed as

$$k_B T \sum_{n=0}^{\infty} n\beta e^{-n\beta} = -k_B T \beta \frac{dZ}{d\beta}. \quad (2.37)$$

To evaluate this, we take the derivative of Eq. (2.35) with respect to β

$$\frac{dZ}{d\beta} = \frac{d}{d\beta} (1 - e^{-\beta})^{-1} = -e^{-\beta} (1 - e^{-\beta})^{-2} = -e^{-\beta} Z^2, \quad (2.38)$$

which leads to the following numerator

$$k_B T \sum_{n=0}^{\infty} n\beta e^{-n\beta} = k_B T \beta e^{-\beta} Z^2. \quad (2.39)$$

Thus we see that the average energy at frequency ν is now given by

$$\langle E(\nu, T) \rangle = \frac{k_B T \beta e^{-\beta} Z^2}{Z} = \frac{h\nu}{e^{\beta} - 1}, \quad (2.40)$$

where we have used the fact that $k_B T \beta = h\nu$. Combining this with the density of states we determined in the previous section, Eq. (2.23), enables us to calculate the spectral energy density for a blackbody following Planck's quantized oscillator hypothesis

$$\rho_P(\nu, T) = g_{3D}(\nu) \langle E(\nu, T) \rangle = \frac{8\pi}{c^3} \nu^2 \frac{h\nu}{e^{h\nu/(k_B T)} - 1}. \quad (2.41)$$

This is Planck's major result. By assuming that the oscillators in the wall with which the electromagnetic field modes are in thermal equilibrium can

only absorb and reemit radiation with energy that is an integer multiple of the oscillator (and thus radiation) frequency, Planck was able to derive the spectral energy density that he had previously obtained from empirical data. With this mathematical “trick” Planck ushered in the quantum era.

Stefan’s law and total energy density: There are a few things to note about the blackbody spectrum. The first is the total energy density emitted by a blackbody at temperature T is given by

$$\rho(T) = \int_0^\infty \rho_P(\nu, T) d\nu = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3}{e^{h\nu/(k_B T)} - 1} d\nu, \quad (2.42)$$

which can be calculated using the following definite integral and change of variables

$$\int_0^\infty \frac{x^3}{e^x - 1} dx = \frac{\pi^4}{15}. \quad (2.43)$$

The resulting total energy density is

$$\rho(T) = \frac{8\pi^5 k_B^4}{15(hc)^3} T^4 = \frac{4}{c} \sigma T^4, \quad (2.44)$$

where σ is known as the *Stefan-Boltzmann constant*. This relationship between total energy density emitted from a blackbody and its temperature was known by from careful experiments and is known as *Stefan’s law*.

Wein’s displacement law: The frequency at which a blackbody emits the peak energy can be determined from the Planck distribution, Eq. (2.41), by differentiating with respect to frequency, setting the result to 0 and solving for the frequency. This gives

$$\frac{d\rho_P(\nu, T)}{d\nu} = \frac{8\pi h}{c^3} \left(3\nu^2(e^\beta - 1) - \nu^3 \frac{h}{k_B T} e^\beta \right) / (e^\beta - 1)^2 = 0, \quad (2.45)$$

where I have made use of the definition of $\beta = h\nu/(k_B T)$. This can be rearranged to give the following transcendental equation

$$\frac{\beta}{3} = 1 - e^{-\beta}. \quad (2.46)$$

This equation has no analytic solution, but can be solved numerically, which gives $\beta_{\max} \approx 2.82144$, or $\nu_{\max} \approx 2.82144 k_B T/h$.

The spectral energy density per unit wavelength, as opposed to frequency, does not peak at the same energy owing to the fact that the wavelength is inversely proportional to frequency. Thus to determine the peak of the wavelength distribution we have to write

$$\tilde{\rho}_P(\lambda, T) d\lambda = \rho_P(c/\lambda, T) \left| \frac{d\nu}{d\lambda} \right| d\lambda = \rho_P(c/\lambda, T) \frac{c}{\lambda^2} d\lambda, \quad (2.47)$$

where $\tilde{\rho}_P(\lambda, T)$ is the spectral energy density per unit wavelength. Note the additional factor of c/λ^2 . Using the expression for the spectral energy density per unit frequency, this gives

$$\tilde{\rho}_P(\lambda, T) = \frac{8\pi}{\lambda^4} \frac{hc/\lambda}{e^{hc/(\lambda k_B T)} - 1}. \quad (2.48)$$

The peak of the wavelength distribution can be found in a similar manner, which leads to a slightly different transcendental equation

$$\frac{x}{5} = e^{-x} - 1, \quad (2.49)$$

where $x = hc/\lambda k_B T$ is the wavelength equivalent of β . This equation can be solved numerically and gives $x_{\max} \approx 4.965$. The peak wavelength is thus given by

$$\lambda_{\max} = \frac{hc}{x_{\max} k_B T}. \quad (2.50)$$

This is known as *Wein's displacement law*, which was known from careful experiments on blackbody radiation. The displacement refers to the way in which the peak wavelength changes with temperature of the blackbody. The product $\lambda_{\max} T$ is a constant, as shown in Fig. 2.14.

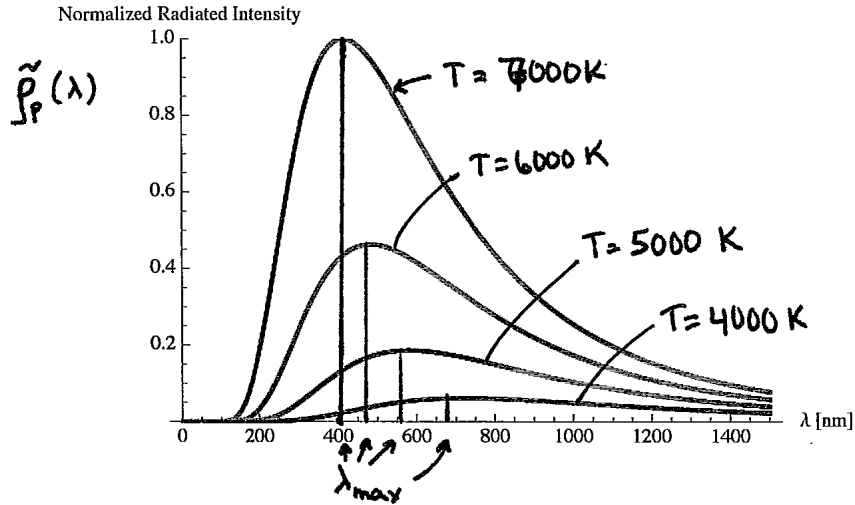


Figure 2.14: Blackbody spectrum as a function of wavelength for four different temperatures. The wavelength at which the peak intensity is emitted scales as $1/T$, which is known as Wein's displacement law.

2.3 Shot noise

Further evidence of the quantized nature of light can be found in the shot-noise associated with fluctuations in the optical intensity of the electromag-

netic field owing to the discrete and independent nature of photons. The term shot noise can refer to either the random, time-dependent fluctuations in an electrical current due to the discreteness of the electron charge or similar random, time-dependent fluctuations in the intensity of electromagnetic radiation (observed on a detector) due to the quantization of the electromagnetic field energy $E = nh\nu$, where n takes on positive integer values. This is illustrated by considering image formation on a photosensitive film when illuminated at extremely low intensity (i.e. very few photons per unit time) as depicted in Fig. 2.15 below.

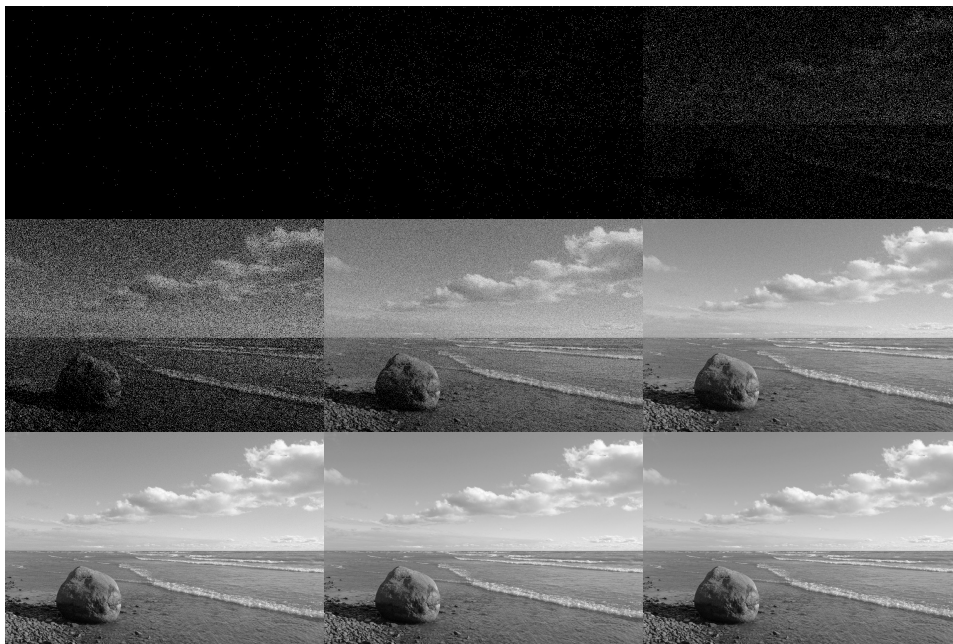


Figure 2.15: Visualization of the shot noise associated with discretization of the electromagnetic field. The image formed on the film is built up over time by low photon-flux illumination. In the first three time windows (top), the shot noise dominates and there is essential no information about the image. The SNR is approximate 1 in this case. As more photons are collected, the SNR improves and the image starts to take shape (middle). Finally, after sufficient illumination, there is little information gain for the collection of additional photons (bottom).

The noise associated with counting a discrete number of objects is not a unique phenomenon to quantum physics though, and pervades many faucets of statistical analysis. For example, the number of rain drops that fall per unit area or the number of nuclear decay events per unit time are also governed by the same statistical description, called Poisson statistics. The events are assumed to be independent and randomly occur with an average of $\langle n \rangle$ per unit time, or space, etc. (Note that the brackets imply an ensemble

average.) Then the average uncertainty in the number counted, quantified by the square root of the variance $\text{var}(n) = \langle (n - \langle n \rangle)^2 \rangle = \langle n^2 \rangle - \langle n \rangle^2$, otherwise called the standard deviation, is just given by

$$\Delta n = \sqrt{\text{var}(n)} = \sqrt{\langle n^2 \rangle - \langle n \rangle^2} = \sqrt{\langle n \rangle}. \quad (2.51)$$

This is characteristic of Poisson processes, that is, the standard deviation from the mean is just the square root of the mean. The signal-to-noise ratio (SNR) is defined as the average signal strength, which in the case of counting photons is the average number of photons detected $\langle n \rangle$, divided by the standard deviation (the uncertainty or noise in our measurement), which in the case of photon counting is just $\sqrt{\langle n \rangle}$. Thus for a Poisson process, the signal-to-noise ratio is

$$\text{SNR} = \langle n \rangle / \sqrt{\langle n \rangle} = \sqrt{\langle n \rangle}. \quad (2.52)$$

As the number of photons detected increases, the better we can determine the actual value of the average intensity (or mean photon number). This concept is nicely demonstrated by considering the formation of an image on a photoreactive film when illuminated at low light levels (much less than one photon “per pixel” per unit time), as depicted in Fig. 2.15.

2.4 Compton Scattering

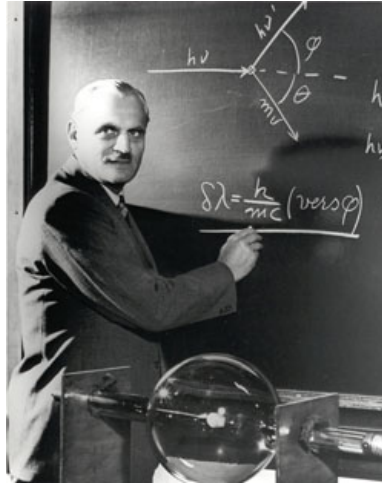


Figure 2.16: Arthur H. Compton

When a light beam encounters a single free electron, some of the radiation can be scattered at an angle that deviates from the initial path. Part of the radiation energy is exchanged for kinetic energy of the electron, and the remainder is reemitted as electromagnetic radiation. In the classical

wave picture of light, the scattered radiation has less energy than the initial beam, but the same frequency (since the only oscillating frequency is that of the incident radiation). However, the photon concept leads to a different conclusion as we will see.

We begin by analyzing the scattering process as an elastic collision between a single photon and an electron, which we assume to be at rest in the laboratory frame. The geometry of the collision is depicted in Fig. 2.17.

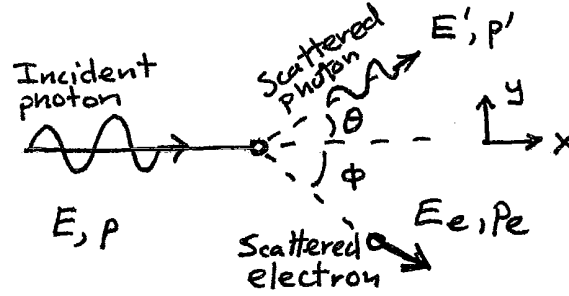


Figure 2.17: Schematic of Compton scattering experimental setup. A source of X rays is collimated into a beam with wavelength λ by two slits and directed toward the scattering target. The scattered X rays are detected at various angles θ by a moveable, spectrally-resolving detector.

The incident photon has energy E given by

$$E = h\nu = \frac{hc}{\lambda}, \quad (2.53)$$

where h is Planck's constant, c is the vacuum speed of light, ν is the frequency of the electromagnetic field, and λ is corresponding wavelength. From Einstein's relativistic formula relating energy, momentum and mass, $E^2 = c^2p^2 + (mc^2)^2$, we note that since the photon has zero rest mass, it will carry linear momentum

$$p = \frac{E}{c} = \frac{h}{\lambda}. \quad (2.54)$$

The electron at rest has energy mc^2 , where m is the electron rest mass. After scattering, the photon has energy E' , momentum p' , and moves at an angle θ with respect to the direction of the incident beam. The electron has total energy E_e , momentum p_e , and travels at an angle ϕ with respect to the incident photon. We use the relativistic relations to account for possible high-energy incident photons. Next we apply energy and momentum conservation to give

$$\begin{aligned} E_i &= E_f \\ E + mc^2 &= E' + E_e, \end{aligned} \quad (2.55)$$

for energy conservation and component-wise for momentum

$$\begin{aligned}(p_x)_i &= (p_x)_f \\ p &= p_e \cos(\phi) + p' \cos(\theta),\end{aligned}\tag{2.56}$$

and

$$\begin{aligned}(p_y)_i &= (p_y)_f \\ 0 &= p_e \sin(\phi) - p' \sin(\theta).\end{aligned}\tag{2.57}$$

There are in general three equations and four unknowns (θ, ϕ, E_e, E') ; p_e and p' are not independent since they are related to the others via the Einstein energy-moment-mass equations for each particle. Thus we cannot uniquely solve the equations. However, we can eliminate any two of the four unknowns. Since we are interested in what happens to the photon after scattering, we will eliminate the electron energy E_e , and scattering angle ϕ . The angle ϕ is eliminated by solving the momentum equations for the electron contributions, and summing the square

$$\begin{aligned}p_e \cos(\phi) &= p - p' \cos(\theta) \\ p_e \sin(\phi) &= p' \sin(\theta).\end{aligned}$$

Squaring and then adding gives

$$p_e^2 = p^2 - 2pp' \cos(\theta) + p'^2.\tag{2.58}$$

The electron energy can be eliminated by using the relativistic energy relation to solve for E_e in terms of p_e , giving

$$E_e^2 = c^2 p_e^2 + m^2 c^4.$$

From Eq. (2.55), we isolate E_e by subtracting E' from both sides, and then squaring to give

$$(E + mc^2 - E')^2 = E_e^2 = c^2 p_e^2 + m^2 c^4,$$

which we can easily solve for p_e^2 . Substituting Eq. (2.58) in for p_e^2 gives

$$(E + mc^2 - E')^2 = E_e^2 = c^2 (p^2 - 2pp' \cos(\theta) + p'^2) + m^2 c^4.\tag{2.59}$$

Using the relativistic energy formula for the initial and final photon energies, $E = cp$ and $E' = cp'$ in Eq. (2.59), we find after a bit of algebra

$$\frac{1}{E'} - \frac{1}{E} = \frac{1}{mc^2} (1 - \cos(\theta)).\tag{2.60}$$

This is the Compton scattering formula relating the incident photon energy to the scattered photon energy and scattering angle. We can rewrite this equation in terms of the photon wavelengths

$$\lambda' - \lambda = \frac{h}{mc} (1 - \cos(\theta)), \quad (2.61)$$

where λ is the wavelength of the incident photon and λ' is the wavelength of the scattered photon. The quantity h/mc is known as the *Compton wavelength of the electron*, which has a value of 0.0002426 nm. However, keep in mind this is not a true wavelength, but rather a change of wavelength. The Compton formulae above give the change in energy (or wavelength) as a function of scattering angle θ of the photon. Since the quantity on the right-hand side is greater than or equal to zero, E' is always less than E , the scattered photon has less energy than the incident photon. The difference $E - E'$ is just the kinetic energy given to the electron, $E_e - mc^2$. Similarly, $\lambda' \geq \lambda$. The change in wavelength ranges from 0 at $\theta = 0$ to twice the Compton wavelength at $\theta = 180^\circ$.

The first experimental observation of this type of scattering was performed by Arthur Compton in 1923. A sketch of the experimental setup is shown in Fig. 2.18. A beam of X rays is incident on a scattering target. Compton used carbon, and although no scattering target contains actual “free” electrons, the outer-shell (valence) electrons in many materials are weakly bound, thus behaving like nearly-free electrons. This assumes the binding energy is negligible compared with the kinetic energy gained from the scattering process. A moveable detector capable of measuring the scattered photon wavelength (energy) is positioned at various observation angles θ .

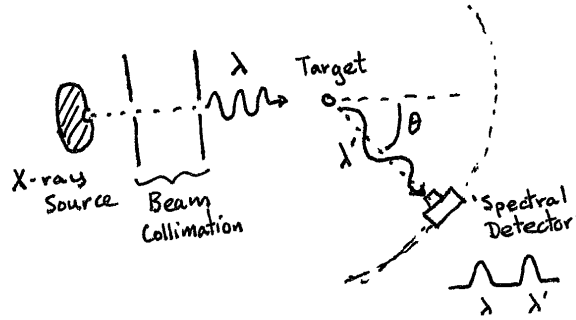


Figure 2.18: Schematic of Compton scattering experimental setup. A source of X rays is collimated into a beam with wavelength λ by two slits and directed toward the scattering target. The scattered X rays are detected at various angles θ by a moveable, spectrally-resolving detector.

The results from Compton’s experiments are illustrated in Fig. 2.19. At each scattering angle, two peaks are observed, corresponding to scattered X-

ray photons with two different energies. The wavelength of one peak does not change as the angle is varied. This peak corresponds to X rays scattered off inner-shell electrons that are so tightly bound to the atom that they cannot be “knocked loose” by the incident X rays. The X rays scattered from these inner-shell electrons do not lose any energy. However, the wavelength of the other peak varies strongly with scattering angle and is accurately described by the Compton formula.

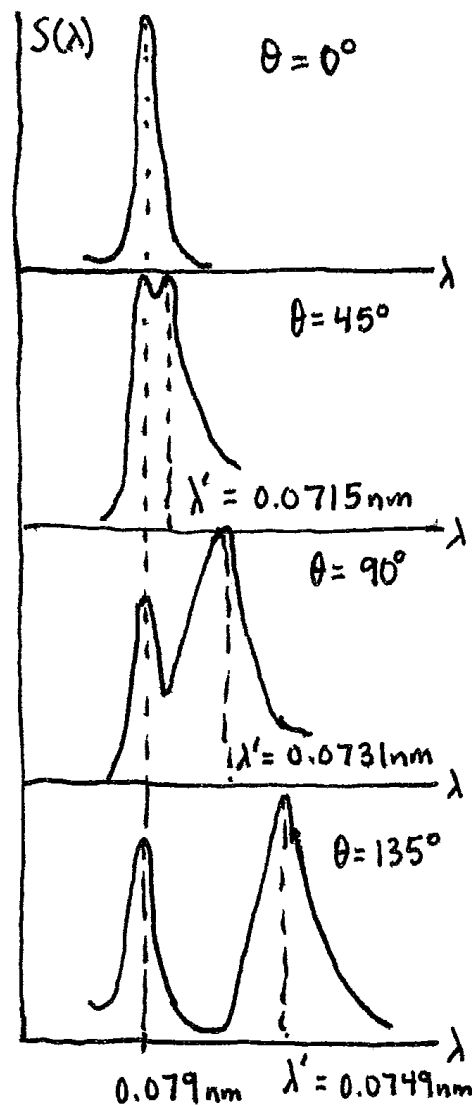


Figure 2.19: Sketch of the results from Compton’s original X-ray scattering experiments.