Golden Rule Number 2

Light and Atoms:

Absorption

Stimulated Emission

Spontaneous Emission

Photo-ionization

Two Forms of the EM Interaction Hamiltonian

Direct Coupling $H_{int} = d \cdot E$

(length gauge)

(electric dipole moment) dot (electric field)

Minimal Coupling $H_{int} = p \cdot A$

(velocity gauge)

(particle momentum) dot (field momentum)

The Unification of the Lorentz and Coulomb Gauges of Electromagnetic Theory

David M. Drury, *Member, IEEE*

***Abstract*—The Lorentz and Coulomb gauges of electromagnetic field theory are shown to be special cases of the same gauge, denoted in this paper as the velocity gauge since it specifies the divergence of the vector potential in terms of the velocity of propagation of the scalar potential. This velocity is the speed of light for the Lorentz gauge and infinite for the Coulomb gauge. With the velocity gauge, the selection of a gauge can be explained in physical terms—the specification of the velocity of propagation of the scalar potential—which engineering students might find more intuitive than a suspiciously arbitrary choice made for mathematical convenience.**

***Index Terms*—Coulomb guage, electromagnetic field theory, guage theory, Lorentz guage.**

I. Introduction

In electromagnetic field theory, the vector potential $\mathbf{A}$ is defined as the field whose curl is the magnetic flux density, i.e., $\mathbf{B} = \nabla \times \mathbf{A}$. This leaves the divergence of the vector potential unspecified, and it is usually chosen for mathematical convenience. This arbitrary choice is known as the *gauge*, *gauge condition*, or simply the *condition* (for example, see [1]–[5]). The most often used gauge is the Lorentz gauge; some textbooks [6]–[9] do not even mention the Coulomb (also known as the transverse, solenoidal, or radiation) gauge. However, at least one commercially available software package used the Coulomb gauge for electromagnetic field simulation [10], so undergraduate electrical engineering students should be made aware of it.

In this paper it is shown that the Lorentz and Coulomb gauges are limiting cases of what is here termed the velocity gauge. The apparently arbitrary choice of gauge is shown to be a specification of the velocity of propagation of the scalar potential V.

II. The Velocity Gauge Condition

In the course of deriving the free-space wave equations for the potential fields from Maxwell's equations (for example, see [5]), the equations

$$\boxed{\nabla^2 V + \frac{\partial \nabla \cdot \mathbf{A}}{\partial t} = -\frac{\rho}{\epsilon_o}} \tag{1}$$

and

$$\boxed{\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_o \mathbf{J} + \nabla(\nabla \cdot \mathbf{A}) + \frac{1}{c^2}\frac{\partial \nabla V}{\partial t}} \tag{2}$$

result. In these equations, V is the scalar potential, $\mathbf{A}$ is the vector potential, ρ is the volume charge density, $\mathbf{J}$ is the volume current density, c is the speed of light in free space, and ϵ_o and μ_o are the permittivity and permeability of free space such that $\epsilon_o \mu_o = c^{-2}$. In order to separate these coupled equations, a gauge must be chosen. The most common choice is the Lorentz gauge

$$\boxed{\nabla \cdot \mathbf{A} = -\frac{1}{c^2}\frac{\partial V}{\partial t}} \tag{3}$$

since it leads to complete separation of the wave equations for $\mathbf{A}$ and V as shown below. Sometimes [1], [10] the Coulomb gauge

$$\boxed{\nabla \cdot \mathbf{A} = 0} \tag{4}$$

is more convenient. However, the Lorentz and the Coulomb gauges are two special cases of a *velocity gauge*

$$\boxed{\nabla \cdot \mathbf{A} = -\frac{1}{u^2}\frac{\partial V}{\partial t}.} \tag{5}$$

If $u = c$ the Lorentz gauge results, while if $u \to \infty$ the Coulomb gauge is obtained.

In terms of the velocity gauge, the wave equations (1) and (2) for the potential fields become

$$\boxed{\nabla^2 V - \frac{1}{u^2}\frac{\partial^2 V}{\partial t^2} = -\frac{\rho}{\epsilon_o}} \tag{6}$$

and

$$\boxed{\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_o \mathbf{J} + \left[\frac{1}{c^2} - \frac{1}{u^2}\right]\frac{\partial \nabla V}{\partial t}.} \tag{7}$$

From these equations (which apply only in free space) we can see that the propagation speed of the vector potential is c while the speed of propagation of the scalar potential is u (for example, see [7]). Therefore the gauge condition, (5), specifies the divergence of the vector potential in terms of the speed of propagation u of the scalar potential.

In the Lorentz gauge, $u = c$ and (6) and (7) become the familiar

$$\nabla^2 V - \frac{1}{c^2}\frac{\partial^2 V}{\partial t^2} = -\frac{\rho}{\epsilon_o} \tag{8}$$

and

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_o \mathbf{J}. \tag{9}$$

Manuscript received January 3, 1996; revised October 4, 1999.
The author is with the Department of Electrical and Software Engineering, University of Wisconsin-Platteville, Platteville, WI 53818 USA (e-mail: drury@uwplatt.edu).
Publisher Item Identifier S 0018-9359(00)01442-4.

These equations are completely uncoupled, and the source of the scalar potential is the charge density while the source of the vector potential is the current density. In the less-familiar Coulomb gauge, $u \to \infty$, and (6) and (7) become

$$\boxed{\nabla^2 V = -\frac{\rho}{\epsilon_o} \tag{10}}$$

and

$$\boxed{\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_o \mathbf{J} + \frac{1}{c^2}\frac{\partial \nabla V}{\partial t}. \tag{11}}$$

These equations are still partially coupled: the scalar potential can be obtained from the charge density, but in order to find the vector potential $\mathbf{A}$, (10) must first be solved for the scalar potential V and the result substituted into (11). This requires more work, so it is not surprising that some undergraduate textbooks do not include the Coulomb gauge. (One way of uncoupling these equations is given in [1] and reproduced here in the Appendix; the complexity and shortcomings of this operation reinforce the tendency to emphasize the Lorentz gauge in undergraduate engineering education.) In the magnetostatic case all time derivatives vanish so that all of the above-mentioned gauges reduce to

$$\nabla \cdot \mathbf{A} = 0$$

and

$$\nabla^2 V = -\frac{\rho}{\epsilon_o} \tag{12a}$$

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{J}. \tag{12b}$$

III. Physicality of the Velocity Gauge

A derivation in [11] is frequently cited to claim that only the Lorentz gauge is compatible with charge conservation. (A simplified version of this derivation is given in [6].) However, the velocity gauge (and hence its special case of the Coulomb gauge) is also compatible with charge conservation. Differentiating (6) with respect to time and dividing the result by c^2 gives

$$\frac{1}{c^2}\nabla^2\frac{\partial V}{\partial t} - \frac{1}{c^2}\frac{1}{u^2}\frac{\partial^2}{\partial t^2}\frac{\partial V}{\partial t} = -\frac{1}{\epsilon_o c^2}\frac{\partial \rho}{\partial t} = -\mu_o\frac{\partial \rho}{\partial t}. \tag{13}$$

We now take the divergence of (7)

$$\nabla^2(\nabla \cdot \mathbf{A}) - \frac{1}{c^2}\frac{\partial^2 \nabla \cdot \mathbf{A}}{\partial t^2} = -\mu_o \nabla \cdot \mathbf{J} + \left[\frac{1}{c^2} - \frac{1}{u^2}\right]\frac{\partial \nabla^2 V}{\partial t} \tag{14}$$

add this equation to (13)

$$\nabla^2\left(\nabla \cdot \mathbf{A} + \frac{1}{c^2}\frac{\partial V}{\partial t}\right) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\nabla \cdot \mathbf{A} + \frac{1}{u^2}\frac{\partial V}{\partial t}\right) = -\mu_o\left(\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t}\right) + \left(\frac{1}{c^2} - \frac{1}{u^2}\right)\frac{\partial \nabla^2 V}{\partial t} \tag{15}$$

and rearrange terms to get

$$\nabla^2\left(\nabla \cdot \mathbf{A} + \frac{1}{u^2}\frac{\partial V}{\partial t}\right) - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left(\nabla \cdot \mathbf{A} + \frac{1}{u^2}\frac{\partial V}{\partial t}\right) = -\mu_o\left(\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t}\right). \tag{16}$$

However, from the velocity gauge condition, (5), the left side of (16) vanishes, and we are left with

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{17}$$

which is the mathematical expression of conservation of charge.

Now consider the speed of propagation of the scalar potential V. The possibility of $u > c$ is certain to prompt at least one student to interrupt a lecture on gauge conditions with the objection that nothing travels faster than c in free space. Actually, it is matter and energy that do not propagate faster than c in free space. The energy of an electromagnetic field is proportional to the squared magnitudes of the electric and magnetic force fields $\mathbf{E}$ and $\mathbf{B}$, and their wave equations [3], [5], [6] show that these fields propagate at speed c in free space no matter what gauge is used to calculate the potential fields V and $\mathbf{A}$. For students who refuse to believe that $\mathbf{E} = -\nabla V - \partial \mathbf{A}/\partial t$ propagates no faster than c even though V is propagating at $u > c$, the following derivation might prove enlightening. Start with (6) and (7), the wave equations for V and $\mathbf{A}$ in the velocity gauge. In free space, $\rho = 0$ and $\mathbf{J} = \mathbf{0}$, so these equations reduce to

$$\nabla^2 V - \frac{1}{u^2}\frac{\partial^2 V}{\partial t^2} = 0 \tag{18}$$

and

$$\nabla^2 \mathbf{A} - \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2} = \left[\frac{1}{c^2} - \frac{1}{u^2}\right]\frac{\partial \nabla V}{\partial t}. \tag{19}$$

Take the gradient of (18) and the partial time derivative of (19) to get

$$\nabla^2 \nabla V - \frac{1}{u^2}\frac{\partial^2 \nabla V}{\partial t^2} = \mathbf{0} \tag{20}$$

and

$$\nabla^2\frac{\partial \mathbf{A}}{\partial t} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\frac{\partial \mathbf{A}}{\partial t} = \left[\frac{1}{c^2} - \frac{1}{u^2}\right]\frac{\partial^2 \nabla V}{\partial t^2}. \tag{21}$$

Adding these two equations together produces

$$\nabla^2\left[\nabla V + \frac{\partial \mathbf{A}}{\partial t}\right] - \frac{\partial^2}{\partial t^2}\left[\frac{1}{u^2}\nabla V + \frac{1}{c^2}\frac{\partial \mathbf{A}}{\partial t}\right] = \left[\frac{1}{c^2} - \frac{1}{u^2}\right]\frac{\partial^2 \nabla V}{\partial t^2}. \tag{22}$$

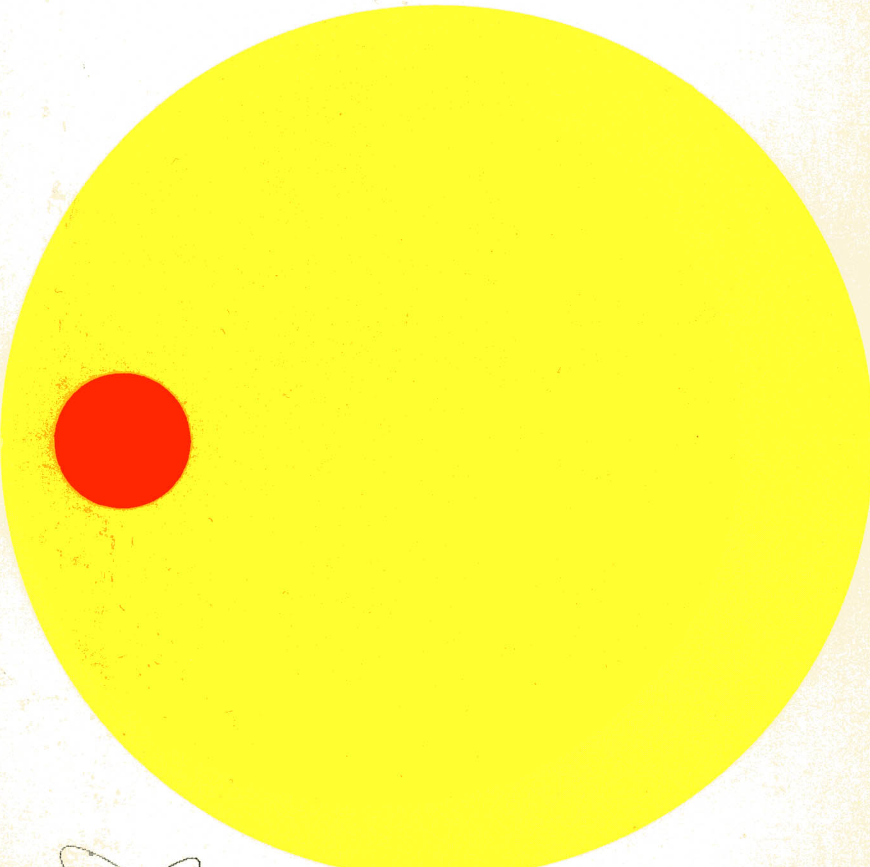
Note that there is a $-(1/u^2)(\partial^2 \nabla V/\partial t^2)$ term on both sides of (22). These terms cancel, and moving the remaining term from the right side to the left side of the equation gives

$$\nabla^2\left[\nabla V + \frac{\partial \mathbf{A}}{\partial t}\right] - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left[\nabla V + \frac{\partial \mathbf{A}}{\partial t}\right] = \mathbf{0}. \tag{23}$$

Great-grandpa's quantum notes

notes on Quantum Mechanics

Enrico Fermi



PSS 512 $1.75
(12s 6d net)

PHOENIX SCIENCE SERIES

SCIENCE

NOTES ON QUANTUM MECHANICS

These are the lecture notes as Fermi prepared them in his own hand the last time he gave the course in quantum mechanics at The University of Chicago in 1954. Fermi wrote them directly on hectograph masters so that copies could be distributed to students who were following the lectures. They have been reproduced in facsimile in order to preserve in as personal a way as possible that unique quality of Fermi's for lecturing simply and clearly on the most essential aspects of a difficult but important subject.

A **PHOENIX SCIENCE SERIES BOOK** *published by The University of Chicago Press*

Cover by ED & JANE BEDNO

Physics/Science

The lecture notes presented here in facsimile were prepared by Enrico Fermi for students of his course at the University of Chicago in 1954. They are vivid examples of Fermi's unique ability to lecture simply and clearly on the most essential aspects of quantum mechanics.

At the close of each lecture, Fermi created a single problem for his students. These challenging exercises were not included in Fermi's notes but were preserved in the notes of his students. This second edition includes a set of Fermi's assigned problems as compiled by a former student, Robert A. Schluter.

"No single individual in this century has contributed so much to physics, through theory as well as experiment, as did Enrico Fermi. Still, in this writer's opinion, his greatest contribution in the Chicago period lay in his teaching. Through his students and their teaching, the Fermi spirit is still alive today. . . . Fermi's way of thinking about, and teaching of, Quantum Mechanics deserves a special mention. His attitude was an entirely pragmatic one: Quantum Mechanics is acceptable *because* its predictions agree so well with experiment. . . . Nothing else counted."—V.L. Telegdi, in *Remembering the University of Chicago*

"[*Notes on Quantum Mechanics*] represents in my opinion much more than historical significance."

–Victor F. Weisskopf, Massachusetts Institute of Technology

The University of Chicago Press

Cover photograph: Enrico Fermi (1901–1954), Ann Arbor. Photograph by S. A. Goudsmit, courtesy of American Institute of Physics, Emilio Segrè Visual Archives.

Cover design: Joan Sommers Design

PREFACE TO THE FIRST EDITION

Enrico Fermi taught courses on quantum mechanics on many occasions.

In the earliest days, when Schrödinger's papers were appearing in the *Annalen der Physik*, Fermi explained them to his students in private seminars; later he recast some of Dirac's papers in more familiar form, in part for didactical reasons. As time went on, his treatment and his courses became more systematic; there must be a number of notebooks of his lectures as recorded by students at the University of Rome, at Columbia University, and at the University of Chicago.

Early in 1954, less than a year before his untimely death, Fermi again gave a course in quantum mechanics at the University of Chicago. This time, however, he prepared the notes for the students himself by writing the outlines of the lectures on duplicator master sheets and delivering copies to the students in advance of each lecture.

The publisher, advised by friends and former pupils of Fermi, has decided to reproduce these lecture outlines in an inexpensive edition, in order to make them available to a larger group of students than those who had the privilege of attending the original lectures.

We hope that young physicists of the coming generation who have never come in direct contact with Fermi, and for whom he must be little more than a name among the great scientists of the century, will enjoy having a notebook on such an important topic as quantum mechanics written for them by such a master in his own hand.

Having pointed out the genesis of these notes, we do not need to emphasize that they cannot be construed in any way as the final presentation of quantum mechanics by Fermi, such as he could have given in a more elaborate text. Heisenberg, Pauli, Dirac, de Broglie, Jordan, Kramers, to mention only some of the creators of quantum mechanics, have all presented their own versions of quantum mechanics in books which are justly famous. The notes by Fermi are not to be compared in any way with these texts. They are written in a spirit and for a purpose completely different from that of the works mentioned above.

Fermi in the last ten or fifteen years of his life scarcely read any book on physics. He kept abreast of scientific developments mainly by hearing the results of investigations and reconstructing them on his own. It is practically certain that he did not consult any text of quantum mechanics while compiling these notes, except perhaps in a very minor fashion. If sections of the notes are very similar to some standard treatments, we must assume that, in rethinking the subject, Fermi arrived in his own way at the formulations contained in these notes.

We repeat that the notes were clearly prepared only for the lectures and that their distribution beyond the class group was not intended by the author. It is only because we know his great interest in teaching that we think it is not irreverent to his memory to publish the notes for the benefit of other students.

E. SEGRÈ

BERKELEY, CALIFORNIA

23– Time dependent perturbation theory, Born approximation.

(1) $\left\{ \quad H = H_0 + \mathcal{H} \right.$ $\quad H_0$ time independent
$\mathcal{H}$ may be time dependent

Unperturbed Scr. eq.

(2) $$i\hbar \dot{\psi}_0 = H_0 \psi_0$$

has solution

(3) $$\psi_0 = \sum \underbrace{a_n^{(0)}}_{\text{constants.}} u_0^{(n)} e^{-\frac{i}{\hbar}E_0^{(n)}t}$$

(4) $$H_0 u_0^{(n)} = E_0^{(n)} u_0^{(n)}$$

Solve Schr eq

(5) $$i\hbar\dot{\psi} = (H_0 + \mathcal{H})\psi$$

by

(6) $$\psi = \sum a_n(t) u_0^{(n)} e^{-\frac{i}{\hbar}E_0^{(n)}t}$$

then multiply by $\widetilde{u_0^{(s)}}$ to left + use orthonormality and (4).

(7) $$\dot{a}_s = -\frac{i}{\hbar}\sum_n a_n \langle s|\mathcal{H}|n\rangle e^{\frac{i}{\hbar}(E_0^{(s)} - E_0^{(n)})t}$$

(8) $$\langle s|\mathcal{H}|n\rangle = \widetilde{u_0^{(s)}}\mathcal{H}u_0^{(n)} = \int u_0^{*(s)}\mathcal{H}u_0^{(n)}dx = \mathcal{H}_{sn}$$

(7) is exact. Use it approximately by substituting in right hand side $a_n(0)$ for $a_n(t)$. Then

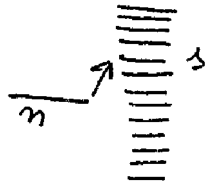
(9) $$a_s(t) \simeq a_s(0) - \frac{i}{\hbar}\sum_n a_n(0)\int_0^t \mathcal{H}_{sn}(t)e^{\frac{i}{\hbar}(E_0^{(s)} - E_0^{(n)})t}dt$$

Important special case. At $t=0$ system in state n. Then $a_n(0)=1$, all other a's are zero.

$$(10)\qquad (s\neq n)\quad a_s(t) = -\frac{i}{\hbar}\int_0^t \mathcal{H}_{sn}(t)\, e^{\frac{i}{\hbar}\left(E_0^{(s)}-E_0^{(n)}\right)t}\,dt$$

Matrix element $\mathcal{H}_{sn}(t)$ causes transitions $n\to s$.

Transitions from n to a continuum of states



Assume $\mathcal{H}_{sn}$ indep. of time, then

$$(11)\quad \begin{cases} a_s(t) = -\mathcal{H}_{sn}\dfrac{e^{\frac{i}{\hbar}(E_0^s-E_0^n)t}-1}{E_0^s-E_0^n} \\[2ex] |a_s(t)|^2 = 4|\mathcal{H}_{sn}|^2\dfrac{\sin^2\frac{t}{2\hbar}\left(E_0^{(s)}-E_0^{(n)}\right)}{\left(E_0^{(s)}-E_0^{(n)}\right)^2} \end{cases}$$

Prob of transition to one state s

$$(12)\quad \begin{cases} P(t) = \sum_s |a_s(t)|^2 = 4\overline{|\mathcal{H}_{sn}|^2}\sum\dfrac{\sin^2\frac{t}{2\hbar}(E^s-E^n)}{(E^s-E^n)^2} = \\[2ex] = 4\overline{|\mathcal{H}_{sn}|^2}\,\rho(E_n)\underbrace{\displaystyle\int\frac{\sin^2\frac{t}{2\hbar}(E^s-E^n)\; d(E^s-E^n)}{(E^s-E^n)^2}}_{\frac{\pi t}{2\hbar}} \\[2ex] = t\,\dfrac{2\pi}{\hbar}|\mathcal{H}_{sn}|^2\rho(E_n) \end{cases}$$

$$\int\frac{\sin^2\alpha x}{x^2}dx = \frac{\pi}{4}$$

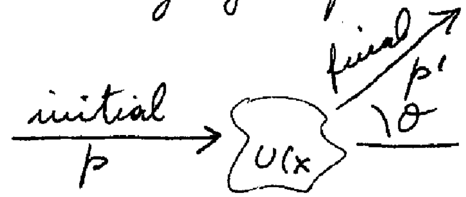
(13) $\rho(E_n)$ = no of states s, close to E_n per unit energy interval.

$$\boxed{\text{Rate of transition} = \frac{2\pi}{\hbar}\overline{|\mathcal{H}_{sn}|^2}\,\rho(E_n)}$$

Discuss: distribution of final states as function of t & relation with uncertainty principle

Example : Born approximation.

(14) Scattering by a potential $U(\vec{x})$



$|p'| = |p|$

$U(x) = \mathcal{H}$ treated as perturbation

(15) initial state $\frac{1}{\sqrt{\Omega}} e^{\frac{i}{\hbar}\vec{p}\cdot\vec{x}}$ (Ω = vol. of box)

final state $\frac{1}{\sqrt{\Omega}} e^{\frac{i}{\hbar}\vec{p'}\cdot\vec{x}}$

$$\langle p'|\mathcal{H}|p\rangle = \frac{1}{\Omega}\underbrace{\int U(x)\, e^{\frac{i}{\hbar}(\vec{p}-\vec{p'})\cdot\vec{x}}\, d^3x}_{\text{Fourier transform of } U}$$

$$= \frac{1}{\Omega} U_{p-p'}$$

(16) No of final states in solid angle $d\omega$ per unit energy interval

$$\rho_{d\omega} = \frac{\Omega\, d\omega}{(2\pi\hbar)^3}\,\frac{p^2 dp}{v\, dp} = \frac{\Omega p^2}{8\pi^3\hbar^3 v}\, d\omega$$

v = velocity $\quad v\,dp = dE$ (correct also relativistically)

Rate of transitions into $d\omega$

$$d\omega\,\frac{v}{\Omega}\,\frac{d\sigma}{d\omega} = \frac{2\pi}{\hbar}\left|\frac{1}{\Omega}U_{p-p'}\right|^2 \frac{\Omega p^2}{8\pi^3\hbar^3 v}\, d\omega$$

(17)
$$\boxed{\frac{d\sigma}{d\omega} = \frac{1}{4\pi^2\hbar^4}\,\frac{p^2}{v^2}\left|U_{p-p'}\right|^2}$$

(18) For non relativistic mechanics $m = \frac{p}{v}$

$$\frac{d\sigma}{d\omega} = \frac{m^2}{4\pi^2\hbar^4}\left|U_{p-p'}\right|^2$$

Limits of validity (discuss)

$$(19) \qquad \frac{1}{\hbar} L \left(\sqrt{p^2 \pm 2mU} - p \right) \ll 1$$

Scattering by Coulomb center

$$(20) \quad \begin{cases} U = \dfrac{zZe^2}{r} \\ U_{p-p'} = zZe^2 \displaystyle\int \frac{e^{\frac{i}{\hbar}(\vec{p}-\vec{p}')\cdot\vec{x}}}{r}\, d^3x = \frac{4\pi z Z e^2}{\frac{1}{\hbar^2}|\vec{p}-\vec{p}'|^2} = \\ \quad = \dfrac{4\pi \hbar^2 z Z e^2}{4 p^2 \sin^2 \frac{\vartheta}{2}} \end{cases}$$

use $\nabla^2 \varphi = -4\pi \dfrac{e^{i\alpha x}}{r}$

$$(21) \quad \left\{ \frac{d\sigma}{d\omega} = \frac{z^2 Z^2}{4} \left(\frac{m e^2}{p^2} \right)^2 \frac{1}{\sin^4 \frac{\vartheta}{2}} \right.$$

(Identical to classical Rutherford formula)

Suggested discussion topics.

Scattering by potential well – Nuclear forces

Limit of long wave length – isotropic scattering

" " short " " – forward "

Role of the mass (neutrino)

Exponential decay of original state in case (11)

24 - Emission and absorption of radiation.

(1) $$\mathcal{H} = eBz\cos\omega t$$

B = amplitude,

at $t=0$ atom in state n. From (23-(10))

(2) $$a_m(t) = -\frac{i}{\hbar} eBz_{mn}\int_0^t \cos\omega t \, e^{i\omega_{mn}t}\,dt$$

$$\omega_{mn} = \frac{E^{(m)} - E^{(n)}}{\hbar} > 0 \qquad \cos\omega t = \frac{e^{i\omega t} + e^{-i\omega t}}{2}$$

$\hbar\omega_{mn}$ m n

this term only important when $\omega \simeq \omega_{mn}$ then

$$a_m(t) \simeq -\frac{ieB}{2\hbar} z_{mn}\int_0^t e^{i(\omega_{mn}-\omega)t}\,dt =$$

$$= +\frac{eB}{2\hbar} z_{mn} \frac{e^{-i(\omega-\omega_{mn})t} - 1}{\omega - \omega_{mn}}$$

(3) $$|a_m(t)|^2 = \frac{e^2B^2}{\hbar^2}|z_{mn}|^2 \frac{\sin^2\frac{t}{2}(\omega-\omega_{mn})}{(\omega-\omega_{mn})^2}$$

Comments on resonance

Light intensity $= \dfrac{cB^2}{8\pi}$

Absorption from continuum overlapping ω_{mn}

(4) $$\frac{cB^2}{8\pi} = \frac{dI}{d\omega}d\omega$$ Substitute in (3), then $\int d\omega$

use $\displaystyle\int \frac{\sin^2\alpha x}{x^2}dx = \pi\alpha$

$$|a_m|^2 = t \times \frac{4\pi^2e^2}{c\hbar^2}|z_{mn}|^2\frac{dI}{d\omega}$$

ω = ang. frequency <u>not</u> solid angle!

(5) $$\boxed{\text{Rate of absorption} = \frac{4\pi^2e^2}{c\hbar^2}|z_{mn}|^2\frac{dI}{d\omega}}$$

For isotropic radiation of volume energy density $u(\omega)\,d\omega$

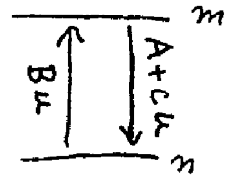
factor 1/3 from averaging over direction of polarization

(6) $$\text{Rate of absorption} = \frac{4\pi^2e^2}{3\hbar^2}|\vec{x}_{mn}|^2 u(\omega_{mn})$$

Relationship between emission & absorption could be derived from quantum electrodynamics — However simpler to use Einsteins A & B method

Rate of $n \to m$ $\quad B\, u(\omega)\, N(n) =$

This B is a coefficient. Has nothing to do with B of page 1



From (6)

(7) $$B = \frac{4\pi^2 e^2}{3\hbar^2}\left|\vec{x}_{mn}\right|^2$$

Rate of $m \to n$) $\quad [A + C\,u(\omega)]\,N(m)$

this is number of atoms in state (n) or (m)

forced transitions

Spontaneous transitions

For thermal equilibrium

(8) $$\frac{N(m)}{N(n)} = e^{-\frac{E^{(m)} - E^{(n)}}{kT}} = e^{-\frac{\hbar\omega_{mn}}{kT}}$$ Boltzmann distribution

At equilibrium: Rate $n \to m$ = Rate $m \to n$

(9) $$\frac{A}{B\,u(\omega)} + \frac{C}{B} = \frac{N_n}{N_m} = e^{\frac{\hbar\omega}{kT}}$$

Planck's law

(10) $$u = \frac{\hbar\omega^3/\pi^2c^3}{e^{\frac{\hbar\omega}{kT}} - 1}$$

$$\frac{\pi^2 c^3}{\hbar\omega^3}\,\frac{A}{B}\left(e^{\frac{\hbar\omega}{kT}} - 1\right) + \frac{C}{B} = e^{\frac{\hbar\omega}{kT}}$$

Must hold at all T's Therefore:

$$\frac{\pi^2 c^3}{\hbar\omega^3}\,\frac{A}{B} = 1 \qquad \frac{C}{B} = 1$$

Einstein's relations

(11) $$\boxed{A = \frac{\hbar\omega^3}{\pi^2 c^3}\,B\ ;\quad C = B}$$ Then from (7)

(12) $$\boxed{\frac{1}{\tau} = A = \frac{4}{3}\,\frac{e^2\omega^3}{\hbar c^3}\left|\vec{x}_{mn}\right|^2}$$ for spontaneous transitions

(12) generalized to many particles by change

(13) $\quad e\vec{x} \rightarrow \sum e_i \vec{x}_i$ (sum to all particles)

(14) $\quad \frac{1}{\tau} = \frac{4}{3}\frac{\omega^3}{\hbar c^3}\left|\sum e_i \langle m|\vec{x}_i|n\rangle\right|^2$

Intensity of radiation proportional to square of matrix element of coordinates (for one electron) or of electric moment (13) for several charged particles.

Discuss – Limitations to validity of (12)

dimensions of atom $\ll \lambda\!\!\!^-$ of radiation

Quadrupole radiation

Case of central forces – Selection rules (See Lect 7)

Spherical harmonics identities

$$(15)\begin{cases}\sqrt{\frac{8\pi}{3}}\,Y_{11}\,Y_{l,m-1} = \sqrt{\frac{(l+m)(l+1+m)}{(2l+1)(2l+3)}}\,Y_{l+1,m} - \sqrt{\frac{(l-m)(l+1-m)}{(2l+1)(2l-1)}}\,Y_{l-1,m} \\ \sqrt{\frac{4\pi}{3}}\,Y_{10}\,Y_{l,m} = \sqrt{\frac{(l+1)^2-m^2}{(2l+1)(2l+3)}}\,Y_{l+1,m} + \sqrt{\frac{l^2-m^2}{(2l+1)(2l-1)}}\,Y_{l-1,m} \\ \sqrt{\frac{8\pi}{3}}\,Y_{1,-1}\,Y_{l,m+1} = \sqrt{\frac{(l-m)(l+1-m)}{(2l+1)(2l+3)}}\,Y_{l+1,m} - \sqrt{\frac{(l+m)(l+1+m)}{(2l+1)(2l-1)}}\,Y_{l-1,m}\end{cases}$$

$$(16)\left\{\quad \sqrt{\frac{8\pi}{3}}\,Y_{11} = -\sin\vartheta\, e^{i\varphi};\quad \sqrt{\frac{4\pi}{3}}\,Y_{10} = \cos\vartheta;\quad \sqrt{\frac{8\pi}{3}}\,Y_{1,-1} = \sin\vartheta\, e^{-i\varphi}\right.$$

Follows: The matrix elements of the coordinates vanish unless

(17) $\quad l' = l \pm 1$ and $m' = m \pm 1$ or m $\quad$ Selection rules

Matrix elements

$$(18)\quad \begin{cases} \langle n', l+1, m+1 \,|\, x+iy \,|\, n, l, m \rangle = -\mathcal{J}\sqrt{\dfrac{(l+m+2)(l+1+m)}{(2l+1)(2l+3)}} \\ \langle n', l+1, m+1 \,|\, x-iy \,|\, n, l, m \rangle = 0 \\ \langle n', l+1, m \,|\, z \,|\, n, l, m \rangle = \mathcal{J}\sqrt{\dfrac{(l+1)^2-m^2}{(2l+1)(2l+3)}} \\ \langle n', l+1, m-1 \,|\, x+iy \,|\, n, l, m \rangle = 0 \\ \langle n', l+1, m-1 \,|\, x-iy \,|\, n, l, m \rangle = \mathcal{J}\sqrt{\dfrac{(l+1-m)(l+2-m)}{(2l+1)(2l+3)}} \end{cases}$$

$$(19)\qquad \mathcal{J} = \int_0^\infty R_{nl}(r)\, R_{n',l+1}(r)\, r^3\, dr$$

Derive

$$(20)\quad \begin{cases} \left|\langle n', l+1, m+1 \,|\, \vec{x} \,|\, n, l, m \rangle\right|^2 + \left|\langle n', l+1, m \,|\, \vec{x} \,|\, n\, l\, m \rangle\right|^2 + \\ + \left|\langle n', l+1, m-1 \,|\, \vec{x} \,|\, n, l, m \rangle\right|^2 = \dfrac{l+1}{2l+1}\,\mathcal{J}^2 \quad (\text{indep. of } m) \end{cases}$$

(21) Therefore: rate of transition

$$(n, l, m) \to (n', l+1 \text{ any } m')$$

$$= \frac{4}{3}\,\frac{e^2\omega^3}{\hbar c^3}\,\frac{l+1}{2l+1}\,\mathcal{J}^2$$

Comments on independence of <u>m</u>

Similarly

$$(22)\quad \begin{cases} \text{Rate}\left(n, l, m \to n', l-1, \text{any } m\right) = \\ = \dfrac{4}{3}\,\dfrac{e^2\omega^3}{\hbar c^3}\,\dfrac{l}{2l-1}\left\{\displaystyle\int_0^\infty R_{nl}(r)\, R_{n',l-1}(r)\, r^3\, dr\right\}^2 \end{cases}$$

Example – Life time of 2p state of hydrogen

$$R_{1s}(r) = \frac{2}{a^{3/2}} e^{-r/a}; \quad R_{2p}(r) = \frac{1}{\sqrt{24a^3}} \frac{r}{a} e^{-r/2a}$$

$$\mathcal{J} = \int R_{1s} R_{2p} r^3 dr = \frac{192\sqrt{2}}{243} a$$

$$\text{Rate}(2p \to 1s) = \frac{294912}{177147} \frac{e^2 \omega^3 a^2}{\hbar c^3} \qquad \omega = \frac{3}{4} \frac{me^4}{2\hbar^3}$$

$$= \frac{1152}{6561} \left(\frac{e^2}{\hbar c}\right)^3 \left(\frac{me^4}{2\hbar^3}\right) \qquad a = \frac{\hbar^2}{me^2}$$

$$= 1.41 \times 10^9 \text{ sec}^{-1}$$

$$\frac{e^2}{\hbar c} = \frac{1}{137}$$

$$= \frac{\text{Ryd}}{\hbar} = 2.067 \times 10^{16} \text{ sec}^{-1}$$

Topics for discussion

Permitted + forbidden lines

Metastable states

Generalization of selection rules

Irradiation by a linear oscillator

Sum rule + effective number of electrons

Polarization of emitted light

Einstein as Armchair Detective: The Case of Stimulated Radiation

Vasant Natarajan

Vasant Natarajan is at the Department of Physics, Indian Institute of Science. His current research involves trapping of atoms to carry out high precision tests of fundamental physics. He has earlier worked on high precision mass spectrometry and on the focussing of atomic beams by laser fields.

Einstein was in many ways like a detective on a mystery trail, though in his case he was on the trail of nature's mysteries and not some murder mystery! And like all good detectives he had a style. It consisted of taking facts that he knew were correct and forcing nature into a situation that would contradict this established truth. In this process she would be forced to reveal some new truths. Einstein's 1917 paper on the quantum theory of radiation is a classic example of this style that enabled him to predict the existence of stimulated radiation starting from an analysis of thermodynamic equilibrium between matter and radiation.

Einstein is rightly regarded as one of the greatest scientific geniuses of all time. Perhaps the most amazing and awe-inspiring feature of his work was that he was an 'armchair' scientist, not a scientist who spent long hours in a darkened laboratory conducting delicate experiments, but one who performed *gedanken* (thought) experiments (see Appendix)while sitting in his favourite chair that nevertheless advanced our understanding of nature by leaps and bounds. Two of his greatest contributions are the special theory of relativity and the general theory of relativity, both abstract creations of his remarkable intellect. They stand out as scientific revolutions that completely changed our perceptions of nature – of space and time in the case of the special theory and of gravity in the case of the general theory. It might be argued that the special theory of relativity was necessitated by experimental facts such as the constancy of the speed of light, but the general theory was almost completely a product of Einstein's imagination. For a person to have achieved one revolution in his lifetime is great enough, but two revolutions seems quite supernatural.

Einstein was like a detective hot on a mystery trail, of course not one solving murder mysteries but one trying to unravel the mysteries of nature.

But is it really so magical? While it is certain that Einstein was a one-of-a-kind genius, is it at least possible to understand the way in which his mind tackled these problems? I think the answer is yes, because deep inside Einstein was like a detective hot on a mystery trail, of course not one solving murder mysteries but one trying to unravel the mysteries of nature. Any keen follower of murder mysteries knows that there are two types of detectives: those who get down on their hands and knees looking for some microscopic piece of clinching evidence at the scene of the crime, and the second type of 'armchair detectives' who seem to arrive at the solution just by thinking logically about the possibilities. Einstein was most certainly of the second kind, and true to this breed, he had his own *modus operandi*. In simple terms, his technique was to imagine nature in a situation where she contradicted established truths, and revealed new truths in the process. As a case in point, we will look at Einstein's 1917 paper titled 'On the quantum theory of radiation' where he predicted the existence of stimulated emission. While Einstein will always be remembered for his revolutionary relativity theories, his contributions to the early quantum theory are certainly of the highest calibre and the 1917 paper is a classic.

It is useful to first set the paper in its historical perspective. By the time Einstein wrote this paper, he had already finished most of his work on the relativity theories. He had earlier done his doctoral thesis on Brownian motion and was a pioneer of what is now called statistical mechanics. He was thus a master at using thermodynamic arguments. He was one of the earliest scientists to accept Planck's radiation law and its light quantum hypothesis. He had already used it in 1905 for his expla-

Atoms in the upper (m) state make a transition to the lower (n) state by spontaneous emission.

nation of the photoelectric effect. He was also aware of Bohr's theory of atomic spectra and Bohr's model of the atom, which gave some explanation for why atoms emitted radiation in discrete quanta. What he did *not know* in 1917 was any of the formalism of quantum mechanics, no Schrödinger equation and not the de Broglie hypothesis for wave nature of particles that we learn in high school these days. Despite this, Einstein was successful in predicting many new things in this paper.

Let us now see what Einstein's strategy in this paper is. He is attempting to understand the interaction between atoms and radiation from a quantum mechanical perspective. For this, he imagines a situation where a gas of atoms is in thermal equilibrium with radiation at a temperature T. The temperature T determines both the Maxwell–Boltzmann velocity distribution of atoms and the radiation density ρ at different frequencies through Planck's law. He assumes that there are two quantum states of the atom Z_n and Z_m, whose energies are ε_n and ε_m, respectively, and which satisfy the inequality $\varepsilon_m > \varepsilon_n$. The relative occupancies W_n, W_m of these states at a temperature T depends on the Boltzmann factor as follows:

$$W_n = p_n \exp(-\varepsilon_n/kT) \qquad (1)$$
$$W_m = p_m \exp(-\varepsilon_m/kT), \qquad (2)$$

where p_n is a number, independent of T and characteristic of the atom and its nth quantum state, called the degeneracy or 'weight' of the particular state. Similarly, p_m is the weight of the mth state.

Einstein then makes the following basic hypotheses about the laws governing the absorption and emission of radiation:

1. Atoms in the upper (m) state make a transition to the lower (n) state by spontaneous emission. The

probability dW that such a transition occurs in the time dt is given by:

$$dW = A_m^n dt. \quad (3)$$

A_n^m in modern terminology is called the Einstein A coefficient. Since this process is intrinsic to the system and is not driven by the radiation field, it has no dependence on the radiation density.

2. Atoms in the lower state make a transition to the upper state by absorbing radiation. The probability that such a transition occurs in the time dt is given by:

$$dW = B_n^m \rho dt. \quad (4)$$

B_m^n is now called the Einstein B coefficient. The absorption process is driven by the radiation field, therefore the probability is directly proportional to the radiation density ρ at frequency ν.

3. The two postulates above seem quite reasonable. Now comes his new postulate, that there is a third process of radiative transition from the upper state to the lower state, namely stimulated emission, *driven by the radiation field.* By analogy with the probability for absorption, the probability for stimulated emission is:

$$dW = B_m^n \rho dt. \quad (5)$$

Einstein calls the processes in both 2 and 3 as "changes of state due to irradiation". We will see below how he is forced to include postulate 3 in order to maintain thermodynamic equilibrium.

The main requirement of thermodynamic equilibrium is that the occupancy of atomic levels given by the equations should not be disturbed by the absorption and emission processes postulated above. Therefore the number of absorption processes (type 2) per unit time from

The main requirement of thermodynamic equilibrium is that the occupancy of atomic levels given by the equations should not be disturbed by the absorption and emission processes.

By substituting this result in (6), Einstein obtains a new, simple derivation of Planck's law:

$$\boxed{\rho = \frac{A_m^n/B_m^n}{\exp[(\varepsilon_m - \varepsilon_n)/kT] - 1}.} \tag{9}$$

Notice that he will not get the correct form of this law if he did not have the stimulated emission term in (6). Another reason for him to be confident that his three hypotheses about absorption and emission are correct. He then compares the above expression for ρ with Wien's displacement law:

$$\rho = \nu^3 f(\nu/T) \tag{10}$$

to obtain

$$\boxed{\frac{A_m^n}{B_m^n} = \alpha\nu^3} \tag{11}$$

and

$$\boxed{\varepsilon_m - \varepsilon_n = h\nu} \tag{12}$$

with constants α and h. The second result is well-known from the Bohr theory of atomic spectra. Einstein is now completely sure that his three hypotheses about radiation transfer are correct since he has been able to derive both Planck's law and Bohr's principle based on these hypotheses.

Einstein does not stop here. He now considers how interaction with radiation affects the atomic motion in order to see if he can predict new features of the momentum transferred by radiation. Earlier he had argued that thermal equilibrium demands that the occupancy of the states remain undisturbed by interaction with radiation. Now he argues that the Maxwell–Boltzmann velocity distribution of the atoms should not be disturbed by the interaction. In other words, the momentum transfer during absorption and emission should result in the same statistical distribution of velocities as obtained from collisions. From kinetic theory, we know that the Maxwell

The momentum transfer during absorption and emission should result in the same statistical distribution of velocities as obtained from collisions.

The momentum transferred to the atom is in the direction of propagation for absorption and in the opposite direction for emission.

Using this approach, Einstein calculates R and $\langle \Delta^2 \rangle$. He shows that (17) is satisfied identically when these values are substituted, which implies that the velocity distribution from kinetic theory is not disturbed if and only if momentum exchange with radiation occurs in units of E/c in a definite direction.

He thus concludes the paper with the following observations. There must be three processes for radiative transfer, namely absorption, spontaneous emission, and *stimulated emission.* Each of these interactions is quantized and takes place by interaction with a single radiation bundle. The radiation bundle (which we today call a photon) carries not only energy of $h\nu$ but also momentum of $h\nu/c$ in a well defined direction. The momentum transferred to the atom is in the direction of propagation for absorption and in the opposite direction for emission. And finally, ever loyal to his dislike for randomness in physical laws ("God does not play dice!"), he concludes that one weakness of the theory is that it leaves the duration and direction of the spontaneous emission process to 'chance'. However, he is quick to point out that the results obtained are still reliable and the randomness is only a defect of the "present state of the theory".

What far reaching conclusions starting from an analysis of simple thermodynamic equilibrium! This is a truly great paper in which we see two totally new predictions. First, he predicts the existence of stimulated emission. And to top that, for the first time since Planck introduced radiation quanta, he shows that each quantum carries well defined momentum. He shows that the directional momentum is present even in the case of spontaneous emission. Thus an atom cannot decay by emitting "outgoing radiation in the form of spherical waves" with no momentum recoil.

Today his conclusions about momentum transfer during absorption and emission of radiation have been abundantly verified. Equally well verified is his prediction of stimulated emission of radiation. Stimulated emission is the mechanism responsible for operation of the laser, which is used in everything from home computers and CD players to long distance communication systems. Stimulated emission, or more correctly stimulated scattering, underlies our understanding of the phenomenon of Bose–Einstein condensation which plays an important role in the explanation of superconductivity and superfluidity. The two predictions, momentum transfer from photons and stimulated emission, are particularly close to my heart because they play a fundamental role in one of my areas of research, namely laser cooling of atoms. In laser cooling, momentum transfer from laser photons is used to cool atoms to very low temperatures of a few millionths of a degree above absolute zero. Perhaps fittingly, it is the randomness or 'chance' associated with the spontaneous emission process which he disliked so much that is responsible for the entropy loss associated with cooling. In other words, as the randomness from the atomic motion gets reduced by cooling, it gets added to the randomness in the radiation field through the spontaneous emission process, thus maintaining consistency with the second law of thermodynamics.

Stimulated emission is the mechanism responsible for operation of the laser, which is used in everything from home computers and CD players to long distance communication systems.

Conclusions

We have seen how Einstein was able to use the principle of thermodynamic equilibrium to imagine a situation, where radiation and matter were in dynamical equilibrium and from that predict new features of the radiative transfer process. As mentioned before, this was a recurring theme in his work, a kind of *modus operandi* for the great 'detective'. In his later writings, he said that he always sought one fundamental governing principle from which he could derive results through these kind of arguments. He found such a principle for thermodynamics, namely the second law of thermodynamics, which states that it is impossible to build a perpetual

Einstein's attempts at geometrizing electromagnetic forces remained an unfulfilled dream, but that is a story for another day.

motion machine. He showed that the second law was a necessary and sufficient condition for deriving all the results of thermodynamics. His quest in the last four decades of his life was to geometrize all forces of nature. In this quest, he felt that he had indeed found the one principle that would allow him to do this uniquely, and this was the *principle of relativity:*

> *the laws of physics must look the same to all observers no matter what their state of motion.*

He had already used this principle to geometrize gravity in the general theory of relativity. His attempts at geometrizing electromagnetic forces remained an unfulfilled dream, but that is a story for another day.

Appendix

Examples of *gedanken* Experiments

We present two examples of *gedanken* experiments that illustrate the Einstein technique for arriving at new results. Both of these experiments yield results associated with the general theory of relativity, but are so simple and elegant that they can be understood without any knowledge of the complex mathematical apparatus needed for the general theory. The first experiment is due to Einstein himself, while the second is due to Hermann Bondi.

Example 1. This is a thought experiment devised by Einstein to arrive at the conclusion that the general theory of relativity is an extension of the special theory which requires curved spacetime, or spacetime in which the rules of plane (Euclidean) geometry do not apply. The 'known' facts are the results of special theory of relativity applicable to inertial systems, and the equivalence principle which states that inertial mass is exactly equal to gravitational mass. Einstein's argument proceeds as follows.

$$H_1 = \vec{d} \cdot \vec{E}$$

$$\vec{d} = e\,\vec{r}$$

$$\vec{E} = \hat{\epsilon}\, E_0\, e^{i(\vec{k}\cdot\vec{r} + \omega t)}$$

$$\langle f |\, \hat{\epsilon}\cdot\vec{r}\; e E_0\, e^{i\vec{k}\cdot\vec{r}} \,| i \rangle$$

$$e^{i\vec{k}\cdot\vec{r}} = 1 + i\vec{k}\cdot\vec{r} + \frac{1}{2!}(i\vec{k}\cdot\vec{r})^2 + \ldots$$

Selection Rules

Selection Rules for Electronic Transitions

In spectral phenomena such as the Zeeman effect it becomes evident that transitions are not observed between all pairs of energy levels. Some transitions are "forbidden" (i.e., highly improbable) while others are "allowed" by a set of selection rules. The number of split components observed in the Zeeman effect is consistent with the selection rules:

$$\Delta \ell = \pm 1 \quad (not \quad zero)$$

$$\Delta m_{\ell} = 0, \pm 1$$

These are the selection rules for an electric dipole transition. One can say that the oscillating electric field associated with the transitions resembles an oscillating electric dipole. When this is expressed in quantum terms, photon emission is always accompanied by a change of 1 in the orbital angular momentum quantum number. The magnetic quantum number can change by zero or one unit.

Another approach to the selection rules is to note that any electron transition which involves the emission of a photon must involve a change of 1 in the angular momentum. The photon is said to have an intrinsic angular momentum or "spin" of one, so that conservation of angular momentum in photon emission requires a change of 1 in the atom's angular momentum. The electron spin quantum number does not change in such transitions, so an additional selection rule is:

$$\Delta m_s = 0$$

The total angular momentum may change be either zero or one:

$$\Delta j = 0, \pm 1$$

An exception to this last selection rule it that you cannot have a transition from j=0 to j=0; i.e., since the vector angular momentum must change by one unit in a electronic transition, j=0 -> 0 can't happen because there is no total angular momentum to re-orient to get a change of 1.

Conservation of Angular Momentum

Photon spin = 1

Therefore $\Delta\ell$ = -1 or +1

Photon Propagation Direction Can Be

parallel

anti-parallel

perpendicular

to the L quantization axis

Therefore Δm = -1, 0, +1

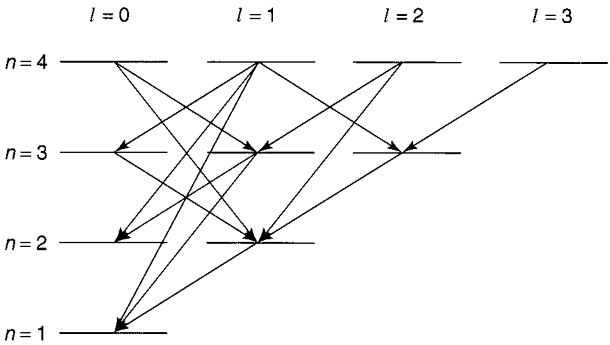


FIGURE 9.6: Allowed decays for the first four Bohr levels in hydrogen.

The Dipole Allowed Decays of n=2 States

Time-Dependent Perturbation Theory: Solved Problems

1. Consider a hydrogen atom in a time-dependent electric field $\mathbf{E} = E(t)\,\mathbf{k}$. Calculate all ten of the dipole matrix elements between the $n = 1$ ground state and the four $n = 2$ excited states. Also calculate the five expectation values of the dipole operator for these five states. Note that "calculate" here means show that fourteen out of the fifteen are zero with a clever argument, so that you only need to do one integral!

First, do the calculation using the even-odd symmetry with respect to z of the three ingredients, namely: (1) the wavefunctions, (2) the dipole term, and (3) the limits of integration. Show which matrix elements must vanish and which ones can survive:

(a) Write down the $n = 1$ ground state wavefunction, and the four $n = 2$ excited state wavefunctions in spherical coordinates:

$$\psi_{nlm}(\mathbf{r}) =< r, \theta, \phi \mid n, l, m >= R_{nl}(r)\; Y_{lm}(\theta, \phi).$$

(b) Show that these five wavefunctions squared $\mid \psi_{nlm}(x, y, z) \mid^2$ are all even functions of z.

(c) Use your result from part b to show that the matrix elements

$$< n, l, m \mid z \mid n, l, m >= \int_{-\infty}^{\infty} z \;\; \mid \psi(x, y, z) \mid^2 \;\; dx\; dy\; dz = 0.$$

(d) Show that four of these five states are even functions of z, namely that ψ_{100}, ψ_{200}, ψ_{211} and ψ_{21-1} are all even functions of z, and that ψ_{210} is an odd function of z.

(e) Use your result from part d to show that all the following dipole matrix elements between pairs of the even states are zero, *i.e.,* show that

$$< 1, 0, 0 \mid z \mid 2, 0, 0 > = < 1, 0, 0 \mid z \mid 2, 1, 1 > = < 1, 0, 0 \mid z \mid 2, 1, -1 > = \; 0,$$

$$< 1, 0, 0 \mid z \mid 2, 1, 0 > = < 2, 1, 1 \mid z \mid 2, 1, 0 > = < 2, 1, -1 \mid z \mid 2, 1, 0 > = \; 0,$$

$$< 2, 0, 0 \mid z \mid 2, 1, 1 > = < 2, 0, 0 \mid z \mid 2, 1, -1 > = < 2, 1, 1 \mid z \mid 2, 1, -1 > = \; 0.$$

(f) Use the even and odd argument in z to explain why the only non-zero matrix elements are

$$< 1, 0, 0 \mid z \mid 2, 1, 0 > = \int_{-\infty}^{\infty} \psi^*_{200}(x, y, z)\; z\; \psi_{210}(x, y, z)\; dx\; dy\; dz,$$

and

$$< 2, 0, 0 \mid z \mid 2, 1, 0 > = \int_{-\infty}^{\infty} \psi^*_{100}(x, y, z)\; z\; \psi_{210}(x, y, z)\; dx\; dy\; dz.$$

(g) Put in the wavefunctions and calculate the two non-zero $H_1 = -eEz$ integrals from part f, *i.e.,* do the integrals. For example, calculate

$$< 1,0,0 \mid H_1 \mid 2,1,0 >= -eE \; \frac{1}{\sqrt{\pi a^3}} \; \frac{1}{\sqrt{32\pi a^3}} \; \frac{1}{a} \; \int_{-\infty}^{\infty} \; e^{-r/a} \; e^{-r/2a} \; z \; d^3r$$

or

$$< 1,0,0 \mid z \mid 2,1,0 >= -eE \; \frac{1}{\sqrt{\pi a^3}} \; \frac{1}{\sqrt{32\pi a^3}} \; \frac{1}{a} \; \int_{-\infty}^{\infty} \; e^{-r/a} \; e^{-r/2a} \; (r\cos\theta) \; \sin\theta \; d\theta \; d\phi \; r^2 \; dr.$$

You should find that

$$< 1,0,0 \mid H_1 \mid 2,1,0 >= -(2^8/3^5\sqrt{2}) \; eEa \; \simeq \; -0.7449 \; eEa,$$

and that

$$< 2,0,0 \mid z \mid 2,1,0 >= -3 \; eEa.$$

Second, do the calculation using the orthonormality of the spherical harmonics and the addition rules for angular momentum:

(h) First show that $z = r\cos\theta \simeq Y_{10}(\theta,\phi)$. Then use the angular momentum addition rules to add Y_{10} to one (or the other) Y_{lm} under the integral. Finally, use the orthonormality of the Y_{lm}'s to show that all the matrix elements except $< 1,0,0 \mid z \mid 2,1,0 >$ must vanish.

(i) Which method do you prefer? Explain why you prefer it! It is very important that you fully understand both methods: they are both extremely powerful and extremely useful!!!

---

1. The wave function expressed in spherical coordinates is given by

$$\psi_{nlm}(\vec{r}) = <r,\theta,\phi\big|n,l,m> \;\; R_{nl}(r)Y_{lm}(\theta,\phi).$$

Using the functional forms of the $R_{nl}s$ and of the spherical harmonics, we find

$$\boxed{\psi_{100} = R_{10}Y_{00} = 2a^{-3/2}e^{-r/a}\left(\frac{1}{4\pi}\right)^{1/2} = \frac{1}{\sqrt{\pi}}a^{-3/2}e^{-r/a},}$$

$$\boxed{\psi_{200} = R_{20}Y_{00} = \frac{1}{\sqrt{2}}a^{-3/2}\left(1-\frac{r}{2a}\right)e^{-r/2a}\left(\frac{1}{4\pi}\right)^{1/2} = \frac{1}{2\sqrt{2\pi}}a^{-3/2}\left(1-\frac{r}{2a}\right)e^{-r/2a},}$$

$$\boxed{\psi_{210} = R_{21}Y_{10} = \frac{1}{\sqrt{24}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\left(\frac{3}{4\pi}\right)^{1/2}\cos(\theta) = \frac{1}{4\sqrt{2\pi}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\cos(\theta),}$$

$$\boxed{\psi_{211} = R_{21}Y_{11} = \frac{1}{\sqrt{24}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\left[-\left(\frac{3}{8\pi}\right)^{1/2}\sin(\theta)e^{i\phi}\right] = -\frac{1}{8\sqrt{\pi}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\sin(\theta)e^{i\phi},}$$

$$\boxed{\psi_{21,-1} = R_{21}Y_{1,-1} = \frac{1}{\sqrt{24}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\left(\frac{3}{8\pi}\right)^{1/2}\sin(\theta)e^{-i\phi} = \frac{1}{8\sqrt{\pi}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\sin(\theta)e^{-i\phi}.}$$

1.(b) Remember, an even function is one for which $f(-x) = f(x)$. If there is more than one independent variable, as we have here, the function may be even with respect to one or more of the variables. Even with respect to z for the function $f(x,y,z)$ means that $f(x,y,-z) = f(x,y,z)$. The wave functions are currently in spherical coordinates $\psi(r,\theta,\phi)$. We need to find their symmetries in Cartesian coordinates

$$r = \left(x^2+y^2+z^2\right)^{1/2}, \quad \cos\theta = \frac{z}{\left(x^2+y^2+z^2\right)^{1/2}}, \quad \sin\theta = \frac{\left(x^2+y^2\right)^{1/2}}{\left(x^2+y^2+z^2\right)^{1/2}}, \quad \text{and} \quad \phi = \tan^{-1}\left(\frac{y}{x}\right).$$

We actually only need to do enough examination to determine the symmetry with respect to z and not a complete change of variables. Using the ψ_{nlm}'s from part a, we find

$$\left|\psi_{100}(x,y,z)\right|^2 = \frac{1}{\pi}a^{-3}e^{-2\left(x^2+y^2+z^2\right)^{1/2}/a}, \qquad \text{where} \qquad (-z)^2 = z^2$$

$$\boxed{\Rightarrow \quad \left|\psi_{100}(x,y,-z)\right|^2 = \left|\psi_{100}(x,y,z)\right|^2 \quad \text{so} \quad \left|\psi_{100}\right|^2 \quad \text{is even wrt } z.}$$

$$\left|\psi_{200}(x,y,z)\right|^2 = \frac{1}{8\pi}a^{-3}\left(1 - \frac{\left(x^2+y^2+z^2\right)^{1/2}}{2a}\right)^2 e^{-\left(x^2+y^2+z^2\right)^{1/2}/a} \quad \text{and} \quad (-z)^2 = z^2 \qquad \text{in both places}$$

$$\boxed{\Rightarrow \quad \left|\psi_{200}(x,y,-z)\right|^2 = \left|\psi_{200}(x,y,z)\right|^2 \quad \text{so} \quad \left|\psi_{200}\right|^2 \quad \text{is even wrt } z.}$$

$$\left|\psi_{210}(x,y,z)\right|^2 = \frac{1}{32\pi}a^{-3}\left(\frac{\left(x^2+y^2+z^2\right)}{a^2}\right)e^{-\left(x^2+y^2+z^2\right)^{1/2}/a}\frac{z^2}{\left(x^2+y^2+z^2\right)}$$

and $(-z)^2 = z^2$ in all four places

$$\boxed{\Rightarrow \quad \left|\psi_{210}(x,y,-z)\right|^2 = \left|\psi_{210}(x,y,z)\right|^2 \quad \text{so} \quad \left|\psi_{210}\right|^2 \quad \text{is even wrt } z.}$$

$$\left|\psi_{211}(x,y,z)\right|^2 = \frac{1}{64\pi}a^{-3}\left(\frac{\left(x^2+y^2+z^2\right)}{a^2}\right)e^{-\left(x^2+y^2+z^2\right)^{1/2}/a}\frac{x^2+y^2}{\left(x^2+y^2+z^2\right)}e^{i\phi(x,y)}$$

and $(-z)^2 = z^2$ in all three places

$$\boxed{\Rightarrow \quad \left|\psi_{211}(x,y,-z)\right|^2 = \left|\psi_{211}(x,y,z)\right|^2 \quad \text{so} \quad \left|\psi_{211}\right|^2 \quad \text{is even wrt } z.}$$

$$\left|\psi_{21,-1}(x,y,z)\right|^2 = \frac{1}{64\pi}a^{-3}\left(\frac{\left(x^2+y^2+z^2\right)}{a^2}\right)e^{-\left(x^2+y^2+z^2\right)^{1/2}/a}\frac{x^2+y^2}{\left(x^2+y^2+z^2\right)}e^{-i\phi(x,y)}$$

and $(-z)^2 = z^2$ in all three places

$$\boxed{\Rightarrow \quad \left|\psi_{211}(x,y,-z)\right|^2 = \left|\psi_{211}(x,y,z)\right|^2 \quad \text{so} \quad \left|\psi_{211}\right|^2 \quad \text{is even wrt } z.}$$

1.(c) Here we use the facts that the product of an even function is an odd function, and that an odd function integrated between symmetric limits is zero. The expectation values of z are given by

$$<n,l,m|z|n,l,m> = \int_{-\infty}^{\infty} z\left|\psi(x,y,z)\right|^2 dx\,dy\,dz = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} z\left|\psi(x,y,z)\right|^2 dz,$$

but z is an odd function, and all of the $\left|\psi_{nlm}(x,y,z)\right|^2$ are even functions, so all of the $z\left|\psi_{nlm}(x,y,z)\right|^2$ are odd functions. The integral with respect to z is between symmetric limits. Therefore

$$\boxed{<n,l,m|z|n,l,m> = \left(\int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy\right)\cdot 0 = 0.}$$

1.(d) Referring to wave functions of part (a) and the Cartesian/spherical relations of part (b),

$$\psi_{100} = \frac{1}{\sqrt{\pi}}a^{-3/2}e^{-\left(x^2+y^2+z^2\right)^{1/2}/a}, \qquad \text{and} \quad (-z)^2 = z^2$$

$$\boxed{\Rightarrow \quad \psi_{100}(x,y,-z) = \psi_{100}(x,y,z) \quad \text{so} \quad \psi_{100} \quad \text{is even wrt } z.}$$

$$\psi_{200}(x,y,z) = \frac{1}{2\sqrt{2\pi}} a^{-3/2} \left(1 - \frac{\left(x^2+y^2+z^2\right)^{1/2}}{2a} \right) e^{-\left(x^2+y^2+z^2\right)^{1/2}/2a},$$

and $(-z)^2 = z^2$ in both places

$$\boxed{\Rightarrow \quad \psi_{200}(x,y,-z) = \psi_{200}(x,y,z) \quad \text{so} \quad \psi_{200} \quad \text{is even wrt } z.}$$

$$\psi_{210} = \frac{1}{4\sqrt{2\pi}} a^{-3/2} \left(\frac{\left(x^2+y^2+z^2\right)^{1/2}}{a} \right) e^{-\left(x^2+y^2+z^2\right)^{1/2}/2a} \frac{z}{\left(x^2+y^2+z^2\right)^{1/2}},$$

This is an odd function. In the three places where $\left(x^2+y^2+z^2\right)^{1/2}$ is substituted for r, $(-z)^2 = z^2$. This portion of the wave function is even. The remaining factor is z, which is an odd function. The product of an even and an odd function is an odd function

$$\boxed{\Rightarrow \quad \psi_{210}(x,y,-z) = -\psi_{210}(x,y,z) \quad \text{so} \quad \psi_{210} \quad \text{is odd wrt } z.}$$

$$\psi_{211} = -\frac{1}{8\sqrt{\pi}} a^{-3/2} \left(\frac{\left(x^2+y^2+z^2\right)^{1/2}}{a} \right) e^{-\left(x^2+y^2+z^2\right)^{1/2}/2a} \frac{\left(x^2+y^2\right)^{1/2}}{\left(x^2+y^2+z^2\right)^{1/2}} e^{i\phi(x,y)},$$

where $(-z)^2 = z^2$ in all three places, and $\phi = \phi(x,y)$ is independent of z,

$$\boxed{\Rightarrow \quad \psi_{211}(x,y,-z) = \psi_{211}(x,y,z) \quad \text{so} \quad \psi_{211} \quad \text{is even wrt } z.}$$

$$\psi_{21,-1} = -\frac{1}{8\sqrt{\pi}} a^{-3/2} \left(\frac{\left(x^2+y^2+z^2\right)^{1/2}}{a} \right) e^{-\left(x^2+y^2+z^2\right)^{1/2}/2a} \frac{\left(x^2+y^2\right)^{1/2}}{\left(x^2+y^2+z^2\right)^{1/2}} e^{-i\phi(x,y)},$$

where $(-z)^2 = z^2$ in all three places, and $\phi = \phi(x,y)$ is again independent of z,

$$\boxed{\Rightarrow \quad \psi_{21,-1}(x,y,-z) = \psi_{21,-1}(x,y,z) \quad \text{so} \quad \psi_{21,-1} \quad \text{is even wrt } z.}$$

1.(e) From part (d), ψ_{100}, ψ_{200}, ψ_{211}, and $\psi_{21,-1}$ are even functions with respect to z. Using the same argument as in part (c),

$$
\begin{aligned}
<\psi_{\text{even wrt } z}\big|z\big|\psi_{\text{even wrt } z}> &= \int_{-\infty}^{\infty} \left(\psi_{\text{even wrt } z}\right)^* z \left(\psi_{\text{even wrt } z}\right)\, dx\, dy\, dz \\
&= \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \int_{-\infty}^{\infty} \left(\psi_{\text{even wrt } z}\right)^* z \left(\psi_{\text{even wrt } z}\right)\, dz.
\end{aligned}
$$

Again, z is an odd function. The product of an even and odd function is odd; this odd function multiplied by another even function yields an odd function overall. The integral with respect to z is between symmetric limits, and an integral of an odd function between symmetric limits is zero. Therefore

$$
\boxed{
\begin{aligned}
&<1,0,0\big|z\big|2,0,0> = <1,0,0\big|z\big|2,1,1> = <1,0,0\big|z\big|2,1,-1> = 0 \\
&<2,0,0\big|z\big|2,1,1> = <2,0,0\big|z\big|2,1,-1> = <2,1,1\big|z\big|2,1,-1> = 0.
\end{aligned}
}
$$

1.(f) The remaining matrix elements are given by

$$
<1,0,0\big|z\big|2,1,0>, \quad <2,0,0\big|z\big|2,1,0>, \quad <2,1,1\big|z\big|2,1,0> \quad \text{and} \quad <2,1,-1\big|z\big|2,1,0>.
$$

These integrals all have the form $\int_{-\infty}^{\infty}$ (even function) (odd function) (odd function) with respect to z, which we would expect to be non-zero. We can examine two at once, using $z = r\cos\theta$, and the volume element in spherical coordinates which is $dv = r^2 \sin\theta dr d\theta d\phi$,

$$
\begin{aligned}
<2,1,\pm1\big|z\big|2,1,0> &= \int_{-\infty}^{\infty} \left(\frac{\mp1}{8\sqrt{\pi}} a^{-3/2} \left(\frac{r}{a}\right) e^{-r/2a} \sin(\theta) e^{\pm i\phi}\right)^* r\cos\theta\, \frac{1}{4\sqrt{2\pi}} a^{-3/2} \left(\frac{r}{a}\right) e^{-r/2a} \cos(\theta)\, dV \\
&= \frac{\mp1}{32\pi\sqrt{2}} \frac{1}{a^5} \int_0^{\infty} dr\, r^5 e^{-r/a} \int_0^{\pi} d\theta\, \sin^2\theta \cos^2\theta \int_0^{2\pi} d\phi\, e^{\mp i\phi}
\end{aligned}
$$

Examining just the azimuthal integral, we find

$$
\begin{aligned}
\int_0^{2\pi} d\phi\, e^{\mp i\phi} &= \int_0^{2\pi} d\phi\, \cos\phi \mp i \sin\phi \\
&= \int_0^{2\pi} d\phi\, \cos\phi \mp i \int_0^{2\pi} d\phi\, \sin\phi \\
&= \sin\phi\Big|_0^{2\pi} \pm i\cos\phi\Big|_0^{2\pi} \\
&= (0-0) \pm i\,(1-1) = 0,
\end{aligned}
$$

therefore, the integral over all space will be zero regardless of the values of the radial and polar integrals, *i.e.*,

$$
\boxed{<2,1,1\big|z\big|2,1,0> = <2,1,-1\big|z\big|2,1,0> = 0.}
$$

1.(g) We have been examining expectation values of z because $H_1 = -eEz$, where $-eE$ is a constant. If the expectation value is non-zero, the value of the integral multiplied by $-eE$ will express the result in energy units.

There are two remaining integrals. Using $z = r\cos\theta$ and $dV = r^2 \sin\theta dr d\theta d\phi$, the integral

$$
\begin{aligned}
<2,0,0\big|z\big|2,1,0> &= \int_{-\infty}^{\infty} \left(\frac{1}{2\sqrt{2\pi}} a^{-3/2} \left(1 - \frac{r}{2a}\right) e^{-r/2a} \right)^* r\cos\theta \, \frac{1}{4\sqrt{2\pi}} a^{-3/2} \left(\frac{r}{a}\right) e^{-r/2a} \cos(\theta)\, dV \\
&= \frac{1}{16\pi a^4} \int_0^{\infty} dr \left(1 - \frac{r}{2a}\right) r^4 e^{-r/a} \int_0^{\pi} d\theta \, \cos^2\theta \sin\theta \int_0^{2\pi} d\phi \\
&= \frac{1}{16\pi a^4} \int_0^{\infty} dr \left(1 - \frac{r}{2a}\right) r^4 e^{-r/a} \int_0^{\pi} d\theta \, \cos^2\theta \sin\theta \, [2\pi] \\
&= \frac{1}{8a^4} \int_0^{\infty} dr \left(1 - \frac{r}{2a}\right) r^4 e^{-r/a} \left[-\frac{\cos^3\theta}{3} \right]_0^{\pi} \\
&= \frac{1}{8a^4} \int_0^{\infty} dr \left(1 - \frac{r}{2a}\right) r^4 e^{-r/a} \left[\frac{2}{3} \right] \\
&= \frac{1}{12a^4} \int_0^{\infty} dr \left(1 - \frac{r}{2a}\right) r^4 e^{-r/a} \\
&= \frac{1}{12a^4} \left[\int_0^{\infty} dr\, r^4 e^{-r/a} - \frac{1}{2a} \int_0^{\infty} dr\, r^5 e^{-r/a} \right].
\end{aligned}
$$

These integrals are evaluated using

$$
\int_0^{\infty} x^n e^{-\mu x}\, dx = n!\, \mu^{-n-1}, \qquad \operatorname{Re}\mu > 0,
$$

with $\mu = 1/a$ for both, and with $n = 4$ and 5 respectively, so we find

$$
\begin{aligned}
<2,0,0\big|z\big|2,1,0> &= \frac{1}{12a^4} \left[4! \left(\frac{1}{a}\right)^{-4-1} - \frac{1}{2a} 5! \left(\frac{1}{a}\right)^{-5-1} \right] \\
&= \frac{1}{12a^4} \left[\frac{4\cdot 3\cdot 2}{(1/a)^5} - \frac{1}{2a} \left(\frac{5\cdot 4\cdot 3\cdot 2}{(1/a)^6} \right) \right] \\
&= \frac{1}{12a^4} \left[24a^5 - \frac{1}{2a} 120a^6 \right] = \frac{1}{12a^4} \left[24a^5 - 60a^5 \right] \\
&= \frac{1}{12a^4} \left(-36a^5 \right) = -3a.
\end{aligned}
$$

Since $H_1 = -eEz$, we find

$$
\boxed{\Rightarrow \quad <2,0,0\big|z\big|2,1,0> = 3eEa.}
$$

The last integral is $<1,0,0\big|z\big|2,1,0>$ which in energy units is given by

$$\begin{aligned}
<1,0,0\big|H_1\big|2,1,0> &= <1,0,0\big|-eEz\big|2,1,0> \\
&= -eE <1,0,0\big|z\big|2,1,0> \\
&= -eE\int_{-\infty}^{\infty}(\psi_{100})^* \, z\,\psi_{210}\,dV \\
&= -eE\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}a^{-3/2}e^{-r/a}\,z\,\frac{1}{4\sqrt{2\pi}}a^{-3/2}\left(\frac{r}{a}\right)e^{-r/2a}\cos(\theta)\,dV \\
&= -\frac{eE}{4\pi\sqrt{2}a^4}\int_{-\infty}^{\infty} re^{-3r/2a}\,z\,\cos(\theta)\,dV.
\end{aligned}$$

Using $z = r\cos\theta$ and $dV = r^2\sin\theta dr d\theta d\phi$, we find

$$\begin{aligned}
<1,0,0\big|H_1\big|2,1,0> &= -\frac{eE}{4\pi\sqrt{2}a^4}\int_{-\infty}^{\infty} r^4e^{-3r/2a}\cos^2\theta\sin\theta\,dr\,d\theta\,d\phi \\
&= -\frac{eE}{4\pi\sqrt{2}a^4}\int_0^{\infty} dr\,r^4e^{-3r/2a}\int_0^{\pi}d\theta\,\cos^2\theta\sin\theta\int_0^{2\pi}d\phi \\
&= -\frac{eE}{4\pi\sqrt{2}a^4}\int_0^{\infty} dr\,r^4e^{-3r/2a}\int_0^{\pi}d\theta\,\cos^2\theta\sin\theta(2\pi) \\
&= -\frac{eE}{2\sqrt{2}a^4}\int_0^{\infty} dr\,r^4e^{-3r/2a}\left[-\frac{\cos^3\theta}{3}\right]_0^{\pi} \\
&= -\frac{eE}{2\sqrt{2}a^4}\int_0^{\infty} dr\,r^4e^{-3r/2a}\left[-\frac{-1-1}{3}\right]_0^{\pi} \\
&= -\frac{eE}{3\sqrt{2}a^4}\int_0^{\infty} dr\,r^4e^{-3r/2a}.
\end{aligned}$$

As before, using

$$\int_0^{\infty} x^ne^{-\mu x}\,dx = n!\mu^{-n-1}, \qquad \mathrm{Re}\,\mu > 0,$$

with $\mu = 3/2a$ and $n = 4$ we find

$$\begin{aligned}
<1,0,0\big|H_1\big|2,1,0> &= -\frac{eE}{3\sqrt{2}a^4}\,4!\left(\frac{3}{2a}\right)^{-5} \\
&= -\frac{eE}{3\sqrt{2}a^4}\,\frac{4\cdot 3\cdot 2(2a)^5}{3^5} \\
&= -\frac{eE}{3\sqrt{2}a^4}\,\frac{3\cdot 2^3\cdot 2^5\cdot a^5}{3^5}
\end{aligned}$$

$$\boxed{\Rightarrow \quad <1,0,0\big|H_1\big|2,1,0> = -\frac{eE}{\sqrt{2}}\,\frac{2^8\cdot a}{3^5} = -0.7449eEa.}$$

1.(h) The wave functions under consideration are $\psi_{nlm} = R_{nl}(r)Y_{lm}(\theta, \phi)$, which are explicitly

$$\psi_{100} = R_{10}Y_{00}, \quad \psi_{200} = R_{20}Y_{00}, \quad \psi_{210} = R_{21}Y_{10}, \quad \psi_{211} = R_{21}Y_{11}, \quad \psi_{21,-1} = R_{21}Y_{1,-1}.$$

The integrals for the expectation values of z are given by

$$\begin{aligned} <n, l, m\big|z\big|n', l', m'> &= <n, l, m\big|r\cos\theta\big|n', l', m'> \\ &= \int_{-\infty}^{\infty} \psi^*_{nlm}\, r\cos\theta\, \psi_{n'l'm'}\, dV \\ &= \int_{-\infty}^{\infty} R^*_{nl}Y^*_{lm}\, r\cos\theta\, R_{n'l'}Y_{l'm'}\, dV \\ &= \int_{0}^{\infty} R^*_{nl}R_{n'l'}\, r^3\, dr \int Y^*_{lm}\cos\theta\, Y_{l'm'}\, d\Omega, \end{aligned}$$

where the factor of r^2 in the radial integral comes from the volume element. The angular momentum addition rules and the integration can be summarized by

$$\int Y^*_{lm}\cos\theta\, Y_{l'm'}\, d\Omega, = \left[\frac{(l'-m'+1)(l'+m'+1)}{(2l'+1)(2l'+3)}\right]^{1/2} \delta_{mm'}\delta_{l,l'+1} + \left[\frac{l'-m')(l'+m')}{(2l'-1)(2l'+1)}\right]^{1/2} \delta_{mm'}\delta_{l,l'-1}.$$

For this integral to be non-zero, l' must differ from l by ± 1. This means that

$$<1,0,0\big|z\big|1,0,0> = <2,0,0\big|z\big|2,0,0> = <2,1,0\big|z\big|2,1,0> = <2,1,1\big|z\big|2,1,1> = <2,1,-1\big|z\big|2,1,-1>$$

$$= <1,0,0\big|z\big|2,0,0> = <2,1,0\big|z\big|2,1,1> = <2,1,0\big|z\big|2,1,-1> = <2,1,1\big|z\big|2,1,-1> = 0.$$

Also, m must equal m' for the integral to be non-zero, so

$$<1,0,0\big|z\big|2,1,1> = <1,0,0\big|z\big|2,1,-1> = <2,0,0\big|z\big|2,1,1> = <2,0,0\big|z\big|2,1,-1> = 0.$$

Only $< 1,0,0\big|z\big|2,1,0 >$ and $< 2,0,0\big|z\big|2,1,0 >$ remain as non-zero possibilities. Knowing that $z = r\cos\theta \sim Y_{10}$, we can see that these two integrals have the form

$$\int Y_{00}Y_{10}Y_{10}\, d\Omega.$$

Parity conservation in angle space can be summarized by $l_1+l_2+l_3+m_1+m_2+m_3 =$ even integer. For our two integrals, this condition is satisfied for the integer 2. For both $<1,0,0\big|z\big|2,1,0>$ and $<2,0,0\big|z\big|2,1,0>$, the integral over solid angle can now be evaluated using

$$\int Y^*_{00}\cos\theta\, Y_{10}\, d\Omega, = \left[\frac{(1-0+1)(1+0+1)}{(2\cdot 1+1)(2\cdot 1+3)}\right]^{1/2} \delta_{00}\delta_{0,1+1} + \left[\frac{(1-0)(1+0)}{(2\cdot 1-1)(2\cdot 1+1)}\right]^{1/2} \delta_{00}\delta_{0,1-1}.$$

Here, the first expression on the right side of the equation will be zero because the indices on the second Kronecker δ are not identical. Both sets of indices on the Kronecker δ of second expression on the right are identical, so we find

$$\int Y^*_{00}\cos\theta\, Y_{10}\, d\Omega, = \frac{1}{\sqrt{3}}.$$

Next, we will evaluate the radial integrals using this angular factor. We find

$$\begin{aligned}
<1,0,0|z|2,1,0> &= \frac{1}{\sqrt{3}}\int_0^\infty R_{10}\, r\, R_{21}\, r^2\, dr \\
&= \frac{1}{\sqrt{3}}\int_0^\infty \left(2a^{-3/2}e^{-r/a}\right)\left(\frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}e^{-r/2a}\right) r^3\, dr \\
&= \frac{1}{3\sqrt{2}a^4}\int_0^\infty r^4 e^{-3r/2a} dr
\end{aligned}$$

This integral can be evaluated using

$$\int_0^\infty x^n e^{-\mu x}\, dx = n!\, \mu^{-n-1}, \qquad \mathrm{Re}\,\mu > 0,$$

with $\mu = 3/2a$ and $n = 4$, so we find

$$\begin{aligned}
<1,0,0|z|2,1,0> &= \frac{1}{3\sqrt{2}a^4} 4\cdot 3\cdot 2\left(\frac{3}{2a}\right)^{-5} \\
&= \frac{1}{3\sqrt{2}a^4}\frac{3\cdot 2^3\cdot 2^5\cdot a^5}{3^5}
\end{aligned}$$

$$\Rightarrow \quad <1,0,0|z|2,1,0> = \frac{2^8}{\sqrt{2}\cdot 3^5}a$$

$$\boxed{\Rightarrow \quad <1,0,0|H_1|2,1,0> = -eE\frac{2^8}{\sqrt{2}\cdot 3^5}a = -0.7449eEa, \quad \text{which is the same as part (g).}}$$

The other integral is given by

$$\begin{aligned}
<2,0,0|z|2,1,0> &= \frac{1}{\sqrt{3}}\int_0^\infty R_{20}\, r\, R_{21}\, r^2\, dr \\
&= \frac{1}{\sqrt{3}}\int_0^\infty \left(\frac{1}{\sqrt{2}}a^{-3/2}\left(1-\frac{r}{2a}\right)e^{-r/2a}\right)\left(\frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}e^{-r/2a}\right) r^3\, dr \\
&= \frac{1}{\sqrt{3}\sqrt{2}\sqrt{24}a^4}\int_0^\infty \left(1-\frac{r}{2a}\right) r^4 e^{-r/a}\, dr \\
&= \frac{1}{12a^4}\int_0^\infty r^4 e^{-r/a}\, dr - \frac{1}{24a^5}\int_0^\infty r^5 e^{-r/a}\, dr.
\end{aligned}$$

We can evaluate this integral using the same procedure, with $\mu = 1/a$ and $n = 4$ and 5 respectively. We find

$$\begin{aligned}
<2,0,0|z|2,1,0> &= \frac{1}{12a^4}\frac{4\cdot 3\cdot 2}{(1/a)^5} - \frac{1}{24a^5}\frac{5\cdot 4\cdot 3\cdot 2}{(1/a)^6} \\
&= \frac{24a^5}{12a^4} - \frac{120a^6}{24a^4}
\end{aligned}$$

$$\boxed{\begin{array}{c} \Rightarrow \quad <2,0,0\big|z\big|2,1,0> = 2a - 5a = -3a, \quad \text{which is the same as part (g)} \\ \text{and} \quad <2,0,0\big|H_1\big|2,1,0> = 3eEa. \end{array}}$$

1.(i) . . .Wow! That spherical harmonic stuff does seem to be a lot less work. . ..

The Dipole Allowed Decays of |3 0 0>

4. An electron in the $n = 3$, $l = 0$, $m = 0$ state of hydrogen decays by a sequence of electric dipole transitions to the ground state. The selection rules for electric dipole transitions are that $\Delta m = \pm 1$ or 0 and that $\Delta l = \pm 1$. In this problem you are only asked to consider the transitions where n changes, so the nine possible transitions are:

$$| \ 3, 0, 0 \ > \Rightarrow | \ 2, 1, 1 \ >$$

$$| \ 3, 0, 0 \ > \Rightarrow | \ 2, 1, 0 \ >$$

$$| \ 3, 0, 0 \ > \Rightarrow | \ 2, 1, -1 >$$

$$| \ 3, 0, 0 \ > \Rightarrow | \ 2, 0, 0 \ >$$

$$| \ 3, 0, 0 \ > \Rightarrow | \ 1, 0, 0 \ >$$

$$| \ 2, 1, 1 \ > \Rightarrow | \ 1, 0, 0 \ >$$

$$| \ 2, 1, 0 \ > \Rightarrow | \ 1, 0, 0 \ >$$

$$| \ 2, 1, -1 > \Rightarrow | \ 1, 0, 0 \ >$$

$$| \ 2, 0, 0 \ > \Rightarrow | \ 1, 0, 0 \ >$$

(a) Which of these nine transitions obey the $\Delta m = \pm 1$ or 0 dipole selection rule?

(b) Which of these nine transitions obey the $\Delta l = \pm 1$ dipole selection rule?

(c) The dipole allowed transitions must obey both rules. Which six of the nine transitions are dipole allowed?

(d) List all of the allowed dipole transition routes, which pass through the $n = 2$ states, from the $| \ 3, 0, 0 \ >$ state to the $| \ 1, 0, 0 \ >$ state, *i.e.,* list the three dipole allowed routes which have the form:

$$| \ 3, 0, 0 \ > \Rightarrow | \ 2, ?, ? \ > \Rightarrow | \ 1, 0, 0 \ > .$$

(e) Write down the integral for the dipole matrix element from the $| \ 3, 0, 0 \ >$ state to the $| \ 2, 1, 0 \ >$ state. Show that this matrix element only depends on the z component of the $\mathbf{r}$ operator, *i.e.,* show that

$$< \ 2, 1, 0 \ | \mathbf{r} | \ 3, 0, 0 \ > = < \ 2, 1, 0 \ | \ z \ | \ 3, 0, 0 \ > \mathbf{k}.$$

(f) Do the integral that you wrote down in part e. You should find $< \ 2, 1, 0 \ | \ z \ | \ 3, 0, 0 \ > =$

$$\left[\sqrt{\frac{3}{4\pi}} \sqrt{\frac{1}{24}} \ a^{-\frac{3}{2}} \right] \left[\sqrt{\frac{1}{4\pi}} \ \frac{2}{\sqrt{27}} \ a^{-\frac{3}{2}} \right] \int_0^\infty \left[r \ \cos(\theta) \ \exp\left(\frac{-r}{2a}\right) \left(1 - \frac{2r}{3a} + \frac{2r^2}{27a^2} \right)\right]$$

$$\times \left[r \cos(\theta) \exp\left(\frac{-r}{3a}\right) \right] r^2 \, dr \, \sin\theta \, d\theta \, d\phi$$

so

$$< 2,1,0 \mid z \mid 3,0,0 > = - \left[\frac{2^8 \, 3^4}{5^6 \sqrt{6}} \right] a.$$

(g) Write down the integrals for the dipole matrix elements from the $\mid 3,0,0 >$ state to the $\mid 2,1,\pm 1 >$ states. Show that these matrix elements only depend on the x and y components of the $\mathbf{r}$ operator, *i.e.*, show that

$$< 2,1,\pm 1 \mid \mathbf{r} \mid 3,0,0 > = < 2,1,\pm 1 \mid x \mid 3,0,0 > \mathbf{i} + < 2,1,\pm 1 \mid y \mid 3,0,0 > \mathbf{j}.$$

(h) Now show that these x and y matrix elements are almost identical, *i.e.,* show that

$$\pm < 2,1,\pm 1 \mid x \mid 3,0,0 > = i < 2,1,\pm 1 \mid y \mid 3,0,0 > .$$

Explain how you can use this to make your life simpler, *i.e.,* explain why you can just calculate one integral and still obtain all four matrix elements!!!

(i) Do the x integral you wrote down in part g. You should find $< 2,1,\pm 1 \mid x \mid 3,0,0 > =$

$$\left[\sqrt{\frac{3}{8\pi}} \sqrt{\frac{1}{24}} \, a^{-\frac{3}{2}} \right] \left[\sqrt{\frac{1}{4\pi}} \frac{2}{\sqrt{27}} \, a^{-\frac{3}{2}} \right] \int_0^\infty \left[r \sin(\theta) \exp(\pm i\phi) \exp\left(\frac{-r}{2a}\right) \left(1 - \frac{2r}{3a} + \frac{2r^2}{27a^2}\right) \right]$$

$$\times \left[r \cos(\theta) \exp\left(\frac{-r}{3a}\right) \right] r^2 \, dr \, \sin\theta \, d\theta \, d\phi$$

so

$$< 2,1,\pm 1 \mid x \mid 3,0,0 > = \pm \left[- \frac{2^7 \, 3^4}{5^6 \sqrt{3}} \right] a.$$

(j) According to Fermi's Golden Rule Number 2, the electric dipole transition rates are proportional to the squares of the matrix elements. Calculate the squares of these matrix elements and show that the two of the three decay routes have identical transition rates and that the third route has twice the transition rate,*i.e.,* show that

$$\frac{1}{2} \mid < 2,1,0 \mid \mathbf{r} \mid 3,0,0 > \mid^2 = \mid < 2,1,1 \mid \mathbf{r} \mid 3,0,0 > \mid^2 = \mid < 2,1,-1 \mid \mathbf{r} \mid 3,0,0 > \mid^2 .$$

So, one half go by one decay route, and one quarter each go by the other two decay routes.

(k) Now the spontaneous emission rates via these three routes are given by

$$A = \frac{\omega^3 \ |<\mathbf{r}>|^2}{3\ \pi\ \epsilon_0\ \hbar\ c^3},$$

so the the total decay rate is given by

$$R = 3\ A = 3 \left(\frac{e^2}{3\ \pi\ \epsilon_0\ \hbar\ c^3} \right) \left(\frac{-5\ E_1}{36\ \hbar} \right)^3 \left(\frac{2^{15}\ 3^7}{5^{12}} \right) a^2 = 6.32 \times 10^6 \text{ seconds}^{-1},$$

and the lifetime of the $|\ 3,0,0\ >$ state is given by $\tau = (1/R) = 1.58 \times 10^{-7}$ seconds.

4.(a) For an electron transition between the $n = 3, l = 0, m = 0$ and ground states, given that it can but does not have to go to the ground state directly, there are nine possible transitions.

All nine possible transitions obey the $\Delta m = \pm 1$ or 0 selection rule.
The nine possible transitions are

$$\begin{aligned}
&|3,0,0> \rightarrow |2,1,1> \\
&|3,0,0> \rightarrow |2,1,0> \\
&|3,0,0> \rightarrow |2,1,-1> \\
&|3,0,0> \rightarrow |2,0,0> \\
&|3,0,0> \rightarrow |1,0,0> \\
\\
&|2,1,1> \rightarrow |1,0,0> \\
&|2,1,0> \rightarrow |1,0,0> \\
&|2,1,-1> \rightarrow |1,0,0> \\
&|2,0,0> \rightarrow |1,0,0>
\end{aligned}$$

4.(b)

Six of these these transitions obey the $\Delta l = \pm 1$ selection rule. These six allowed transitions are

$$
\begin{aligned}
&|3,0,0> \,\rightarrow\, |2,1,1> \\
&|3,0,0> \,\rightarrow\, |2,1,0> \\
&|3,0,0> \,\rightarrow\, |2,1,-1> \\
\\
&|2,1,1> \,\rightarrow\, |1,0,0> \\
&|2,1,0> \,\rightarrow\, |1,0,0> \\
&|2,1,-1> \,\rightarrow\, |1,0,0>
\end{aligned}
$$

4.(c)

The six transitions listed in part b obey both dipole transition rules.

4.(d)

The three allowed transitions via an intermediate state are

$$
\begin{aligned}
&|3,0,0> \,\rightarrow\, |2,1,1> \,\rightarrow\, |1,0,0> \\
&|3,0,0> \,\rightarrow\, |2,1,0> \,\rightarrow\, |1,0,0> \\
&|3,0,0> \,\rightarrow\, |2,1,-1> \,\rightarrow\, |1,0,0>
\end{aligned}
$$

4.(e) The transition

$$
\begin{aligned}
<2,1,0\big|\,\vec{r}\,\big|3,0,0> \;&= \;<\psi_{210}\big|\,\vec{r}\,\big|\psi_{300}> \\
&= \;<R_{21}Y_{10}\big|\,\vec{r}\,\big|R_{30}Y_{00}> \\
&= \int R_{21}Y_{10}\,\vec{r}\,R_{30}Y_{00}\,dV \\
&= \int R_{21}R_{30}r^2\,dr \int Y_{10}\,\vec{r}\,Y_{00}\,d\Omega
\end{aligned}
\qquad .
$$

The angular part of this equation is

$$\int \left(\frac{3}{4\pi}\right)^{1/2} \cos\theta\,\vec{r}\,\left(\frac{1}{4\pi}\right)^{1/2} d\Omega = \frac{\sqrt{3}}{4\pi}\int \vec{r}\,\cos\theta\,d\Omega.$$

Remember $\vec{z} = \vec{r}\cos\theta = z\hat{k}$ so generalizing back into Dirac notation,

$$\boxed{<2,1,0|\,\vec{r}\,|3,0,0> \;=\; <2,1,0|\,z\,|3,0,0>\,\hat{k}.}$$

4.(f) Evaluating the integral by inserting the appropriate radial and angular functions, we find that the matrix element we seek $<2,1,0|\,z\,|3,0,0>$ is equal to the integral

$$I \;=\; \int_{-\infty}^{\infty}\left(\frac{1}{\sqrt{24}}a^{-3/2}\frac{r}{a}e^{-r/2a}\left(\frac{3}{4\pi}\right)^{1/2}\cos\theta\right)^{*}\,(\,z\,)\,\frac{2}{\sqrt{27}}a^{-3/2}\left(1-\frac{2r}{3a}+\frac{2}{27}\frac{r^2}{a^2}\right)e^{-r/3a}\left(\frac{1}{4\pi}\right)^{1/2}\,dV.$$

Factoring out the constants and simplifying, we find:

$$\begin{aligned}
I \;&=\; \frac{1}{\sqrt{24}}\frac{2}{\sqrt{27}}\frac{1}{a^4}\left(\frac{3}{4\pi}\right)^{1/2}\left(\frac{1}{4\pi}\right)^{1/2}\int_{-\infty}^{\infty} r\,e^{-r/2a}\cos\theta\,(\,r\cos\theta\,)\,\left(1-\frac{2r}{3a}+\frac{2}{27}\frac{r^2}{a^2}\right)e^{-r/3a}\,dV\\
&= \frac{1}{12\pi\sqrt{6}a^4}\int_{-\infty}^{\infty} r^2\cos^2\theta\left(1-\frac{2r}{3a}+\frac{2}{27}\frac{r^2}{a^2}\right)e^{-5r/6a}\,dV
\end{aligned}$$

And by doing the angular integrals, we can reduce the problem to the radial integrals that we must do

$$\begin{aligned}
I \;&=\; \frac{1}{12\pi\sqrt{6}a^4}\int_0^{\infty} r^4e^{-5r/6a}\left(1-\frac{2r}{3a}+\frac{2}{27}\frac{r^2}{a^2}\right)dr\int_0^{\pi}\cos^2\theta\sin\theta\,d\theta\int_0^{2\pi}d\phi.\\
&= \frac{1}{12\pi\sqrt{6}a^4}\int_0^{\infty} r^4e^{-5r/6a}\left(1-\frac{2r}{3a}+\frac{2}{27}\frac{r^2}{a^2}\right)dr\int_0^{\pi}\cos^2\theta\sin\theta\,d\theta\,(2\pi)\\
&= \frac{1}{6\sqrt{6}a^4}\int_0^{\infty}\left(r^4e^{-5r/6a}-\frac{2}{3a}r^5e^{-5r/6a}+\frac{2}{27a^2}r^6e^{-5r/6a}\right)dr\,\left(\frac{\cos^3\theta}{3}\Big|_0^{\pi}\right)\\
&= \frac{1}{6\sqrt{6}a^4}\int_0^{\infty}\left(r^4e^{-5r/6a}-\frac{2}{3a}r^5e^{-5r/6a}+\frac{2}{27a^2}r^6e^{-5r/6a}\right)dr\,\left(\frac{-1-1}{3}\right)\\
&= -\frac{1}{9\sqrt{6}a^4}\left(\int_0^{\infty} r^4e^{-5r/6a}\,dr-\frac{2}{3a}\int_0^{\infty} r^5e^{-5r/6a}\,dr+\frac{2}{27a^2}\int_0^{\infty} r^6e^{-5r/6a}\,dr\right).
\end{aligned} \tag{1}$$

We can evaluate all three radial integrals using form 3.381.4 on page 317 of Gradshteyn and Ryzhik, which is

$$\int_0^{\infty} x^{\nu-1}e^{-\mu x}\,dx = \frac{1}{\mu^{\nu}}\Gamma(\nu), \qquad \text{Re }\mu>0, \qquad \text{Re }\nu\;>0.$$

For the first integral, $\nu=5$ and $\mu=5/6a$, so

$$\int_0^{\infty} r^4e^{-5r/6a}\,dr = \frac{1}{(5/6a)^5}\Gamma(5) = \frac{6^5a^5}{5^5}4\cdot3\cdot2 = 24\frac{6^5a^5}{5^5}.$$

For the second integral, $\nu=6$ and $\mu=5/6a$, so

$$\frac{2}{3a}\int_0^{\infty} r^5e^{-5r/6a}\,dr = \frac{1}{(5/6a)^6}\Gamma(6) = \frac{2}{3a}\frac{6^6a^6}{5^6}5\cdot4\cdot3\cdot2 = 80\frac{6^6a^5}{5^6}.$$

For the third integral, $\nu = 7$ and $\mu = 5/6a$, so

$$\frac{2}{27a^2}\int_0^\infty r^6 e^{-5r/6a}\,dr = \frac{1}{(5/6a)^7}\Gamma(7) = \frac{2}{27a^2}\frac{6^7a^7}{5^7}6\cdot5\cdot4\cdot3\cdot2 = \frac{160}{3}\frac{6^7a^5}{5^7}.$$

Substituting these into equation (1),

$$\begin{aligned}
<2,1,0\big|\,z\,\big|3,0,0> &= -\frac{1}{9\sqrt{6}a^4}a^5\left(24\frac{6^5}{5^5} - 80\frac{6^6}{5^6} + \frac{160}{3}\frac{6^7}{5^7}\right)\\
&= -\frac{a}{9\sqrt{6}}\frac{6^5}{5^6}\left(120 - 80\cdot6 + \frac{160\cdot6^2}{3\cdot5}\right)\\
&= -\frac{6^5a}{5^63^2\sqrt{6}}\left(120 - 480 + 384\right)\\
&= -\frac{6^5a}{5^63^2\sqrt{6}}\left(24\right)\\
&= -\frac{2^53^5a}{5^63^2\sqrt{6}}\left(2^3\cdot3\right)
\end{aligned}$$

$$\boxed{\Rightarrow \quad <2,1,0\big|\,z\,\big|3,0,0> = -\frac{2^83^4}{5^6\sqrt{6}}a}$$

4.(g) The integrals for $< 2,1,\pm1\big|\,\vec{\mathbf{r}}\,\big|3,0,0 >$ are easier. These integral depend only on the x and y components of the $\vec{\mathbf{r}}$ operator. Here

$$\begin{aligned}
<2,1,\pm1\big|\,\vec{\mathbf{r}}\,\big|3,0,0> &= \int R_{21}^*Y_{1,\pm1}^*\big(\vec{\mathbf{r}}\big)R_{30}Y_{00}\,dV\\
&= \int R_{21}R_{30}r^2\,dr\int Y_{1,\pm1}^*\big(\vec{\mathbf{r}}\big)Y_{00}\,d\Omega.
\end{aligned}$$

The angular integral is

$$\begin{aligned}
\int Y_{1,\pm1}^*\big(\vec{\mathbf{r}}\big)Y_{00}\,d\Omega &= \int\left(\mp\left(\frac{3}{8\pi}\right)^{1/2}\right)\sin\theta e^{\mp i\phi}\big(\vec{\mathbf{r}}\big)\left(\frac{1}{4\pi}\right)^{1/2}\,d\Omega\\
&= \mp\left(\frac{3}{8\pi}\right)^{1/2}\left(\frac{1}{4\pi}\right)^{1/2}\int\sin\theta e^{\mp i\phi}\big(\vec{\mathbf{r}}\big)\,d\Omega\\
&= \mp\frac{1}{4\pi}\sqrt{\frac{3}{2}}\int\big(\vec{\mathbf{r}}\big)\sin\theta\left(\cos\phi\mp i\sin\phi\right)\,d\Omega\\
&= \mp\frac{1}{4\pi}\sqrt{\frac{3}{2}}\int\left(\vec{\mathbf{r}}\sin\theta\cos\phi\mp i(\vec{\mathbf{r}}\sin\theta\sin\phi)\right)\,d\Omega.
\end{aligned}$$

Realizing $\vec{\mathbf{r}}\sin\theta\cos\phi = \vec{\mathbf{x}} = x\hat{\mathbf{i}}$ and $\vec{\mathbf{r}}\sin\theta\sin\phi = \vec{\mathbf{y}} = y\hat{\mathbf{j}}$, we can write this

$$\int Y_{1,\pm1}^*\big(\vec{\mathbf{r}}\big)Y_{00}\,d\Omega = \mp\frac{1}{4\pi}\sqrt{\frac{3}{2}}\int\left(x\hat{\mathbf{i}}\mp i(y\hat{\mathbf{j}})\right)\,d\Omega,$$

i.e., we can look at directional or angular dependence as a function of $\vec{\mathbf{x}}$ and $\vec{\mathbf{y}}$ only. Generalizing back into Dirac notation, which is representation free so the constants are irrelevant,

$$\begin{aligned} <2,1,\pm 1\big|\,\vec{\mathbf{r}}\,\big|3,0,0> &= <2,1,\pm 1\big|\,x\hat{\mathbf{i}} \mp i(y\hat{\mathbf{j}})\,\big|3,0,0> \\ &= <2,1,\pm 1\big|\,x\,\big|3,0,0>\hat{\mathbf{i}} \;\mp\; <2,1,\pm 1\big|\,iy\,\big|3,0,0>\hat{\mathbf{j}}\,. \end{aligned}$$

The sign "$\mp$" between the two elements reflects only a phase convention, and we will choose without loss of generality the "+" sign for our phase so

$$\boxed{<2,1,\pm 1\big|\,\vec{\mathbf{r}}\,\big|3,0,0> = <2,1,\pm 1\big|\,x\,\big|3,0,0>\hat{\mathbf{i}} \;+\; <2,1,\pm 1\big|\,iy\,\big|3,0,0>\hat{\mathbf{j}}.}$$

4.(h) To show

$$<2,1,\pm 1\big|\,x\,\big|3,0,0> = \;i<2,1,\pm 1\big|\,y\,\big|3,0,0>$$

consider the commutator $[L_z, x] = i\hbar y$, and the eigenvalue equation $L_z|n,l,m> = m\hbar|n,l,m>$. In general

$$\begin{aligned} <n',l',m'\big|\,[L_z,x]\,\big|n,l,m> &= <n',l',m'\big|\,i\hbar y\,\big|n,l,m> \\ &= i\hbar <n',l',m'\big|\,y\,\big|n,l,m>\,. \end{aligned}$$

This must be the same as $<n',l',m'\big|\,[L_z,x]\,\big|n,l,m>$ when the commutator is evaluated explicitly, *i.e.*,

$$\begin{aligned} i\hbar <n',l',m'\big|\,y\,\big|n,l,m> &= <n',l',m'\big|\,[L_z,x]\,\big|n,l,m> \\ &= <n',l',m'\big|\,L_z x - xL_z\,\big|n,l,m> \end{aligned}$$

where L_z can operate to the left or right. So

$$\begin{aligned} i\hbar <n',l',m'\big|\,y\,\big|n,l,m> &= <n',l',m'\big|\,m'\hbar x - xm\hbar\,\big|n,l,m> \\ &= <n',l',m'\big|\,(m'-m)\hbar x\,\big|n,l,m> \\ &= (m'-m)\hbar <n',l',m'\big|\,x\,\big|n,l,m> \end{aligned}$$

$$\Rightarrow \quad (m'-m)<n',l',m'\big|\,x\,\big|n,l,m> = \;i<n',l',m'\big|\,y\,\big|n,l,m>\,.$$

For the specific states of interest

$$(1-0)<2,1,1\big|\,x\,\big|3,0,0> = \;i<2,1,1\big|\,y\,\big|3,0,0>$$

$$\Rightarrow \quad <2,1,1\big|\,x\,\big|3,0,0> = \;i<2,1,1\big|\,y\,\big|3,0,0>,$$

and

$$(-1-0)<2,1,-1\big|\,x\,\big|3,0,0> = \;i<2,1,-1\big|\,y\,\big|3,0,0>$$

$$\Rightarrow \quad -<2,1,-1\big|\,x\,\big|3,0,0> = \;i<2,1,-1\big|\,y\,\big|3,0,0>,$$

so

$$\boxed{\pm<2,1,\pm 1\big|\,x\,\big|3,0,0> = \;i<2,1,\pm 1\big|\,y\,\big|3,0,0>\,.}$$

There are four matrix elements here. If we evaluate the two integrals in x though, we have the two integrals in y from the above relation. Also, because of the symmetry in ϕ, we can do both integrals in x at the same time, so in effect, we have only one integral to evaluate to get all four matrix elements.

4.(i) To evaluate the integrals in x, remember $x = r\sin\theta\cos\phi$, and

$$\pm <2,1,\pm 1\big|\, x\,\big|3,0,0> \; = \pm <2,1,\pm 1\big|\, r\sin\theta\cos\phi\,\big|3,0,0> \quad \text{so}$$

$$\begin{aligned}
\pm <2,1,\pm 1\big|\, x\,\big|3,0,0> &= \int_{-\infty}^{\infty} R_{21}^* Y_{1,\pm 1}^*\, r\sin\theta\cos\phi\, R_{30}Y_{00}\, r^2\Omega \\
&= \int_0^{\infty} R_{21}R_{30}r^3\, dr \int Y_{1,\pm 1}^*\, \sin\theta\cos\phi\, Y_{00}\, d\Omega
\end{aligned}$$

$$= \int_0^{\infty} \frac{1}{\sqrt{24}} a^{-3/2}\frac{r}{a}e^{-r/2a}\frac{2}{\sqrt{27}}a^{-3/2}\left(1 - \frac{2r}{3a} + \frac{2}{27}\frac{r^2}{a^2}\right)e^{-r/3a}r^3\, dr \int \mp\sqrt{\frac{3}{8\pi}}\sin\theta e^{\mp i\phi}(\sin\theta\cos\phi)\sqrt{\frac{1}{4\pi}}\, d\Omega$$

$$= \mp\frac{1}{\sqrt{24}}\frac{2}{\sqrt{27}}\frac{1}{a^4}\sqrt{\frac{3}{8\pi}}\sqrt{\frac{1}{4\pi}}\int_0^{\infty} r^4 e^{-5r/6a}\left(1 - \frac{2r}{3a} + \frac{2}{27}\frac{r^2}{a^2}\right)dr \int_0^{\pi}\sin^3\theta\, d\theta \int_0^{2\pi}\cos\phi e^{\mp i\phi}\, d\phi, \tag{1}$$

where the third factor of $\sin\theta$ is from $d\Omega = \sin\theta\, d\theta\, d\phi$. The constants are

$$\mp\frac{1}{\sqrt{24}}\frac{2}{\sqrt{27}}\frac{1}{a^4}\sqrt{\frac{3}{8\pi}}\sqrt{\frac{1}{4\pi}} = \mp\frac{1}{\sqrt{2^3\cdot 3}}\frac{2}{\sqrt{3^3}}\frac{1}{a^4}\frac{1}{4\pi}\frac{\sqrt{3}}{\sqrt{2}} = \mp\frac{2\sqrt{3}}{\sqrt{2^4\cdot 3^4}}\frac{1}{4\pi a^4} = \mp\frac{1}{24\pi a^4\sqrt{3}}.$$

The azimuthal integral is

$$\begin{aligned}
\int_0^{2\pi}\cos\phi e^{\mp i\phi}\, d\phi = \int_0^{2\pi}\cos\phi(\cos\phi \mp i\sin\phi)\, d\phi &= \int_0^{2\pi}\cos^2\phi\, d\phi \mp i\int_0^{2\pi}\cos\phi\sin\phi)\, d\phi \\
&= \left[\frac{1}{2}\phi + \frac{1}{4}\sin(2\phi)\right]_0^{2\pi} \mp i\left[\frac{1}{2}\sin^2\phi\right]_0^{2\pi} = \left[\frac{1}{2}2\pi - 0 + 0 - 0\right] \mp i\,[0-0] = \pi.
\end{aligned}$$

The polar integral is

$$\int_0^{\pi}\sin^3\theta\, d\theta = -\frac{1}{3}\Big[(\cos\theta)(\sin^2\theta + 2)\Big]_0^{\pi} = -\frac{1}{3}\left[(-1)(0+2) - (1)(0+2)\right] = -\frac{1}{3}\left[-2-2\right] = \frac{4}{3}.$$

The radial integral becomes three integrals

$$\int_0^{\infty} r^4e^{-5r/6a}\left(1 - \frac{2r}{3a} + \frac{2}{27}\frac{r^2}{a^2}\right)dr = \int_0^{\infty} r^4e^{-5r/6a}\, dr - \frac{2}{3a}\int_0^{\infty} r^5e^{-5r/6a}\, dr + \frac{2}{27a^2}\int_0^{\infty} r^6e^{-5r/6a}\, dr$$

and we have already evaluated these integrals Using the results of part (f),

$$\int_0^{\infty} r^4e^{-5r/6a}\, dr = 24\frac{6^5a^5}{5^5},$$

$$\frac{2}{3a}\int_0^{\infty} r^5e^{-5r/6a}\, dr = 80\frac{6^6a^5}{5^6},$$

$$\frac{2}{27a^2}\int_0^{\infty} r^6e^{-5r/6a}\, dr = \frac{160}{3}\frac{6^7a^5}{5^7}.$$

Compiling these six results, equation (1) becomes

$$\begin{aligned}
\pm <2,1,\pm1\big|\, x\,\big|3,0,0> &= \mp\frac{1}{24\pi a^4\sqrt{3}}\pi\frac{4}{3}\left(24\frac{6^5a^5}{5^5}-80\frac{6^6a^5}{5^6}+\frac{160}{3}\frac{6^7a^5}{5^7}\right)\\
&= \mp\frac{1}{18a^4\sqrt{3}}\frac{6^5a^5}{5^6}\left(24\cdot5-80\cdot6+\frac{160}{3}\frac{6^2}{5}\right)\\
&= \mp\frac{a}{2\cdot3^2\sqrt{3}}\frac{6^5}{5^6}\left(120-480+384\right)\\
&= \mp\frac{a}{2\cdot3^2\sqrt{3}}\frac{2^5\cdot3^5}{5^6}\left(24\right)\\
&= \mp\frac{2^5\cdot3^5\cdot2^3\cdot3}{2\cdot3^2\cdot5^6\sqrt{3}}a
\end{aligned}$$

$$\boxed{\begin{aligned}
&\Rightarrow\quad <2,1,\pm1\big|\, x\,\big|3,0,0> = \pm\left[-\frac{2^7\cdot3^4}{5^6\sqrt{3}}\right]a,\\
&\qquad\text{and}\\
&<2,1,\pm1\big|\, y\,\big|3,0,0> = \pm i\left[-\frac{2^7\cdot3^4}{5^6\sqrt{3}}\right]a.
\end{aligned}}$$

4.(j) According to Fermi's Golden Rule Number 2, the electric dipole transition rates are proportional to the squares of the matrix elements. We have all three matrix elements so we can calculate the relative rates of decays for the three paths. From part f, we have

$$\big|<2,1,0\big|\,\vec{\mathbf{r}}\,\big|3,0,0>\big|^2 = \left[-\frac{2^8\cdot3^4}{5^6\sqrt{6}}a\right]^2 = \frac{2^{16}\cdot3^8}{5^{12}\cdot6}a^2 = \frac{2^{15}\cdot3^7}{5^{12}}a^2,$$

and from parts g and i, we have

$$<2,1,\pm1\big|\,\vec{\mathbf{r}}\,\big|3,0,0>=<2,1,\pm1\big|\, x\,\big|3,0,0> \pm i <2,1,\pm1\big|\, y\,\big|3,0,0>,$$

so the total transition rate is the sum of the x and y induced rates, and is twice as large as the individual x and y matrix elements squared:

$$\big|<2,1,\pm1\big|\,\vec{\mathbf{r}}\,\big|3,0,0>\big|^2 = 2\left[\mp\frac{2^7\cdot3^4}{5^6\sqrt{3}}a\right]^2 = \frac{2^{15}\cdot3^7}{5^{12}}a^2.$$

Consequently, we conclude that the three decay rates are equal:

$$\boxed{\big|<2,1,0\big|\,\vec{\mathbf{r}}\,\big|3,0,0>\big|^2 = \big|<2,1,1\big|\,\vec{\mathbf{r}}\,\big|3,0,0>\big|^2 = \big|<2,1,-1\big|\,\vec{\mathbf{r}}\,\big|3,0,0>\big|^2.}$$

4.(k) The spontaneous emission rates are given by

$$A = \frac{\omega^3 \left| \, q < \psi_b \right| \vec{\mathbf{r}} \left| \psi_a > \, \right|^2}{3\pi\epsilon_0 \hbar c^3} \qquad \text{where} \qquad \omega = \frac{E_b - E_a}{\hbar}.$$

These are given by

$$\begin{aligned}
A_{3,0,0\to 2,1,0} &= \frac{\left[13.6/2^2 - 13.6/3^2\right]^3 \frac{1}{\hbar^3} e^2}{3\pi\epsilon_0 \hbar c^3} \left| \; <3,0,0 \right| \vec{\mathbf{r}} \left| 2,1,0> \; \right|^2 \\
&= \left(\frac{e^2}{4\pi\epsilon_0} \right) \frac{4}{3} \frac{1}{\hbar^4 c^3} \left[\frac{13.6}{4} - \frac{13.6}{9} \right]^3 \frac{2^{15} \cdot 3^7}{5^{12}} a^2 \\
&= (1.440\,\text{eV} \cdot \text{nm}) \, \frac{4}{3} \frac{(2\pi)^4}{h^4 c^3} \left[3.40 - 1.51\right]^3 (\text{eV})^3 \, 0.294 a^2 \\
&= \frac{(1.440\,\text{eV} \cdot \text{nm})}{(hc)^3} \frac{64\pi^4}{3h} \left[1.89\right]^3 \; \text{eV}^3 \, (0.294)(0.0529\,\text{nm})^2 \\
&= \frac{(1.440\,\text{eV} \cdot \text{nm})}{(1.240 \times 10^3\,\text{eV} \cdot \text{nm})^3} \frac{2078.06}{h} \left[6.75\right] \; (0.294)(0.00280) \, \text{eV}^3 \text{nm}^2 \\
&= \frac{(1.440\,\text{eV} \cdot \text{nm})}{1.907 \times 10^9 \, \text{eV}^3 \cdot \text{nm}^3} \frac{11.547}{h} \, \text{eV}^3 \text{nm}^2 \\
&= \frac{8.72^{-9}\,\text{eV}}{4.136 \times 10^{-15} \text{eV} \cdot \text{s}} \\
&= 2.11 \times 10^6 \, \text{s}^{-1}
\end{aligned}$$

$$\boxed{\Rightarrow \qquad \tau_{3,0,0\to 2,1,0} = \frac{1}{A} = 4.75 \times 10^{-7}\,\text{s}.}$$

The spontaneous emission rates for $< 2,1,1 \left| \, \vec{\mathbf{r}} \, \right| 3,0,0 >$ and $< 2,1,-1 \left| \, \vec{\mathbf{r}} \, \right| 3,0,0 >$ are calculated similarly, and since the matrix elements are identical in value, we find:

$$A_{3,0,0\to 2,1,1} = A_{3,0,0\to 2,1,-1} = A_{3,0,0\to 2,1,0} = 2.11 \times 10^6 \, \text{s}^{-1}.$$

So the rate via each path is the same:

$$\boxed{\begin{aligned}
\tau_{3,0,0\to 2,1,1} &= \frac{1}{A} = 4.75 \times 10^{-7}\,\text{s} \\
\tau_{3,0,0\to 2,1,-1} &= \frac{1}{A} = 4.75 \times 10^{-7}\,\text{s} \\
\tau_{3,0,0\to 2,1,0} &= \frac{1}{A} = 4.75 \times 10^{-7}\,\text{s},
\end{aligned}}$$

and the total decay rate is set by

$$A_T = 3(2.11 \times 10^6\,\mathrm{s}^{-1}) = 6.33 \times 10^6\,\mathrm{s}^{-1},$$

which gives us the lifetime

$$\boxed{\tau_T = \frac{1}{A_T} = 1.58 \times 10^{-7}\,\mathrm{s}.}$$

Fermi's Theory of Beta Decay

Fermi Theory of Beta Decay

In 1930, Wolfgang Pauli postulated the existence of the neutrino to explain the continuous distribution of energy of the electrons emitted in beta decay. Only with the emission of a third particle could momentum and energy be conserved. By 1934, Enrico Fermi had developed a theory of beta decay to include the neutrino, presumed to be massless as well as chargeless.

Treating the beta decay as a transition that depended upon the strength of coupling between the initial and final states, Fermi developed a relationship which is now referred to as Fermi's Golden Rule:

$$\lambda_{if} = \frac{2\pi}{\hbar} \left| M_{if} \right|^2 \rho_f$$

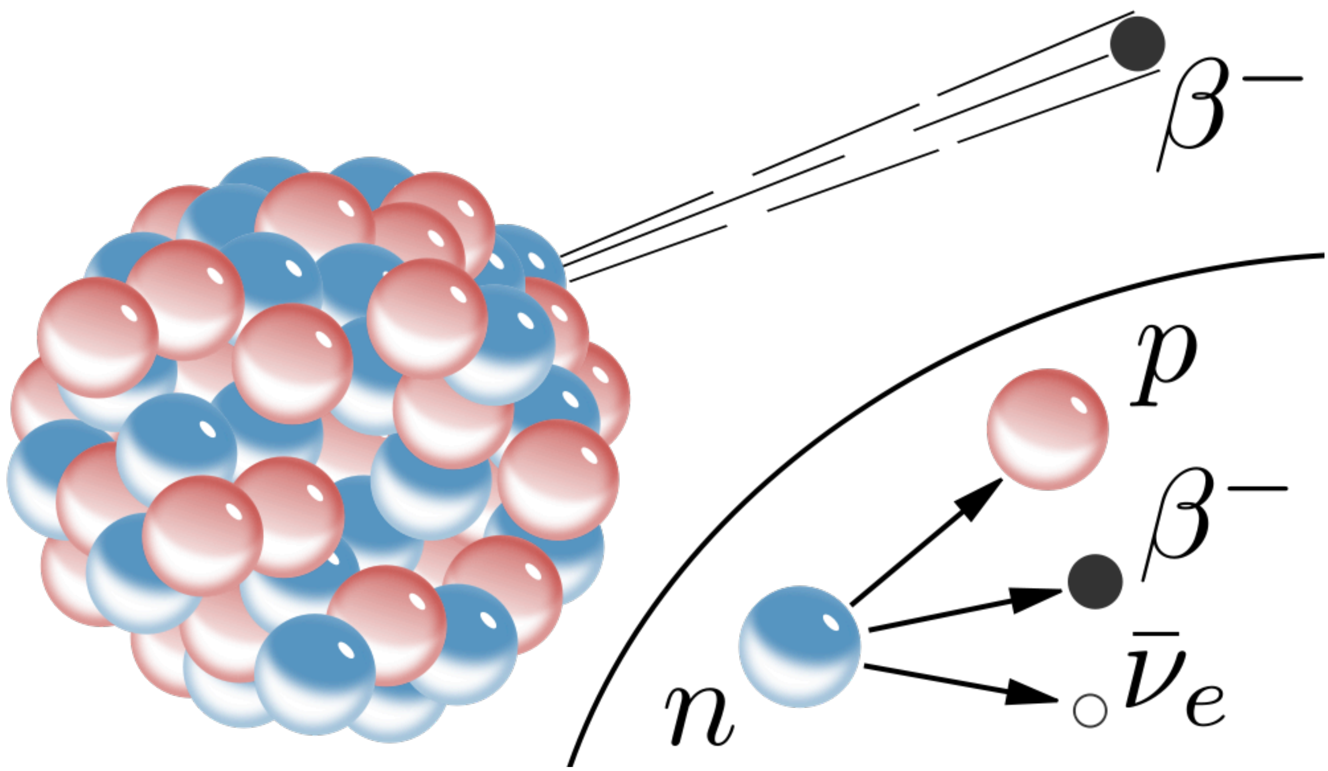
Fermi's Golden Rule

Transition probability

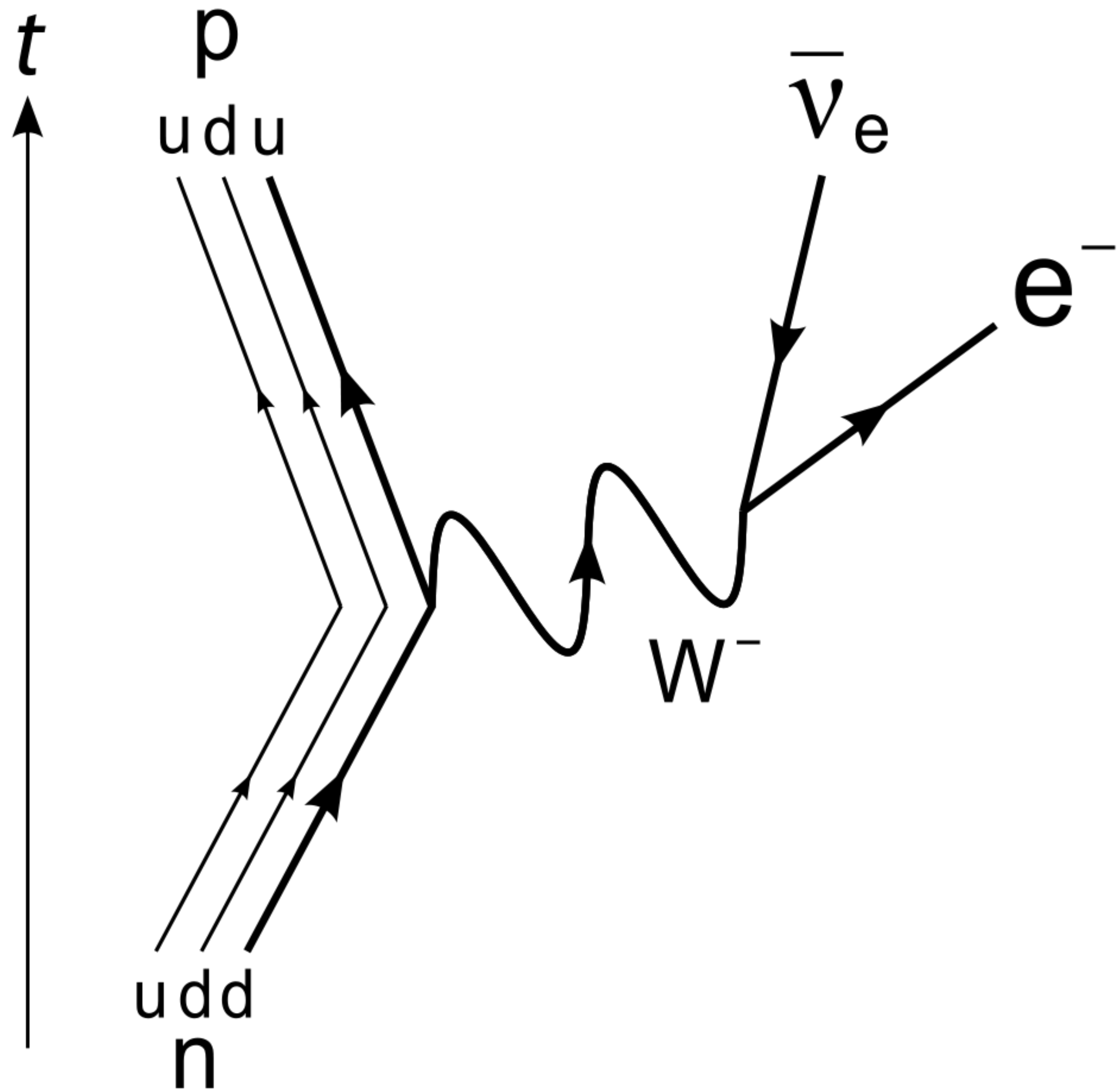
Matrix element for the interaction

Density of final states

Straightforward in concept, Fermi's Golden Rule says that the transition rate is proportional to the strength of the coupling between the initial and final states factored by the density of final states available to the system. But the nature of the interaction which led to beta decay was unknown in Fermi's time (the weak interaction). It took some 20 years of work (Krane) to work out a detailed model which fit the observations. The nature of that model in terms of the distribution of electron momentum p is summarized in the relationship below.



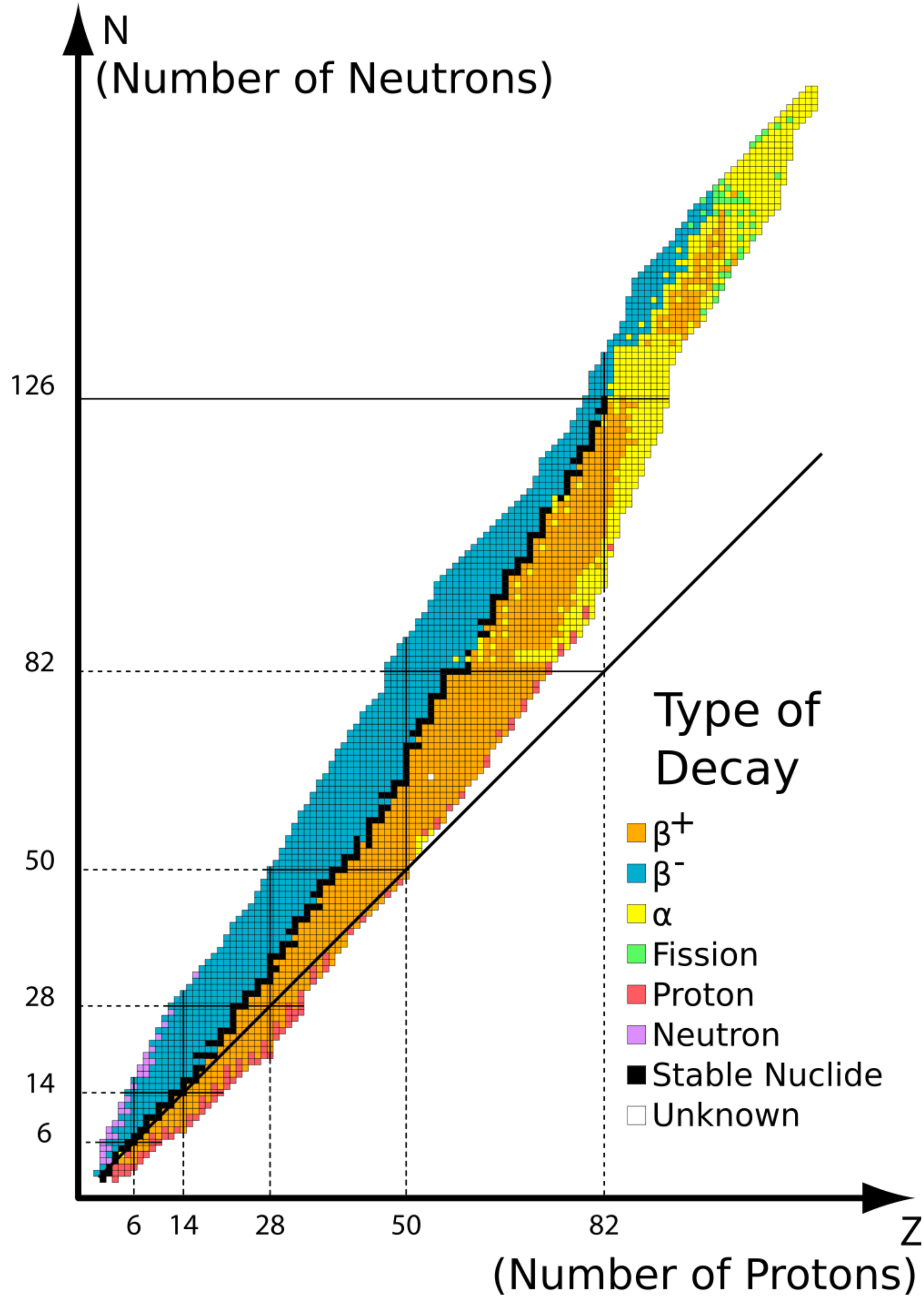
β^-
p
β^-
n
$\bar{\nu}_e$


t
p
u d u
$\bar{\nu}_e$
e^-
W^-
u d d
n

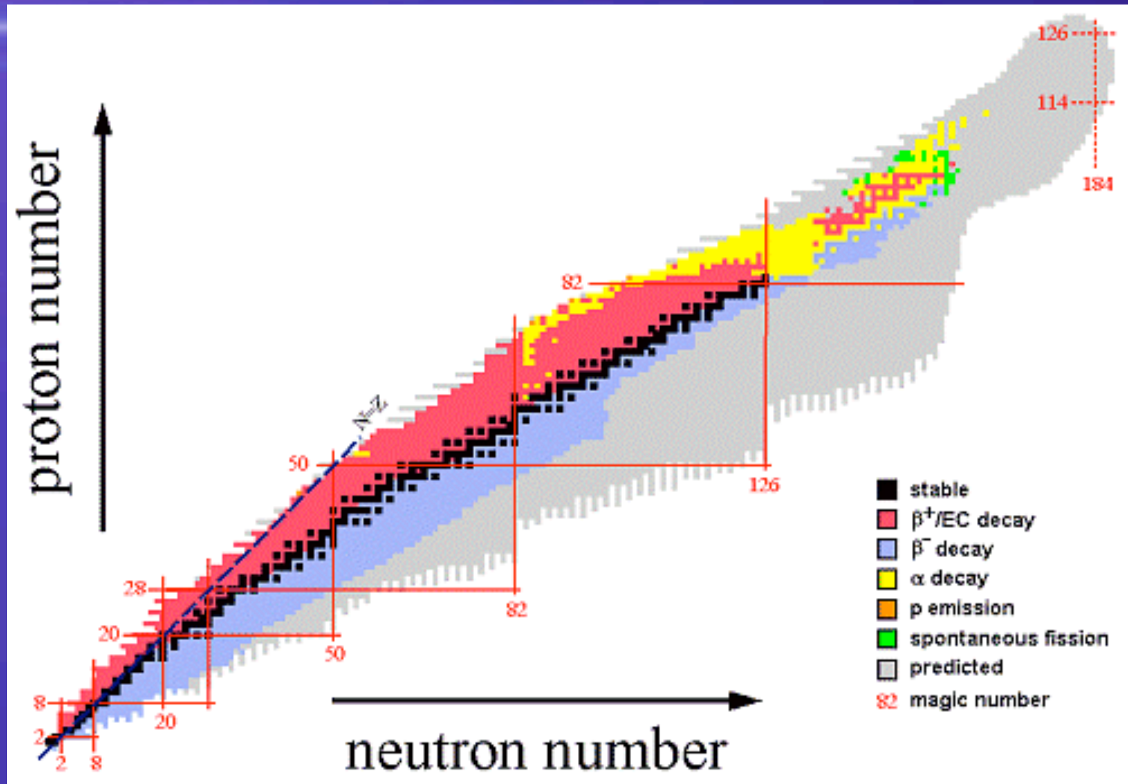
The Fermi Selection Rule for Beta Decay

(`fer·me si`lek·shen `rül)

There is no change in the total angular momentum or the parity of the nucleus


N
(Number of Neutrons)
126
82
50
28
14
6
6
14
28
50
82
Z
(Number of Protons)
Type of Decay
β+
β-
α
Fission
Proton
Neutron
Stable Nuclide
Unknown

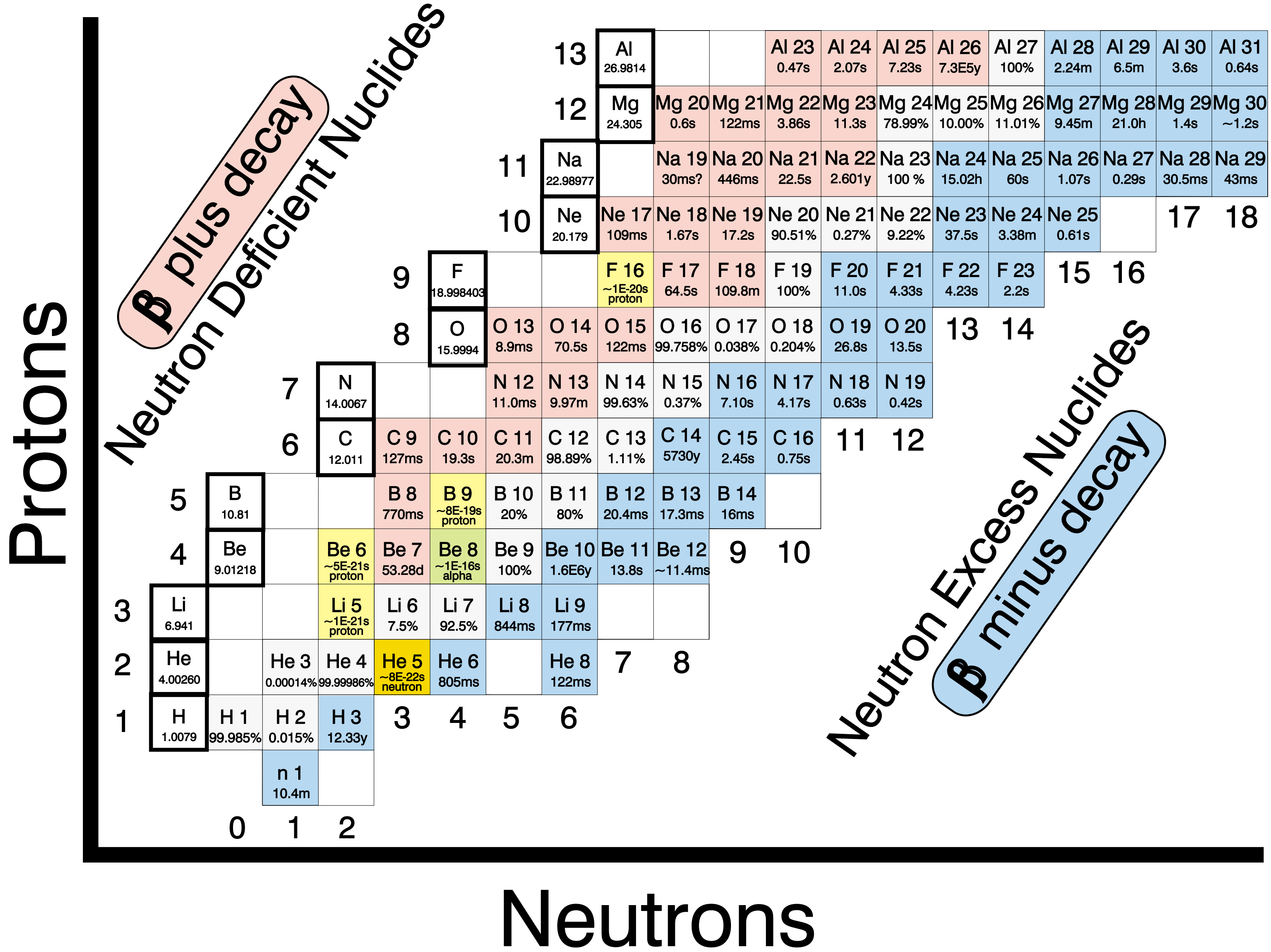
Requirements - I



1.) m(A,Z) > m(A,Z+2)

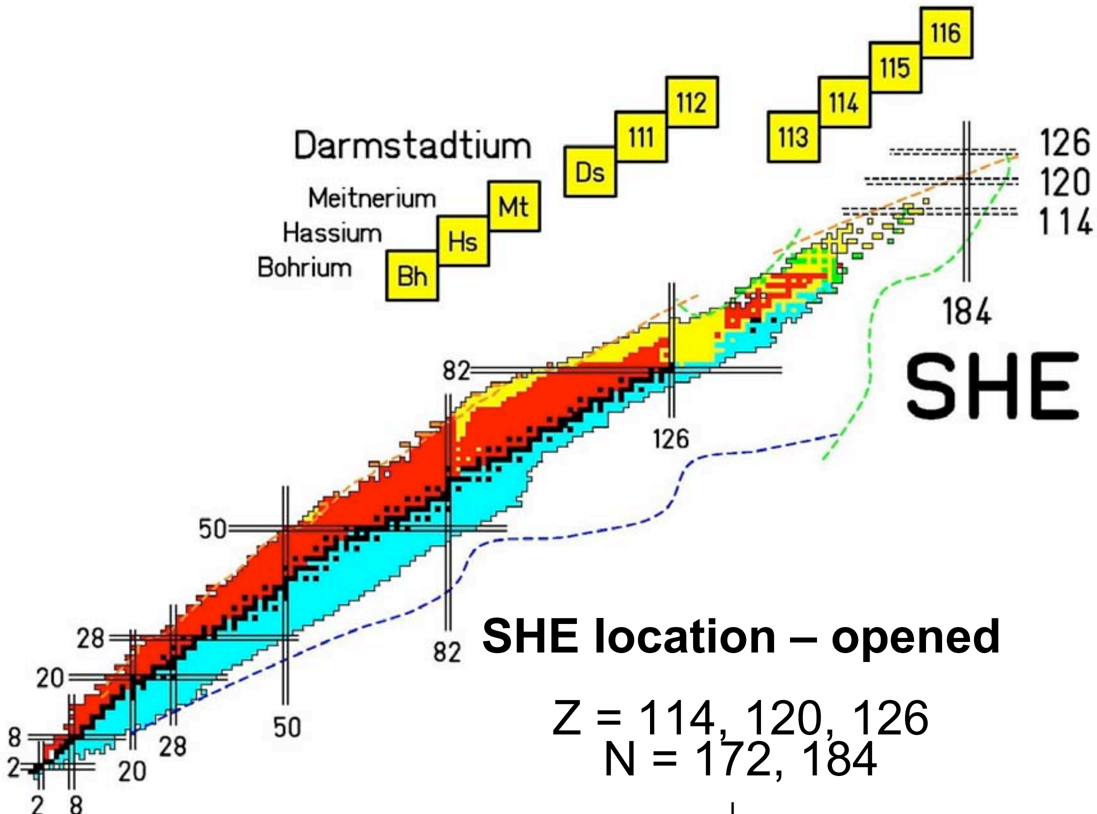
2.) Single beta decay must be forbidden (m (A,Z) < m (A,Z+1)) or at least strongly suppressed (large change in angular momentum)

CHART OF THE NUCLIDES



Physical Motivation

Predictions of the island of SHE



SHE location – opened

Z = 114, 120, 126
N = 172, 184

↓

fundamentals of spin-orbit force

- nuclei beyond Fm exist only due to shell effects
- predictions of highly stabilized SHE (T_{theo} ~ min -y)
- failure to synthesize SHE by reactions of the type Pb + Pb (U + U)

1) Production of SHE via "hot" and "cold" fusion

2) Systematic study of nuclear structure of transfermium isotopes

Our Understanding of Beta Decay

Pauli: missing momentum => undetected particle

Fermi: Point interaction theory

Sudarshan-Marshak: VA theory
Feynman-Gellmann

Neutrino mass: no theory yet

10 Theories of Neutrino Masses and Mixings

Rabindra N. Mohapatra

10.1 Introduction

The history of weak-interaction physics has to a large extent been a history of our understanding of the properties of the elusive spin-half particles called neutrinos. Evidence for only left-handed neutrinos being emitted in beta decay was the cornerstone of the successful V–A theory of weak interactions suggested by Sudarshan, Marshak, Feynman and Gell-Mann; evidence for the neutral-current interactions in the early 1970s provided brilliant confirmation of the successful gauge unification of the weak and electromagnetic interactions proposed by Glashow, Salam and Weinberg.

Today, as we enter a new millennium, we again have evidence for a very important new property of neutrinos, i.e. they have mass and, as a result, like the quarks, they mix with each other and lead to the phenomenon of neutrino oscillation. This is contrary to the expectations based on the Standard Model as well as on the old V–A theory (in fact one may recall that one way to make the V–A theory plausible was to use invariance of the weak Lagrangian under the so-called γ_5 invariance of all fermions, a principle which was motivated by the assumption that neutrinos have zero mass). The simple fact that neutrino masses vanish in the Standard Model is proof that the nonzero neutrino mass is an indication of new physics at some higher scale (or shorter distances). Study of details of neutrino masses and mixings is therefore going to open up new vistas in our journey towards a deeper understanding of the properties of the weak interactions at very short distances. This, no doubt, will have profound implications for the nature of the final theory of particles, forces and the Universe.

We are, of course, far from a complete picture of the masses and mixings of the various neutrinos and cannot therefore have a full outline of the theory that explains them. However, there exist enough information and indirect indications that constrain the masses and mixings among the neutrinos that we can see a narrowing of the possibilities for the theories. Many clever experiments now under way will soon clarify or rule out many of the allowed models. It is one of the goals of this chapter to give a panoramic view of the most likely scenarios for new physics that explain what is now known about neutrino masses.[1] We hope to emphasize the several interesting ideas for understanding the small neutrino masses and discuss in general terms how they can lead to the scenarios for neutrinos currently being discussed in order to understand the observations. These ideas have a very good chance of being part of the final theory of neutrino masses. We then touch briefly on some specific models that are based on the above general framework but attempt to provide an understanding of the detailed mass and mixing patterns. These works are instructive for several reasons: first, they provide proof of the detailed workability of the general ideas described above (they provide a sort of existence proofs that things will work); second, they often illustrate the kind of assumptions needed and through doing so a unique insight into which direction the next step should be in; and finally of course, nature may be generous in picking one of those models as the final message bearer.

As discussed in an earlier section, the neutrino mass can be of either Dirac or Majorana type. In this article we shall discuss our understanding of neutrino masses assuming that they are of Majorana type.

10.2 Experimental Indications of Neutrino Masses

As has been extensively discussed elsewhere in this book, while the direct-search experiments for neutrino masses using tritium beta decay and neutrinoless double beta decay have only yielded upper limits, the searches for neutrino oscillation, which can occur only if neutrinos have masses and mixings, have yielded positive evidence. There is now clear evidence from one experiment and strong indications from other experiments of neutrino oscillations and hence neutrino masses. The evidence comes from the atmospheric neutrino data in the SuperKamiokande experiment [10.4], which confirms the indications of oscillations in earlier data from the Kamiokande [10.5] and IMB [10.6] experiments. More recent data from the Soudan II [10.7] and MACRO [10.8] experiments provide further confirmation of this evidence.

From the existing data, several important conclusions can be drawn: (i) the data cannot be fitted assuming oscillation between ν_μ and ν_e; (ii) two oscillation scenarios that fit the data are ν_μ–ν_τ and ν_μ–ν_s oscillations (where ν_s is a sterile neutrino that does not couple to the W or Z bosons in the basic Lagrangian), although, at the 2σ level, the first scenario is a better fit than the latter. The more precise values of the oscillation parameters at a 90% confidence level are

$$\Delta m^2_{\nu_\mu\nu_\tau} \simeq (2\text{–}8) \times 10^{-3}\ \mathrm{eV}^2\ , \qquad \sin^2 2\theta_{\mu\tau} \simeq 0.8\text{–}1\ . \tag{10.1}$$

The second evidence for neutrino oscillation comes from the five experiments that have observed a deficit in the flux of neutrinos from the sun as

Pauli’s “Neutrino”

Dear Radioactive Ladies and Gentlemen: Zurich, December 4, 1930

I beg you to receive graciously the bearer of this letter who will report to you in detail how I have hit on a desperate way to escape from the problems of the "wrong" statistics of the N and Li6 nuclei and of the continuous beta spectrum in order to save the "even-odd” rule of statistics and the law of conservation of energy. Namely the possibility that electrically neutral particles, which I would like to call neutrons might exist inside nuclei; these would have spin 1/2, would obey the exclusion principle, and would in addition differ from photons through the fact that they would not travel at the speed of light. The mass of the neutron ought to be about the same order of magnitude as the electron mass, and in any case could not be greater than 0.01 proton masses. The continuous beta spectrum would then become understandable by assuming that in beta decay a neutron is always emitted along with the electron, in such a way that the sum of the energies of the neutron and electron is a constant. Now, the question is, what forces act on the neutron? The most likely model for the neutron seems to me, on wave mechanical grounds, to be the assumption that the motionless neutron is a magnetic dipole with a certain magnetic moment μ (the bearer of this letter can supply details). The experiments demand that the ionizing power of such a neutron cannot exceed that of a gamma ray, and therefore μ probably cannot be greater than e (10^{-13}cm). [e is the charge of the electron].

At the moment I do not dare to publish anything about this idea, so I first turn trustingly to you, dear radioactive friends, with the question: how could such a neutron be experimentally identified if it possessed about the same penetrating power as a gamma ray or perhaps 10 times greater penetrating power?

I admit that my way out may look rather improbable at first since if the neutron existed it would have been seen long ago. But nothing ventured, nothing gained. The gravity of the situation with the continuous beta spectrum was illuminated by a remark by my distinguished predecessor in office, Mr. DeBye, who recently said to me in Brussels, "Oh, that’s a problem like the new taxes; one had best not think about it at all." So one ought to discuss seriously any way that may lead to salvation. Well, dear radioactive friends, weigh it and pass sentence! Unfortunately, I cannot appear personally in Tubingen, for I cannot get away from Zurich on account of a ball, which is held.here on the night of December 6-7

With best regards to you and to Mr. Baek,

Your most obedient servant,

W. Pauli

Contribution to the Centennial Celebration in 2001 of the One-Hundredth Birthday of Enrico Fermi on September 29, 1901

(This paper was sent to the Rome conference in September 2001. Yang was not present and it was not part of the program. Part of it is presented here. See end of Chapter 18 for other comments by Yang.—J. O.)

Enrico Fermi was, of all the great physicists of the twentieth century, among the most respected and admired. He was respected and admired because of his contributions to both theoretical and experimental physics, because of his leadership in discovering for mankind a powerful new source of energy, and above all, because of his personal character: He was always reliable and trustworthy. He had both of his feet on the ground all the time. He had great strength but never threw his weight around. He did not play to the gallery. He did not practice one-upmanship. He exemplified, I always believe, the perfect Confucian gentleman.

Fermi's earliest interests in physics seem to be in general relativity. Starting from around 1923 he began to think deeply about the "Gibbs paradox" and the "absolute entropy constant" in statistical mechanics. This research led to his first monumental work and to the "Fermi distribution," "Fermi sphere," "Fermi liquid," "Fermi statistics," "Fermions," etc.

It was characteristic of Fermi's style in research that he should follow this abstract contribution with an application to the heavy atom, leading to what is now known as the Thomas-Fermi method. The differential equation involved in this method was solved by Fermi numerically with a small and primitive hand calculator. This numerical work took him probably a week. E. Majorana, who was a lightning-fast calculator and a very skeptical man, decided to check the numerical work. He did this by transforming the equation into a Riccato equation and solving the latter numerically. The result agreed exactly with the one obtained by Fermi. Fermi's love of the use of computers. small and large, which we graduate students at Chicago observed and admired, began evidently early in his career and lasted throughout his entire life.

Fermi's next major contribution was in quantum electrodynamics, where he succeeded in eliminating the longitudinal field to arrive at the Coulomb interaction. Fermi was very proud of this work as his students at the University of Chicago in the years 1946 to 1951 knew. (But it seems today that few theorists under the age of 65 know about this contribution of Fermi's.) It again was characteristic of Fermi's style that in this work he saw through complicated formalisms to arrive at the basics, in this case a collection of harmonic oscillators, and to proceed to solve a simple Schrodinger-like equation. The work was first presented in April 1929 in Paris and later at the famous Summer School at Ann Arbor in the summer of 1930. G. Uhlenbeck told me in the late 1950s that before this work of Fermi nobody really understood the quantum theory of radiation and that this work had established Fermi as among the few top field theorists in the world.

I shall skip describing his beautiful contribution in 1930 to the theory of hyperfine structure, and come to the theory of beta-decay. According to Segré, Fermi had considered, throughout his life, that this theory was his most important contribution to theoretical physics. I had read Segré's remarks in this regard, but was puzzled. One day in the 1970s, I had the following conversation with Eugene Wigner in the cafeteria of Rockefeller University:

Yang: What do you think was Fermi's most important contribution to theoretical physics?

Wigner: beta-decay theory.

Yang: How could that be? It is being replaced by more fundamental ideas. Of course it was a very important contribution which had sustained the whole field for some 40 years: Fermi had characteristically swept what was unknowable at that time under the rug and focused on what can be calculated. It was beautiful and agreed with experiment. But it was not permanent. In contrast, the Fermi distribution is permanent.

Wigner: No, no, you do not understand the impact it produced at the time. Von Neumann and I had been thinking about beta-decay for a long time as did everybody else. We simply did not know how to create an electron in a nucleus.

Yang: Fermi knew how to do that by using a second quantized psi?

Wigner: Yes.

Yang: But it was you and Jordan who had first invented the second quantized psi.

Wigner: Yes, yes. But we never dreamed that it could be used in real physics.

Nuclear Physics

A Course Given by ENRICO FERMI *at the University of Chicago. Notes Compiled by Jay Orear, A. H. Rosenfeld, and R. A. Schluter*

Revised Edition

THE UNIVERSITY OF CHICAGO PRESS

CHAPTER IV BETA DECAY

A. Introduction

β decay is the process in which an electron or a positron is emitted by a nucleus. The term is extended to include absorption of electrons.

The most remarkable feature of β decay phenomena is the apparent failure of energy conservation. In other nuclear processes, such as α decay, energy is clearly conserved. For example, if nucleus A decays to nucleus B, producing an α, the energy equation is $E_A=E_B+E_\alpha$. If the nuclei are in excited states before and after, the energy equation may be different: $E_A'=E_B'+E_\alpha'$, but always energy isconserved. In β emission, such an energy equation involving only the observed particles cannot be written. The reason is that the energies of β's from a single type of process have a continuous distribution of values. Empirically, the relative number of β's of a given energy, N(E), as a function of energy is a curve like FIG. IV.1. No β is emitted having energy greater than some value E_β^{max}.

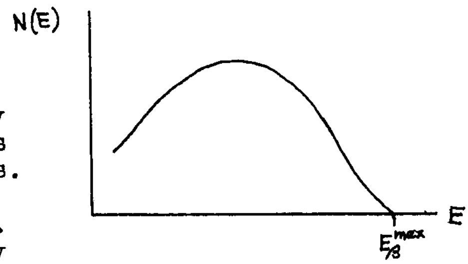


FIG. IV.1

A conceivable explanation that retains the energy conservation law is that the states of the final nucleus are very closely spaced, and the various β energies correspond to various final states. But we must account for the extra energy of the final excited state. The only conceivable mode of decay to the ground state is by gamma emission. Since for some cases of β emission there is no gamma radiation at all, and in any case no gamma radiation with a nearly continuous energy spectrum, the hypothesis of many final states of the nucleus must be discarded.

There is no alternative to admitting that the final and initial states of the nucleus are definite, but that the β may have any energy below E_β^{max}.

Experimentally, to within about 20 kev., $E_\beta^{max}=E_A-E_B$, showing that although energy may disappear, none is ever gained. The law of energy conservation might be forsaken, and replaced by: $E_{initial} \geq E_{final}$, a law that energy either is conserved or disappears, but never increases. Such a law would forbid perpetual motion machines of the first kind and accord with β decay phenomena.

The most favored explanation of this apparent non-conservation of energy is the neutrino hypothesis first suggested by Pauli. It postulates that an additional particle, the neutrino, (or perhaps more than one) is produced in β decay and carries away the missing energy. To accord with experiment, the neutrino, denoted by ν, must be made very difficult to detect. This is done by postulating that it is electrically neutral, (conservation of charge also imposes this), and of very small mass. Under the neutrino hypothesis, the energy balance equation is: $E_A-E_B=E_\beta+E_\nu$.

field of large gradient has not been observed.

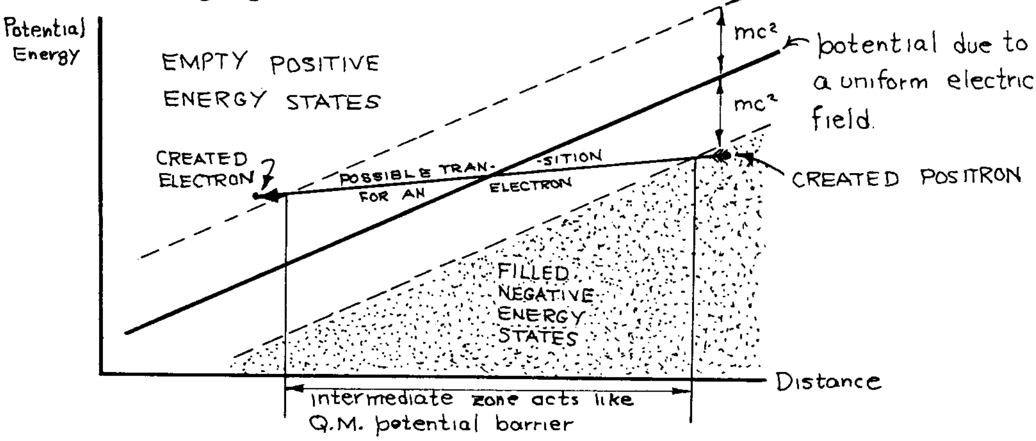


FIG. IV. 7

(b) Statistical arguments favor the proton-neutron theory of nuclear constitution. A nucleus having an odd number of elementary particles has Fermi statistics; a nucleus having an even number has Bose-Einstein statistics. The different hypotheses for the composition of nuclei lead to different numbers of particles. For the nucleus $_Z(\,)^A$,

Electrons-in-nucleus	Neutron hypothesis
A protons	Z protons
A-Z electrons	A-Z neutrons
2A-Z elem. particles	A elem. particles

For example, the $_7N^{14}$ nucleus has 14 particles under the neutron hypothesis, but 21 under the electron-in-nucleus hypothesis. Experiments in molecular spectroscopy* show that $_7N^{14}$ has Bose-Einstein statistics, therefore an even number of particles, confirming the neutron-in-nucleus hypothesis.

(c) The spin of the nucleus, whether integral or half-odd, depends on the number of elementary particles, and can be determined experimentally. The evidence again favors the neutron hypothesis.

Since the electron does not exist in the nucleus, it must be formed at the moment of its emission just as a photon is formed at the moment of its emission from an atom. The neutrino is created at the moment of emission, also. These particles are created into states represented by the wave-functions Ψ_β and Ψ_ν. Assume these are functions for plane waves with momenta $\underline{p}_\beta$ and $\underline{p}_\nu$, respectively,

$$\psi_\beta = N_\beta e^{i \underline{p}_\beta \cdot \underline{r} \hbar^{-1}}, \qquad \psi_\nu = N_\nu e^{i \underline{p}_\nu \cdot \underline{r} \hbar^{-1}} \qquad \text{IV.1}$$

where N is a normalization factor. Ψ_β is actually more complicated than given here because it is affected by the nuclear charge Z. The plane wave Ψ_β is a good approximation if the energy of the electron is much larger than Zx(Rydberg). For a low energy electron, say 200 Kev, near a nucleus of high Z, Ψ_β is strongly perturbed.

The probability of emission will be assumed to depend on the

* See Chapter I, sec. D.

expectation value for the electron and the neutrino to be at the nucleus, i.e., on the factor $|\Psi_\beta(0)|^2|\Psi_\nu(0)|^2$. It also depends on other factors, whose nature is uncertain.

One factor is the square modulus of a matrix element $\mathcal{M}$ taken between the initial and final states of the nucleus. This matrix element is analogous to the matrix element in the theory of emission of photons. In photon emission the matrix element is definitely known and has the form (for dipole radiation)

$$\int \psi^*_{final}\,(\text{electric moment})\,\psi_{initial}\,d\tau$$

$\mathcal{M}$, in β theory, is not known. There are several possible forms. For the case $N \rightarrow P$, the simplest is

$$\mathcal{M} = \int \psi_P^* \psi_N \, d\tau \qquad \text{IV.2}$$

assuming just one nucleon participates. Ψ_N represents the initial state of the nucleon, Ψ_P the final, i.e., proton, state of the nucleon. According to another form of the theory, $\mathcal{M}$ is a vector having x component

$$\mathcal{M}_x = \int \psi_P^* \sigma_x \psi_N \, d\tau \qquad \text{IV.3}$$

where σ_x is the x component of a (relativistic) spin operator. Then

$$|\mathcal{M}|^2 = |\mathcal{M}_x|^2 + |\mathcal{M}_y|^2 + |\mathcal{M}_z|^2 \qquad \text{IV.4}$$

The choice of $\mathcal{M}$ determines the selection rules, discussed in ~~section~~ section H.

The expression for the probability of emission includes also a constant factor g^2 which represents the strength of the coupling giving rise to emission, and is a universal constant. Experimentally,

$$g = 10^{-48} \text{ to } 10^{-49} \text{ g cm}^5 \text{ sec}^{-2} \qquad \text{IV.5}$$

Altogether, the probability of emission per unit time is*

$$\frac{2\pi}{\hbar}\left(|\psi_\beta(0)||\psi_\nu(0)||\mathcal{M}|\,g\right)^2 \frac{dn}{dE} \qquad \text{IV.6}$$

where dn/dE=energy density of final states, and 0 refers to the location of the nucleus. The Ψ functions are normalized in a volume Ω so that $\int_\Omega \Psi^*\Psi \, d\tau = 1$. Therefore $N = 1/\sqrt{\Omega}$ IV.7

and $$\psi_\beta = \frac{1}{\sqrt{\Omega}} e^{i\hbar^{-1} p_\beta \cdot r} \qquad \psi_\nu = \frac{1}{\sqrt{\Omega}} e^{i\hbar^{-1} p_\nu \cdot r} \qquad \text{IV.8}$$

It has meaning to say the nucleus is at $\underline{r} = 0$ only if Ψ changes little over the dimension of the nucleus. The rapidity of variation of Ψ is measured by $\lambda\!\!\!^{-} = \hbar/p \approx 10^{-11}$ cm, for a usual value of p. But the nucleus is about 10^{-12} cm in diameter. Therefore it is permissible to say that the nucleus is at $\underline{r} = 0$.

---

* This is analogous to the usual Q.M. formula for transition probability per unit time. This formula, "Golden Rule No. 2", is

prob. per second $= \frac{2\pi}{\hbar}|H_{21}|^2\frac{dn}{dE}$, where $H_{21} = \int \psi_2^* H' \psi_1 \, d\tau$.

Ψ_1 and Ψ_2 are the wave functions of the initial and final states, resp. This is derived, for example, in Schiff, Q.M., p. 193. It is discussed in more detail in Ch. VIII, sec. B.

For $\underline{r} = 0$

$$\psi_\beta(0) = \frac{1}{\sqrt{\Omega}} \qquad \psi_\nu(0) = \frac{1}{\sqrt{\Omega}} \tag{IV.9}$$

The number of plane wave states having magnitude of momentum between p and p + dp, with the particle anywhere in Ω, is*

$$\frac{p^2 dp\,\Omega}{2\pi^2\hbar^3} \tag{IV.10}$$

Therefore $$dn = \frac{p_\beta^2\,dp_\beta}{2\pi^2\hbar^3} \times \frac{p_\nu^2\,dp_\nu}{2\pi^2\hbar^3} \times \Omega^2 = \Omega^2\,\frac{p_\beta^2\,p_\nu^2}{4\pi^4\hbar^6}\,dp_\beta\,dp_\nu$$

$dp_\beta\,dp_\nu = J\,dp_\beta\,dE$ where J is the Jacobian. Using the relation $E = cp_\nu + E_\beta$, J is found to be 1/c.* (Mass of ν assumed zero for this derivation.)

Thus $$\frac{dn}{dE} = \Omega^2\,\frac{p_\beta^2\,p_\nu^2}{4\pi^4\hbar^6 c}\,dp_\beta$$

Using this to express IV.6, the probability of emission per unit time, $P(p_\nu,p_\beta)dp_\beta$, is

$$P(p_\nu,p_\beta)dp_\beta = \frac{2\pi}{\hbar}\left(\frac{1}{\Omega}|M|\,g\right)^2 \frac{\Omega^2\,p_\beta^2\,p_\nu^2\,dp_\beta}{4\pi^4\hbar^6 c} \tag{IV.14}$$

Using the relation $p_\nu c = E_\nu = E_\beta^{max} - E_\beta$ to eliminate p_ν, and writing p for p_β from now on,

$$P(p)dp = \frac{g^2|M|^2}{2\pi^3\hbar^7 c^3}\left(E_\beta^{max} - E_\beta\right)^2 p^2\,dp \tag{IV.15}$$

Using the equation $E_\beta^{max} = \sqrt{m^2c^4 + c^2p_{max}^2}$ to define p_{max}, (IV.16) we get:

$$\boxed{P(p)dp = \frac{g^2|M|^2}{2\pi^3\hbar^7 c^3}\left(\sqrt{m^2c^4 + c^2p_{max}^2} - \sqrt{m^2c^4 + p^2c^2}\right)^2 p^2\,dp} \tag{IV.17}$$

E. Rate of Decay

The lifetime τ is found by integrating over all possible p.

---

* For simplicity, assume Ω is a cubical box of side L. $\Omega = L^3$. That the particle is confined to the box means that the potential rises to ∞ at the sides. Schrödinger's equation within the box is

$$-\nabla^2 u = \frac{2mE}{\hbar^2}u = \left(p_x^2 + p_y^2 + p_z^2\right)\hbar^{-2}$$

Solutions satisfying the boundary condition u = 0 at x.y.z = 0, and u = 0 at x,y,z = L are $u = \sin\frac{p_x}{\hbar}x\,\sin\frac{p_y}{\hbar}y\,\sin\frac{p_z}{\hbar}z$ where $p_x/\hbar$ is restricted to the values $n_x\pi/L$, etc. The number of states, n′, representing momentum less than p equals the number of combinations of p_x, p_y, p_z such that $p_x^2+p_y^2+p_z^2 < p^2$, or, using the condition on the p's, $n_x^2+n_y^2+n_z^2 < p^2L^2/\pi^2\hbar^2$. $n_x, n_y, n_z > 0$ since − values give no new independent solutions. The number of sets of n_x, n_y, n_z satisfying the condition above equals 1/8 the number of points of a cubical lattice enclosed by a sphere of radius $\sqrt{\frac{p^2L^2}{\pi^2\hbar^2}}$. The 1/8 comes from restricting $n_x, n_y, n_z > 0$. Then the number of states, $n' = \frac{1}{8}\,\frac{4\pi}{3}\left(\frac{pL}{\pi\hbar}\right)^3$. (IV.18)

The number of states having momentum between p and p+dp is

$$\frac{dn'}{dp}\,dp = \Omega\,p^2dp/2\pi^2\hbar^3$$

---

* $$J = \begin{vmatrix} \frac{\partial p_\beta}{\partial p_\beta} & \frac{\partial p_\beta}{\partial E} \\ \frac{\partial p_\nu}{\partial p_\beta} & \frac{\partial p_\nu}{\partial E} \end{vmatrix} = \begin{vmatrix} 1 & \frac{\partial p_\beta}{\partial E} \\ 0 & \frac{1}{c} \end{vmatrix} = \frac{1}{c}$$

E. Fermi's publications on the Weak Interaction

E. Fermi, "Tentative Theory of Beta Rays"
Letter Submitted to Nature (1933)

REJECTED

ANNO IV - VOL. II - N. 12 QUINDICINALE 31 DICEMBRE 1933 - XII

LA RICERCA SCIENTIFICA

ED IL PROGRESSO TECNICO NELL'ECONOMIA NAZIONALE

Tentativo di una teoria dell'emissione dei raggi "beta"

Nota del prof. ENRICO FERMI

Riassunto: Teoria della emissione dei raggi β delle sostanze radioattive, fondata sull'ipotesi che gli elettroni emessi dai nuclei non esistano prima della disintegrazione ma vengano formati, insieme ad un neutrino, in modo analogo alla formazione di un quanto di luce che accompagna un salto quantico di un atomo. Confronto della teoria con l'esperienza.

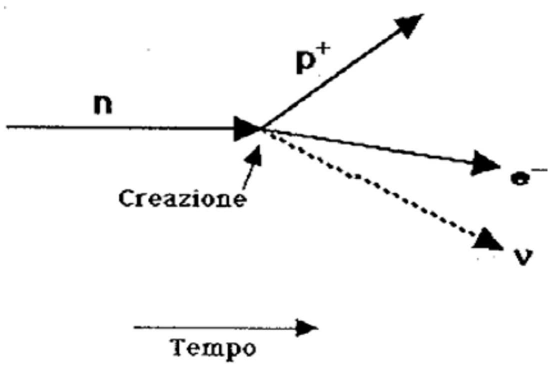
Published in Nuovo Cimento and Zeitschrift fur Physik

Fermi's paper on beta decay:

- Established a predictive realization of Pauli's proposal
- Established the connection between quantum field theory and particles.
- Predicted the statistical shape of the beta spectrum and the consequences of finite neutrino mass.
- Anticipated the most likely experimental distortions to beta spectrum.
- Discussed the dominate electromagnetic corrections to the beta decay spectrum.
- Established a theory that remains the (essentially) correct description of beta decay.

Fermi's theory remains the "correct" description of beta decay except:

- As pointed out by Gamow and Teller in 1936 another component of the Hamiltonian is required to account for decays like ^{6}He
- Neutrons and protons are not elementary particle and there are forbidden contributions (induced terms) due to their structure



The fundamental process

•In analogy with the theory of radiation Fermi applied the creation and distruction operators of Dirac-Jordan-Klein-Wigner and Dirac's relativisitic theory for spin 1/2 particles

•Of the possibilities in Dirac invariant interactions (S, V, A, T, P) Fermi chose a vector interaction for the nucleon current and the lepton current.

$$H(x) = g\,[p^{+}(x)\,\gamma^{o}\gamma^{\mu}n(x)]\;\;[e^{+}(x)\,\gamma_{o}\gamma_{\mu}\nu(x)]$$

Three Nobel Prizes for Neutrinos

Frederick Reines (1995)
Neutrinos from Reactors

Raymond Davis (2002)
Neutrinos from the Sun

Masatoshi Koshiba (2002)
Neutrinos from the Stars

Dear Enrico, October 4, 1952

We thought that you might be interested in the latest version of our experiment to detect the free neutrino, hence this letter. as you recall, we planned to use a nuclear explosion for the source because of the background difficulties. Only last week it occurred to us that background problems could be reduced to the point where a Hanford pile would suffice by counting only delayed coincidences between the positron pulse and neutron capture pulse. You will remember that the reaction we plan to use is $p + \nu \rightarrow n + \beta^+$. Boron loading a liquid scintillator makes it possible to adjust the mean time T between these two events and we are considering $T \sim 10\ \mu$sec. Our detector is a 10 cubic foot fluor filled cylinder surrounded by about 90 5819's operating as two large tubes of 45 5819's each. These two banks of ganged tubes isotropically distributed about the curved cylindrical wall are in coincidence to cut tube noise. The inner wall of the chamber will be coated with a diffuse reflector and in all we expect the system to be energy sensitive, and not particularly sensitive to the position of the event in the fluor. This energy sensitivity will be used to discriminate further against background. Cosmic ray anti-coincidence will be used in addition to mercury of low background lead for shielding against natural radioactivity. We plan to immerse the entire detector in a large borax water solution for further necessary reduction of pile background below that provided by the Hanford shield.

Fortunately, the fast reactor here at Los Alamos provides the same leakage flux as Hanford so that we can check our gear before going to Hanford. Further, if we allow enough fast neutrons from the fast reactor to leak into our detector we can simulate double pulses because of the proton recoil pulse followed by the neutron capture which occurs in this case. We expect a counting rate at Hanford in our detector about six feet from the pile face of ~1/min with a background somewhat lower than this.

As you can imagine, we are quite excited about the whole business, have canceled preparations for use of a bomb, and we are working like mad to carry out the ideas sketched above. Because of the enormous simplification in the experiment. We have already made rapid progress with the electronic gear and associated equipment and expect that tin the next few months we shall be at Hanford reaching for the slippery particle.

We would of course appreciate any comments you might care to make.

Sincerely,

Fred Reines, Clyde Cowan

Dear Fred, October 8, 1952

Thank you for your letter of October 4th by Clyde Cowan and yourself. I was very much interested in you new plan for the detection of the neutrino. Certainly your new method should be much simpler to carry out and have the great advantage that the measurement can be repeated any number of times. I shall be very interested seeing how your 10 cubic foot scintillaton counter is going to work, but I do not know of any reason why it should not.

Good Luck.

Sincerely yours,

Enrico Fermi

“I shall be very interested seeing how your 40,624 cubic foot scintillaton counter is going to work, but I do not know of any reason why it should not.”

The Neutrino: From Poltergeist to Particle

Nobel Lecture, December 8, 1995
Frederick Reines

The Second World War had a great influence on the lives and careers of very many of us for whom those were formative years. I was involved during, and then subsequent to, the war in the testing of nuclear bombs, and several of us wondered whether this man-made star could be used to advance our knowledge of physics. For one thing this unusual object certainly had lots of fissions in it, and hence, was a very intense neutrino source. I mulled this over somewhat but took no action.

Then in 1951, following the tests at Eniwetok Atoll in the Pacific, I decided I really would like to do some fundamental physics. Accordingly, I approached my boss, Los Alamos Theoretical Division Leader, J. Carson Mark, and asked him for a leave in residence so that I could ponder. He agreed, and I moved to a stark empty office, staring at a blank pad for several months searching for a meaningful question worthy of a life's work. It was a very difficult time. The months passed and all I could dredge up out of the subconscious was the possible utility of a bomb for the direct detection of neutrinos. Afterall, such a device produced an extraordinarily intense pulse of neutrinos and thus the signals produced by neutrinos might be distinguishable from background. Some handwaving and rough calculations led me to conclude that the bomb was the best source. All that was needed was a detector measuring a cubic meter or so. I thought, well, I must check this with a real expert.

It happened during the summer of 1951 that Enrico Fermi was at Los Alamos, and so I went down the hall, knocked timidly on the door and said, "I'd like to talk to you a few minutes about the possibility of neutrino detection." He was very pleasant, and said, "Well, tell me what's on your mind?" I said, "First off as to the source, I think that the bomb is best." After a moment's thought he said, "Yes, the bomb is the best source." So far, so good! Then I said, "But one needs a detector which is so big. I don't know how to make such a detector." He thought about it some and said he didn't either. Coming from the Master that was very crushing. I put it on the back burner until a chance conversation with Clyde Cowan. We were on our way to Princeton to talk with Lyman Spitzer about controlled fusion when the airplane was grounded in Kansas City because of engine trouble. At loose ends we wandered around the place, and started to discuss what to do that's interesting in physics. "Let's do a real challenging problem," I said. He said, "Let's work on positronium." I said, "No, positronium is a very good thing but Martin Deutsch has that sewed-up. So let's not work on positronium." Then I said, "Clyde let's work on the neutrino." His immediate response was, “GREAT IDEA.” He knew as little about the neutrino as I did but he was a good experimentalist with a sense of derring do. So we shook hands and got off to working on neutrinos.

Need for Direct Detection

Before continuing with this narrative it is perhaps appropriate to recall the evidence for the existence of the neutrino at the time Clyde and I started on our quest. The neutrino of Wolfgang Pauli[l] was postulated in order to account for an apparent loss of energy-momentum in the process of nuclear beta decay. In his famous 1930 letter to the Tübingen congress, he stated: “I admit that my expedient may seem rather improbable from the first, because if neutrons* existed they would have been discovered long since.
*When the neutron was discovered by Chadwick, Fermi renamed Pauli’s particle the “neutrino”.
Nevertheless, nothing ventured nothing gained... We should therefore be seriously discussing every path to salvation.”

All the evidence up to 1951 was obtained “at the scene of the crime” so to speak since the neutrino once produced was not observed to interact further. No less an authority than Niels Bohr pointed out in 1930[2] that no evidence “either empirical or theoretical” existed that supported the conservation of energy in this case. He was, in fact, willing to entertain the possibility that energy conservation must be abandoned in the nuclear realm. However attractive the neutrino was as an explanation for beta decay, the proof of its existence had to be derived from an observation at a location other than that at which the decay process occurred - the neutrino had to be observed in its free state to interact with matter at a remote point.
It must be recognized, however, that, independently of the observation of a free neutrino interaction with matter, the theory was so attractive in its explanation of beta decay that belief in the neutrino as a “real” entity was general. Despite this widespread belief, the free neutrino’s apparent undetectability led it to be described as “elusive, a poltergeist.”

So why did we want to detect the free neutrino? Because everybody said, you couldn’t do it. Not very sensible, but we were attracted by the challenge. After all, we had a bomb which constituted an excellent intense neutrino source. So, maybe we had an edge on others. Well, once again being brash, but nevertheless having a certain respect for certain authorities, I commented in this vein to Fermi, who agreed. A formal way to make some of these comments is to say that, if you demonstrate the existence of the neutrino in the free state, i.e. by an observation at a remote location, you extend the range of applicability of these fundamental conservation laws to the nuclear realm. On the other hand, if you didn't see this particle in the predicted range then you have a very real problem.

As Bohr is reputed to have said, "A deep question is one where either a yes or no answer is interesting." So I guess this question of the existence of the "free" neutrino might be construed to be deep. Alright, what about the problem of detection? We fumbled around a great deal before we got to it. Finally, we chose to look for the reaction Te + p + n + e'. If the free neutrino exists, this inverse beta decay reaction has to be there, as Hans Bethe and Rudolf Peierls recognized, and as I'm sure did Fermi, but they had no occasion to write it down in the early days. Further, it was not known at the time whether *V,* and *V,* were different. We chose to consider this reaction because if you believe in what we today call "crossing symmetry" and use the measured value of the neutron half life then you know what the cross section has to be - a nice clean result. (In fact, as we learned some years later from Lee and Yang, the cross section is a factor of two greater because of parity nonconservation and the handedness of the neutrino.) Well, we set about to assess the problem of neutrino detection. How big a detector is required? How many counts do we expect? What features of the interaction do we use for signals? Bethe and Peierls in 1934 [3], almost immediately after the Fermi paper on beta decay[4], estimated that if you are in the few MeV range the cross section with which you have to deal would be ~ l0-44 cm2. To appreciate how minuscule this interaction is we note that the mean free path is ~ 1000 light years of liquid hydrogen. Pauli put his concern succinctly during a visit to Caltech when he remarked: "I have done a terrible thing. I have postulated a particle that cannot be detected." No wonder that Bethe and Peierls concluded in 1934 "there is no practically possible way of observing the neutrino." I confronted Bethe with this pronouncement some 20 years later and with his characteristic good humor he said, "Well, you shouldn't believe everything you read in the papers."

Reflecting on the trail that took us from bomb to reactor, it is evident that it was our persistence which led us from a virtually impossible experiment to one that showed considerable promise. The stage had been set for the detection of neutrinos by the discovery of fission and organic scintillators - the most important barrier was the generally held belief that the neutrino was undetectable.

Absorption Test

The only known particles, other than ie produced by the fission process, were discriminated against by means of a gamma-ray and neutron shield. When a bulk shield measured to attenuate gamma rays and neutrons by at least an order of magnitude was added, the signal was observed to remain constant; that is the reactor-associated signal was 1.74 ± 0.12/hour with, and 1.69 ± 0.17/hour without the shield.

Telegram to Pauli

The tests were completed and we were convinced. It was a glorious feeling to have participated so intimately in learning a new thing, and in June of 1956 we thought it was time to tell the man who had started it all when, as a young fellow, he wrote his famous letter in which he postulated the neutrino, saying something to the effect that he couldn't come to a meeting and tell them about it in person because he had to go out to a dance! The message was forwarded to him at CERN, where he interrupted the meeting he was attending to read the telegram to the conferees and then made some impromptu remarks regarding the discovery. That message reads, "We are happy to inform you that we have definitely detected neutrinos from fission fragments by observing inverse beta decay of protons. Observed cross section agrees well with expected six times ten to minus forty four square centimeters." We learned later that Pauli and some friends consumed a case of champagne in celebration! Many years later (~ 1986) C.P. Enz, a student of Pauli's, sent us a copy of a night letter Pauli wrote us in 1956, but which never arrived. It says, "Thanks for the message. Everything comes to him who knows how to wait. Pauli"

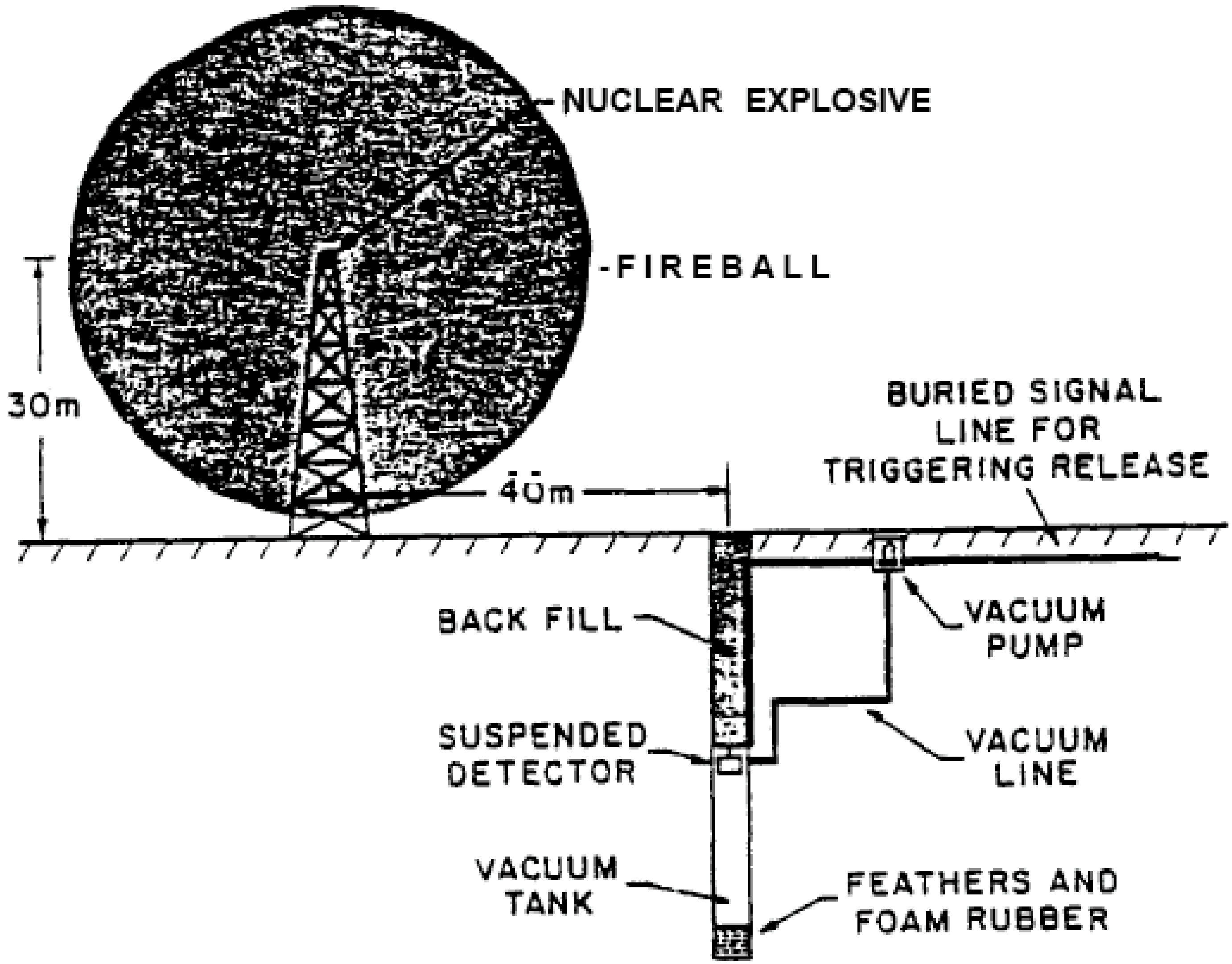


Figure 1. Sketch of the originally proposed experimental setup to detect the neutrino using a nuclear bomb. This experiment was approved by the authorities at Los Alamos but was superceded by the approach which used a fission reactor.

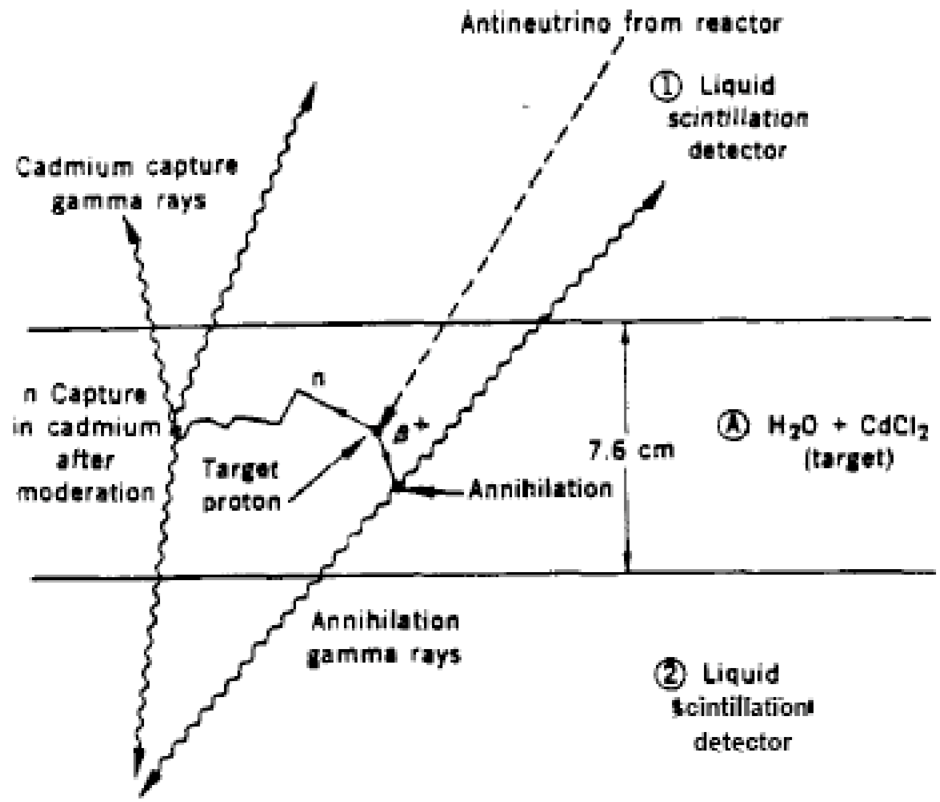
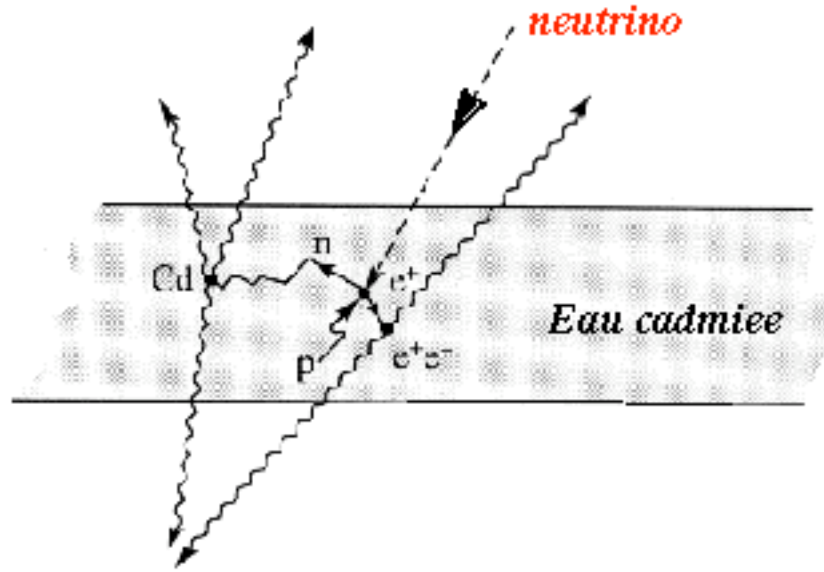
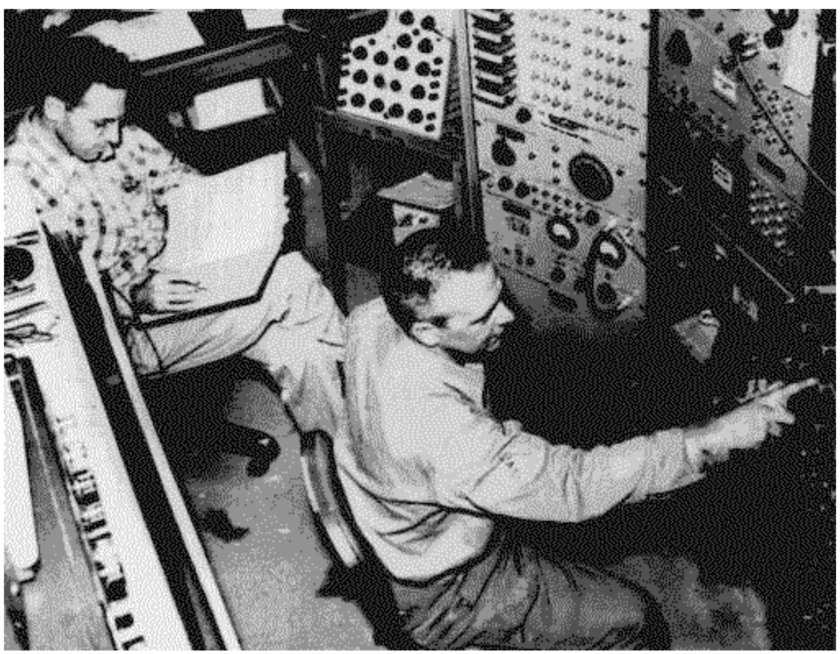
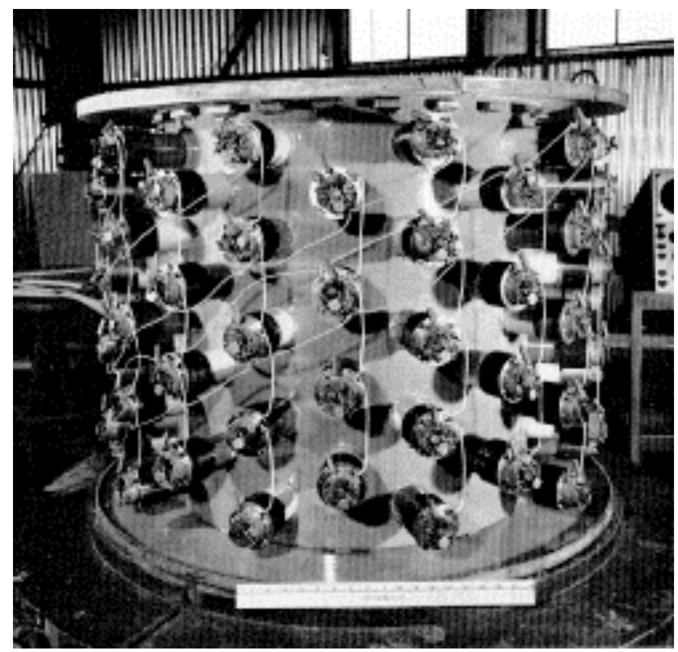


Figure 4. Schematic of the detection scheme used in the Savannah River experiment. An antineutrino from the reactor interacts with a proton in the target, creating a positron and a neutron. The positron annihilates on an electron in the target and creates two gamma rays which are detected by the liquid scintillators. The neutron slows down (in about 10 microseconds) and is captured by a cadmium nucleus in the target; the resulting gamma rays are detected in the liquid scintillators.

Direct Detection
of the Neutrino







Neutrino Detection

Neutrinos are elusive. A low energy neutrino has some chance of passing through 1000 light-years of lead without interacting!

Cosmic Gall

-John Updike-

Neutrinos, they are very small.
They have no charge and have no mass
And do not interact at all.
The earth is just a silly ball
To them, through which they simply pass,
Like dustmaids through a drafty hall
Or photons through a sheet of glass.
They snub the most exquisite gas,
Ignore the most substantial wall,
Cold-shoulder steel and sounding brass,
Insult the stallion in his stall,
And scorning barriers of class,
Infiltrate you and me! Like tall
And painless guillotines, they fall
Down through our heads into the grass.
At night, they enter at Nepal
And pierce the lover and his lass
From underneath the bed-you call
It wonderful; I call it crass.

The New Yorker Magazine, Inc. , 1960

YOU are now being invaded by about 10^{14} neutrinos each second!

A HALF-CENTURY WITH SOLAR NEUTRINOS

Nobel Lecture, December 8, 2002

by

RAYMOND DAVIS, JR.

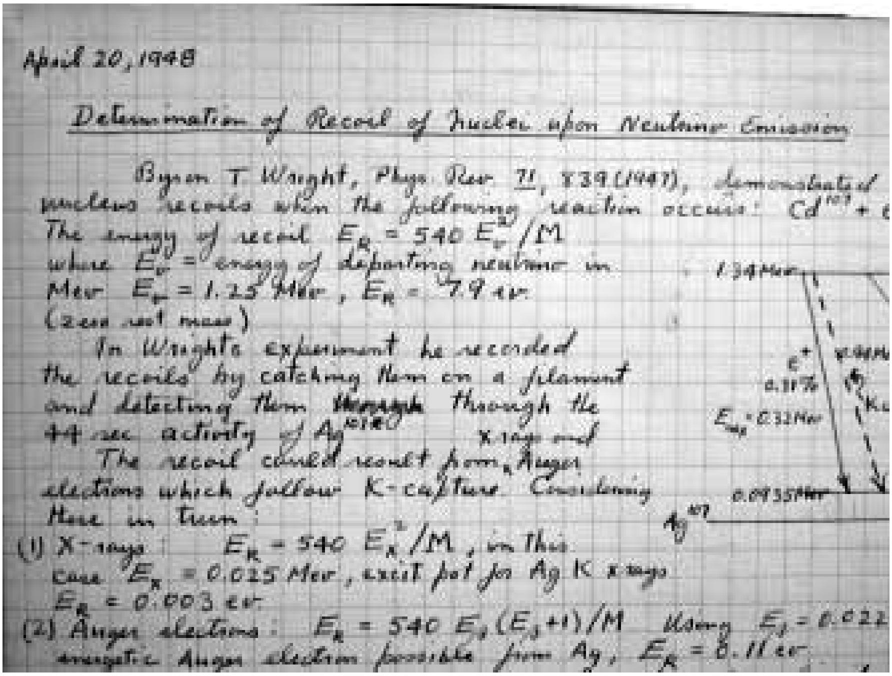
Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104 and Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973, USA.

Neutrinos are neutral, nearly massless particles that move at nearly the speed of light and easily pass through matter. Wolfgang Pauli (1945 Nobel Laureate in Physics) postulated the existence of the neutrino in 1930 as a way of carrying away missing energy, momentum, and spin in beta decay. In 1933, Enrico Fermi (1938 Nobel Laureate in Physics) named the neutrino ("little neutral one" in Italian) and incorporated it into his theory of beta decay.

The Sun derives its energy from fusion reactions in which hydrogen is transformed into helium. Every time four protons are turned into a helium nucleus, two neutrinos are produced. These neutrinos take only two seconds to reach the surface of the Sun and another eight minutes or so to reach the Earth. Thus, neutrinos tell us what happened in the center of the Sun eight minutes ago. The Sun produces a lot of neutrinos, 1.8×10^{39} per second: even at the Earth, 150 million kilometers from the Sun, about 100 billion pass through an average fingernail (1 cm^2) every second. They pass through the Earth as if it weren't there and the atoms in the human body capture a neutrino about every seventy years, or once in a lifetime. As we will see, neutrinos captured me early in my career.

I received my Ph.D. from Yale in 1942 in physical chemistry (Davis, 1942) and went directly into the Army as a reserve officer. After the war, I decided to search for a position in research with the view of applying chemistry to studies in nuclear physics. After two years with the Monsanto Chemical Company in applied radiochemistry of interest to the Atomic Energy Commission, I was very fortunate in being able to join the newly created Brookhaven National Laboratory. Brookhaven was created to find peaceful uses for the atom in all fields of basic science: chemistry, physics, biology, medicine, and engineering.

When I joined the Chemistry Department at Brookhaven, I asked the chairman, Richard Dodson, what he wanted me to do. To my surprise and delight, he told me to go to the library and find something interesting to work on. I found a stimulating review on neutrinos (Crane, 1948). This quote from Crane shows that neutrino physics was a field that was wide open to exploration: "Not everyone would be willing to say that he believes in the existence of the neutrino, but it is safe to say that there is hardly one of us who is not


April 20, 1948

Determination of Recoil of Nuclei upon Neutrino Emission

Byron T. Wright, Phys. Rev. 71, 839 (1947), demonstrated nucleus recoils when the following reaction occurs: Cd^{107} + e
The energy of recoil $E_R = 540\ E_\nu^2/M$
where E_ν = energy of departing neutrino in Mev $E_\nu = 1.25$ Mev, E_R = 7.9 ev.
(zero rest mass)
In Wrights experiment he recorded the recoils by catching them on a filament and detecting them ~~through~~ through the 44 sec activity of Ag^{107}
The recoil could result from x rays and Auger electrons which follow K-capture. Considering these in turn:
(1) X-rays: $E_R = 540\ E_x^2/M$, in this case E_x = 0.025 Mev, excit pot for Ag K x rays
E_R = 0.003 ev
(2) Auger electrons: $E_R = 540\ E_a(E_a+1)/M$ Using E_a = 0.022 energetic Auger electron possible from Ag, E_R = 0.11 ev

Figure 1. The first page in my first laboratory notebook at Brookhaven National Laboratory. I was hooked on neutrinos from the beginning.

served by the neutrino hypothesis as an aid in thinking about the beta-decay hypothesis". Neutrinos also turned out to be suitable for applying my background in physical chemistry. Crane had quite an extensive discussion on the use of recoil experiments to study neutrinos. I immediately became interested in such experiments (Fig. 1). I spent the first year working on the recoil of ^{107}Ag from the electron-capture decay of ^{107}Cd, but these experiments were inconclusive.

My first successful experiment was a study of the recoil energy of a ^{7}Li nucleus resulting from the electron-capture decay of ^{7}Be. In ^{7}Be decay, a single monoenergetic neutrino is emitted with an energy of 0.862 MeV, and the resulting ^{7}Li nucleus should recoil with a characteristic energy of 57 eV. A measurement of this process provides evidence for the existence of the neutrino. In my experiment, the energy spectrum of a recoiling ^{7}Li ion from a surface deposit of ^{7}Be was measured and found to agree with that expected from the emission of a single neutrino (Davis, 1952). This was a very nice result, but I was scooped by a group from the University of Illinois (Smith and Allen, 1951).

In 1951, I began working on a radiochemical experiment for detecting neutrinos using a method that was suggested by Pontecorvo (1946): capturing neutrinos with the reaction: $^{37}Cl + \nu_e \rightarrow {}^{37}Ar + e^-$. Bruno Pontecorvo's short paper was quite detailed, and the method he described, removing

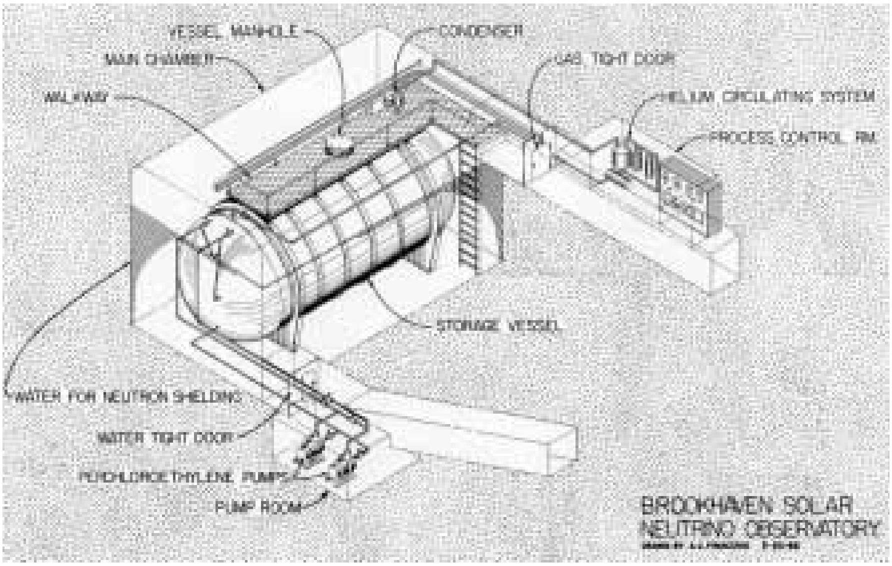


Figure 9. A drawing of the chlorine experiment. From *Sharp Bits,* Spring 1969.

on the enclosed sheet. As you see, there is no visible peak at Ar^{37} (3.0 keV). The background for this counter run just before the sample is shown also on the enclosed sheet. Comparing these we can obtain the following results:

$$\begin{array}{rl} \text{Argon from } 10^5 \text{ gal tank} & = 16 \pm 4 \text{ counts (tot. 39.7 d)} \\ \text{Background} = 4 \times (39.7/11.5) & = 14 \pm 4 \text{ counts (for 39.7 d)} \\ \text{Increase} & = 2 \pm 5 \text{ counts} \end{array}$$

$$\text{Using: } 2.1 \times 10^{30}\ Cl^{37} \text{ atoms in tank}$$

$$\text{Counter efficiency} \approx 0.50$$

$$\begin{array}{rl} \text{Then, } \Sigma\phi\sigma & = (0.2 \pm 0.4) \times 10^{-35} \text{ sec}^{-1} \\ & \leq 0.6 \times 10^{-35} \text{ sec}^{-1} \\ \text{Using } \phi(B^8) & = 1.35 \times 10^{-42} \text{ (Bahcall)} \\ \phi B^8 & \leq 0.5 \times 10^7 \text{ cm}^{-2} \text{ sec}^{-1} \end{array}$$

This limit is quite low, but according to the latest opus from Bahcall and Shaviv the B^8 flux is $1.4\ (1 \pm 0.6) \times 10^7$ cm^{-2} sec^{-1}. I hope to improve these results by improving the counter background, statistics, and longer irradiations.

Please regard these results as very preliminary. There are several points that must be checked before we are certain this is a bonafide observation. I will collect another sample in September–we are ready now, turn on the sun.

I have hopes of showing you the apparatus sometime in the future. The scenery is not be be compared with the English countryside, but it has its attractions.

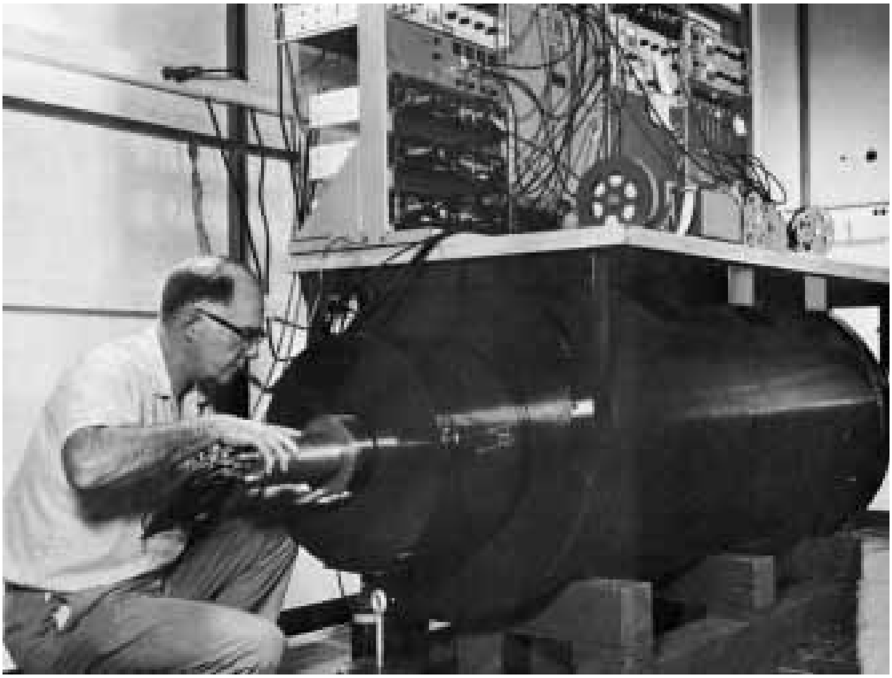


Figure 13. In the first few years of the experiment, the counters were placed in sections of pre-bomb battleship gun barrels for shielding. I am shown loading a counter into the barrel. From *Sharp Bits,* Spring 1969.

sured. These tests and more, as well as the standard operating procedure for the experiment, are described in Cleveland *et al.* (1998).

The solar neutrino problem lasted from 1967–2001. Over this period neither the measured flux nor the predicted flux changed significantly. I never found anything wrong with my experiment. John Bahcall never found anything wrong with the standard solar model, in fact, the advent of helioseismology confirmed the temperature profile in his model. The discrepancy between theory and experiment was a robust factor of three. Cleveland *et al.* (1998) summarized all of the data from the Homestake experiment. One hundred and eight runs were made after rise-time counting was implemented (Fig. 15). Over a period of 25 years, we counted a total of 2200 ^{37}Ar atoms and obtained a solar neutrino flux of 2.56 ± 0.16 (statistical error) $\pm$ 0.16 (systematic error) SNU. The current prediction from the standard solar model (Bahcall *et al.*, 2001) is $7.6^{+1.3}_{-1.1}$ SNU.

The results from the Homestake experiment provoked a great deal of activity among theorists. Here are some of the more interesting and, in retrospect amusing, alternatives to the standard solar model. Fowler (1968, 1972) and Sheldon (1969) suggested that there was a secular instability in energy production in the center of the Sun. Since light takes about 10 million years to reach the surface of the Sun, while neutrinos sample the core eigth minutes ago, the energy production could be low at the present time. Neutrino

BIRTH OF NEUTRINO ASTROPHYSICS

Nobel Lecture, December 8, 2002

by

MASATOSHI KOSHIBA

International Center for Elementary Particle Physics, University of Tokyo, 7-3-1-Hongo, Bunkjo-Ku, Tokyo 113-0033, Japan.

In giving this talk I am very much helped by the preceding talk because I can skip some of the topics. If you want further information, please refer to my review article, "Observational Neutrino Astrophysics," [1].

I am to talk about the birth of the neutrino astrophysics, but before the birth, there was a very important event, which was just described by Prof. Davis. [2]. It was the radiochemical work using the reaction $\nu_e + {}^{37}Cl$ going to $e + {}^{37}Ar$. He found that the observed neutrino flux was only 1/3 of the theoretically expected. This could be considered as the conception of the neutrino astrophysics and was in fact the impetus for us to begin seriously working on the solar neutrinos.

I will talk about two experiments. The first is the original KamiokaNDE, which might be called an Imaging Water Cerenkov detector with a surface coverage of 20% by photomultipliers and the total mass of the water inside this detector is 3,000 tons. It costed about 3 million U.S. dollars. This, mind you, was meant to be the feasibility experiment on the astrophysical detection of solar neutrinos. The second experiment is called Super-KamiokaNDE, the same type of detector but with a better light sensitivity, that is, 40% of the entire surface was covered by the photocathode and the total mass of the water was 50,000 tons. It costed about 100 million U.S. dollars. This was considered to be the full-scale solar neutrino observatory.

Both the experiments are situated about 1,000 meters underground in Kamioka Mine. The capital letters NDE at the end of the two experiments originally implied "Nucleon Decay Experiment." However, because of our detection of various neutrinos by these detectors, people started calling it, "Neutrino Detection Experiment".

Fig.1 shows the interior of KamiokaNDE. You can see arrays of photomultipliers on the sidewalls as well as on the top and at the bottom. When we were preparing for this KamiokaNDE experiment, we heard that a much bigger, experiment but of similar type, was being planned in the United States. [3]. We had to think very seriously about the competition with this bigger detector. Both experiments aimed at the detection of a certain type of proton decay, i.e., $e^+ + \pi^0$ mode. If we were aiming only for the detection of such particular types of proton decays, certainly the much bigger U.S. experiments would

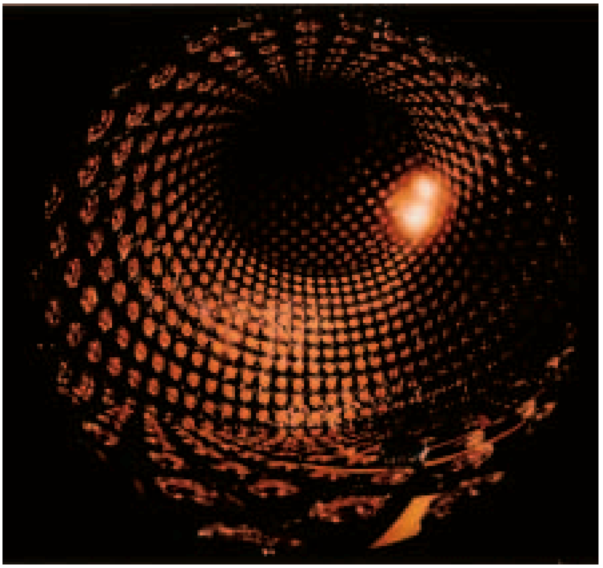


Figure 1. The interior of KamiokaNDE.

Figure 2. The newly developed large photomultiplier.

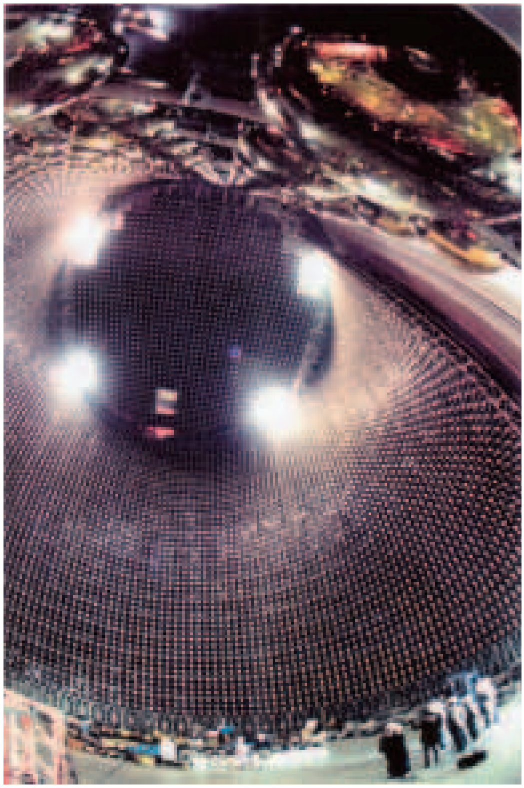


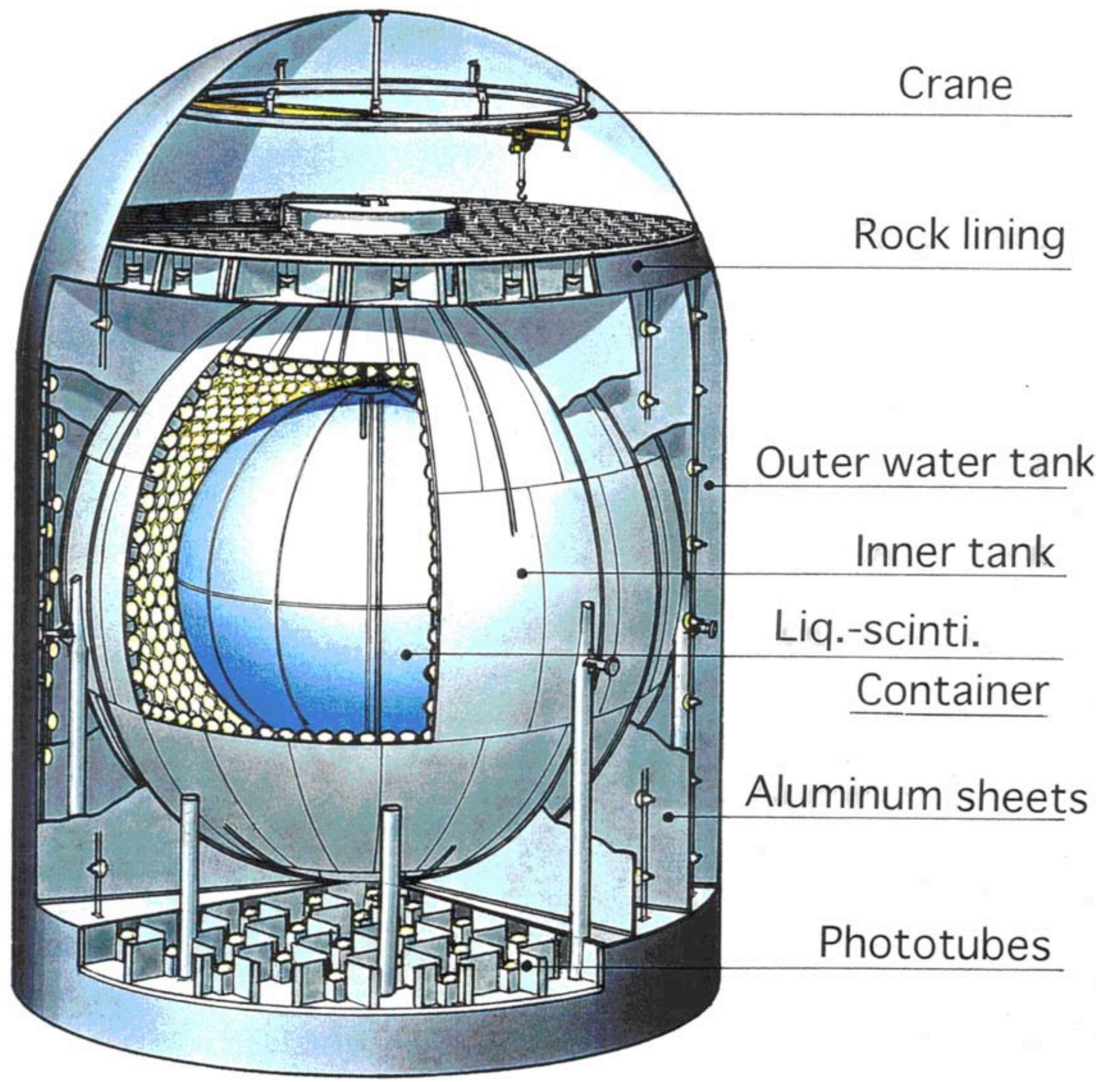
Figure 3. The interior of S-K through fish-eye lens.

find it first. Then, what could we do with a smaller detector? We thought very seriously about this competition and we came to the conclusion that the only possible way to compete with this bigger detector would be to make our detector much more sensitive than the U.S. competitors so that we could not only detect the easiest proton decay mode, but also measure other types of proton decays. Then eventually we could say that the proton decays into this mode with this branching ratio and into that mode with that branching ratio and so forth. Then our experiment would be able to point the way to the possible future, what is called the Grand Unified Theory, which is a new type of theory combining strong forces, weak forces, and electromagnetic forces.

Thanks to the cooperation of Hamamatsu Photonics Co., we jointly developed very large photomultiplier tubes [4]. I was so happy, as you can see in Fig. 2 that this tube was successfully developed.

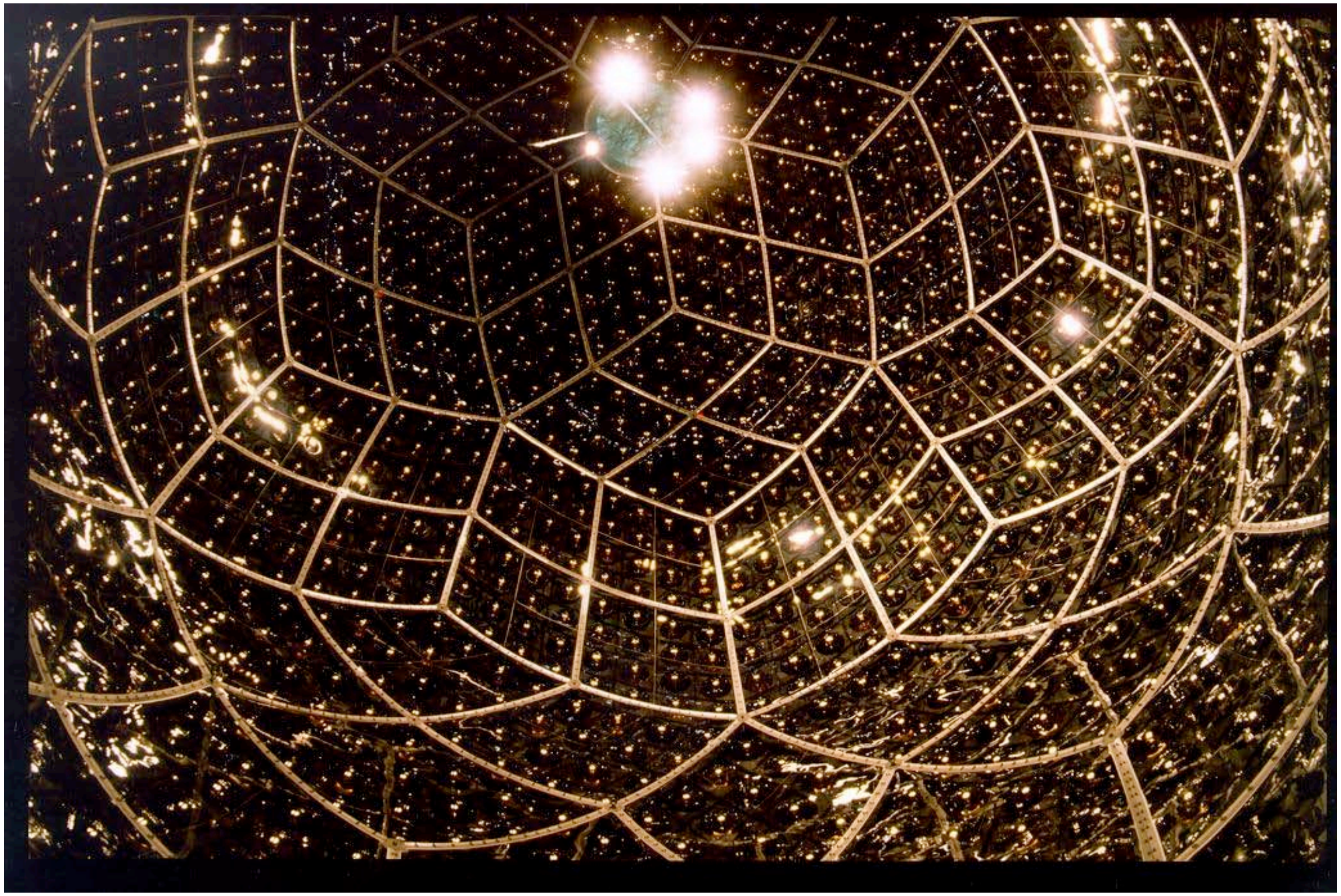
Fig. 3 shows the fish-eye view of the Super-KamiokaNDE interior. You can see many more phototubes, a total of about 11,000 big phototubes.

Since I suppose that not many people are familiar with this type of detector, I want to show you the performance of Super-KamiokaNDE. The first example is a very slow motion picture of a cosmic ray muon passing through the detector.

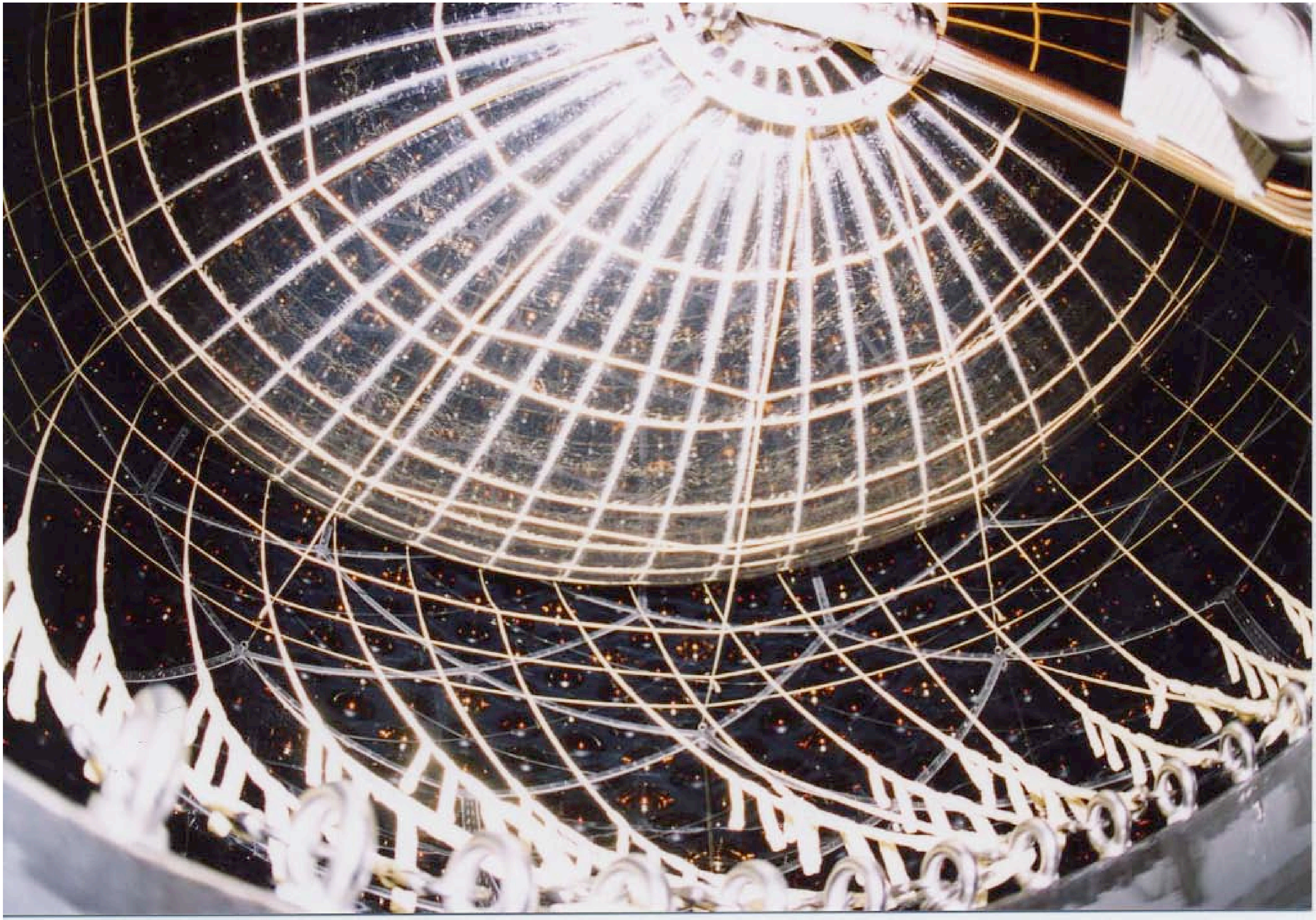


Cutaway view of the KamLAND detector

Exterior view of KamLAND sphere



Interior of KamLAND sphere October 2000



KamLAND Detector Ready for Fill May 2001