MMPP—NPC—1962—1

MICHIGAN MEMORIAL PHOENIX PROJECT
THE UNIVERSITY OF MICHIGAN

LECTURES FROM THE SECOND NEUTRON PHYSICS CONFERENCE

Boyne Mountain, Michigan

June, 1962

ANN ARBOR MICHIGAN

MMPP-NPC-1962-1

LECTURES FROM THE SECOND NEUTRON PHYSICS CONFERENCE

Boyne Mountain, Michigan

June, 1962

Sponsored By

The National Science Foundation

and

The Michigan Memorial-Phoenix Project

INTRODUCTION

These notes have been prepared from lectures delivered by A. Simon, A. C. Kolb, R. Zelazny, N. Corngold, L. L. Foldy, and S. N. Gupta at the Second Neutron Physics Conference at Boyne Mountain, Michigan, June 11 to 16, 1962.

The lecture notes were taken by the following persons:

Simon-------Richard Nicholson

Kolb--------Ziyaeddin Akcasu

Zelazny-----Ralph Nossal

Corngold----Sanford Cohen

Foldy-------Ronald Hayborn

Gupta-------George Summerfield

The Conference was sponsored by the National Science Foundation and the Michigan Memorial-Phoenix Project. It was organized by members of The University of Michigan Faculty. Chairman of the Conference was Kenneth Case of the Physics Department, Program Chairman was Paul Zweifel of the Department of Nuclear Engineering, Secretary was John King of the Department of Nuclear Engineering.

Table of Contents

PLASMA PHYSICS AND THERMONUCLEAR THEORY

by

Albert Simon

General Atomic

INTRODUCTION

The emphasis in this series of four lectures is upon plasma physics rather than on thermonuclear reactors. The first lecture does treat thermonuclear reactors but only from a basic point of view, i.e. the discussion is limited to the factors that are basic requirements for a successful controlled fusion reactor and little time is devoted to the question of economics, efficiency, size, etc. The era of optimism about producing an economic thermonuclear reactor in the immediate future has passed. The luxury of those kinds of engineering fantasies has vanished in the light of the harsh realities which we have encountered in trying simply to create a small, but hot, dense plasma which exists for reasonable lifetimes. Thus, most of the discussion is concerned with basic experiments and theories that are aimed at clarification of the nature of the difficulties that now limit us, particularly the instabilities which cause the plasma to escape containment in very short times.

I. Basic Requirements for a Controlled Fusion Reactor

The fusion reactions that can be considered in a controlled fusion reactor are the deuteron-deuteron (D-D) reaction and the deuteron-triton (D-T) reaction. Figure 1 shows the cross section as a function of energy for those two reactions.

One can see in Figure 1 that the cross section is much higher for the D-T reaction than it is for the D-D reaction. The D-T reaction has a disadvantage in that tritium must be

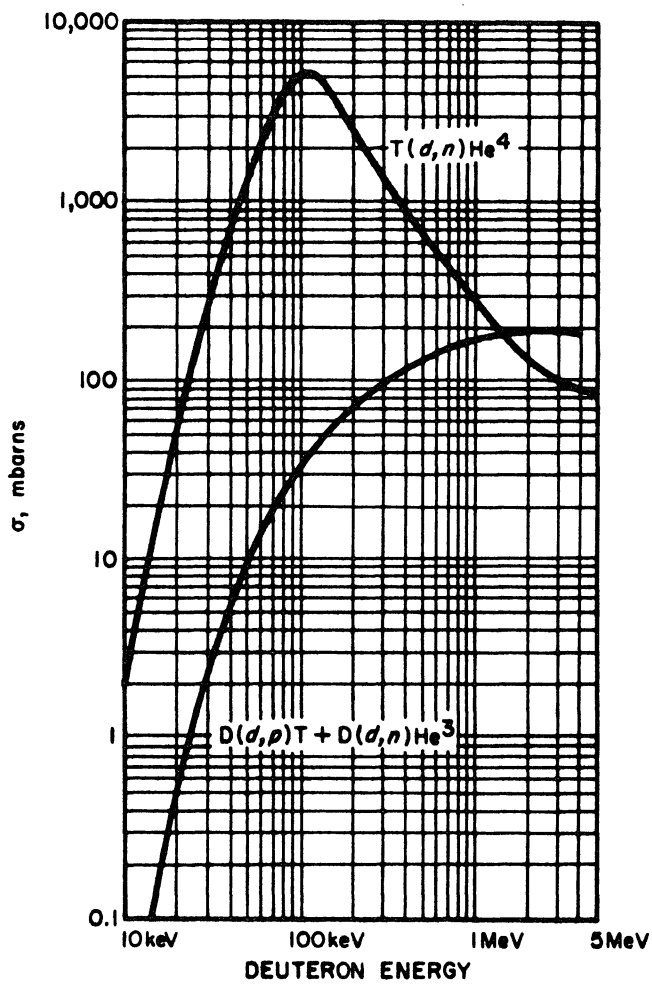

10,000
1,000
100
10
1
0.1
σ, mbarns
T(d,n)He4
D(d,p)T + D(d,n)He3
10keV
100keV
1MeV
5MeV
DEUTERON ENERGY

FIGURE 1

bred whereas deuterium occurs quite abundantly in nature. There is no real difficulty here, however, since tritium can be bred in a blanket surrounding the reactor at the expense of lithium-6. Breeding ratios greater than unity can easily be obtained.

Figure 1 shows the typical behavior of charged particle reactions having a potential barrier. The cross section rises rapidly above the barrier to a maximum of 0.2 barn for the D-D reaction and 5 barns for the D-T reaction. In both cases the cross sections are very small below 10 kev. How then might one build a thermonuclear reactor with deuterium and tritium? An obvious suggestion is to accelerate tritium or deuterium ions to 20 kev and then fire them at a deuterium target. It is easy to show that this will not be profitable since most of the ions would be slowed down below the barrier by collision with cold electrons before a reaction could occur. The cross section for collision with electrons (large angle coulomb scattering) is

$$\sigma_{coul} \cong \left(\frac{e^2}{m\, v_{ion}^2} \right)^2 .$$

The average energy transferred to the electron is roughly the energy available in the center of mass which is

$$E_{c.M.} \cong \tfrac{1}{2} m v_{ion}^2 = \frac{m}{M} E_{ion} ,$$

where m and M are the electron and ion masses respectively and E_{ion} is the energy of the ion in the laboratory. The resulting rate of energy loss which is proportional to the product of the coulomb cross section and the average loss per collision is

$$\frac{dE_{ion}}{dt} = n_e \sigma_{coul} v_{ion} \frac{m}{M} E_{ion}$$

$$= n_e \left[\left(\frac{e^2}{m v_{ion}^2} \right)^2 \frac{m}{M} \right] v_{ion} E_{ion} ,$$

where n_e is the electron density in the solid target. Another way of putting this is in terms of an energy loss cross section

$$\sigma_{en\ loss} = \sigma_{coul} \frac{m}{M} = \left(\frac{e^2}{M v_{ion}^2} \right)^2 \frac{M}{m} ,$$

which is proportional to the ion-ion cross section and larger than it by the ratio of the ion mass to the electron mass. At 50 kev this is

$$\sigma_{en\ loss} = \left(\frac{(4.8)^2 \times 10^{-20}}{10^5 (1.6 \times 10^{-12})} \right)^2 (3600) \cong 7.5 \times 10^{-21} .$$

This is to be compared with about 10 *mb* for the cross section for a D-D reaction at that energy. The ratio is

$$R \cong \frac{10^{-26}}{7.5 \times 10^{-21}} \cong 10^{-6}$$

That is, only about one ion in a million will succeed in producing a nuclear reaction. This is even too great by about a factor of 100. The above analysis considered only those collisions which were nearly head-on, i.e. where the energy transfer was near a maximum. It turns out that most of the energy loss comes from repeated small angle scattering and the net result is that only one in 100 million ions will produce a nuclear reaction. It is easy to show that this is not sufficient to give a net production of energy, since 50 kev was expended in accelerating the ion.

The energy production in the D-D reaction is as follows: about half the time a triton is produced,

$$D + D \longrightarrow T\,(1\,Mev) + p\,(3.00\,Mev);$$

and half the time a He^3 is produced,

$$D + D \longrightarrow He^3\,(0.8\,Mev) + n\,(2.45\,Mev).$$

The triton would ultimately react with another deuteron,

$$D + T \longrightarrow He^4\,(3.5\,Mev) + n\,(14.1\,Mev).$$

If the containment time were long enough the He^3 could also react with another deuteron,

$$He^3 + D \longrightarrow He^4\,(3.6\,Mev) + p\,(14.7\,Mev).$$

The net result would then be the conversion

$$6D \rightarrow 2He^4\,(7.1\ Mev) + 2p\,(17.7\ Mev) + 2n\,(16.6\ Mev) + 1.8\ Mev.$$

If we optimistically count all the energy, the net production is seen to be 7.1 Mev. per deuteron reaction. The deuterons were accelerated to 50 kev, so that the ratio of energy from the nuclear reaction to energy required to accelerate the deuteron is

$$\frac{E_{out}}{E_{in}} = \frac{7.1}{.05} \cong 140.$$

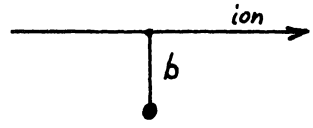
Obviously this is not sufficient to off-set the ratio of 10^{-8} in the cross section for a reaction compared to the cross section for energy loss. This scheme is hopeless as a source of power.

The trouble came from the energy lost to the cold electrons. How about heating the target so the electrons do not dissipate the energy of all the ions? For this to be successful, we must raise the electron temperature at least to the order of 1 kev $\cong (10^7)^\circ K$. At these temperatures the target is no longer a solid or a liquid or a gas. It becomes a plasma, and the problem becomes one of containing a hot plasma.

If one can't use solid targets then how about using two colliding beams of ions? In order to have a reasonably high density beam of ions it is necessary that the beam be neutralized

by electrons, and again there is the problem of energy loss to the electrons. Suppose, however, that the electrons are either absent or hot. In this case we find that ion-ion scattering out of the beam dominates the nuclear reaction. Let us calculate this scattering cross section.

Suppose an ion is passing a fixed scattering center at an impact parameter b .



The momentum change imparted to the ion will be the integral with respect to time of the coulomb force. This can be approximated by the product of the maximum force, i.e. the force at the point of closest approach, times the time during which the ion is in the vicinity of the electron.

$$\Delta(Mv) \cong \frac{ze^2}{b^2} \cdot \frac{2b}{v} = \frac{2ze^2}{bv} ,$$

where z is the charge on the ion (z = 1 for a deuteron).

For small deflections the angular deflection in radians is just equal to the fractional change in momentum

$$\Delta\theta = \frac{\Delta(Mv)}{Mv} = \frac{2ze^2}{Mv^2} \cdot \frac{1}{b} .$$

As an ion moves along, the deflections it suffers will be random with respect to the beam direction so that

$$\overline{\Delta\theta} = 0$$

However, $\overline{(\Delta\theta)^2}$ is not zero and there is a random walk, in angle, away from the original direction. The mean square deflection is found by integrating over all scattering centers. When the particle has traveled a distance λ .

$$\overline{(\Delta\theta)^2} = n\lambda \int \left(\frac{2Ze^2}{Mv^2}\right)^2 \frac{1}{b^2} \cdot 2\pi b\, db$$

$$= 8\pi n\lambda \left(\frac{Ze^2}{Mv^2}\right)^2 \ln \frac{b_{max}}{b_{min}} .$$

It is of course necessary to choose some maximum and minimum cutoff values for b since the integral is otherwise divergent. The physically significant maximum cutoff is determined by coulomb screening and is usually taken to be the Debye radius,

$$b_{max} = \sqrt{\frac{kT}{4\pi ne^2}} .$$

The lower limit is the classical distance of closest approach

$$b_{min} = \frac{e^2}{\frac{1}{2}Mv^2} ,$$

or the DeBroglie wave length

$$b_{min} = \frac{\hbar}{Mv} ,$$

whichever is larger. In most cases of interest it is found that $\ln \frac{b_{max}}{b_{min}}$ is somewhere between 10 and 20. Because of the logarithmic dependence the result will not be strongly sensitive to the choice of b_{max} and b_{min} . One can now define crudely a "cross section" for deflection through an angle

of $\pi/2$.

$$\text{"}\sigma_{\pi/2}\text{"} = \pi\left(\frac{Ze^2}{Mv^2}\right)^2 \cdot 8 \log \frac{b_{max}}{b_{min}} ,$$

which is essentially 8 $\log \frac{b_{max}}{b_{min}}$ times the coulomb cross section. This is the basis for the earlier statement that the energy loss by the soft small angle deflections is about 100 times as great as that from close collisions. (This is not quite an ordinary cross section since we were calculating the mean squared deflection, which means that the path length needed to give, on the average, two successive ninety degree deflections is about 4 times as great as that for the first ninety degrees.)

This effective "cross section" for deflection through ninety degrees evaluated at 50 kev is about 520 barns compared to the 10 mb nuclear reaction cross section. Because the ratio of cross sections is 52,000, the particles are deflected through 90^o hundreds of times before a reaction occurs. Since the energy gain per reaction is a factor of 142, the idea of colliding beams begins to look like a possibility. However, in practice it still won't work for quite a different reason, the low densities that can be attained in ion beams. Typical beams can have 10^8 or 10^9 ions per cm^3 and at most 10^{10}. Then the mean free path for scattering out of the beam is of the order of

$$\lambda \cong \frac{1}{10^{10} \times 5 \times 10^{-22}} \cong 10^{11} \text{ cm}.$$

Thus the particles would have to be contained and passed by one another again and again. This essentially brings us back to the problem of containing a plasma.

Let us now concentrate our attention on the problem of containing a plasma at high temperatures. Just what do we mean by high temperature? A minimum condition is that the temperature be high enough so that the rate of energy production from the nuclear reactions exceeds the rate of energy loss by bremsstrahlung. Let us calculate the bremsstrahlung loss.

Suppose that an electron of mass m is passing an ion with charge Z_i and that the impact parameter is b. The power radiated is classically

$$P = \frac{2}{3}\frac{e^2}{c^3}|a|^2 ,$$

where the acceleration a is given by

$$a = \frac{Z_i e^2}{m b^2} .$$

The total energy radiated in the collision is this power times the time during which the two particles are near one another,

$$t \cong \frac{2b}{v} .$$

Then,

$$Pt = \frac{2}{3}\frac{e^2}{c^3}\left(\frac{Z_i e^2}{m b^2}\right)^2 \cdot \frac{2b}{v} .$$

This must be integrated over all impact parameters and multiplied by the flux of electrons nv and density of targets to obtain the net power radiated per unit volume. Since the plasma is electrically neutral,

$$n = NZ ,$$

where N and Z are the ion density and charge respectively.

$$P_{total} = N^2 Z v \cdot \frac{2}{3} \frac{e^2}{c^3} \left(\frac{Ze^2}{m}\right)^2 \frac{4\pi}{v} \int \frac{db}{b^2} \quad ,$$

so

$$P_{total} \cong \frac{8}{3} \frac{\pi e^6 Z^3 N^2}{m^2 c^3} \cdot \frac{1}{b_{min}} \quad .$$

For fast electrons, the appropriate b_{min} is the de Broglie wave length

$$b_{min} = \frac{\hbar}{mv} \quad ,$$

and the total power radiated is

$$P_{total} = \frac{\frac{8}{3}\pi Z^3 e^6 N^2}{\hbar m c^3} \sqrt{\frac{8kT}{\pi m}} \quad ,$$

which shows that the energy radiated is a slowly increasing function of the temperature, being proportional to the square root of T . Also notice that for a given ion density the power radiated by collisions with impurity ions would be proportional to the density of impurity ions and to the square of their charge. Therefore, the concentration of high impurities must be kept very small or the energy losses will be drastically increased.

For the reaction to proceed in steady state the energy production from nuclear reactions must balance the energy losses. Figure 2 shows the reaction rate as a function of temperature for the D-T and D-D reactions.

FIGURE 2

Thermonuclear Reaction Rates

kT (kev)	$\overline{\sigma v}_{D-T}$ ($cm^3/sec.$)	$\overline{\sigma v}_{D-D}$ (cm^3/sec)
0.05	7×10^{-35}	2×10^{-35}
0.5	6×10^{-23}	2×10^{-24}
2.0	3×10^{-19}	5×10^{-21}
10.0	1.1×10^{-16}	8.6×10^{-19}
100.0	8.1×10^{-16}	3×10^{-17}

The nuclear reaction rate is proportional to $\overline{\sigma v}$.

$$P_{nuc} \propto N^2 \overline{\sigma v} .$$

Figure 3 is a plot of the bremsstrahlung losses and nuclear reaction yield as a function of temperature. The point at which the curves cross is the ignition temperature. It can be seen that the temperature must be at least 10 kev before ignition takes place.

The calculation of the bremsstrahlung loss assumed that the radiation escaped from the plasma freely. This will surely be the case in any system which we can realize in the laboratory. If the radiation were strongly self-absorbed the plasma would

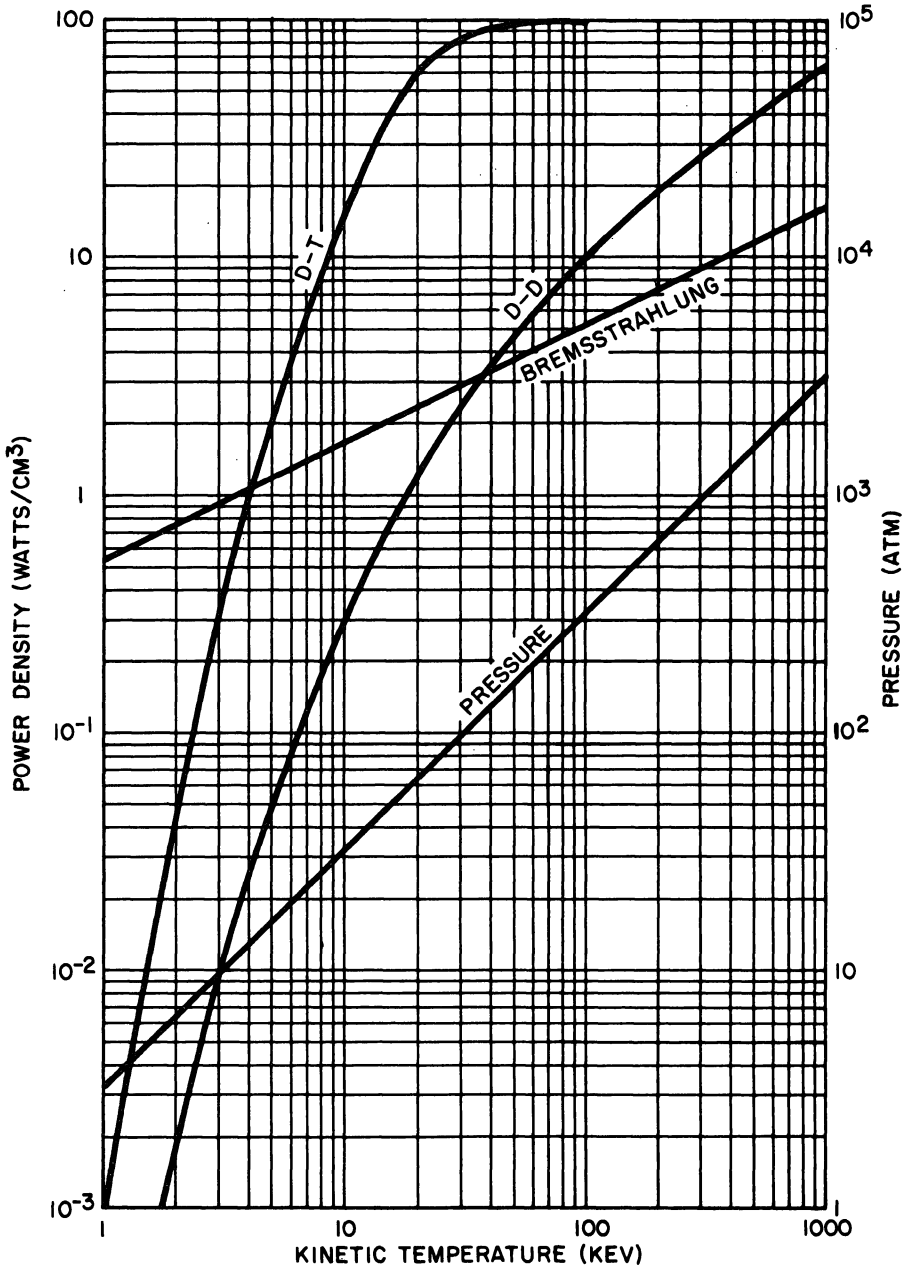

100
10
1
10^{-1}
10^{-2}
10^{-3}
POWER DENSITY (WATTS/CM^3)
10^5
10^4
10^3
10^2
10
1
PRESSURE (ATM)
D-T
D-D
BREMSSTRAHLUNG
PRESSURE
1
10
100
1000
KINETIC TEMPERATURE (KEV)

FIGURE 3

be equivalent to a black body. The radiation pressure at a temperature of 10 kev is 10^{11} atmospheres.

We have now seen that ignition can be achieved if a plasma can be held at temperatures greater than 10 kev. The next question is how one could contain such a plasma for a sufficient period of time to produce some useful amount of energy. It obviously can't be done by material walls. Even if one just decided to make the walls thick enough to accomplish containment for a short time, the poisoning effect of high Z materials discussed above would extinguish the reaction. One's thoughts turn immediately to fields, and of these the only two which seem to be reasonable are electric and magnetic fields. Electric fields are not very promising since a field which contains particles of one charge will tend to pull out those of the opposite sign. The most promising solution seems to be containment by magnetic fields. Let us now calculate how much plasma can be contained by a reasonable size field.

The equation of motion in the hydrodynamic approximation is

$$\rho \frac{d\vec{v}}{dt} + \vec{\nabla} P = \vec{j} \times \vec{B},$$

where ρ is the density, $\vec{v}$ the velocity, P the pressure, $\vec{j}$ the current density and $\vec{B}$ the magnetic field. One of Maxwell's equations states that

$$curl\ \vec{B} = 4\pi \vec{j} \quad .$$

This can be used to eliminate $\vec{j}$, so that in the steady state, neglecting displacement currents and quadratic terms in the velocity, the equation reduces to

$$\vec{\nabla}\left(P + \frac{B^2}{8\pi}\right) = \frac{1}{4\pi}(\vec{B}\cdot\vec{\nabla})\vec{B} \quad .$$

If, in addition, the field is uncurved, $\vec{B}\cdot\vec{\nabla}$ vanishes and $P + B^2/8\pi$ is a constant. If P vanishes on some surface, which is equivalent to stating that the plasma is contained, then we can write

$$\frac{B_{ext.}^2}{8\pi} = P_{int} + \frac{B_{int}^2}{8\pi} \quad .$$

The maximum possible value of P_{int} occurs when B_{int} goes to zero. Then

$$\frac{B_{ext}^2}{8\pi} = P_{int} \quad .$$

This magnetic pressure is about 15 atm for a field B = 20 kilogauss. Thus a field of this magnitude can contain a pressure of about 15 atm, which determines the maximum possible ion density since the minimum temperature has been fixed at about 10 kev.

$$15\ atm = N k T_i$$

$$N = 10^{15}\ ions/cm^3 .$$

It can be seen that in order to obtain a high plasma density we want the ratio

$$\beta \equiv \frac{P}{B^2/8\pi}$$

to be as near to unity as possible.

For the D-D reaction higher temperatures near 100 kev are needed. Then the 20 kilogauss field could contain 10^{14} ions/cm^3. Of course, one can increase the magnetic field strength and thus increase the maximum density. However, material stress problems become very severe for fields above approximately 50 kilogauss.

Having determined the plasma density and temperature, we require one last basic parameter. This is the minimum containment time. How long must the plasma be held together at these conditions for a reasonable fraction of the ions to undergo a nuclear reaction? The nuclear reaction rate is

$$\frac{dN}{dt} = -N^2 \overline{\sigma v} \ ,$$

and hence the effective life time for a nuclear reaction is

$$t \cong \frac{1}{N\bar{\sigma}\bar{v}}$$

For the D-T reaction at a density of 10^{15} and a temperature of 10 kev,

$$t_{DT} = \frac{1}{N\,\bar{\sigma}\bar{v}_{DT}} \cong \frac{1}{10^{15} \times 10^{-16}} = 10 \text{ sec.}$$

For D-D the result is more like 1000 sec.

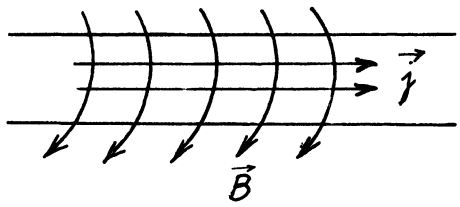
It is not necessary economically to react all of the ions, and in fact it turns out that a containment time 1/100 as great as this would do. The times we are aiming for are about

t = 1/10 sec D-T
t = 10 sec D-D.

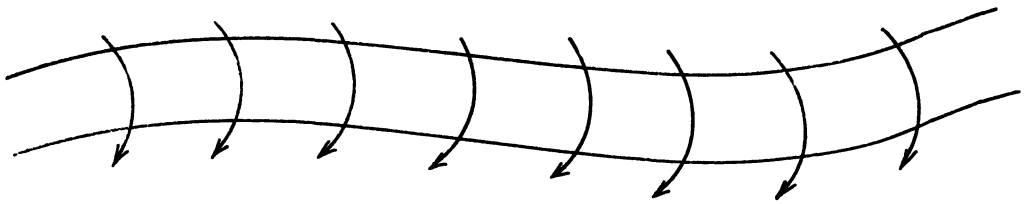
To summarize, the following basic requirements must be met to achieve an economical thermonuclear reactor. For the D-T reaction temperatures of the order of 10 kev and densities of the order of 10^{15} must be maintained for roughly 0.1 sec. Containment is by a magnetic field of 20 kilogauss. For the same magnetic field and the D-D reaction a density of 10^{14} must be contained at a temperature of 100 kev for about 10 sec. If higher fields can be attained then the densities can be higher and the requirement on containment time is correspondingly reduced.

II. Plasma Stability

At the present time the major problem facing those who are trying to achieve controlled fusion is plasma stability. There are a great variety of possible instabilities. One of the most familiar is that which occurs in the pinch. In the simplest variation of the pinch a high electric current is passed down a column of plasma. The current heats the plasma, and the self field squeezes it down. The self field is mostly outside the plasma and confines it.



Suppose the column of plasma begins to kink.

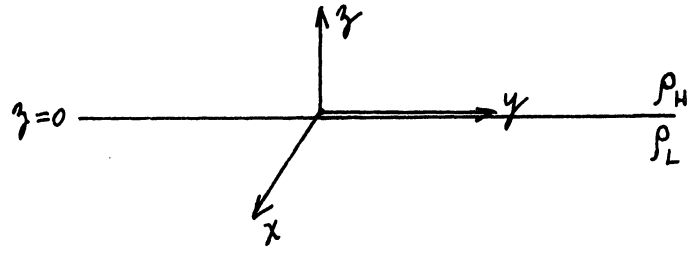


Simply from the geometrical effect, $B^2/8\pi$ is increased on the inside of the bend and decreased on the outside. The magnetic pressure tends to cause the kink to grow and the plasma to be unstable. The growth time of the instability is of the order of the thermal transit time, so the plasma disassembles in a microsecond or less.

There is a more general type of instability called the flute instability. We will start out by analyzing a simple problem which will demonstrate most of the basic principles that

are involved in the more pertinent situation to be discussed later. Consider the problem of a heavy fluid supported by a light fluid. It will be assumed that there is initially a plane horizontal boundary separating two semi-infinite fluids of different density, the higher density being on top. Such a configuration is unstable but it is well known that it can persist for an appreciable length of time if it is prepared carefully. We ask what the behavior of this system will be if a perturbation is applied to the boundary. Will the perturbation grow or not?

We choose the following coordinate system,



The gravitational force per unit mass is in the negative direction and has strength g ,

$$F = -g\vec{i}_z .$$

We will deal with this problem in the hydrodynamic approximation which will later be applied to a plasma.

The continuity equation states

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}\cdot(\rho\vec{v}) = 0 . \qquad (1)$$

Assuming that the pressure is a scalar, the force equation is

$$\rho \frac{d\vec{v}}{dt} = -\vec{\nabla} P - \rho g \vec{i}_z \,. \tag{2}$$

An equation of state is also needed, and for this we take the simplest one, that which applies to incompressible fluids. This insures that

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \rho = 0 \,, \tag{3}$$

and the continuity equation combined with this gives

$$\vec{\nabla} \cdot \vec{v} = 0. \tag{4}$$

An initial state is chosen in which there is no mass velocity, $\vec{v}(0) = 0$, and no x or y dependence.

Then the force balance equation gives for the initial state

$$\frac{dP_0}{dz} = -\rho_0 g \,, \tag{5}$$

where ρ_0 is either ρ_H or ρ_L . The boundary conditions at the interface are that pressure and velocity are continuous. The gradient of the pressure will be discontinuous because of the change in density.

Now, as usual in a problem of this type, we will linearize for a small disturbance about equilibrium. The linearized force equation becomes

$$\rho_0 \frac{d\vec{v}_1}{dt} = \rho_0 \frac{\partial \vec{v}_1}{\partial t} = -\vec{\nabla} P_1 - \rho_1 g \vec{i}_z \,, \tag{6}$$

and the continuity equation is

$$\frac{\partial \rho_1}{\partial t} + (\vec{v}_1 \cdot \vec{\nabla}) \rho_0 = 0, \tag{7}$$

and

$$\vec{\nabla} \cdot \vec{v}_1 = 0. \tag{8}$$

In these equations the subscript 1 refers to the change in the quantity whose initial value is designated by subscript 0.

Next we Fourier transform in the variables time, x and y, i.e. we seek solutions of the form $e^{i\omega t + ik_x x + ik_y y}$. When this is done, the x, y and z components of the force equation are

$$i\omega \rho_0 v_x = -ik_x P_1 \tag{9}$$

$$i\omega \rho_0 v_y = -ik_y P_1 \tag{10}$$

$$i\omega \rho_0 v_z = -\frac{\partial P_1}{\partial z} - \rho_1 g \,, \tag{11}$$

and from the continuity equation and incompressibility we get

$$i\omega \rho_1 + v_z \frac{d\rho_0}{dz} = 0 \tag{12}$$

$$i(k_x v_x + k_y v_y) + \frac{\partial v_z}{\partial z} = 0. \tag{13}$$

We can substitute (9) and (10) into (13) to obtain

$$\rho_1 = \frac{\omega \rho_0}{i k^2} \frac{\partial v_z}{\partial z}, \tag{14}$$

where

$$k^2 = k_x^2 + k_y^2 . \tag{15}$$

Now put (12) and (14) into (11) to obtain a single equation in v_z .

$$\frac{\partial}{\partial z}\left(\rho_0 \frac{\partial v_z}{\partial z}\right) - \left(k^2 \rho_0 + \frac{k^2 g}{\omega^2} \frac{\partial \rho_0}{\partial z}\right) v_z = 0. \tag{16}$$

This is now a homogeneous second order differential equation. We need boundary conditions at infinity. These are provided by the fact that the perturbation in all quantities will vanish at infinity. In addition there are two boundary condition at the interface. One is that v_z must be continuous as follows immediately from Fig. (8). Using the symbol $[\quad]$ to designate the jump in a quantity across the interface, we have

$$[v_z] = 0. \tag{17}$$

Then integration of equation (16) over a little pillbox enclosing the surface gives

$$\left[\rho_0 \frac{\partial v_z}{\partial z} \right] = \frac{k^2 g [\rho_0] v_z}{\omega^2} . \qquad (18)$$

Now if it is assumed that ρ_0 is a constant in each region, equation (16) reduces to

$$v_z'' - k^2 v_z = 0.$$

Since v_z must vanish at $\pm \infty$ the solution is

$$v_{zU} = A e^{-kz}$$
$$v_{zL} = B e^{+kz},$$

where U and L refer to the upper and lower regions respectively. But from (17) A must equal B, and (18) gives

$$\rho_H(-k) - \rho_L(+k) = \frac{k^2 g}{\omega^2}(\rho_H - \rho_L).$$

This can be solved for ω and gives

$$\omega = \pm i \sqrt{kg} \sqrt{\frac{\rho_H - \rho_L}{\rho_H + \rho_L}} .$$

Thus there are two modes, one decaying in time and the other growing. This is the well-known Rayleigh-Taylor instability, with

$$i\omega \approx \sqrt{kg} .$$

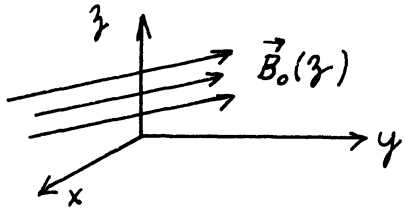
This linear result says that the shortest wave lengths grow most rapidly. It is of course valid only for small displacements and fails when the amplitude becomes large.

Indeed it is observed experimentally that when the displacements become large, the longest wave lengths grow most rapidly and the velocity approaches a constant,

$$v \sim \sqrt{\frac{g}{k}} \ .$$

Now we return to plasma physics and consider a problem that is similar in most respects to the simple problem treated above. The problem to be solved is that of a plasma supported against gravity by a magnetic field, and we again investigate the stability of such a situation.

The coordinate system is chosen as for the previous problem, and the magnetic field is taken to be in the x - y plane,



and is independent of x and y. Thus

$$\vec{B}_0(z) \cdot \vec{i}_z = 0.$$

Again we assume that the fluid is incompressible,

$$\vec{\nabla} \cdot \vec{v} = 0 \tag{19}$$

$$\frac{\partial \rho}{\partial t} + (\vec{v} \cdot \vec{\nabla}) \rho = 0. \tag{20}$$

The force equation is

$$\rho \frac{d\vec{v}}{dt} = -\nabla P + \vec{j} \times \vec{B} - \rho g \vec{i}_z \,. \tag{21}$$

The Maxwell equations for the electric and magnetic fields are

$$curl\ \vec{B} = 4\pi \vec{j} \tag{22}$$

$$curl\ \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t} \,. \tag{23}$$

Now the assumption of infinite conductivity is introduced in the form of the equation.

$$\vec{E} + \frac{\vec{v} \times \vec{B}}{c} = 0 \,. \tag{24}$$

Equations (23) and (24) are combined to give

$$\begin{aligned} \frac{\partial B}{\partial t} &= curl\,(v \times B) \\ &= \vec{v}(\vec{\nabla} \cdot \vec{B}) - \vec{B}(\vec{\nabla} \cdot \vec{v}) + (\vec{B} \cdot \vec{\nabla})\vec{v} - (\vec{v} \cdot \vec{\nabla})\vec{B}. \end{aligned} \tag{25}$$

Next (22) is substituted in (21) to eliminate $\vec{j}$. Then after applying some vector identities we get

$$\rho \frac{d\vec{v}}{dt} = -\vec{\nabla}\left(P + \frac{B^2}{8\pi}\right) + \frac{1}{4\pi}(\vec{B} \cdot \vec{\nabla})\vec{B} - \rho g \vec{i}_z \,. \tag{26}$$

The equations in this problem can be seen to be similar to those governing the Taylor instability, except that there is one

more variable, $\vec{B}$, and one more equation, (25). Once again we linearize about the equilibrium steady state. The linearized equations are

$$\rho_0 \frac{\partial \vec{v}_1}{\partial t} = -\vec{\nabla}\left(P_1 + \frac{\vec{B}_0 \cdot \vec{B}_1}{4\pi}\right) + \frac{1}{4\pi}(\vec{B}_0 \cdot \vec{\nabla})\vec{B}_1 + \frac{1}{4\pi}(\vec{B}_1 \cdot \vec{\nabla})\vec{B}_0 - \rho_1 g \vec{i}_z \tag{27}$$

$$\frac{\partial \rho_1}{\partial t} + (\vec{v}_1 \cdot \vec{\nabla})\rho_0 = 0 \tag{28}$$

$$\vec{\nabla} \cdot \vec{v}_1 = 0 \tag{29}$$

$$\frac{\partial \vec{B}_1}{\partial t} = (\vec{B}_0 \cdot \vec{\nabla})\vec{v}_1 - (\vec{v}_1 \cdot \vec{\nabla})\vec{B}_0 \, . \tag{30}$$

The subscript 0 denotes the steady state value and 1 denotes the departure from the steady state. Once again we take the Fourier transforms in time, x, and y. Define the quantity

$$\Lambda \equiv \frac{\vec{B}_0 \cdot \vec{B}_1}{4\pi} + P_1 \, . \tag{31}$$

Just to show what the transformed equations look like we exhibit the x component,

$$i\omega \rho_0 v_x = -ik_x \Lambda + \frac{i\vec{k} \cdot \vec{B}_0}{4\pi} B_{1x} + \frac{B_{1z}}{4\pi} \frac{\partial B_{0x}}{\partial z} \, . \tag{32}$$

There is a similar one for the y direction, and in the z direction we get

$$i\omega \rho_0 v_z = -\frac{\partial}{\partial z}\Lambda + \frac{i\vec{k} \cdot \vec{B}_0}{4\pi} B_{1z} - \rho_1 g \, . \tag{33}$$

By combining the various equations to eliminate Λ , the components of $\vec{B}_1$, and v_x and v_y we get a second order equation in v_z ,

$$\frac{d}{dz}\left\{\left[\rho_0 - \frac{(\vec{k}\cdot\vec{B}_0)^2}{4\pi\omega^2}\right]\frac{dv_z}{dz}\right\} - \frac{k^2}{\omega^2}\left[\rho_0\omega^2 - \frac{(\vec{k}\cdot\vec{B}_0)^2}{4\pi} + g\frac{d\rho_0}{dz}\right]v_z = 0, \tag{34}$$

where $k^2 = k_x^2 + k_y^2$, i.e.

$$\vec{k} = k_x\vec{i}_x + k_y\vec{i}_y \ . \tag{35}$$

Now remember that we must consider arbitrary perturbations which, in this normal mode analysis, means that we have to look at all possible directions of $\vec{k}$ with respect to the direction of $\vec{B}_0$. Suppose that the direction of $\vec{B}_0$ is independent of z and that we look at the case where $\vec{k}\cdot\vec{B}_0 = 0$, i.e. modes which have $\vec{k}$ perpendicular to the direction of $\vec{B}_0$. Suppose also that ρ_0 is independent of z. Then the equation reduces to that which we had in the Taylor instability, and the plasma is completely unstable, i.e. it cannot be supported against gravity in a magnetic field whose direction is independent of z.

The way to remedy the situation is to introduce shear in the magnetic field. We will analyze the simple case where the magnetic field is constant in the upper and lower half spaces but is discontinuous in direction across the boundary. Take

$$B_{Ux} = 0 \quad ; \quad B_{Lx} = B_0$$

$$B_{Uy} = \alpha_p B_0 \; ; \quad B_{Ly} = \alpha_v B_0 \; .$$

Under these conditions we still get the exponential solutions found in the Taylor instability, but the expression for ω^2 is

$$\omega^2 = -kg + \frac{B_0^2}{4\pi\rho_0}\left[(\alpha_p k_y)^2 + (k_x + \alpha_v k_y)^2\right] .$$

Now a stabilizing effect has been added because the second quantity on the right hand side is positive definite. Only if $\alpha_p = 0$ is it possible to find a direction of $\vec{k}$ for which this term vanishes. Of course the plasma will still be unstable unless the second term on the right exceeds the first. For given values of the other parameters, it is always possible to find a small enough value of k (long enough wave length) to make ω^2 negative. It would appear that the plasma will always be unstable to sufficiently long wave lengths. However, in practice the plasma has a finite extent in the x and y directions and it is improper to consider wave lengths larger than the plasma dimensions. When this effect is included, the plasma can be made stable.

So far we have been considering the stability of a plasma supported against gravity by a magnetic field. Of course, the gravitational force itself is not the real problem. Instead

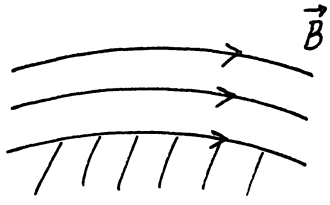
it is accelerative forces produced by particle motion along curved magnetic lines or by acceleration of the plasma as a whole. The effect is entirely equivalent to that of a gravitation, of course.

An example of the latter is provided by the linear pinch with an included B_z field. During the initial phase the plasma shell is accelerated inward by the pressure of the field produced by the pinch current (B_θ). In the frame of the plasma the "gravitational" force points out of the central core which is filled with the B_z field and into the plasma. Hence one expects stability. Continued compression of the trapped B_z field in the center finally stops this process and then produces a deceleration. During this part of the cycle, one expects "gravitational" instability to occur.

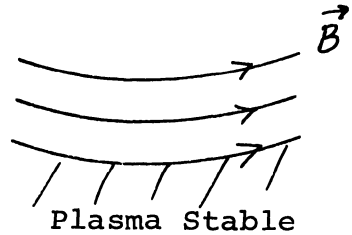
Figure 4 is a photograph of an experimental pinch during the deceleration phase, showing the characteristic flute instability. The discussion above indicated that the addition of sufficient shear should stabilize the flute instability. This can be done by running a wire down the center of the tube and creating a B_θ field in the central region. Figure 5 illustrates the stabilizing effect of doing this.

But it is neither the force of gravity nor the accelerations of implosion that we are most worried about, but rather the accelerations experienced by the particles as they move along curved magnetic lines. For example, if we have an electron

moving along a curved magnetic line, in its rotating frame of reference it is subject to a centrifugal force that plays the same role as gravity does. The effect is such that a plasma confined by a magnetic field with a center of curvature in the plasma (such as in the center of a mirror machine) is unstable, whereas if the curvature of the lines is reversed (as in a cusp) the effect is stabilizing.



Plasma Unstable



Plasma Stable

III. The Motion of Single Particles

In the previous section, plasma stability was investigated from the hydromagnetic point of view, which regards the plasma as a continuous fluid rather than as a collection of particles. Of course, the plasma actually consists of discrete particles and it should be possible, in principle, to investigate stability by studying the motion of single particles.

In this section the motion of single particles will be calculated and then in section IV they will be put together to determine the collective motion of the plasma. The procedure is a self-consistent one in which the fields in which the individual particles move are determined by the externally applied fields plus the fields produced by the collective motion of all the particles. It is interesting and encouraging that the two vastly different approaches to the problem that we shall consider both give the same answer for the growth rate of the flute instabilities.

Consider the motion of a single particle in a magnetic field. Suppose that the particle is moving in a nearly uniform field such that the amount of field variation in one Larmor gyration is small compared to the field itself:

$$\frac{\Delta B}{B} \ll 1 \qquad \text{over a distance } \rho_L = \frac{m v_\perp c}{e B} .$$

It is assumed that the magnetic field is not time dependent and that other fields such as the gravitational field are of

the same order of smallness as $\frac{\Delta B}{B}$. Then the motion to zeroth order, in the magnetic field gradient, is given by

$$m \frac{d v_0}{dt} = \frac{e}{c} \vec{v}_0 \times \vec{B}_0 \quad . \tag{36}$$

where $\vec{B}_0$ is a spatially constant vector.

There is no acceleration parallel to the field. The co-ordinate system is chosen with the z axis in the direction of the magnetic field. Then

$$v_{z0} = \text{constant},$$

and

$$v_{x0} = A \sin \omega_c t + B \cos \omega_c t \tag{37}$$

$$v_{y0} = A \cos \omega_c t - B \sin \omega_c t \tag{38}$$

where ω_c is given by

$$\omega_c = \frac{e B_0}{m c} \tag{39}$$

The corresponding coordinates of the particle as a function of time are

$$z = dt + e \tag{40}$$

$$x = -\frac{A}{\omega_c}\cos\omega_c t + \frac{B}{\omega_c}\sin\omega_c t + f \qquad (41)$$

$$y = \frac{A}{\omega_c}\sin\omega_c t + \frac{B}{\omega_c}\cos\omega_c t + g\,, \qquad (42)$$

where e, f and g are constants. The energy in the plane perpendicular to the z axis is a constant of the motion, since

$$v_\perp^2 = v_x^2 + v_y^2 = A^2 + B^2. \qquad (43)$$

The time averages of v_x^2 and v_y^2 are

$$\overline{v_x^2} = \overline{v_y^2} = \frac{A^2+B^2}{2}. \qquad (44)$$

All other quadratic products of the velocities vanish.

The time averaged products of the velocity and coordinates are

$$\overline{z v_x} = 0 \;;\quad \overline{z v_y} = 0 \;;\quad \overline{z v_z} = 0$$
$$\overline{v_x z} = 0 \;;\quad \overline{v_y z} = 0 \qquad (45)$$

$$\overline{v_x y} = \frac{v_\perp^2}{2\omega_c} \;;\quad \overline{v_x x} = 0 \qquad (46)$$

$$\overline{v_y x} = -\frac{v_\perp^2}{2\omega_c} \;;\quad \overline{v_y y} = 0\,. \qquad (47)$$

In the above the time averaging has been chosen so that the

uniform motion along the z-axis leads to the particle being at $z=0$ at the midpoint in time. We are always free to do this.

To solve the problem with slightly curved fields, the usual procedure is to transform to a coordinate system which moves with the center of gyration but maintains a constant orientation. We find it more convenient to introduce a curving coordinate system that rotates in such a way that the z direction is always parallel to the actual magnetic field.

If $\vec{q}$ is the position co-ordinate of the particle then the velocity is

$$\vec{v} = \dot{\vec{q}} = v_x(t)\vec{i}_x(\vec{q}) + v_y(t)\vec{i}_y(\vec{q}) + v_z(t)\vec{i}_z(\vec{q}), \tag{48}$$

where we have explicitly noted that in our curved coordinate system the unit vectors $\vec{i}_x$, $\vec{i}_y$ and $\vec{i}_z$ are functions of the position of the particle. Since we have assumed that the field is only slightly curved, the unit vectors can be expanded in a Taylor series about the zeroth order position, and we will keep only two terms,

$$\vec{i}_x = \vec{i}_{xo} + (\vec{q}\cdot\vec{\nabla})\vec{i}_x\Big]_0 \qquad \text{etcetera.}$$

Then to first order, the velocity can be written

$$\vec{v}_1 = \left\{ v_{x1}(t)\vec{i}_{xo} + v_{y1}(t)\vec{i}_{yo} + v_{z1}(t)\vec{i}_{zo} \right\}$$

$$+ \left\{ v_{xo}(t)(\vec{q}\cdot\vec{\nabla})\vec{i}_x \right]_o + v_{yo}(t)(\vec{q}\cdot\vec{\nabla})\vec{i}_y \Big]_o \cdot$$

$$+ v_{zo}(t)(\vec{q}\cdot\vec{\nabla})\vec{i}_z \Big]_o \Big\} \quad . \tag{49}$$

The first curly bracket represents the perturbed motion to first order in the uncurved geometry. The quantity in this bracket will be called $\tilde{v}_1(t)$.

$$\tilde{v}_1(t) \equiv v_{x1}(t)\vec{i}_{xo} + v_{y1}(t)\vec{i}_{yo} + v_{z1}(t)\vec{i}_{zo} \cdot$$

The second curly bracket gives to first order the effect of the curving geometry.

Now we will write the equations governing the motion of a particle in terms of this coordinate system. The equation of motion is

$$m\frac{d\vec{v}}{dt} = \frac{e}{c}\vec{v}\times\vec{B} + \vec{F} \quad , \tag{50}$$

where F is any other constant force exerted on the particle. The procedure after expressing quantities in the curving coordinate system will be to average the first order terms over a time that is long compared to the Larmor gyration time. First note that

$$\frac{d}{dt}\vec{v} = \frac{\partial}{\partial t}\vec{v} + (\vec{v}\cdot\vec{\nabla})\vec{v} \tag{51}$$

is to first order

$$\frac{d}{dt}\vec{v}_1 = \frac{\partial}{\partial t}\vec{v}_1 + (\vec{v}_0\cdot\vec{\nabla})\vec{v}_1 \tag{52}$$

since $\vec{v}_0$ is space independent. Further

$$\frac{\partial\vec{v}_1}{\partial t} = \frac{\partial}{\partial t}\Big[\tilde{v}_1(t) + \frac{\partial v_{x0}}{\partial t}(\vec{q}\cdot\vec{\nabla})\vec{i}_x\Big]_0 + \frac{\partial v_{y0}}{\partial t}(\vec{q}\cdot\vec{\nabla})\vec{i}_y\Big]_0 + \frac{\partial v_{z0}}{\partial t}(\vec{q}\cdot\vec{\nabla})\vec{i}_z\Big]_0 . \tag{53}$$

From the zeroth order calculation above we have

$$\frac{\partial v_{x0}}{\partial t} = \omega_c v_{y0};\quad \frac{\partial v_{y0}}{\partial t} = -\omega_c v_{x0};\quad \frac{\partial v_{z0}}{\partial t} = 0 , \tag{54}$$

so

$$\frac{\partial\vec{v}_1}{\partial t} = \frac{\partial\tilde{v}_1(t)}{\partial t} + \omega_c v_{y0}(\vec{q}\cdot\vec{\nabla})\vec{i}_x\Big]_0 - \omega_c v_{x0}(\vec{q}\cdot\vec{\nabla})\vec{i}_y\Big]_0 . \tag{55}$$

Now we take our time average, and with the help of equations (45), (46), and (47) we get

$$\frac{\partial\vec{v}_1}{\partial t} = \frac{\partial\tilde{v}_1(t)}{\partial t} - \frac{v_\perp^2}{2}\frac{\partial\vec{i}_x}{\partial x}\Big]_0 - \frac{v_\perp^2}{2}\frac{\partial\vec{i}_y}{\partial y}\Big]_0 \tag{56}$$

(we do not distinguish $\tilde{v}_1(t)$ from its orbit average, throughout)

Similarly from equations (43) and (44) the time average of $(\vec{v}_0 \cdot \vec{\nabla})\vec{v}_1$ is

$$\overline{(\vec{v}_0 \cdot \vec{\nabla})\vec{v}_1} = \overline{v_{0x}^2 \frac{\partial \vec{i}_x}{\partial x}}\Big]_0 + \overline{v_{0y}^2 \frac{\partial \vec{i}_y}{\partial y}}\Big]_0 + \overline{v_{0z}^2 \frac{\partial \vec{i}_z}{\partial z}}\Big]_0$$

$$= \frac{v_\perp^2}{2} \frac{\partial \vec{i}_x}{\partial x}\Big]_0 + \frac{v_\perp^2}{2} \frac{\partial \vec{i}_y}{\partial y}\Big]_0 + v_{0z}^2 \frac{\partial \vec{i}_z}{\partial z}\Big]_0 \quad (57)$$

When (56) and (57) are combined in (52) the result is

$$\frac{d\vec{v}_1}{dt} = \frac{d\tilde{v}_1}{dt} + v_{0z}^2 \frac{\partial \vec{i}_z}{\partial z}\Big]_0 . \quad (58)$$

Now note that

$$\frac{\partial \vec{i}_z}{\partial z} = -\frac{\vec{R}}{R^2} , \quad (59)$$

where $\vec{R}$ is the radius of curvature of magnetic field (pointing away from the center of curvature), and (58) becomes

$$\frac{d\vec{v}_1}{dt} = \frac{d\tilde{v}_1}{dt} - \frac{v_{0z}^2 \vec{R}}{R^2} . \quad (60)$$

Thus the left-hand side of the equation of motion (50), when written in the curving coordinate system, has one additional term which is due to the centrifugal effect.

Next consider the $\frac{e}{c}\vec{v}\times\vec{B}$ term on the right hand side of equation (50). Again to first order,

$$\frac{e}{c}\vec{v}\times\vec{B} = \frac{e}{c}\vec{v}_1\times\vec{B}_0 + \frac{e}{c}v_0\times(\vec{q}\cdot\vec{\nabla})\vec{B}\Big]_0 \tag{61}$$

The first term on the right can be written

$$\frac{e}{c}\vec{v}_1\times\vec{B}_0 = \frac{e}{c}\tilde{\vec{v}}_1\times\vec{B}_0 + \frac{e}{c}\left[v_{x0}(t)(\vec{q}\cdot\vec{\nabla})\vec{i}_x + v_{y0}(t)(\vec{q}\cdot\vec{\nabla})\vec{i}_y + v_{z0}(t)(\vec{q}\cdot\vec{\nabla})\vec{i}_z\right]\times\vec{B}_0 . \tag{62}$$

When the bracket is replaced by its time average and equations (45), (46), and (47) are used the result is

$$\frac{e}{c}\vec{v}_1\times\vec{B}_0 = \frac{e}{c}\tilde{\vec{v}}_1\times\vec{B}_0 + \frac{eB_0}{c}\frac{v_1^2}{2\omega_c}\left\{\frac{\partial\vec{i}_x}{\partial y}\Big]_0\times\vec{i}_{z0} - \frac{\partial i_y}{\partial x}\Big]_0\times\vec{i}_{z0}\right\}, \tag{63}$$

where we have used also the fact that

$$\vec{B}_0 = B_0 i_{z0} . \tag{64}$$

The remaining Lorentz term is

$$\frac{e}{c}\vec{v}_0\times(\vec{q}\cdot\vec{\nabla})B\vec{i}_z\Big]_0 =$$

$$\frac{e}{c}\vec{v}_0\times\left(x\frac{\partial}{\partial x}+y\frac{\partial}{\partial y}\right)B\vec{i}_z\Big]_0 + \frac{e}{c}\vec{v}_0 z\frac{\partial}{\partial z}B\vec{i}_z\Big]_0 . \tag{65}$$

(since $\vec{B} = B\vec{i}_z$ by definition of our coordinate system)

The last term on the right averages to zero since

$$\overline{v_{0\alpha}z} = 0 \qquad \alpha = x, y, \text{ or } z .$$

The other terms do not vanish and if equations (45), (46), and (47) are applied again here the result is

$$\overline{\frac{e}{c}\vec{v}_0 \times (\vec{q}\cdot\vec{\nabla})\vec{B}} = -\frac{e}{c}\frac{v_\perp^2}{2\omega_c}\left[\frac{\partial B}{\partial x}\Big|_0 \vec{i}_{ox} + \frac{\partial B}{\partial y}\Big|_0 \vec{i}_{oy}\right]$$

$$-\frac{e}{c}\frac{v_\perp^2}{2\omega_c} B_0 \left[\vec{i}_{y_0} \times \frac{\partial \vec{i}_y}{\partial x}\Big|_0 - \vec{i}_{x_0} \times \frac{\partial \vec{i}_y}{\partial y}\Big|_0\right]. \qquad (66)$$

Now if we introduce the operator

$$\nabla_\perp \equiv \vec{i}_{ox}\frac{\partial}{\partial x} + \vec{i}_{oy}\frac{\partial}{\partial y} \quad , \qquad (67)$$

The equations (66), (63), (61), (60) can be substituted back into equation (50) to obtain the result

$$m\frac{d}{dt}\vec{\tilde{v}}_1 = \frac{e}{c}\vec{\tilde{v}}_1 \times \vec{B}_0 + \vec{G} \quad , \qquad (68)$$

where $\vec{G}$ is defined by

$$\vec{G} \equiv \frac{m v_{y_0}^2}{R^2}\vec{R} - \frac{m v_\perp^2}{2B}\vec{\nabla}_\perp B - \frac{m v_\perp^2}{2}\left(\frac{\partial \vec{i}_x}{\partial x} + \frac{\partial \vec{i}_y}{\partial y}\right)_0 + \vec{F} \; .$$

Since all second order terms are neglected, $\vec{G}$ represents a force which is constant in both space and time. The solution of this equation is immediate and gives the first curly bracket in equation (49). The second bracket can be evaluated directly from the zeroth order calculation. The sum of these two is then $\vec{v}_1(t)$. Therefore, the complete solution of the problem

to first order and in the approximation of replacing the correction terms by their time averages reduces to solving equation (68) for $\vec{\tilde{v}}_1$

We will not restrict the discussion to situations which in the external force $\vec{F}$ is either zero or has no z component. Then the z component of equation (68) is

$$m \frac{d\tilde{v}_{1z}}{dt} = - \frac{m v_\perp^2}{2} \vec{i}_z \cdot \left(\frac{\partial \vec{i}_x}{\partial x} + \frac{\partial \vec{i}_y}{\partial y} \right). \tag{70}$$

now $$\vec{\nabla} \cdot \vec{B} = \vec{\nabla} \cdot (B \vec{i}_z) = 0$$

or $$\left[\vec{i}_x \frac{\partial}{\partial x} + \vec{i}_y \frac{\partial}{\partial y} + \vec{i}_z \frac{\partial}{\partial z} \right] \cdot \left[B \vec{i}_z \right] = 0$$

or $$\frac{\partial B}{\partial z} + B \left(\vec{i}_x \cdot \frac{\partial \vec{i}_z}{\partial x} + \vec{i}_y \cdot \frac{\partial \vec{i}_z}{\partial y} \right) = 0 .$$

This can be written as

$$\frac{\partial B}{\partial z} + B \left(\frac{\partial}{\partial x} (\vec{i}_x \cdot \vec{i}_z) - \vec{i}_z \cdot \frac{\partial \vec{i}_x}{\partial x} + \frac{\partial}{\partial y} (\vec{i}_y \cdot \vec{i}_z) - \vec{i}_z \cdot \frac{\partial \vec{i}_y}{\partial y} \right) = 0$$

or

$$\frac{\partial B}{\partial z} - B \vec{i}_z \cdot \left(\frac{\partial \vec{i}_x}{\partial x} + \frac{\partial \vec{i}_y}{\partial y} \right) = 0 .$$

Hence,

$$\vec{i}_z \cdot \left(\frac{\partial \vec{i}_x}{\partial x} + \frac{\partial \vec{i}_y}{\partial y} \right) = \frac{1}{B} \frac{\partial B}{\partial z} , \tag{71}$$

and equation (70) can be written

$$m \frac{d\tilde{v}_{1z}}{dt} = - \frac{m v_\perp^2}{2B} \frac{\partial B}{\partial z} . \tag{72}$$

Now taking the orbit time average of the z-component of equation (49), we have

$$v_{z1} = \tilde{v}_{z1}(t) + \frac{v_\perp^2}{2\omega_c}\vec{i}_z \cdot \left(\frac{\partial \vec{i}_x}{\partial y} - \frac{\partial \vec{i}_y}{\partial x}\right).$$

But

$$\vec{i}_z \cdot (\vec{\nabla}\times\vec{B}) = \vec{i}_z \cdot \left\{\left[\vec{i}_x\frac{\partial}{\partial x} + \vec{i}_y\frac{\partial}{\partial y} + \vec{i}_z\frac{\partial}{\partial z}\right]\times(B\vec{i}_z)\right\}$$

$$= \vec{i}_z \cdot \left\{-\vec{i}_y\frac{\partial B}{\partial x} + \vec{i}_x\frac{\partial B}{\partial y} + B\left(\vec{i}_x \times \frac{\partial \vec{i}_z}{\partial x}\right) + B\left(\vec{i}_y \times \frac{\partial \vec{i}_z}{\partial y}\right) + B\left(\vec{i}_z \times \frac{\partial \vec{i}_z}{\partial z}\right)\right\}$$

$$= B\vec{i}_z \cdot \left[\frac{\partial}{\partial x}(\vec{i}_x \times \vec{i}_z) - \frac{\partial \vec{i}_x}{\partial x}\times\vec{i}_z + \frac{\partial}{\partial y}(\vec{i}_y\times\vec{i}_z) - \frac{\partial \vec{i}_y}{\partial y}\times\vec{i}_z\right]$$

$$= B\vec{i}_z \cdot \left(\frac{\partial \vec{i}_x}{\partial y} - \frac{\partial \vec{i}_y}{\partial x}\right).$$

Now if there are no appreciable currents flowing in the particles vicinity, then $\vec{\nabla}\times\vec{B} = 0$ and we find that

$$v_{z1} = \tilde{v}_{z1}$$

Hence, to first order

$$m\frac{dv_z}{dt} = -\frac{m v_\perp^2}{2B}\frac{\partial B}{\partial z}. \qquad (73)$$

Equation (73) is the basis of containment of particles in a magnetic mirror geometry. If we multiply this equation by v_z we have

$$\frac{d(v_z^2)}{dt} = -\frac{v_\perp^2}{B}\frac{\partial B}{\partial z}v_z$$

$$= -\frac{v_\perp^2}{B}\frac{\partial B}{\partial z}\frac{\partial z}{\partial t}$$

$$= -\frac{v_\perp^2}{B}\frac{dB}{dt},$$

neglecting second order effects. Now, the total energy of a particle is a constant for motion in a time-independent magnetic field. Hence

$$v_{\parallel}^2 + v_{\perp}^2 = \text{constant} \tag{74}$$

$$\therefore \frac{d(v_{\perp}^2)}{dt} = \frac{v_{\perp}^2}{B}\frac{dB}{dt} \tag{75}$$

which has the solution

$$\frac{v_{\perp}^2}{B} = \text{constant} \tag{76}$$

The constant above is an adiabatic invariant (since it depends on second order effects being small). If a particle now spirals into a region of increasing magnetic field, we see that $v_{\perp}^2$ must increase. Since total energy must be conserved this can only occur at the expense of v_z^2. If the increase in B is sufficient, we will reduce v_z^2 to zero. In that case the forward motion of the particle ceases and it is reflected.

Returning now to equation (68) we consider the components of $\tilde{v}_1$ in the x and y directions. Denoting the corresponding components of $\vec{G}$ by G_x and G_y, we find at once that

$$\tilde{v}_{x1} = \alpha \sin \omega_c t + \beta \cos \omega_c t + \frac{G_y}{m\omega_c} \tag{77}$$

$$\tilde{v}_{y1} = \alpha \cos \omega_c t - \beta \sin \omega_c t - \frac{G_x}{m\omega_c} \quad , \tag{78}$$

where α and β are constants.

Taking the orbit time averages of the x and y components of equation (49) we find:

$$v_{x1} = \tilde{v}_{x1} + \frac{v_\perp^2}{2\omega_c} \vec{i}_x \cdot \left(\frac{\partial \vec{i}_x}{\partial y} - \frac{\partial \vec{i}_y}{\partial x} \right)$$

$$v_{y1} = \tilde{v}_{y1} + \frac{v_\perp^2}{2\omega_c} \vec{i}_y \cdot \left(\frac{\partial \vec{i}_x}{\partial y} - \frac{\partial \vec{i}_y}{\partial x} \right).$$

Hence, by equations (77) and (78)

$$v_{x1} = \frac{G_y}{m\omega_c} + \frac{v_\perp^2}{2\omega_c} \vec{i}_y \cdot \frac{\partial \vec{i}_x}{\partial x}$$

$$v_{y1} = -\frac{G_x}{m\omega_c} - \frac{v_\perp^2}{2\omega_c} \vec{i}_x \cdot \frac{\partial \vec{i}_y}{\partial y} \quad ,$$

and substituting from equation (69), we have:

$$v_{x1} = \frac{G'_y}{m\omega_c}$$

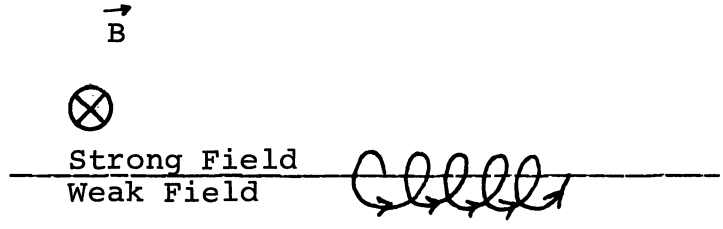
$$v_{y1} = -\frac{G'_x}{m\omega_c}$$

where $\vec{G}' = \frac{m v_{\parallel 0}^2}{R^2} \vec{R} - \frac{m v_\perp^2}{2B} \vec{\nabla}_\perp B + \vec{F}_\perp$. (79)

Hence the entire perpendicular motion of the particle can be described as a uniform rotation about the magnetic field direction, plus an added drift velocity $\vec{v}_{\perp 1}$ given by

$$\vec{v}_{\perp} = \frac{c}{eB^2} \vec{G}' \times \vec{B} . \qquad (80)$$

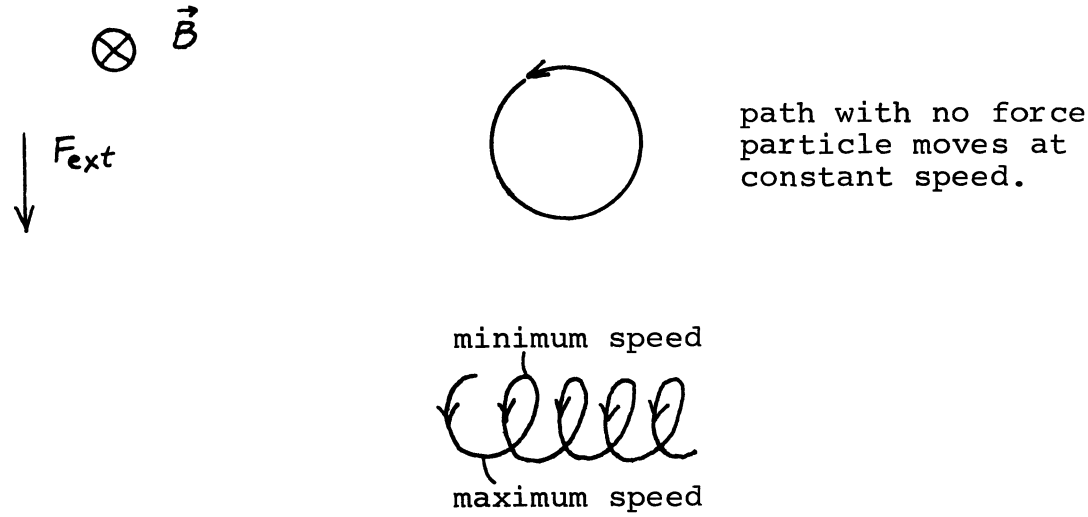
Thus there will be drifts due to external forces, centrifugal effects from the curved field and the so-called grad B drift. The grad B drift can be understood easily by the following picture in which there is an abrupt change in B across a plane boundary with B constant on either side. Assume $\vec{B}$ is directed into the paper. Then in the strong field the particles move in circles of small radii, and in the weak field the radius is larger, resulting in the drift.



The same drift occurs if the field varies continuously, except that the curvature of the path varies continuously rather than abruptly.

The effect of an external or centrifugal force is similar. Again assume the magnetic field is directed into the paper. In the absence of any external force the particle would move in a circle. But in the presence of the force shown the particle will attain a maximum velocity at the bottom of its trajectory

and a minimum velocity at the top of its trajectory resulting in a drift to the right.

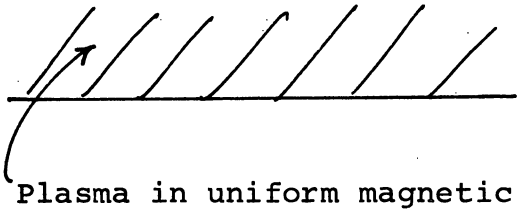


The drift is seen to be in the direction $\vec{F}$ X $\vec{B}$. A gravitational force will produce such a drift as will a centrifugal or inertial force. The expression (80) for the drift velocity contains the electric charge; so for forces which are independent of charge, electrons and ions drift in opposite directions. But if the force is due to an electric field then $\vec{F} = e\vec{E}$ and is in the opposite direction for electrons and ions. Then e drops out of the expression for $\bar{v}_\perp$. In such a field the electrons and ions drift in the same direction with a velocity $\vec{v} = c\vec{E}/\vec{B}$.

IV. The Single Particle Approach to Plasma Stability

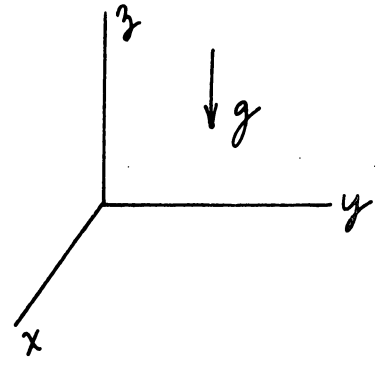
Having derived the drift velocities of the individual particles in given fields, we now want to essentially sum over the motion of all particles to determine the behavior of the plasma. We will be interested in the relatively long time behavior of the plasma so that we can average the motion of the individual particles over a time that is long compared to the Larmor gyration time. Then it is just the drift velocities calculated in Section III that will determine the plasma behavior.

Suppose we try to repeat the analysis of the stability of a plasma supported against the force of gravity(1) that was done from the hydromagnetic point of view in Section II.



Plasma in uniform magnetic field

Directed out of the paper



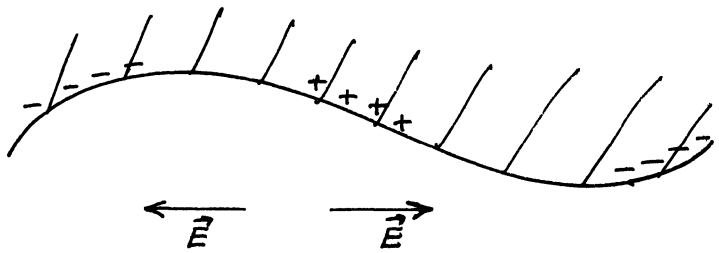
According to the single particle picture the ions will drift left and the electrons will drift right.

(1) M. N. Rosenbluth and C. L. Longmire, Annals of Physics 1, 120 (1957).

The drift velocity from equation (79) is

$$\vec{v}_D = \frac{-mgc\,\vec{i}_z \times \vec{B}}{eB^2} = -\frac{mgc}{eB}\vec{i}_y \,. \qquad (81)$$

If the plasma is infinite and there is no surface perturbation, there will be no effects from this drift. What happens when there is a surface perturbation?



Magnetic field out of the paper

Due to the drifts there tends to be a charge separation as shown, and this produces an electric field. The electric field in turn will influence the drifts. In Section III the fields were all assumed known. Thus, to utilize those results we must adopt a self-consistent procedure in which the fields that occur in those equations must be the net fields including the effect of particle drifts. We will find that this self-consistent single particle approach will predict that the perturbation will grow exponentially and at the same rate as was calculated in Section II using the hydromagnetic equations.

The electric field produced by the ion drift in the presence of the surface perturbation shown is parallel to the surface and perpendicular to the magnetic field. $\vec{E} \times \vec{B}$ in this case is upward where the surface is perturbed upward and downward where the surface is perturbed downward.

Therefore the effect of the charge separation is to cause drifts that tend to increase the perturbation, and the perturbation will grow. Let us calculate the growth rate.

Consider a boundary perturbation of the form

$$\Delta z = a \sin ky \,. \tag{82}$$

The separation of charges produces a net electric charge density that will vary with position on the surface and which is described by the continuity equation

$$\frac{\partial \sigma}{\partial t} = -\vec{\nabla} \cdot (\eta \vec{v}_D) \,, \tag{83}$$

where η is the total ionic charge per unit area of surface that is contained within a surface thickness equal to the perturbation,

$$\eta = -eN\Delta z \,. \tag{84}$$

Here N is the volumetric ion density. The electron drift will be neglected since we see from (79) that the drift velocity is directly proportional to the mass. Substituting (82) into (83) we get

$$\begin{aligned} \frac{\partial \sigma}{\partial t} &= \vec{\nabla} \cdot [eN\vec{v}_D a \sin ky] \\ &= eNv_D ak \cos ky. \end{aligned} \tag{85}$$

To first order, v_D is independent of y, so σ is separable in space and time.

$$\sigma = \sigma_0(t) \cos ky \,, \tag{86}$$

and we have

$$\frac{d\sigma_0}{dt} = eNv_D ak \,. \tag{87}$$

What are the electric fields that arise from a charge distribution like this? They can be found in a straightforward manner. The potential is

$$V(\vec{r},t) = \iiint \frac{\sigma(\vec{r}',t)}{\epsilon|\vec{r}-\vec{r}'|} d^3r' \,, \tag{88}$$

where

$$\sigma(\vec{r}',t) = \sigma_0(t) \cos ky\, \delta(z') \,, \tag{89}$$

when the integration is performed the result is:

$$V(\vec{r},t) = \frac{4\pi\sigma_0(t)}{\epsilon k} \cos ky\, e^{-kz} \,. \tag{90}$$

Taking the gradient of this potential we immediately find the components of the electric field:

$$E_x = 0 \,; \quad E_y = \frac{4\pi\sigma_0}{\epsilon} \sin ky\, e^{-kz}$$
$$E_z = \frac{4\pi\sigma_0}{\epsilon} \cos ky\, e^{-ky} \,. \tag{91}$$

For a small perturbation the E_z field will merely alter slightly the surface charge distribution and can be ignored as a second order effect. The E_y field, however, will produce a drift in the z direction given by equation (80).

$$v_z = -\frac{cE_y}{B} = -\frac{c}{B}\frac{4\pi\sigma_0}{\epsilon}\sin ky \,. \tag{92}$$

But from equation (82)

$$v_z = \frac{d\Delta z}{dt} = \frac{da}{dt}\sin ky \,, \tag{93}$$

and we can identify $\frac{da}{dt}$ as

$$\frac{da}{dt} = -\frac{c}{B}\frac{4\pi}{\epsilon}\sigma_0(t) \,. \tag{94}$$

Differentiating (94) and combining it with (87) we get

$$\frac{d^2a}{dt^2} = -\frac{c}{B}\frac{4\pi}{\epsilon} eN v_D ka \,. \tag{95}$$

Then v_D can be replaced by means of equation (81):

$$\frac{d^2a}{dt^2} = \frac{4\pi NMc^2 kg\, a(t)}{B^2\epsilon} \tag{96}$$

All the quantities in the coefficient of $a(t)$ on the right hand side of equation (96) are positive constants, so we know immediately the system is unstable. The solution of equation (96) is the sum of a positive and negative exponential.

Everything in the time constant has been well defined except ϵ, the dielectric constant. We need to evaluate it for an oscillating electric field that is perpendicular to the magnetic field. Because the frequency of oscillation that is involved here is small compared to the frequency of collisions we can use hydrodynamics.

$$\rho \frac{d\vec{v}}{dt} = \vec{j} \times \vec{B} = \rho_i \frac{d\vec{v}_i}{dt} , \tag{97}$$

where the electrons have again been ignored because of their low mass. $\vec{v}_i$ is given by equation (80):

$$\vec{v}_i = c \frac{\vec{E} \times \vec{B}}{B^2} , \tag{98}$$

which is substituted into (97) to give

$$\vec{j} \times \vec{B} = \rho_i \frac{c}{B^2} \frac{d\vec{E} \times \vec{B}}{dt} . \tag{99}$$

The polarization current $\vec{j}$ can be identified from this equation to be

$$\vec{j} = \rho_i \frac{c}{B^2} \frac{d\vec{E}}{dt} . \tag{100}$$

This is the typical equation for a polarizable medium. If P is the polarizability

$$\rho = PE , \tag{101}$$

then the polarization current is related to P by

$$j = \frac{e}{c}\frac{d\rho}{dt} = \frac{eP}{c}\frac{dE}{dt} \quad . \tag{102}$$

From (102) and (100) P can be identified as

$$P = \frac{\rho_i c^2}{eB^2} \quad . \tag{103}$$

The dielectric constant is

$$\epsilon = 1 + 4\pi e P = 1 + \frac{4\pi \rho_i c^2}{B^2} \quad . \tag{104}$$

Usually $\frac{4\pi\rho_i c^2}{B^2} >> 1$, so the electric field will not penetrate far into the plasma. In this case

$$\epsilon \approx \frac{4\pi \rho_i c^2}{B^2} \quad . \tag{105}$$

We can evaluate this for a typical situation to see how good the approximation is. Take $N = 10^{14}$ and $B = 10^4$ gauss. Then

$$\frac{4\pi\rho_i c^2}{B^2} \cong \frac{4\pi \cdot 10^{14} \cdot 10^{-24} \cdot 10^{21}}{10^8} \cong 10^4 .$$

Equation (105) can be substituted into (96) to give

$$\frac{d^2 a}{dt^2} = kg\,a \quad . \tag{106}$$

This is equivalent to the result found from the hydromagnetic point of view. So again we find that the perturbation is

unstable with an exponential growth constant,

$$\omega = \sqrt{gk} \tag{107}$$

The agreement between these two vastly different approaches leads one to wonder if the single particle approach might also be used for other problems. It is not known whether the two methods will always agree in other situations. Unfortunately, the single particle method does not appear to be adapted for problems involving complicated geometry, and so far has been restricted to simple problems like that treated here.

V. The δW Formalism

In more complicated geometries the method of normal modes and the single particle approach are too cumbersome. There is another powerful technique called the δW formalism that is more adapted to complicated geometries. We will develop the essential features of this formalism as applied to the magnetohydrodynamic equations.

The system of equations that will be used to describe a plasma supported against gravity by a magnetic field are:

$$\rho \frac{d\vec{v}}{dt} = -\vec{\nabla}P + \vec{j} \times \vec{B} - \rho \vec{\nabla}\varphi, \tag{108}$$

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0, \tag{109}$$

$$\vec{E} + \frac{\vec{v}}{c} \times \vec{B} = 0, \tag{110}$$

$$\frac{d}{dt}\left(P/\rho^{\gamma}\right) = 0, \tag{111}$$

$$\text{curl } \vec{E} = -\frac{1}{c}\frac{\partial \vec{B}}{\partial t}, \tag{112}$$

$$\text{curl } \vec{B} = 4\pi \vec{j}, \tag{113}$$

$$\text{div } \vec{B} = 0. \tag{114}$$

In equation (108) the gravitational force has been expressed through the gravitational potential φ. The analysis will then

apply to any gravitational-like force. Instead of assuming the plasma is incompressible as we did in the normal mode analysis, we have assumed that any compression or expansion of the plasma is adiabatic as expressed in equation (111).

Let us go to a Lagrangian description of the fluid motion. In that description all quantities are functions of $\vec{r}_o$ (the initial location of the fluid element) and time t. We define the displacement vector (the displacement of each fluid element from its initial position $\vec{r}_o$)

$$\vec{\xi}(\vec{r}_o, t) = \vec{r} - \vec{r}_o \quad , \tag{115}$$

where $\vec{r}$ is the location of the fluid element at time t. Clearly $\vec{\xi}$ satisfies the initial condition

$$\vec{\xi}(\vec{r}_o, 0) = 0. \tag{116}$$

Now

$$\left.\frac{\partial}{\partial \vec{r}}\right)_t = \frac{\partial}{\partial \vec{r}}(\vec{r}_o) \cdot \frac{\partial}{\partial \vec{r}_o} \quad , \tag{117}$$

and from (115)

$$\frac{\partial \vec{r}_o}{\partial \vec{r}} = 1 - \frac{\partial \vec{\xi}}{\partial r} \tag{118}$$

To first order in $\vec{\xi}$ i.e. for small displacements from equilibrium

$$\frac{\partial \vec{\xi}}{\partial \vec{r}} = \frac{\partial \vec{\xi}}{\partial \vec{r}_0} \quad , \tag{119}$$

and therefore

$$\vec{\nabla} = \vec{\nabla}_0 - (\vec{\nabla}_0 \vec{\xi}) \cdot \vec{\nabla}_0 \tag{120}$$

Let us now consider systems which are in a condition of static equilibrium ($\vec{v} = 0$ and $\frac{\partial}{\partial t} = 0$) at t = 0. The equations describing this equilibrium are

$$\vec{\nabla}_0 P_0 = \vec{j}_0 \times \vec{B}_0 - \rho_0 \vec{\nabla} \varphi_0 \quad , \tag{121}$$

$$\vec{E}_0 = 0 \, , \tag{122}$$

$$curl_0 \vec{B}_0 = 4\pi \vec{j}_0 \quad , \tag{123}$$

$$div_0 \vec{B}_0 = 0 \, , \tag{124}$$

and $\vec{\xi}_0 = 0$ for all time.

If we now make a small displacement of each fluid element by an amount $\vec{\xi}(\vec{r}_0, 0^+)$, $\vec{\xi}$ then will begin to change with time by obeying the equations of motion. Hence,

$$\frac{d\vec{\xi}}{dt} = \vec{v}_1 \quad , \tag{125}$$

where $\vec{v}_1$ is also small. We can also calculate the change in the other quantities due to the small displacement. We have, by equations (110) and (112)

$$\frac{\partial \vec{B}}{\partial t} = \text{curl}\ \vec{v} \times \vec{B}\ , \tag{126}$$

or

$$\frac{d\vec{B}}{dt} - (\vec{v}\cdot\vec{\nabla})\vec{B} = \text{curl}\ \vec{v}\times\vec{B}\ . \tag{127}$$

Linearizing about the unperturbed state we get

$$\frac{d\vec{B}_1}{dt} - (\vec{v}_1\cdot\vec{\nabla})\vec{B}_0 = \text{curl}\ \vec{v}_1 \times \vec{B}_0\ , \tag{128}$$

which can be written

$$\frac{d}{dt}\left\{\vec{B}_1 - (\vec{\xi}\cdot\vec{\nabla})_0 \vec{B}_0\right\} = \frac{d}{dt}\left[\vec{\nabla}_0 \times (\vec{\xi}\times\vec{B}_0)\right]\ . \tag{129}$$

This can be integrated and solved for $\vec{B}_1$. The integration constant vanishes because $\vec{B}_1$ is zero at t = 0.

$$\vec{B}_1 = (\vec{\xi}\cdot\vec{\nabla})_0 \vec{B}_0 + \vec{\nabla}_0 \times \left(\vec{\xi}\times\vec{B}_0\right), \tag{130}$$

or

$$\vec{B} = \vec{B}_0 + (\vec{\xi}\cdot\vec{\nabla})\vec{B}_0 + \vec{\nabla}_0 \times (\vec{\xi} \times \vec{B}_0) \ . \tag{131}$$

From (113) and (120) we get

$$\begin{aligned} 4\pi\vec{j} &= 4\pi\vec{j}_0 + 4\pi\vec{j}_1 \\ &= 4\pi\vec{j}_0 + 4\pi(\vec{\nabla}_0 \times \vec{B}_1) - 4\pi(\vec{\nabla}_0\cdot\vec{\xi})\vec{\nabla}_0 \times \vec{B}_0 \ . \end{aligned} \tag{132}$$

This is combined with (130) to give

$$\begin{aligned} \vec{j} = \vec{j}_0 + \vec{\nabla}_0 \times \big[(\vec{\xi}\cdot\vec{\nabla})\vec{B}_0 + \vec{\nabla}_0 \times (\vec{\xi} \times \vec{B}_0) \big] \\ - \big[(\vec{\nabla}_0 \vec{\xi}) \cdot \vec{\nabla}_0 \big] \times \vec{B}_0 \ . \end{aligned} \tag{133}$$

Since

$$\frac{\partial \rho}{\partial t} = \frac{d\rho}{dt} - (\vec{v}\cdot\vec{\nabla})\rho \ , \tag{134}$$

we have from (109)

$$\frac{d\rho}{dt} + \rho(\vec{\nabla}\cdot\vec{v}) = 0 \ . \tag{135}$$

This we will linearize to obtain

$$\frac{d\rho_1}{dt} + \rho_0 (\vec{\nabla}\cdot\vec{v}_1) = 0, \tag{136}$$

or

$$\frac{d}{dt}\big[\rho_1 + \rho_0(\vec{\nabla}_0\cdot\vec{\xi})\big] = 0 \ . \tag{137}$$

Again because $\rho_1 = 0$ at t = 0 we can drop the integration constant and write

$$\rho_1 + \rho_0 (\vec{\nabla}_0 \cdot \vec{\xi}) = 0, \tag{138}$$

or

$$\rho = \rho_0 - \rho_0 (\vec{\nabla}_0 \cdot \vec{\xi}) \ . \tag{139}$$

From the adiabatic law, equation (111)

$$P \rho^{-\gamma} = P_0 \rho_0^{-\gamma} \ . \tag{140}$$

This we combine with (139) to obtain

$$P = P_0 \left(\frac{\rho_0}{\rho_0 - \rho_0 (\vec{\nabla}_0 \cdot \vec{\xi})} \right)^{-\gamma} \tag{141}$$

Now equation (141) is expanded and to first order

$$P = P_0 \left[1 - \gamma (\vec{\nabla}_0 \cdot \vec{\xi}) \right] \ . \tag{142}$$

The first order expression for the gravitational potential is

$$\varphi = \varphi_0 + (\vec{\xi} \cdot \vec{\nabla})_0 \varphi_0 \ . \tag{143}$$

The linearized force equation can now be written down by substituting equations (125), (131), (133), (139), (142) and (143) into equation (1) and dropping all second order terms. We have

$$\rho_0 \ddot{\vec{\xi}} = \vec{F}(\vec{\xi}), \tag{144}$$

where $\vec{F}(\vec{\xi})$ can be reduced to

$$\vec{F}(\vec{\xi}) = \vec{\nabla}_0 \left[\gamma P_0 (\vec{\nabla}_0 \cdot \vec{\xi}) + (\vec{\xi} \cdot \vec{\nabla}_0) P_0 \right] + \vec{j}_0 \times \vec{Q} - \vec{B}_0 \times \left(\vec{\nabla}_0 \times \vec{Q} \right) + \left[\vec{\nabla}_0 \cdot (\rho_0 \vec{\xi}) \right] \vec{\nabla}_0 \varphi_0 . \tag{145}$$

and where

$$\vec{Q} = \vec{\nabla}_0 \times (\vec{\xi} \times \vec{B}_0) . \tag{146}$$

Notice $\vec{F}$ is purely a function of position and independent of velocity $\dot{\vec{\xi}}$.

We can show with proper boundary conditions that $\vec{F}(\vec{\xi})$ is a self-adjoint operator. That is

$$\int \vec{\eta} \cdot \vec{F}(\vec{\xi}) \, d\tau_0 = \int \vec{\xi} \cdot \vec{F}(\vec{\eta}) \, d\tau_0 , \tag{147}$$

where $\int d\tau_0$ is integration over all regions of fluid elements $\vec{r}_0$. Note that we do not write F* since it is a real operator.

Let $\vec{\xi}_i$ be a normal mode

$$\vec{\xi}_i (\vec{r}_0, t) = \vec{\xi}(\vec{r}_0) e^{i\omega_i t} . \tag{148}$$

The self-adjoint property of F determines that ω_i^2 is real. Substitute equation (148) into (144).

$$-\omega_i^2 \rho_0 \vec{\xi}_i = \vec{F}(\vec{\xi}_i) \quad . \tag{149}$$

Similarly for the conjugate

$$-\omega_i^{2*} \rho_0 \vec{\xi}_i^* = \vec{F}(\vec{\xi}_i^*) \tag{150}$$

we multiply (149) and (150) by $\vec{\xi}_i^*$ and $\vec{\xi}_i$ respectively, subtract the two and integrate over volume.

$$-(\omega_i^2 - \omega_i^{2*}) \rho_0 \int \vec{\xi}_i \cdot \vec{\xi}_i^* \, d\tau_0 = \int \vec{\xi}_i^* \cdot \vec{F}(\vec{\xi}_i) d\tau_0 - \int \vec{\xi}_i \cdot \vec{F}(\vec{\xi}_i^*) \, d\tau_0 . \tag{151}$$

But by equation (147) the right hand side of (151) vanishes and we have

$$\omega_i^2 = \omega_i^{2*} \quad . \tag{152}$$

which proves our assertion that ω_i^2 is real. This agrees with our previous normal mode analysis.

Now consider the change in potential energy.

$$\delta W = -\frac{1}{2} \int \vec{\xi} \cdot \vec{F}(\vec{\xi}) d\tau_0 \quad . \tag{153}$$

Assume that we have a complete set of orthonormal modes, satisfying equation (149). Then $\vec{\xi}$ can be expanded

$$\vec{\xi} = \sum_n a_n \vec{\xi}_n \, , \tag{154}$$

and δW can be expressed in terms of the expansion coefficients a_n

$$\delta W = -\frac{1}{2}\int \sum_n a_n \vec{\xi}_n \cdot \left(-\rho_0 \sum_m \omega_m^2 \vec{\xi}_m a_m\right) d\tau_0$$

$$= \frac{1}{2}\rho_0 \sum_n a_n^2 \omega_n^2 \,. \qquad (155)$$

We see now that δW can be negative if and only if there exists at least one unstable mode, i.e. a mode with imaginary ω_n. Hence we have the essence of the δW formalism. In practice one minimizes δW using all possible $\vec{\xi}$'s that satisfy the boundary conditions and discovers the conditions under which δW can become negative.

This method is clearly related to Rayleigh's principle which says that if we minimize

$$\omega^2 = \frac{\delta W(\vec{\xi},\vec{\xi})}{K(\vec{\xi},\vec{\xi})} \,, \qquad (156)$$

where

$$K(\vec{\xi},\vec{\xi}) = \frac{1}{2}\int d\tau_0 \rho_0 \xi^2 \qquad (157)$$

Then functions which make this stationary are normal modes, and the ω^2 values calculated from (156) are the eigenvalues. Of course this is equivalent to minimizing δW with the extra condition that $\vec{\xi}$ be normalized by

$$K(\vec{\xi},\vec{\xi}) = 1$$

This is a serious restraint. We drop this in the δW formalism and gain advantage of easier determination of stability. We pay the price, however, of losing the ability to calculate the actual value of ω^2 . We can still establish stability criteria but cannot study the detailed behavior of an unstable situation.

The δW formalism (or the energy principle) has been used to analyze more complicated geometries than can be handled by normal mode techniques. In this way, it has been shown that the stellarator (at low β) should be stable and mirror machines should be unstable (without shear). Neither of these prophesies has proven entirely correct. In fact, the history of the pinch supplies an excellent example of the failure of the MHD predictions. On the other hand, one does see flute instabilities, as predicted, in the linear pinch.

COMMENTS ON THE FAST MAGNETIC COMPRESSION OF PLASMAS

Alan C. Kolb

U. S. Naval Research Laboratory

I. Introduction

Besides the fundamental problem for producing controlled thermonuclear reactions--that of finding a stable configuration for the confinement and heating of a high energy plasma in a magnetic field--there are many other questions of equal importance for achieving a basic understanding of plasma phenomena. The ramifications of laboratory investigations of the fundamental processes which govern the behavior of ionized gases in a magnetic field are relevant to astrophysics, space physics, energy conversion, atomic theory, to name but a few of the more important areas that are receiving increased attention in recent years.

In this series of lectures we shall make the questionable assumption that it will be possible to thermally isolate a high temperature plasma by means of magnetic forces. The discussion will therefore be limited to problems connected with understanding processes in a stable plasma, except in the section on the structure of collision-free shockwaves. Even with this drastic restriction, the subject is extremely complicated and the difficulties are compounded by the fact that the relevant equations are generally of a highly nonlinear character, even when there is equilibrium among the kinetic degrees of freedom of the charged particles.

There are many approaches to the problem of actually producing stable high temperature plasmas. Broadly speaking, the electrotechnical difficulties depend strongly on the

following factors: magnetic field configuration and strength; characteristic times, which range from fractions of a microsecond to steady state; charged particle densities, which differ by at least ten orders of magnitude in different experiments now in progress; and the method of heating.

The number of experimental possibilities is accordingly enormous. However, one can roughly distinguish two avenues, one involving high pulsed magnetic fields and high plasma densities ($\sim 10^{15}$ to 10^{17} particles /cm^3) and the other with moderate quasi-stationary fields and densities (a "few" to $\sim 10^{15}$ particles/cm^3). It would be inappropriate to attempt a review of all the attempts and failures to reach the ultimate goals. In any of the particular investigations, one would like to know if there are common factors which influence the heating and confinement; the role of impurities (radiation losses); and the distribution of densities, temperature and fields in the plasma. Intimately related to such questions is the experimental goal of refining diagnostic measurements to the point where it can be determined whether or not the observations can be fitted into any theoretical picture and to determine the range of validity of the simplifying assumptions which are built into the present theories. From this point of view, it should be emphasized from the outset that plasma physics is still in its infancy and a great deal remains to be done before one can understand the most important mechanisms underlying a great amorphous mass of what are, with rare exceptions, generally crude experimental observations.

Since the one common feature of all the experimental approaches to obtain a high temperature plasma is the presence of a magnetic field, it will serve our purpose here to choose a specific geometrical configuration on which to base the following discussion. The aim is to illustrate what is involved in the interpretation of the data for a particular and more or less typical experiment and to indicate the nature of the diagnostic techniques.

The interpretative sections will be based on the magneto-hydrodynamic (abbreviated by MHD) approximation with the restriction that all considerations apply to the case where the charged particle densities are sufficiently high so that coulomb collisions play an important role. Then one can expect the usual transport coefficients for viscosity, thermal and electrical conduction to be meaningful. Furthermore, in such a fluid description one can hope to treat nonlinear wave phenomena, i.e. shockwaves, semi-quantitatively.

In addition, if the electron velocity distribution is Maxwellian, it should be possible to take into account spectral line and continuous radiation from impurities in an otherwise pure low-Z gas, i.e. hydrogen or deuterium. This is important since radiation losses are known to be of great importance in the energy balance of laboratory plasmas as well as in stellar atmospheres. To complete the analogy between laboratory plasma physics and astrophysics, the MHD equations can provide the basis for "model atmosphere" calculations which are the foundation

for the interpretation of stellar spectra. This intimate connection can be expected to be of significance for fundamental radiation studies which will aid in the interpretation of vacuum ultraviolet radiation from the solar corona, where million degree excitation temperatures are encountered.

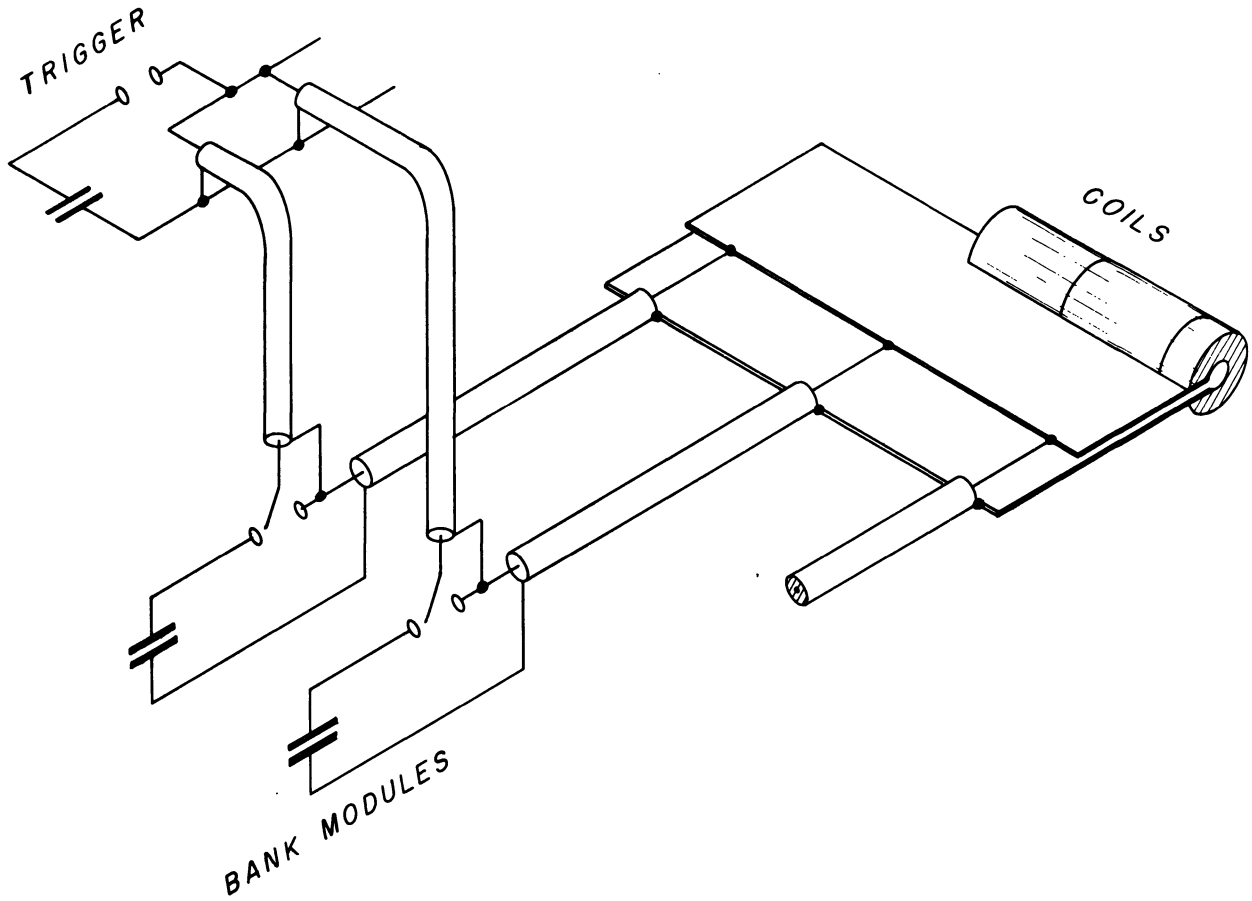
For the experimental verification of the MHD equations, cylindrical systems are most easily studied theoretically (with the aid of computers) and, therefore, only these will be considered here.

Finally, it should be remarked that the theoretical analysis of plasma processes in a hypothetical thermonuclear reactor is much more involved. At the extremely high temperatures and low densities expected in a reactor, the importance of charged particle collisions is reduced. One can expect that a more sophisticated theory will be required than that based on the fluid equations for a collision-dominated plasma, especially if some form of microturbulence manifests itself in the plasma.

II. Cylindrically Symmetic, High-Current Electrodeless Discharges -- the θ-Pinch.

The θ-pinch configuration shown in Figure 1 has been selected for discussion because of its geometrical simplicity (and because of the experimental data available to the writer from the N.R.L.).

The idea of the experiment is to compress and heat a low density deuterium or hydrogen plasma inside a quartz or ceramic



Schematic Diagram of the θ-Pinch Configuration

FIGURE 1

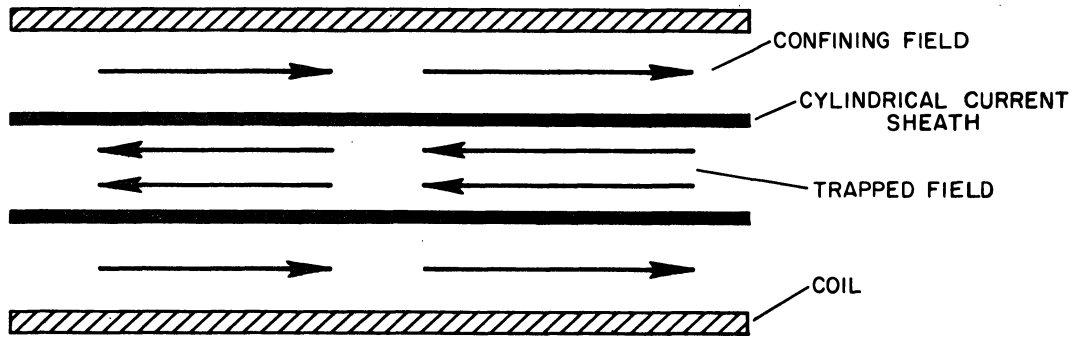
tube by a rapidly rising axial magnetic field. The field is generated by discharging a capacitor bank into a low inductance single coil.

Assume for a moment that the gas is fully ionized initially. The currents in the coil will produce a rapidly rising axial magnetic field within the tube which induces aximuthal gas currents, i.e. in the θ-direction. These currents flow first in a thin cylindrical sheath of plasma near the walls. The force $\vec{J}_\theta \times \vec{B}_z$ resulting from these currents drives a plasma shell inwards with velocities up to about 20 cm./μsec., depending on the electrical parameters and initial gas density. The imploding plasma sweeps up the gas ahead, thereby producing a radial compression of the plasma. After this first implosion the plasma then undergoes a series of damped radial oscillations. These phenomena are completely analogous to the behavior of a dynamic linear-pinch with axial currents and azimuthal magnetic fields; a configuration which has been under investigation for over ten years and which preceded studies of the θ-pinch. The main disadvantage of the ordinary linear-pinch are violent instabilities which appear during the first few bounces. In this latter respect, the θ-pinch seems to be somewhat superior.

In practice, the magnetic compression involves the sequential discharging of at least three separate capacitor banks. A low voltage, high-inductance bank is switched to the coil, producing a slowly rising, quasi-static magnetic field everywhere within the gas. The second discharge which is produced by a fast

capacitor bank, i.e. with a shorter time constant, is started after the quasi-static magnetic field builds up to a desired value. It produces an oscillating magnetic field superimposed on the quasi-static magnetic field, which in turn generates an azimuthal electric field. This induced electric field causes the plasma to break down and then Joule heats it up to $\sim 1 - 5 \times 10^4$ ^{o}K, providing a high initial conductivity. This "preheating" discharge is used to prepare a fully ionized gas for the third and main discharge which compresses the plasma as described earlier.

For reasons which will be explained later, the polarity of the main discharge is such that the rapidly rising compression field is in the opposite direction to the quasi-static field. As will be seen, the cylindrical current sheath moving inward provides a diamagnetic layer which isolates the external magnetic field from the interior of the plasma. Therefore, while the confining main magnetic field remains outside the imploding plasma, the quasi-static magnetic field which is present during the preheating cycle is trapped within the imploding plasma shell. It will be shown that the presence of the trapped field provides a very effective heating mechanism and also confines the plasma axially, because the field lines must be closed at the ends of the plasma column.



θ-Pinch With A Trapped Reverse Axial Magnetic Field

FIGURE 2

III. The Hydromagnetic Equations[1] (two fluid model)

These fluid equations can be derived from the Boltzmann equation. Because we are taking a macroscopic point of view, the following analytical treatment should be regarded as a discussion rather than a rigorous derivation.

Consider a fully ionized plasma filling an infinitely long cylinder. Assume that all quantities depend only on r, the radial distance from the magnetic axis. In the θ-pinch, only axial magnetic fields and aximuthal currents are allowed. We denote them by B and J respectively. It is obvious that by the rotational symmetry condition any hydromagnetic instabilities involving turbulance are excluded from the theory by definition.

One recognizes that in a fully ionized plasma there are two coupled fluids, the ions and the electrons. When the Debye length is much smaller than the characteristic dimensions of the plasma, there will be large electrostatic forces if the plasma departs very much from quasineutrality. In our discussion of a hydrogen or deuterium plasma we assume quasineutrality which implies the number densities of the ions and electrons are equal:

$$n_e = n_i = n . \qquad (1)$$

[1]This section is based on K. Hain, G. Hain, K.V. Roberts, J.J. Roberts and W. Koppendorfer, Naturforsch., 154, 1039, (1960); K. Hain and A.C. Kolb, Nuclear Fusion, 1962 Supplement, Part 2., p. 561.

Since we assume that there is no radial current in a θ-pinch, i.e. $J = n\,(\vec{V}_i - \vec{V}_e)_r = 0$, the radial components of the flow velocities are equal:

$$(\vec{V}_e)_r = (\vec{V}_i)_r = V . \tag{2}$$

Neglecting the electron mass, one obtains the mass density as

$$\rho = n m_i , \tag{3}$$

where m_i is the mass of an ion. The usual continuity equation which describes the conservation of mass is

$$\frac{\partial \rho}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}(r \rho V) = 0 \tag{4}$$

Euler's equation for the conservation of momentum can be written as

$$\rho \frac{d\vec{V}}{dt} + \nabla(p_e + p_i + q_i) + \frac{1}{c}\vec{J} \times \vec{B} = 0 \tag{5}$$

where p_e and p_i are the electron and ion pressures. The viscous term ∇q_i is introduced[2] to include a purely artificial dissipative mechanism of the right form and strength so that the shock transition, if it exists, will be smooth. By this method, unphysical discontinuities at the shock front are avoided and, therefore, no extra boundary conditions are needed to make the solutions single-valued. The conditions which are required to make the solution unique, if the artificial viscosity term were not present, are known as the "Rankine-Hugoniot"

[2]R. D. Richtmyer, Difference Methods for Initial Value Problems, Interscience Publishers, New York, 1957.

equations, relations which insure that mass, momentum and energy are conserved. These are the jump conditions across the shock front, relating the values of V, $\frac{1}{\rho}$, p, B and internal energy $\mathcal{E}$ on each side of the shock front. We shall discuss these conditions explicitly in a later section. What we want to point out here is that by introducing an artificial viscosity, the Rankine-Hugoniot conditions are automatically satisfied at the shock front, which is now a continuous transition extending over a small region. In other words, by this method which is due to J. von Neumann and Richtmyer (1950) the shocks are automatically taken care of whenever and wherever they arise without any particular attention.

The quantity q_i is the pseudo-viscous pressure which is given by

$$q_i = \begin{cases} L^2 \rho \left(\frac{\partial V}{\partial r}\right)^2, & \text{if } \frac{\partial V}{\partial r} < 0 \\ 0, & \text{if } \frac{\partial V}{\partial r} \geq 0 \end{cases}, \qquad (6)$$

where L is a characteristic length associated with the shock thickness, to be adjusted according to the specific problem with which one is dealing (see VII. 3-VII. 5).

In the case of a plasma, one may have a continuous shock transition even when the viscosity is negligible. The dissipation which removes the discontinuities in the shock front, in this case, is due to the ohmic heating resulting from the finite conductivity of the plasma and the currents flowing in the shock front. The shock thickness is then determined, as we shall discuss later in detail, by the skin effect. But this dissipa-

tion mechanism alone is not always sufficient to explain the shock structure, and therefore, one has to keep the artificial viscosity term to represent the other dissipation mechanisms involving the ions which are not yet well understood.

As the index i in equation (6) indicates, we assume that the artificial viscosity applies only to the ions which transport the momentum. This implies that the shock heating resulting from the artificial viscosity is confined only to the ions, whereas the electron motion behaves adiabatically in the shock zone because of the small Larmor radii.

The current density J and the magnetic field B in equation (5) are related to each other by Maxwell's equations. If one ignores the displacement current which is small by a factor $\left(\frac{v}{c}\right)^2$, one obtains in gaussian units

$$\vec{J} = \frac{c}{4\pi} \vec{\nabla} \times \vec{B} \ . \tag{7a}$$

In addition to (7a), Maxwell's equations yield the following relation

$$\frac{1}{c} \frac{\partial \vec{B}}{\partial t} = - \nabla \times \left[\vec{E} + \frac{\vec{v} \times \vec{B}}{c} \right] \ , \tag{7b}$$

where $\vec{E}$ is the electric field in the coordinate system moving with the plasma. Ohm's Law relates $\vec{E}$ to the current density $\vec{J}$ by

$$E_i = \sum_k \eta_{ik} \ J_k \ , \tag{7c}$$

where η_{ik} is the resistivity tensor. In a plasma with a magnetic field, the resistivity in the direction parallel to the magnetic field is different from the transverse resistivity in the direction perpendicular to the field.

In the present problem, Maxwell's equations reduce to the following:

$$J = -\frac{c}{4\pi}\frac{\partial B}{\partial r} \tag{7d}$$

and

$$\frac{\partial B}{\partial t} = \frac{1}{r}\frac{\partial}{\partial r} r\left(\frac{\eta_\perp c^2}{4\pi}\frac{\partial B}{\partial r} - vB\right), \tag{7e}$$

where $\eta_\perp$ is the transverse resistivity.

The temperatures are expressed in terms of an equation of state

$$T_{e,i} \equiv \frac{P_{e,i}}{\rho}, \tag{8}$$

which is realistic if the densities are not very high so that the kinetic energy density is large compared to the coulomb-interaction energy density. In our model, T_e and T_i are allowed to be different to include relaxation effects, i.e. the energy transfer from one fluid to the other.

Since the particles in a fully ionized gas have only translational degrees of freedom, the internal energy is given by the usual ideal gas equation:

$$\mathcal{E}_{e,i} = \frac{p_{e,i}}{\rho} \frac{1}{\gamma - 1} = T_{e,i} \frac{1}{\gamma - 1} \quad . \tag{9}$$

The specific heat ratio γ is

$$\gamma \equiv \frac{f + 2}{f} \quad , \tag{10}$$

where f is the number of translational degrees of freedom of the particles. If the plasma density is relatively high and collisions among the particles are fairly frequent so that the time between collisions is short in comparison with the compression times, then the velocity distribution is spatially isotropic. In this case $f = 3$ and in equation (9), $\gamma = 5/3$. On the other hand, for relatively low plasma densities, where collisions are rare and the collision time may be long compared to the compression time, the particles gain energy only in the two directions perpendicular to the magnetic field. In this case, $f = 2$ and $\gamma = 2$.

Using the definitions of $T_{e,i}$ and J, and introducing the local derivative, i.e.

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + v \frac{\partial}{\partial r} \quad ,$$

one can rewrite equation (5) finally as

$$\frac{Dv}{Dt} + \frac{1}{\rho}\left(T_e + T_i\right)\frac{\partial \rho}{\partial r} + \frac{\partial}{\partial r}\left(T_e + T_i\right) + \frac{1}{\rho}\frac{\partial p_i}{\partial r} = -\frac{1}{4\pi\rho} B \frac{\partial B}{\partial r} \tag{11}$$

In our general formulation we have assumed that only the ions are heated directly by viscous dissipation, and only the electrons are heated directly by ohmic dissipation. It follows, therefore, that the ions and electrons will have different temperatures and two more equations are required to describe the heat balance.

To obtain the two energy equations, we consider the first law of thermodynamics, i.e. $d\mathcal{E} = dQ - p\,d(1/\rho)$ where dQ is the heat given and lost from the system and $p\,d(1/\rho)$ is the work done by the pressure forces, $(1/\rho)$ being the specific volume. Defining a source term S as

$$\frac{DQ}{Dt} = \frac{S}{\rho} \tag{12}$$

so that S is the net heat energy supplied to the plasma per unit volume per second, and using the continuity equation which can also be written as

$$\frac{D}{Dt}(1/\rho) = \frac{1}{\rho}\nabla\cdot\vec{V} \tag{13}$$

in the first law of thermodynamics one obtains the energy balance equation:

$$\frac{D}{Dt}\mathcal{E} + \frac{p}{\rho}\nabla\cdot\vec{V} = \frac{S}{\rho} \quad . \tag{14}$$

Substituting (8) and (9) for $\mathcal{E}$ and p, (14) yields two equations, one for the electrons and one for the ions:

$$\frac{DT_e}{Dt} + (\gamma-1)\,T_e\,\frac{1}{r}\,\frac{\partial}{\partial r}(rv) = (\gamma-1)\,\frac{S_e}{\rho} \tag{15}$$

$$\frac{DT_i}{Dt} + (\gamma - 1) T_e \frac{1}{r} \frac{\partial}{\partial r} (rV) = (\gamma - 1) \frac{S_i}{\rho} \quad , \qquad (16)$$

where S_e and S_i are the energy source terms associated with the electrons and ions, respectively. They lump the thermal conduction losses, the energy transfer between ions and electrons, the ohmic heating of electrons, and the shock heating of ions. In principle, one can include also the radiation losses. The second term in equations (15), (16) represent the heating ratio due to the radial compression.

Equations (7e), (11), (13), (15) and (16) form a closed set of equations which can be solved numerically if not analytically, for V, ρ or n, T_e, T_i and B as a function of t and r. They contain all the information about the plasma within the limitations of the two-fluid hydromagnetic model. The results of the numerical calculations will be presented in a later section.

IV. Source Terms

The source terms in the energy balance equations have yet to be expressed explicitly in terms of the four unknowns V, p, T_e and T_i in order to make the four hydromagnetic equations a closed set. All the physics of the problem of investigating the plasma theoretically enters the formulas through the source terms. It includes at least six different mechanisms:

1. thermal conduction;
2. relaxation between T_e and T_i;
3. ohmic heating of electrons;
4. shock heating of the ions;
5. radiation losses;
6. ionization losses due to the impurities.

The diffusion of the plasma across the field is taken care of automatically due to the finite conductivity.

1. Thermal Conduction Losses (S_{th})

The thermal conduction loss is due to the flow of the heat energy across a temperature gradient. Denoting the thermal flux by ϕ , one can express the thermal conduction loss S_{th}, in ergs/(cm^3) (sec), as

$$S_{th}^{e,i} = -\nabla \cdot \phi^{e,i} \quad . \tag{17}$$

The thermal flux in turn is given by

$$\phi^{e,i} = -K_{e,i} \frac{1}{\gamma - 1} \nabla T_{e,i} \quad , \tag{18}$$

where K is the thermal conductivity of the plasma. The factor $1/\gamma-1$ stems from the relation between the internal energy and temperature given by (9).

In the absence of a magnetic field, an expression for K can be derived by means of the kinetic theory of gases, since the charged particles in the plasma behave like neutral particles to a first approximation. For our purpose, however, one has to include the effect of a magnetic field on the thermal conductivity. One obtains theoretically[3,4] that the thermal conductivity perpendicular to the magnetic field, which is the thermal conductivity of interest in θ-pinch, is given by

$$K_{e,i} = K_{i_{e,i}} T_{e,i}^{5/2} \frac{1}{1+\alpha(\tau_{e,i}\,\omega_{ge,i})^2} \quad , \tag{19}$$

where

K_1 = the mobility,

α = a numerical constant ($\alpha \simeq 1$),

τ = time between the collisions,

ω_g = gyrofrequency.

[3]M. N. Rosenbluth and A. B. Kaufmann, Physical Review, 109, 1 (1958)

[4]W. Marshall, "Kinetic Theory of an Ionized Gas III." AERE T/R 2419

The effect of the magnetic field is contained in $(\omega_g \tau)^2$ in the denominator.[5,6] The gyrofrequency for an electron is

$$\omega_{ge} = \frac{eB}{m_e c} \quad esu. \quad , \tag{20}$$

and the collision time (See Simon's lectures)

$$\tau \approx 2 \frac{T_e^{3/2}}{n_e} \quad sec. \quad . \tag{21}$$

For ions one has approximately

$$\omega_{gi} = \frac{m_e}{m_i} \omega_{ge} \tag{22a}$$

and

$$\tau_i \approx \sqrt{\frac{m_i}{m_e}} \; \tau_e \tag{22b}$$

Using these relations

$$\omega_{ge} \tau_e \approx 3 \times 10^7 \left(T_e^{3/2} \frac{B}{n_e} \right) \quad . \tag{23}$$

For a plasma with $n_e = 10^{17}$ electron/cm^3 and $T = 10^6$ ^oK,

$$\omega_{ge} \tau_e \approx 3 \times 10^{-1} B \quad , \tag{24}$$

which indicates that $\omega_{ge} \tau_e$ becomes large compared to unity even at high densities with low magnetic fields of the order of 100 gauss. Thus, the thermal conduction is inhibited sharply

[5] See for example; S. Glasstone and Ralph H. Lovberg, Controlled Thermonuclear Reactions, p. 96, D. Van Nostrand, Inc., Princeton, New York (1960).

[6] A. Simon, An Introduction to Thermonuclear Research, p. 30, International Series of Monographs on Nuclear Energy, Pergamon Press, New York (1959).

with only a few hundred gauss and decreases with the magnetic field and temperature as $1/T^3B^2$. In general, it can be concluded that the thermal conduction perpendicular to the magnetic field would probably not be a serious source of energy loss for a stable plasma. The thermal conduction parallel to the magnetic field does not exist in an infinitely long cylinder because there is no temperature gradient in that direction. However, it may be a serious source of energy loss through the ends in an actual experimental device with a finite length.

2. Thermal Relaxation (S_{rel})

There is a net energy transfer between the ions and electrons through the ion-electron collisions when the electron and ion temperatures are different. Assuming that both ions and electrons have a Maxwellian velocity distribution, the approximate rate of energy transfer is given by Spitzer[7] as

$$S_{rel}^{e,i} = \pm \frac{T_e - T_i}{\tau_{eq}} \frac{\rho}{\gamma - 1} \tag{25}$$

where τ_{eq} is the equipartition time. The factor $\gamma - 1$ comes from the fact that $(T_e - T_i)/\tau_{eq}$ gives the rate of energy transfer per unit mass, whereas S in (12) or (16) is defined as the heating rate per cm^3. The minus sign is for electrons and the plus sign is for ions. When T_e and T_i are widely different, then one should compute the relaxation in velocity space, for

[7]L. Spitzer, Physics of Fully Ionized Gases, p. 80, Interscience Publ., New York (1956).

example, via the Fokker-Planck equation. Such a more refined theory would take into account the departures from the Maxwellian velocity distribution but would be enormously more complicated.

3. Ohmic heating (S_{ohm})

This is due to the finite conductivity of the ionized gas. The conductivity of the plasma in the presence of a magnetic field depends on the direction. Since only azimuthal currents are present in a θ-pinch experiment, the ohmic heating involves only the transverse resistivity $\eta_\perp$ in the direction perpendicular to the magnetic field lines. For a strong magnetic field, the transverse direct current resistivity has been calculated to be[8]

$$\eta_\perp = 1.29 \times 10^{13} \frac{\ln \Lambda}{T_e^{3/2}} \quad e.m.u. \tag{26}$$

where the logarithmic term is of the order 10. The value of $\eta_\perp$ is about twice the value of the resistivity of the plasma in the absence of a magnetic field. The latter resistivity is also equal to the resistivity in the direction parallel to the magnetic field lines since the presence of a magnetic field influences only the transverse resistivity. The reason for this preferential influence is that in a magnetic field the electron velocity distribution normal to the field is quite different from the electron velocity distribution in the direction parallel to the field. In particular, the fraction of high-velocity electrons which contribute to the current is decreased as a

[8]Spitzer, p. 86.

result of the gyration, and thus, the effective ion-electron scattering cross section is increased. Therefore, the rate of momentum transfers to the ions, and thus the resistivity is increased in the normal direction. The motion of the electrons parallel to the field is not affected, and hence the parallel resistivity remains unaffected.

The resistivity of the plasma depends also on the frequency of the applied signal. However, if the electron-ion collision time is much less than the period of the applied signal, as it is in the case of the θ-pinch experiments, the D. C. resistivity is applicable.

Since one has only azimuthal currents with a density $J = \frac{c}{4\pi} \frac{\partial B}{\partial r}$ which follows from (7), the ohmic heating per cm^3 is $\eta_\perp J^2$. Thus,

$$S_{ohm} = \eta_\perp \left(\frac{c}{4\pi}\right)^2 \left(\frac{\partial B}{\partial r}\right)^2 \quad . \tag{27}$$

Combining equations (17), (25) and (27), one obtains the expressions for S_e and S_i appearing in (15) and (16) as:

$$S_e = \frac{1}{\gamma-1} \frac{1}{r} \frac{\partial}{\partial r} \left(K_e r \frac{\partial T_e}{\partial r}\right) - \frac{\rho}{\gamma-1} \frac{T_e - T_i}{\tau_{eq}} + \eta_\perp \left(\frac{c}{4\pi} \frac{\partial B}{\partial r}\right)^2 , \tag{28}$$

and

$$S_i = \frac{1}{\gamma-1} \frac{1}{r} \frac{\partial}{\partial r} \left(K_i r \frac{\partial T_i}{\partial r}\right) + \frac{\rho}{\gamma-1} \frac{T_e - T_i}{\tau_{eq}} - q_i \frac{1}{r} \frac{\partial}{\partial r} (rV) . \tag{29}$$

Note that the ohmic heating is included only in the source term for electrons. The reason for this is that the rate at which a charged particle gains energy from an electric field is proportional to the velocity of the particle. Since the electrons can move much more rapidly than the ions because of their smaller mass, nearly all of the energy absorbed by the field will be taken up by electrons. Hence, the ohmic heating of the ions can be neglected. Also one can see that the electrons carry most of the current so that the Lorentz magnetic force is exerted on the electron gas, while the ions are carried along with the electrons due to space charge.

Note again that the shock heating, the last term in equation (29), is confined only to the ions. This will be discussed in detail in later sections concerned with the shock structure.

It must be born in mind that one can have a shock wave even $q_i = 0$, as mentioned earlier when q_i was first introduced (cf. equation (6)). In this case the dissipation is due to the ohmic heating from currents flowing in the shock front. This dissipation mechanism is automatically included in the ohmic heating term. When $q_i = 0$, the shock thickness is determined by the skin effect. We shall return to the subject of shock heating later.

V. Some Comments Concerning the MHD Equations

Before going further, it seems appropriate to summarize the results which we have so far presented. We have a theory for a collision dominated ($\gamma = 5/3$) plasma which is suitable for investigating one-dimensional radial magnetic compression. The theory includes the ohmic heating of electrons, the shock heating of ions, the relaxational heating of ions (or cooling of electrons), and compressional heating of both ions and electrons. Although we have ignored the radiation losses of electrons and ionization losses of impurities, they can be included in the theory in principle. The theory is formulated mathematically by a set of non-linear coupled differential equations of (7e), (11), (13), (15) and (16) with the addition of (28) and (29), where the coefficients are not constant. One can investigate the radial distribution of V, T_e, T_i, B and n as a function of time by means of these equations under a variety of conditions.

When the collisions are infrequent during the times of interest, one can take $\gamma = 2$ (cf. equation 10) to approximate this situation. However, one loses rigor rapidly because all the transport coefficients in the theory are calculated for a Maxwellian plasma. When collisions are not important, the ion and electron distributions will depart from Maxwellian, and could become almost anything. For this reason, it is of interest to work at high enough densities so that the plasma is collision dominated up to kev temperatures. Under these conditions the present theory can be tested or at least can be used as a guide for designing experiments.

Another restriction on the validity of the present theory is that it is developed for an infinitely long plasma column. The finite length of the solenoid is expected to play a role for times $t \gtrsim (\ell/\bar{v}_i)$, where ℓ is the length of the plasma column and $\bar{v}_i$ is the average thermal velocity of the ions. For deuterons, $\bar{v}_i$ is given by

$$\bar{v}_i \approx 10^4 \sqrt{T_i} \quad \text{cm./sec.}$$

Thus, for an ℓ of a few meters, and $T_i = 10^6$ °K ($\bar{v}_i \sim 10$ cm/μ sec), the theory might be expected to describe the plasma state for times of the order of 5 to 10 μ sec.

Instabilities which could lead to turbulent diffusion and the like are also not taken into account in this theory. The types of instabilities observed in θ-pinches will be discussed briefly later.

VI. Alfvén Waves

In this section, we shall deal with the simplest case of wave propagation in the linearized magnetohydrodynamic theory, in particular with the Alfvén waves. The results of this discussion will be useful later when we deal with the subject of large amplitude shock waves in the nonlinear theory.

Consider the continuity equations (4) and Euler's equation (5), linearize them and eliminate the velocity between them. Then, in the absence of a magnetic field, one obtains a wave equation for the density and pressure perturbations in a compressible field,

$$\frac{\partial^2 p}{\partial t^2} = \frac{\partial^2 \rho}{\partial x^2} \quad .$$

Using this equation to express the polytropic relation $p = p(\rho)$, one gets

$$\frac{1}{c^2}\frac{\partial^2 \rho}{\partial t^2} = \frac{\partial^2 \rho}{\partial x^2} \quad ,$$

where the wave velocity is given by

$$c^2 = \frac{dp}{d\rho} \quad . \tag{30}$$

For adiabatic, isotropic waves, for which p/ρ^γ = constant, one obtains the ordinary sound waves with

$$c = \sqrt{\frac{\gamma p}{\rho}} = \sqrt{\frac{5}{3}\frac{kT}{m}} \quad , \tag{31}$$

where k is the Boltzmann constant.

In the presence of a magnetic field and in an incompressible fluid, one gets so-called Alfvén waves. If the fluid is compressible and the magnetic field is present, one gets coupled Alfvén waves and sound waves. A pure Alfvén wave results from the motion of a conducting fluid, which induces currents if there is a magnetic field present. These currents generate fields and reaction forces opposing the original motion. This leads to a wave motion where energy is transferred back and forth between the field and the fluid.

We shall now discuss Alfvén waves very briefly. Consider a plasma which is initially at rest. Suppose that an external magnetic field B_o is present in the z-direction and that the displacement velocity is in the y-direction. As mentioned earlier, we assume that the fluid is incompressible, i.e. ρ = constant. We shall look for a wave progagating parallel to the steady magnetic field, i.e. in the z-direction. Thus, all the time dependent quantities will be functions of z and t. The components of the total magnetic field are then given by:

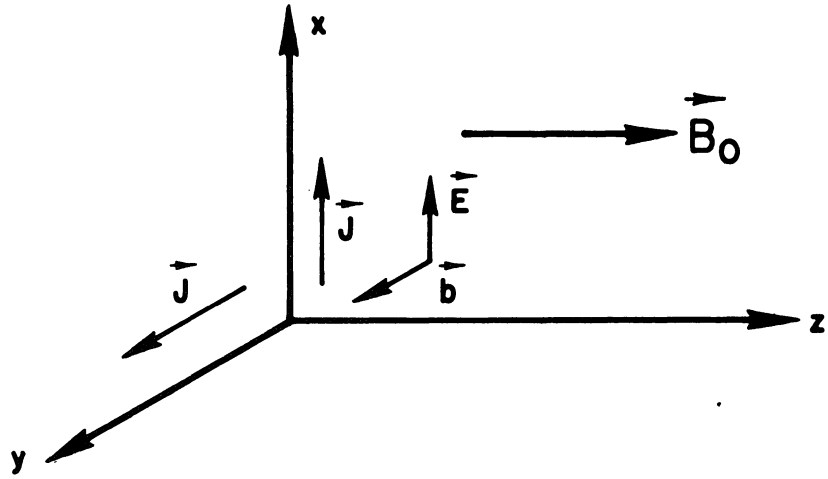
$$B_z = B_o$$

$$B_y = b(z,t)$$

$$B_x = 0$$

Maxwell's equations are:

$$-\frac{\partial b}{\partial z} = \frac{4\pi}{c} J_x \,, \quad J_y = J_z = 0 \tag{32}$$



Field Configuration for Discussion of Alfvén Waves

FIGURE 3

$$\frac{\partial E_x}{\partial z} = -\frac{1}{c}\frac{\partial b}{\partial t} \quad , \quad E_y = E_z = 0 \tag{33}$$

The generalized form of Ohm's law gives

$$J_x = \sigma\left[E_x + \frac{1}{c}\left(B_z v_y - B_y v_z\right)\right] \tag{34}$$

where the second term $\vec{V} \times \vec{B}$ is the electric field arising from the hydrodynamic motion of the plasma in a magnetic field.[9] The term $B_y V_z$ vanishes because in an incompressible fluid $\nabla \cdot \vec{V} = 0$ and thus V_z = constant. We can consider a case where no streaming in the z-direction is present, i.e. $V_z = 0$. Note also that σ refers to the transverse conductivity normal to the magnetic field.

The y-component of Euler's equation (5) gives

$$\rho \frac{\partial V_y}{\partial t} = \frac{J_x B_o}{c} \quad , \tag{35}$$

and its z-component is

$$\frac{\partial p}{\partial z} + \frac{B_y J_x}{c} = 0 \quad . \tag{36}$$

Substituting J_x from equation (32), one obtains

$$p + \frac{b^2}{8\pi} = \text{constant}, \tag{37}$$

[9] Spitzer.

which expresses pressure balance in the direction of B_o.

Again, substitute J_x from (32) into (34), differentiate it once with respect to z, substitute $\partial E_x / \partial z$ in the resulting equation from (33), differentiate now with respect to t, eliminate ($\partial^2 V_y / \partial z \partial t$) by means of (35), and get

$$\frac{\partial^2 b}{\partial t^2} = \frac{B_o^2}{4\pi} \frac{\partial^2 b}{\partial z^2} + \frac{c^2}{4\pi\sigma} \frac{\partial^3 b}{\partial z^2 \partial t} \quad . \tag{38}$$

In the case of infinite conductivity, this equation reduces to an ordinary wave equation with a speed of propagation

$$V_A = \frac{B_o}{\sqrt{4\pi\rho}} \tag{39}$$

and with a solution for $B_x = b(z,t)$ as

$$b(z,t) = A \sin \omega \left(t - \frac{z}{V_A} \right) \quad , \tag{40}$$

where A is a constant which determines the amplitude of the wave. The time average of the magnetic energy density is

$$\frac{\langle b^2 \rangle}{8\pi} = \frac{A^2}{16\pi} \quad , \tag{41}$$

where $\langle \; \rangle$ indicates the time average. One can also solve for the velocity V_y

$$V_y = \frac{-A}{\sqrt{4\pi\rho}} \sin \omega \left(t - \frac{z}{V_A} \right) \quad . \tag{42}$$

The time average of the kinetic energy density is found to be

$$\frac{\rho \langle V_y^2 \rangle}{2} = \frac{A^2}{16\pi} \quad . \tag{43}$$

Equations (42) and (43) indicate that the Alfvén waves correspond to a wave motion where the energy is transferred back and forth between the magnetic field and the motion of fluid. The Alfvén waves are often referred to loosely as hydromagnetic (or magnetohydrodynamic) waves.

When the conductivity is finite, the solution of equation (38) can be shown to be

$$b = A e^{i(\omega t - Kz)} , \tag{44}$$

where K is the root of the dispersion relation

$$\omega^2 = (V_A K)^2 \left(1 + i \frac{\omega c^2}{4\pi\sigma V_A^2}\right) . \tag{45}$$

Finding the real and imaginary parts of K, one can rewrite (44) as

$$b = A e^{-z/z_0} e^{i\omega\left(t - \frac{k}{\omega} z\right)} , \tag{46}$$

where z_o and k are the attenuation length and the wave number respectively. When the attenuation is small (σ is large), z_o and k can be approximated by

$$k = \frac{\omega}{V_A} , \quad z_0 = \frac{8\pi\sigma V_A^2}{\omega^2 c^2} . \tag{47}$$

One observes from (47) that the attenuation length decreases with increasing Alfvén velocity and increasing conductivity. This behavior is qualitatively different than for the electromagnetic waves, which are governed by a differential equation of the following form:

$$\frac{\partial^2 b}{\partial z^2} = \frac{\epsilon\mu}{c^2}\frac{\partial^2 b}{\partial t^2} + \frac{4\pi\sigma\mu}{c^2}\frac{\partial b}{\partial t} \quad .$$

The attenuation length, which is called skin depth in electromagnetic theory, is given by

$$z_0 = \frac{c}{\sqrt{2\pi\omega\sigma\mu}} \quad , \tag{48}$$

which becomes shorter with increasing conductivity.

The physical significance of the Alfvén speed, which is proportional to the magnetic field strength and inversely proportional to the square root of $4\pi\rho$, where ρ is the density of the conducting fluid, can be seen by taking the ratio of the Alfvén velocity to the sound speed. From (31) and (39), this ratio is

$$\frac{V_A}{c} = \sqrt{\frac{B_0^2}{4\pi\rho\gamma k T/m_i}} = \sqrt{\frac{B_0^2/8\pi}{nkT(\gamma/4)}} \quad . \tag{49}$$

It follows that (v_A/c) is of the order of magnitude of the square root of the ratio of the energy density of the magnetic field to the kinetic energy density of the plasma. Thus, if the Alfvén velocity is greater than the sound velocity, then the energy stored in the magnetic field is the dominant term in the total energy of the system. Hence, one may expect that the magnetic effects will be very important.

Experimental studies of Alfvén wave propagation are much easier, in principle, using plasmas instead of liquid conductors such as mercury or sodium. This follows because the mass density of a plasma is generally low so that the Alfvén velocity is very large and the attenuation length is correspondingly small.

VII. Semi-Quantitative Discussion of Heating Rates

We shall now discuss the various heating mechanisms and their relative importance in heating the plasma under various conditions.

1. Compressional Heating (Adiabatic Heating)

Adiabatic heating results from the compression of the plasma at a relatively slow rate, which makes it different from shock heating where the compression is rapid and inertial forces are significant. The compression is brought about by the radial force $\vec{J} \times \vec{B}$ where J is azimuthal and B is axial in a θ-pinch. The increase in temperature due to the compression can be estimated simply when there is no trapped magnetic field, in other words when the magnetic field inside the plasma is small. Since the compression takes place at a low rate, one can consider the steady state equations. From Euler's equation (5) and from (7), one obtains by ignoring the artificial viscosity term

$$\frac{\partial p}{\partial r} + \frac{1}{4\pi} B \frac{\partial B}{\partial r} = 0$$

which yields upon integration with respect to r

$$p + \frac{B^2}{8\pi} = \text{constant} \qquad (50)$$

The quantity $\frac{B^2}{8\pi}$, which is the energy density in the magnetic field, can be regarded as a "magnetic pressure" associated with the field. Then, equation (50) states that the sum of the kinetic gas pressure $p = p_e + p_i$ and the magnetic pressure is constant at any point within the plasma. If the plasma is confined completely, the kinetic pressure must fall to zero

outside the plasma. Hence, denoting the external magnetic field by B_e one obtains from (50)

$$p + \frac{B^2}{8\pi} = \frac{B_e^2}{8\pi} \quad . \tag{51}$$

It is observed that the magnetic field within the plasma is always less than the external field. Consequently, a plasma confined by a magnetic field appears to be diamagnetic. It is convenient to define a dimensionless ratio β as

$$\beta = \frac{p}{B_e^2/8\pi} = 1 - \frac{B^2}{B_e^2} \quad , \tag{52}$$

which is the relative magnitude of the kinetic gas pressure at a given point inside the plasma with respect to the external magnetic pressure.

If one ignores the trapped field inside the plasma (β = 1 everywhere), equation (51) becomes,

$$p = nkT = \frac{B_e^2}{8\pi} \tag{53}$$

which expresses pressure balance.

Now suppose that the external field is increased from B_o to B (we drop the subscript e). Then the pressure increases and the plasma is compressed from an initial radius R_o to R. The adiabatic law is

$$\frac{T}{T_o} = \left(\frac{n}{n_o}\right)^{\gamma-1}$$

The conservation of particles requires

$$\frac{n}{n_o} = \left(\frac{R_o}{R}\right)^2$$

Combining these two equations with the pressure balance relation (53), one finds

$$\frac{T}{T_o} = \left(\frac{R_o}{R}\right)^{2(\gamma-1)} = \left(\frac{R_o}{R}\right)^2 \quad \text{for } \gamma = 2 \qquad (54)$$

$$= \left(\frac{R_o}{R}\right)^{4/3} \quad \text{for } \gamma = 5/3 \,,$$

or in terms of the initial and final values of the external magnetic field

$$\frac{T}{T_o} = \left(\frac{B}{B_o}\right)^{2(1-1/\gamma)} = \frac{B}{B_o} \quad \text{for } \gamma = 2 \qquad (55)$$

$$= \left(\frac{B}{B_o}\right)^{4/5} \quad \text{for } \gamma = 5/3$$

For γ = (5/3) and for a volume compression of ten, i.e. $(R_o/R)^2 = 10$, one obtains

$$T = T_o \times 10^{2/3} \sim 5T_o$$

and

$$B = B_o \left(\frac{R_o}{R}\right)^{5/3} = B_o \times 10^{5/6} \sim 8B_o \,. \qquad (56)$$

Thus, if one had a plasma of 2 kev, a volume compression of 10 with a field increase of about a factor 8 would raise the temperature to 10 kev. For γ = 2 the above relations also apply if $\beta < 1$ since the magnetic field behaves essentially as a fluid with two degrees of freedom.

To actually produce a 10 kev plasma in a θ-pinch one would need to "preheat" initially to a few kev. Because of the relatively high plasma densities the initial fields would also have to be fairly high, perhaps 40 or 50 k gauss. Then with final fields of 200-400 k gauss it might be possible to reach temperatures of thermonuclear interest with time-scales of the order milliseconds. The technical and scientific problems connected with such an enterprise would be tremendous, especially since the basic physics is not very well understood. Here we shall content ourselves with the "preheating" phase, where the more modest question is how to reach plasma "temperatures" of 20 or 30 million degrees for some tens of microseconds. This can be done in at least two ways:

a) inject a plasma with energies in the kev range;

b) use a combination shock and ohmic heating with subsequent compression.

We shall only discuss the second possibility here although θ-pinch injection experiments are also in progress at Los Alamos. First, we have to consider the structure of shock waves.

2. Shock Waves and the Rankine-Hugoniot Relations

Suppose that one raises the magnetic pressure very rapidly, that is, in a time short as compared to the time sound waves or Alfvén waves traverse the medium. Then shock waves are formed which can lead directly to non-adiabatic, irreversible heating of the ions because of viscous stresses and of the electrons due to finite electrical conductivity. If there is preferential heating of either ions or electrons under special circumstances

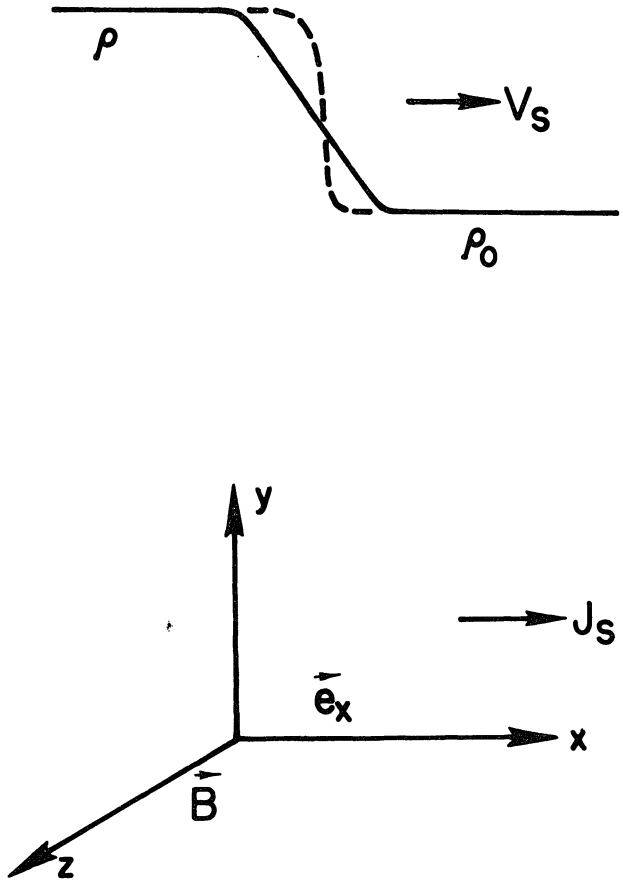
then the other component is heated indirectly by relaxation processes.

The mechanism of the formation of a shock wave in a θ-pinch can be described as follows: the sudden increase in the external magnetic field produces high diamagnetic currents and a large pressure jump in a localized region near the outer surface of the plasma. A compression wave is then propagated through the plasma. The propagation velocity increases with the pressure, with the result that the following portion of the pressure wave travels faster and tends to catch up with the leading edge. Thus, the wave front becomes increasingly steeper, forming eventually a shock front.

At the leading edge of the shock front the flow is not adiabatic. The entropy can increase due to ion collisions and ohmic dissipation, ionization or charge-exchange collisions, and very likely by instabilities for sufficiently high Mach numbers where the mean-free-paths are very large and the usual viscosity is small.

Suppose that there is a pressure discontinuity, and ρ_0 and ρ are the densities in front and behind a shock wave propagating with velocity V_s to the right. If there is a magnetic field B present; associated with this discontinuity, currents will be induced in the plasma near the shock front.

It is clear that the width of the shock transition will influence the current density J in the shock front, and in



Shock Wave in a Transverse Magnetic Field

FIGURE 4

turn this will influence the rate of the ohmic heating, $J^2\eta$. Thus, it is of importance to know, or at least to estimate, the connection between the shock thickness and the relative heating of ions and electrons. In the hydromagnetic equations, as we have written them down, we have simply put an artificial viscosity term in the ion-energy equation. But this, as we have mentioned earlier, really hides our ignorance of the factors which really determine the structure of the shock wave in the presence of a magnetic field. If one first ignores the dissipative mechanisms which determine the detailed structure of the shock wave, one can relate the values of the flow variables p, V, ρ, B, $\mathcal{E}$ well behind the shock discontinuity to those ahead of the shock discontinuity from the stationary solutions of the magneto-hydrodynamic equations. These relations, as mentioned earlier, are the generalizations of the famous Rankine-Hugoniot relations for the conservation of mass, momentum, and energy in the absence of a magnetic field. Since the derivation of the Rankine-Hugoniot relations with a magnetic field can be found in the literature[10,11,12], we shall just state the results for a shock wave propagating through a transverse magnetic field, since this is the situation of most interest in a θ-pinch experiments.

[10] DeHoffman and Teller, Phys. Rev., 80, 692 (1950).

[11] R. Lüst, Z. Naturforsch., 8a, 277 (1953).

[12] H. L. Helfer, Astrophys. J., 117, 177 (1953).

Assume that the plasma has an infinite conductivity. This assumption is consistent with our neglect of all dissipative mechanisms. Thus, all electric field components vanish for an observer moving with the shock front and the flow will appear time-independent. Therefore, one can use the stationary solutions of the MHD equations in a frame of reference moving with the shock front. Owing to the low shock velocities involved, the relativistic effect will be ignored. Hence, the magnetic field in the moving system will be identical to that in the coordinate system at rest. Let $\vec{u}$ be the flow velocity in the moving frame and $\vec{v}$ in the laboratory coordinate system. Then $\vec{u} = \vec{V} - \vec{V}_s$ follows. Denoting the jump of a quantity across the shock front by the bracket $\langle\ \rangle$, one obtains the following conservation laws:

$$\langle \rho u_x \rangle = 0 \quad , \qquad \text{(conservation of mass)} \qquad (57)$$

$$\langle M\vec{u} + \left(p + \frac{B^2}{8\pi}\right)\vec{e}_x \rangle = 0 \quad , \quad \text{(conservation of momentum)} \qquad (58)$$

$$\langle M\frac{u^2}{2} + M\frac{T}{\gamma - 1} + M\frac{B^2}{8\pi\rho} \quad \text{(conservation of energy)} \qquad (59)$$

$$+ u_x\left(p + \frac{B^2}{8\pi}\right)\rangle = 0 \quad ,$$

$$\langle B u_x \rangle = 0 \qquad \text{(conservation of flux)} \qquad (60)$$

where $\vec{e}_x$ is the unit vector in the x-direction, and M is the conserved value of ρu_x, i.e.

$$M \equiv \rho_0 u_{x_0} = \rho u_x \quad ,$$

where ($\rho_o u_{x_o}$) and (ρu_x) are the quantities in front of and behind the shock front. Since the flow velocity is zero in front of the shock in a stationary plasma in the laboratory frame, $u_{x_o} = V_s$. Thus, M is the mass swept over by the shock front per unit time per unit area.

Taking the y-z components of (58), one finds the tangential flow velocity components are the same on both sides of the shock front, i.e. $\langle u_y \rangle = \langle u_z \rangle = 0$, which also implies $\langle v_y \rangle = \langle v_z \rangle = 0$. One recognizes in the energy equation (59) that the first term is the kinetic energy in the flow, the second term is the internal energy, the third is the energy in the magnetic field, and the fourth term is the work done in transporting the fluid.

The conservation equations (57), (58), (59), and (60) can be solved simultaneously to yield the flow parameters behind the shock front. In the limit of strong shocks, where the pressure jump p/p_o is much larger than unity, as is the case in our experiments where the shock heating can carry the temperature to quite high values, one finds that the density jump is approximately given by

$$\frac{\rho - \rho_o}{\rho_o} \approx \frac{2}{\gamma - 1} \quad \text{or} \quad \frac{\rho}{\rho_o} = \frac{\gamma + 1}{\gamma - 1} = 4 \quad \text{for} \quad \gamma = 5/3 \,. \tag{61}$$

Combining (57) and (60) yields

$$\frac{B}{\rho} = \frac{B_o}{\rho_o} \,, \tag{62}$$

which is a general relation not depending on the strong shock approximation. It then follows from (61) and (62) that

$$\frac{B-B_0}{B_0} \simeq \frac{2}{\gamma-1} \quad . \tag{63}$$

These are two results of the Rankine-Hugoniot relations which we shall presently need.

3. Relative Heating Rates of Ions and Electrons in a Shock Front

We shall now give a qualitative discussion of the factors that determine the relative heating of the ions and electrons. It will be apparent presently that the heating rates depend on the factors which determine the shock thickness, i.e. various dissipation mechanisms both heat the plasma and broaden the shock front.

Let us denote the ohmic heating rate per unit area of electrons by $\mathcal{E}_j$. This is of the order of the integral of $J_y^2 \eta_\perp$ across a shock front with a thickness L, i.e.

$$\mathcal{E}_j = \int_0^L J_y^2 \eta_\perp \, dx \quad , \tag{64}$$

where $\eta_\perp$ is the transverse resistivity. The current density is given by Maxwell's equations (Gaussian units)

$$\frac{4\pi}{c} J_y = -\frac{\partial B}{\partial x} \simeq -\frac{\Delta B}{L} \quad . \tag{65}$$

The shock thickness may be determined primarily by dissipative mechanisms associated with either the ions or electrons, or by a combination of both. With (65), $\mathcal{E}_j$ becomes

$$\mathcal{E}_j \simeq \frac{c^2}{(4\pi)^2} \frac{\eta_\perp}{L} (\Delta B)^2 = \frac{4c^2}{(4\pi)^2} \frac{B_0^2}{(\gamma-1)^2} \frac{\eta_\perp}{L} \quad , \tag{66}$$

where B is substituted from equation (63). We see that the thicker the shock the less is the rate of ohmic heating. We have assumed that the resistivity and current density in equation (64) represent average values.

Let $\mathcal{E}_s$ denote the flow energy transported per second per unit area by the ions. This is of the order of

$$\mathcal{E}_s \simeq \frac{\rho V^2}{2} V_s \quad , \tag{67}$$

where V and V_s are the flow and shock velocities as defined earlier, and where ρ is the density behind the shock front. Using equation (61) for ρ gives

$$\mathcal{E}_s \simeq \frac{\gamma+1}{\gamma-1} \frac{\rho_0 V^2}{2} V_s \quad . \tag{68}$$

We now introduce a local Alfvén Mach number

$$M_A \equiv \frac{V}{V_A} = \frac{V\sqrt{4\pi\rho_0}}{B_0} \quad . \tag{69}$$

If there is negligible dissipation for the ions then $V \sim V_s$. We then deal with an ion beam coupled by space charge forces to the electrons. If a strong dissipation exists then the shock velocities are still numerically comparable to the flow velocities according to the conservation equations.

The ratio $\mathcal{E}_s / \mathcal{E}_j$ follows from equations (64) and (67) as

$$\frac{\mathcal{E}_s}{\mathcal{E}_j} \simeq \frac{(\gamma+1)(\gamma-1)}{8} M_A^2 \frac{L}{L_r} \tag{70}$$

where we have introduced L_r, which is the characteristic length (skin depth) associated with the finite resistivity,

$$L_r = \frac{c^2}{4\pi}\frac{\eta_\perp}{V} \simeq \frac{10^{13}}{\bar{T}_e^{3/2} V} \quad cm., \tag{71}$$

where we take the logarithmic term in the resistivity equal to 10. The average electron temperature in the shock front $\bar{T}_e$ is in ^oK and v is in cm/sec. To give a numerical example, which is typical for many present experiments: for $\bar{T}_e \sim 10^5$ ^oK in the shock front and $v \sim 10^7$ cm/sec, then $L_r \simeq 0.3$mm. We see then that the thickness of a resistive shock can be quite small.

For $\gamma = 5/3$, the ratio (70) becomes

$$\frac{\mathcal{E}_s}{\mathcal{E}_j} \simeq \frac{2}{9} M_A^2 \frac{L}{L_r} \tag{72}$$

One observes from this expression that if the shock thickness L is equal to L_r, i.e. if the shock thickness is determined by electron collisions and the resistivity, then the ions transport relatively more energy in the shock zone than the electrons for Alfvén Mach numbers greater than about $\sqrt{9/2}$. Thus, for Alfvén Mach numbers greater than about 2, one can also expect that thermal heating of the ions in the shock front will dominate over ohmic heating of electrons under certain circumstances, provided $L \gtrsim L_r$.

This discussion is so far based on the assumption that the electron velocity distribution is Maxwellian and the resistivity is given by the usual formula. It must be borne in mind that if the currents are too high, then the electron drift velocity

could exceed the electron thermal velocities. This in itself could lead to two-stream instabilities resulting in anomalous diffusion rates which could broaden the front. However, complete numerical solutions of the resistive shock structure show that the drift velocities are usually small compared to the electron thermal velocities for most cases of laboratory interest.

4. Influence of Ion-Ion Collisions (Viscosity) on Shock Structure

If we consider first the case where the shock thickness L is controlled by the viscosity of the ions, i.e. by ion-ion collisions, then the shock thickness will be of the order of the flow velocity times the relaxation time for ion-ion collisions, $L = L_{coll} \simeq \tau_{ii} V \simeq 1.6 \bar{T}_i^{3/2} V/n$, where T_i is in ^{o}K and n is ions/cm^3. Again the logarithmic term is taken to be 10 in Spitzer's result for the ion-ion collision time τ_{ii} .[13]

When ion-ion collisions are frequent then the ordered flow kinetic energy is partially dissipated and a thermal component of the ion velocity distribution develops. Thus the ions are heated by work done against the viscous stresses and the internal energy $\bar{T}_i / \gamma - 1$ is comparable to the flow kinetic energy according to the conservation equations. If the shock heating of ions dominates the ohmic heating of electrons, then the electrons are first heated by adiabatic compression in the shock front and subsequently by ion-electron collisional relaxation. Under such circumstances one would then expect the shock thickness to be

13 Spitzer, p. 78.

controlled by the viscosity and not by the skin effect due to the currents. This is confirmed by solutions obtained from the nonlinear equations. In terms of the particle density, the ion and electron dissipation lengths are comparable when $L_{coll} = L_r$, or when

$$n = n_c \equiv 2 \times 10^{-13} \left(\bar{T}_e \bar{T}_i\right)^{3/2} V^2/cm^3 \quad . \tag{73}$$

If we neglect the exchange of energy between the ions and electrons to first order, then according to the conservation equations $m_i V^2 \simeq 3k \bar{T}_i$. Using this in the above relation for the critical density n_c, for which the two dissipative lengths are about equal, one finds

$$n_c \simeq 10^{-25} \bar{T}_e^{3/2} V^5/cm^3 \quad . \tag{74}$$

which depends strongly on the velocity. For example, with $\bar{T}_e \sim 10^5$ °K and $V \sim 10^7$ cm/sec, $n_c \simeq 10^{17}/cm^3$; and for $\bar{T}_e \sim 5 \times 10^4$ °K and $V \sim 7 \times 10^6$ cm/sec, corresponding to a somewhat slower shock, one finds $n_c \sim 6 \times 10^{15}$ /cm^3. At higher densities, for these particular examples, the shock thickness is likely to be connected with the resistive skin depth. Said another way, at high densities the spatial extent of the viscous heating zone becomes very small and ohmic dissipation determines the physical processes in the shock transition.

Finally, one concludes that the production of shock waves in the laboratory with a thickness determined by the viscous stress should not be too difficult for densities in the range of 10^{15}

particles/cm^3 and velocities $\sim 10^7$ cm/sec.

Consider now the ratio $\mathcal{E}_s/\mathcal{E}_j$ for the case where $L = L_{coll}$. The equation (72) becomes

$$\frac{\mathcal{E}_s}{\mathcal{E}_j} \simeq 10^{-49} \frac{\bar{T}_e^{3/2} V^7}{B_o^2} \quad . \tag{75}$$

Note first that this ratio does not depend on the density. Here we have again used $m_i v^2 \simeq 3\ k\bar{T}_i$. Therefore, equation (75) gives the ratio of the thermal energy transported by the ions and electrons. For the previous example ($\bar{T}_e \sim 10^5$ ^oK and $V \sim 10^7$ cm/sec) with $B_o \sim 1000$ gauss, this ratio is about 30. For slightly lower velocities of 7×10^6 cm/sec, the ratio is reduced by about a factor of 10. These conclusions are substantiated from numerical solutions of the full hydromagnetic equations as discussed later.

5. Comments on Collision-Free Shock Waves

Let us turn now to the situation at high Alfvén Mach numbers where $L_{coll} \gg L_r$. Under such circumstances one would conclude from equation (72) that shock heating of the ions is generally the important factor at high Alfvén Mach numbers and that the shock width is connected with the ion-mean-free path. However, it could turn out that with very long ion mean-free-paths for high temperature, that other mechanisms determine the dissipation of the ion flow energy. This is likely since according to equation (49) the magnetic energy density is then large compared to the kinetic energy density of the plasma, and instabilities in the flow may develop. If some kind of turbu-

lence (about which next to nothing is known) exists, which is driven by some sort of instability, then the ordered flow energy could conceivably be randomized over a distance which is short compared to the ion mean-free-path but which is large compared to the resistive skin depth. Such shock waves are a class of so-called "collision-free" shock waves.

It is still an open and controversial question whether or not such collision-free shock waves exist at all. If shock waves of this kind actually occur, then a knowledge of this fact could have great significance in understanding the heating mechanisms associated with shock waves propagating into the solar corona and the interplanetary "solar wind", and the heating mechanisms associated with high energy shock waves in controlled thermonuclear fusion experiments. A large number of theories have been proposed for collision-free shock waves. It is not our purpose here to review these theories, but to discuss some of the experimental consequences of their existence in a qualitative way.

Let us suppose as an exercise that there is such a thing as a collision-free shock wave whose thickness is connected with the ion cyclotron radius r_{ig} . (There may well be other lengths which are more important for collision-free shock waves. Our purely intuitive choice of r_{ig} as such a length was only made to illustrate the general ideas.) One can then obtain an expression for L/L_r as

$$\frac{L}{L_r} = \frac{r_{ig}}{L_r} \simeq \frac{m_i c}{e B_o} \frac{V}{L_r} \qquad (76)$$

Taking $v \sim v_s$ for a deuterium plasma and using equation (71) for L_r, this ratio is approximately

$$\frac{r_{ig}}{L_r} \approx 2 \times 10^{-17} \frac{v^2 T_e^{3/2}}{B_o} \qquad (77)$$

Suppose a collision-free shock wave propagates into a fully ionized gas with $T_e \sim 10^5$ °K, $v \sim 10^7$ cm/sec and $B_o \sim 1000$ gauss. One finds $(r_{ig}/L_r) \sim 60$, so that shock heating is very effective. However, for such a shock wave to be collision-free, the characteristic length chosen here, mcv/eB_o, should be small compared to the collisional length $\tau_{ii} V$. This condition reduces to

$$\omega_{ig} \tau_{ii} >> 1$$

If the collision frequency is large compared to the gyrofrequencies, i.e. if there are many ion-ion collisions during a gyro-period, then the situation should be similar to the ordinary magnetohydrodynamic case with an ordinary collisional viscosity. Therefore, one would expect here also the thickness would be connected with the mean-free-path. Now as we have indicated, the situation could become much more complicated if the ion mean-free-path is large compared to other possible dissipation lengths such as the ion cyclotron radius. Then one can conceive of a situation where the density and the temperature discontinuities can occur over lengths which are quite small compared to a

mean-free-path, forming a collision-free shock front.

To estimate when one might expect collision-free shocks of this particular type, we consider the collision time for 90° multiple scattering of a deuteron beam with velocity v

$$\tau = \frac{1}{n \sigma_c V} \tag{78}$$

For multiple coulomb scattering[14]

$$\sigma_c \simeq \frac{80 \pi e^4}{m_i^2 V^4} \simeq 10^{12}/V^4 \ cm^2$$

Therefore

$$\omega_{ig} \tau_{ii} \simeq 5 \times 10^{-9} \frac{V^3 B_0}{n} \gg 1 \ . \tag{79}$$

According to this inequality, the ion density should be smaller than $5 \times 10^{15}/cm^3$ for $v \sim 10^7$ cm/sec and $B \sim 10^3$ gauss, which are arbitrarily taken as typical experimental parameters. Collision-free shocks could also exist at higher densities for correspondingly higher flow velocities.

From the experimental point of view, one might phrase the question of such collision-free (no ion-ion collision) shock waves as follows: is there a dissipation mechanism for the ions which is associated with ion pressure jumps, satisfies the Rankine-Hugoniot relations and has a transition length less than the ion mean-free path and greater than the resistive skin depth? These crude estimates given above serve to indicate the kind of experimental conditions that might result in the production of collision-free shock waves in the laboratory.

[14] Simon p. 15

One can appreciate the highly nonlinear nature of collision-free shock waves and inquire by what general process they will form in a plasma. There are many speculative theories in this connection, but one can still give a qualitative picture within the framework of the magnetohydrodynamic theory. However, only experiments can provide the final justification for such an approach.

Suppose that one accelerates a plasma with an initial electron temperature of a few ev by means of $\vec{J} \times \vec{B}/c$ forces. The ions will be carried along at a high velocity due to the strong space-charge fields. The conductivity of the electrons will usually be sufficient to result in a narrow skin layer at the front. Since the electron-electron collision time is usually small compared to the other times of interest, the usual formula for the electrical conductivity may have some validity even at quite low densities. Such a resistive shock will result in large gradients in the flow velocity, electron temperature and density at the front. Then because of these steep gradients, an instability of some sort may develop which dissipates some fraction of the flow kinetic energy and increases the entropy of the ions.

The two-fluid MHD equations might still be used to describe the general structure of such shock waves if the characteristic length for the dissipation of the ion flow energy were known either experimentally or theoretically. This length could then be used in the von Neumann term to compute the ion heating. Such

a calculation would be of especial significance (1) when the electron gas is Maxwellian, since then one then knows the electrical and thermal transport coefficients of the electrons; and (2) since one of the most important factors for the determination of the relative heating rates of the ions and electrons is the ratio of a characteristic randomization length for the ions and ohmic skin depths, which can be estimated from the full nonlinear theory.

Finally we should mention that we have not considered here energy transfer mechanisms in a fully-ionized, pure deuterium plasma which are associated with such things as the absorption and emission of Cerenkov and cyclotron radiation, the influence of electron inertia on the resistivity (which is a factor in the theory of collision-free "solitary waves"), the acceleration of ions by voltages associated with instabilities or the finite resistivity and high currents, departures from a Maxwell electronvelocity distribution ("run-away" electrons in high induction fields), or the nonlinear interaction and scattering of waves in the shock transition zone if there are hydromagnetic instabilities (which are excluded from the present numerical solutions which only depend on one spatial coordinate). This final comment is intended to emphasize the extraordinary complexity of the collision-free shock problem and to qualify the preceding remarks and estimates. This problem is one of the most challenging in plasma physics.

6. Estimates of Shock Velocities

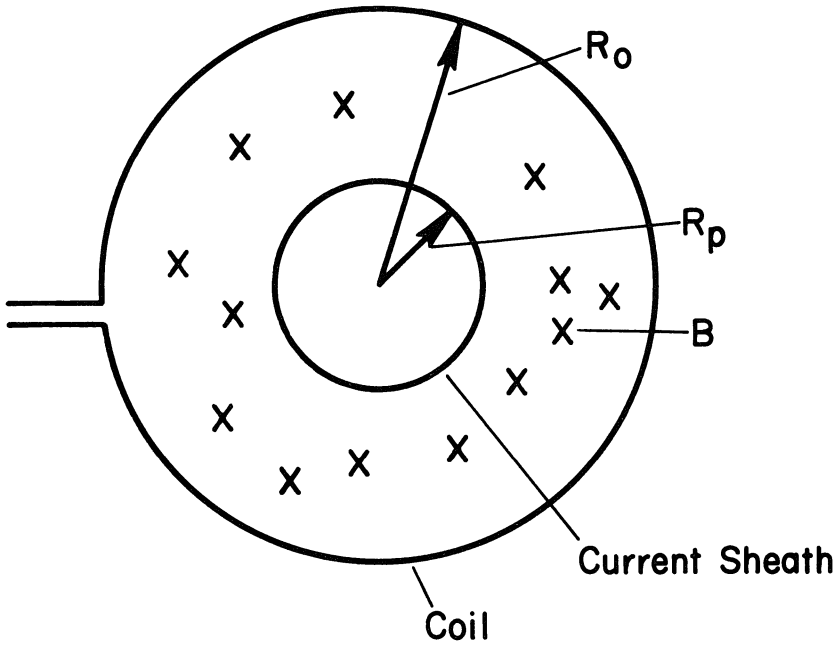
In order to estimate the shock velocities for a given rate of current rise, it is not necessary to solve the full set of hydromagnetic equations. This is because the electrical conductivity rises very rapidly during the early stage of the discharge, making a very thin current sheath which separates the plasma in the major volume of the discharge tube from the magnetic field outside. If the magnetic field or the magnetic pressure is made to rise sufficiently fast to values which are large compared to the gas kinetic pressure at the start of the discharge, then one can neglect the gas pressure in the hydromagnetic equations for estimating the gross dynamic behavior. In other words, one can estimate the dynamic behavior by simply balancing the inertial forces against the magnetic forces. If the situation is such that the magnetic pressure rises to a value large compared to the gas pressure in a time short compared to the time it takes the plasma to propagate in the center, a shock wave can form.

On the basis of these simplifications, one can write

$$\frac{dmv}{dt} = \frac{B^2}{8\pi} 2\pi R_p \quad , \tag{80}$$

where R_p is the radius of the plasma.

The situation is illustrated in figure 5 where R_o is the initial radius. The magnetic field B outside the plasma is shielded from the inside by the currents induced in the plasma as a result of the high electric conductivity. One can partially justify this simple model by photographing the implosion of the



Snowplow Model of a Dynamic θ-Pinch

FIGURE 5

plasma in an actual experiment. One sees, in fact, that the plasma does implode in a thin shell with a thickness of 1/10 or even smaller of the radius of the discharge tube. This is the basis of what is known as the "snowplow"[15, 16] model of the constricted discharge. The mass in the shell accumulates as the shell sweeps in over the undisturbed gas initially in the tube. Thus, the mass of the shell at any instant of time corresponding to a radius R_p is

$$m = \pi \rho_o (R_o^2 - R_p^2) \tag{81}$$

For a solenoid of length ℓ with a single turn, the magnetic field is

$$B = \frac{4\pi I}{c\ell} \tag{82}$$

Defining $y = (R_p/R_o)$, one can obtain a differential equation for y by combining equations (80), (81), and (82):

$$\frac{d}{dt}(1-y^2)\frac{dy}{dt} = -\frac{4\pi I^2}{\rho R_o^2 c^2 \ell} \tag{83}$$

Thus, if one knows the time behavior of the current I, one can then solve this equation numerically for y, i.e. for the position of the imploding shell as a function of time. One can obtain

15 M. Rosenbluth, R. Garwin, A. Rosenbluth, USAEC Report LA-1850 (1954).

16 M. Leontovich and S. M. Osovets, J. Nuclear Energy, 4, 209 (1957).

from the circuit equations

$$V_p = V(o) - L_e \frac{dI}{dt} \quad , \tag{84}$$

where V_p is the plasma voltage, V(O) is the initial voltage of the capacity bank and L_e is the inductance of the external circuit. In this equation, we neglect the term $(1/C) \int I\, dt$, i.e. we assume that the implosion time is short compared to the discharge time of the capacitor bank, so that the voltage across the capacitor bank does not change very much during the implosion time. This is usually a good approximation. In a typical experiment, the discharge time may be 10-20 μ sec., and the implosion time is usually less than 0.5 μ sec.

The plasma voltage is

$$V_p = \frac{d(LI)}{dt} \quad , \tag{85}$$

where LI is the magnetic flux and L is the inductance associated with the discharge tube with a moving conductor in it. The moving conductor is, of course, the imploding plasma. Now L is time dependent because of the motion of the plasma, and can be calculated from

$$\int \frac{B^2}{8\pi} dv = \frac{1}{2} L I^2 \quad ,$$

where the integral is extended over the region not occupied by the plasma. The magnetic field is uniform in the solenoid, so B can be taken outside the integral. Using the solenoid formula, one obtains for L,

$$L = \frac{4\pi^2}{c^2 \ell}\left(R_o^2 - R_p^2\right) \tag{86}$$

Substituting for L and V_p in equation (84), one obtains the current

$$I(t) = \frac{V(o)}{L(o)} \frac{1}{(1-y^2)+\left[L_e/L(o)\right]} t \tag{87}$$

where L(O) is the inductance of the discharge tube when it contains no plasma, i.e.

$$L(o) = \frac{4\pi^2 R_o^2}{c^2} .$$

Substituting for I(t) in equation (83) from equation (87), one obtains

$$\frac{d}{d\tau}(1-y^2)\frac{dy}{d\tau} = \frac{-y\tau^2}{\left\{(1+y^2)+\left[L_e/L(o)\right]\right\}^2} ,$$

where

$$\tau \equiv \left[\frac{4\pi^2}{M\ell}\frac{V^2(o)}{L^2(o)}\right]^{1/4} t \tag{88}$$

and where M is the mass per unit length,

$$M = \pi R_o^2 \rho .$$

By numerical integration, one finds for $\left[L_e/L(O)\right] = .10$ so that

$$v_p = \frac{dR_p}{dt} = 0.7\left[\frac{\pi}{25}\frac{1}{\rho}\left[\frac{V(o)}{L(o)}\right]^2\left(\frac{R_o}{\ell}\right)^2\right]^{1/4} \tag{89a}$$

where v_p is the speed of the plasma just before it reaches the center. Measuring L(O) in henries, V(O) in volts and lengths in cm, one obtains

$$v_p = 2.8 \times 10^3 \frac{\sqrt{V(0)}}{M^{1/4}} \quad \text{cm./sec.} \quad . \qquad (89b)$$

One gets a similar result for the linear Z-pinch but with a different numerical coefficient. This result has been verified quite accurately, particularly at the Institute for Atomic Energy in Moscow where the implosion velocities were measured over a wide range of pressures with deuterium, xenon and argon, in other words, with different masses per unit length. The $V^{\frac{1}{2}}$ and $M^{-\frac{1}{4}}$ dependence was verified. Even more remarkably, the implosion times one calculates with this simple-minded theory agree within 10-20% with the measured values. All these indicate that the snowplow model has some utility during the first dynamic phase when the inertial term in the hydromagnetic equations is large. This simple equation is very useful for designing experiments and is an example of a situation where some crude order of magnitude estimates yield quite good results. One can also calculate the velocities by numerical integration from the full hydromagnetic equations, but the results are similar within 10-15%.

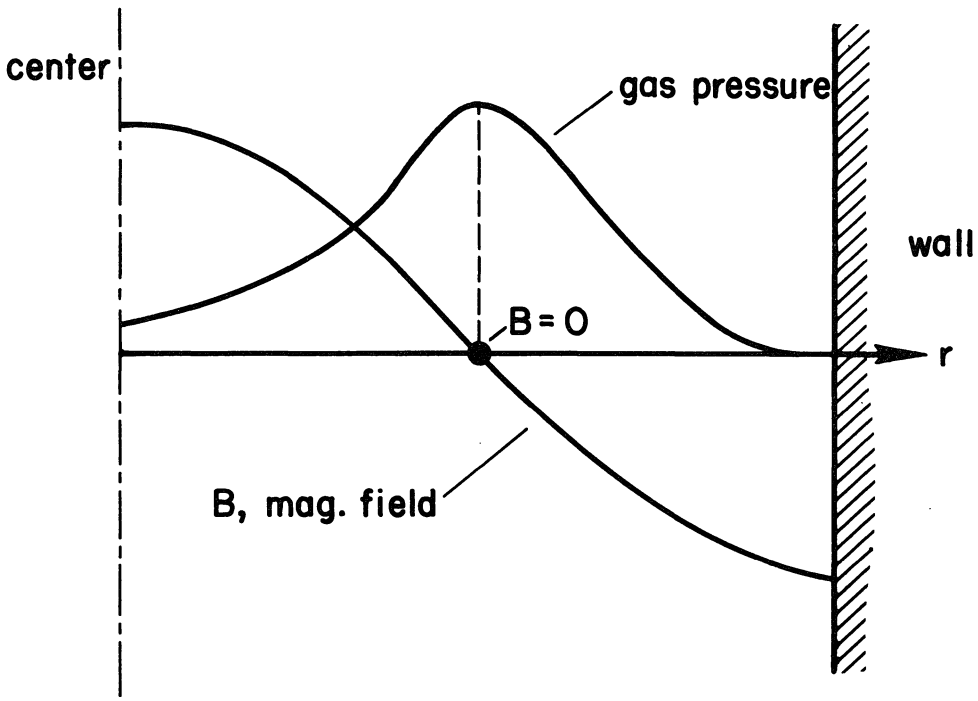
As a numerical example, suppose we apply 40kv to the capacitor bank, which is a reasonable figure. Let the initial density be 10^{15} particles/cm^3. which corresponds approximately to 10 micron initial deuterium pressure. Take the initial

radius to be R_o = 5cm. The calculated velocity for this case is about 10^8 cm/sec., which corresponds to about 70ev. With a reversed, trapped magnetic field, one also has to take into account the magnetic pressure of the field ahead of the shock wave, which gives rise to a restoring force. The calculations in the case of a reversed field have been done by numerical integration of the full hydromagnetic equations.

7. Reversed Field Heating

Reversed field heating turns out to be very important in the θ-pinch experiment. The magnetic field configuration when there is a reversed field present is shown on the following page. There is a current sheath somewhere in the plasma to support the reversed field which is initially produced by the quasi-static magnetic field. The vacuum magnetic field near the walls is produced by the external coil. This magnetic field configuration is similar to that in the ASTRON experiments at the Lawrence Radiation Laboratory at Livermore. There, however, the currents are provided by relativistic electrons supplied by an accelerator.

The gas particles pile up at the point where the magnetic field vanishes, i.e. B = 0. Simple pressure balance considerations would show that the plasma is driven into a region where the magnetic field has a minimum. Thus, the gas pressure distribution looks like the indicated curve. This distribution can be observed experimentally. It can also be found in detail by numerical integration of the hydromagnetic equations. The



Field and Pressure Distribution in a θ-Pinch With a Trapped Reverse Field

FIGURE 6

quantity β , which was defined as the ratio of the gas pressure to the pressure of the confining magnetic field, becomes unity at the point where $B = 0$.

There are several advantages to this configuration.

a. There is a region in the plasma where $\beta = 1$ ($B_{inside} = 0$). Thus, from the definition of β , one has

$$P_{gas} = \frac{B^2_{outside}}{8\pi}$$

which represents the ideal condition for confinement. In this region, the plasma is perfectly diamagnetic.

b. The energy stored in the reversed magnetic field is stored as a result of the currents in the plasma. That means if the energy in the reversed field is dissipated by some mechanism such as ohmic heating or even by some fast turbulent diffusion, then the magnetic energy must be transferred to the plasma. Consequently, the reversed field dissipation can provide a very powerful heating mechanism, and indeed it does seem to do so experimentally.

One can make a little order of magnitude estimate of this heating by using flux conservation during the initial compression, i.e. the initial field B_i times the initial volume V_i is of the order of the final field B_f times the final volume V_f. Thus,

$$B_f V_f \simeq B_i V_i \quad ,$$
$$\mathcal{E}_f = \frac{B_f^2}{8\pi} V_f = \left(\frac{B_i^2 V_i}{8\pi}\right) \frac{V_i}{V_f} \quad , \tag{90}$$

where $\mathcal{E}_f$ is the final energy in the trapped field. The ratio V_i/V_f is, of course, the volume compression. The factor $(B_i^2 V_i/8\pi)$ is the energy stored initially in the quasi-magnetic field. In other words, equation (90) shows that the energy in the trapped field increases as a result of the work done on the trapped magnetic field during the compression and it roughly increases by the volume compression ratio.

If one starts out with $B_i = 1000$ gauss and there is a volume compression ratio of 30, one winds up with a trapped field of about 30 k gauss. The temperature attained if the energy in the trapped field manages somehow to get into the plasma can be estimated by

$$\mathcal{E}_f \simeq V_f n_f k T \quad . \tag{91}$$

Combining this equation with equation (57), one obtains:

$$kT \simeq \frac{B_i^2}{8\pi n_i} \frac{V_i}{V_f} \tag{92}$$

For an initial density of the order of $n_i = 10^{15}$, one comes out with $T \sim 10^7$ ^oK. So, potentially, reversed field dissipation does provide an important heating mechanism.

8. A Remark About Field Diffusion

Some comments are in order concerning the diffusion of a plasma in a magnetic field. The MHD equations, as we have written them, automatically take into account the inter-diffusion of the magnetic fields, or the diffusion of the particles across the fields. To obtain an estimate of the diffusion coefficient,

consider the quasi-stationary case of a confined plasma at a constant temperature. The pressure gradient in this case is

$$\nabla P = \frac{\vec{J} \times \vec{B}}{c} \quad .$$

The electric field generated by the plasma diffusion across the field lines is

$$\vec{E} = \frac{\vec{v} \times \vec{B}}{c} \quad .$$

Finally, with Ohm's Law

$$\vec{J} = \sigma \vec{E} \quad ,$$

and using $\nabla P = kT\nabla n$, one finds

$$kT\nabla n = \frac{\sigma}{c^2}(\vec{v} \times \vec{B}) \times \vec{B} = \frac{\sigma}{c^2}\left[(\vec{v} \cdot \vec{B})\vec{B} - B^2\vec{v}\right] \quad .$$

For a flow across the magnetic field, $\vec{v} \cdot \vec{B} = 0$, so that

$$kT\nabla n = -\frac{B^2\sigma}{c^2} v_\perp \tag{93}$$

The classical effective diffusion coefficient is defined by $n\vec{v}_\perp = -D_\perp \nabla n$. Thus,

$$D_\perp = \frac{nkTc^2}{\sigma B^2} \tag{94}$$

This important result of the classical diffusion theory led to optimism ten years ago about the possibilities for building thermonuclear reactors. One can associate a characteristic time τ to the diffusion of the plasma a distance L

$$\tau \sim \frac{L^2}{D_\perp} = \frac{L^2 \sigma B^2}{nkTc^2} \tag{95}$$

Using $\beta = \left(nkT / B^2/8\pi\right)$, one obtains a characteristic time of

$$\tau = \left(\frac{4\pi\sigma L^2}{c}\right)\left(\frac{2}{\beta}\right) \tag{96}$$

where L is a characteristic dimension. The first factor is just the ordinary skin effect time. In other words, if one has a copper rod with a characteristic dimension L, the field penetrates in times which can be estimated from the first factor. The second factor $2/\beta$ comes from the fact that the plasma, contrary to the copper rod, can expand and contract.

For L = 1cm, β = 0.2, the characteristic diffusion time would be

$$\tau = .1\ \mu\text{sec. at } 10^4\ ^\circ\text{K},$$
$$\tau = .1\ \text{sec. at } 10^8\ ^\circ\text{K}.$$

In a θ-pinch with a reversed trapped field and field gradients which extend over a few millimeters, the characteristic diffusion times would be of the order 10 μ sec. for $T_e \sim 10^6$ $^\circ$K, according to these crude estimates.

We mention this result in passing just to indicate that if one had a stable plasma where the MHD equations describe the diffusion one would expect, in the end, times of this order

of magnitude. The whole problem is to generate a plasma where the classical diffusion coefficient depends on the magnetic field as $1/B^2$.

Equation (83) has been verified with low density plasmas confined in the magnetic mirror geometry and heated by radio frequency fields. The $1/B^2$ dependence, as well as the numerical coefficient within the experimental accuracies of about a factor of 2, has been verified. Thus, plasmas confined by magnetic fields exist where classical diffusion governs the confinement.

However, one finds that if one drives in high enough currents along the field lines, for example, toroidal discharges, the diffusion coefficient is observed to depend on $1/B$, indicating some kind of instability. This instability gives rise to an anomalous fast diffusion called Bohm diffusion. This is one of the main problems in controlled fusion research. One has to be careful with the relative direction of currents in the magnetic field to avoid this anomalously fast diffusion.

VII. Numerical Solutions of the Magnetohydrodynamic Equations

The full set of hydromagnetic equations (Sec. III) which describe a fully ionized plasma in an infinite cylinder have been solved on a digital computer. An implicit method for solving the set of finite difference equations was used which does not bound the time-step by the characteristics.[17] At each time-step T_e, T_i J, n, v, B are computed as a function of the radial coordinate r. The fluid equations are coupled to the equations which describe the external circuit

$$V_\theta = V_c + L \frac{dI}{dt} + IR \quad , \tag{97}$$

where V_c is the capacitor voltage, C the capacitance, R the resistance, I the total current and L represents any stray inductance not connected with the plasma region. The quantity V_θ is the voltage around the plasma column given by

$$V_\theta = 2\pi \frac{d}{dt} \int_0^{r_o} B_z(r) dr \quad , \tag{98}$$

where $B_z(r)$ is given by the hydromagnetic equations and r_o is the initial plasma radius before compression. The total current which appears in equation (97) is expressed in terms of the field by the usual formula

[17] See reference 1 and also K. Hain, UKAEA report AERE-R3383 "Pinch Collapse" (1961).

$$B_z(r_0) = \frac{4\pi}{10}\frac{I}{l} \qquad (99)$$

where l is the length of a single turn solenoid.

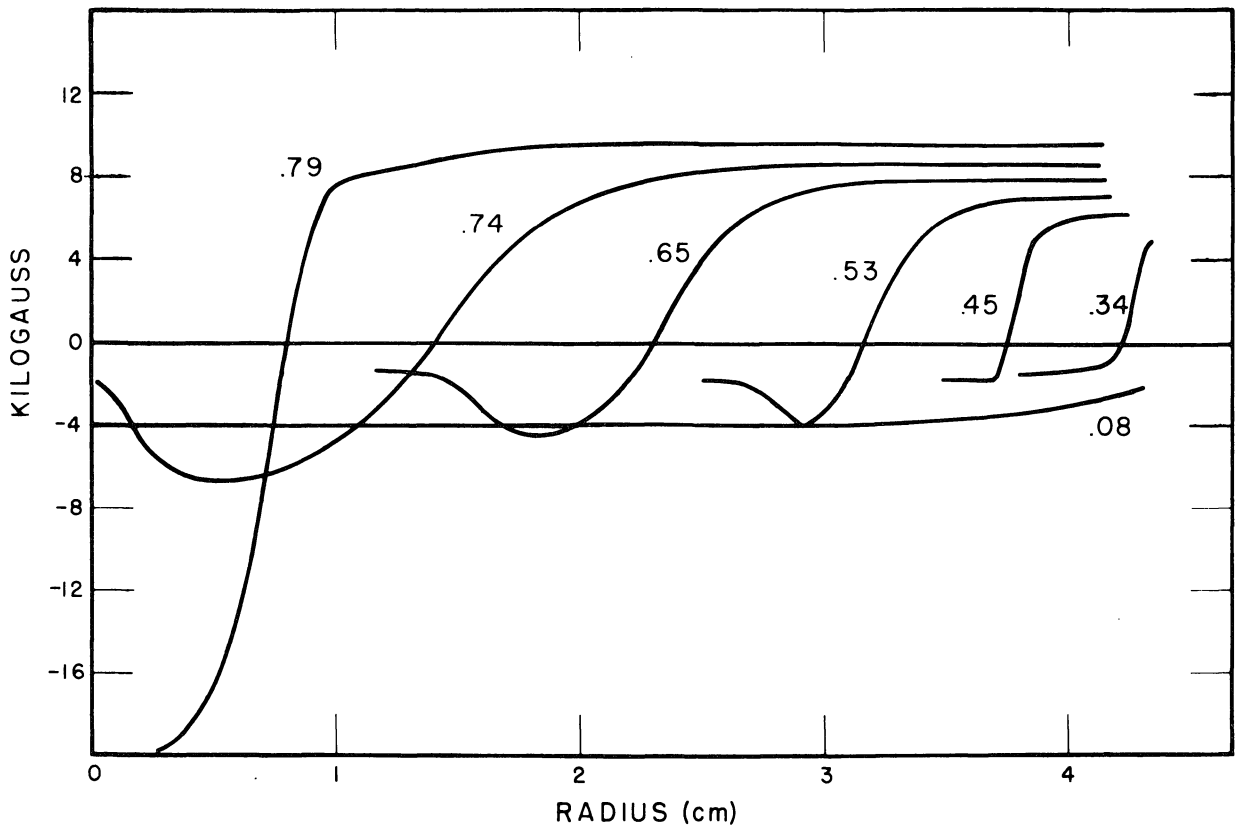
The boundary conditions are discussed in reference 1 and do not influence the state of the compressed plasma because of the rapid heating rates. This was demonstrated by direct computation.

Figures 7, 8, 9 show some typical results[18] obtained for a large θ-pinch with an initial reverse field of -4 k gauss. The other experimental parameters are described further in VIII. The times (in μ sec) corresponding to the various curves are shown.

The magnetic field rises rapidly near the boundary of the plasma column and reverses the direction of the vacuum magnetic field in about 0.3 μ sec. In this case, when the external magnetic pressure exceeds the internal pressures the plasma begins to contract and a large amplitude compression wave moves rapidly toward the center of the discharge tube.

The density and temperature maxima in figures 8, 9 correspond to the region where the magnetic field passes through zero.

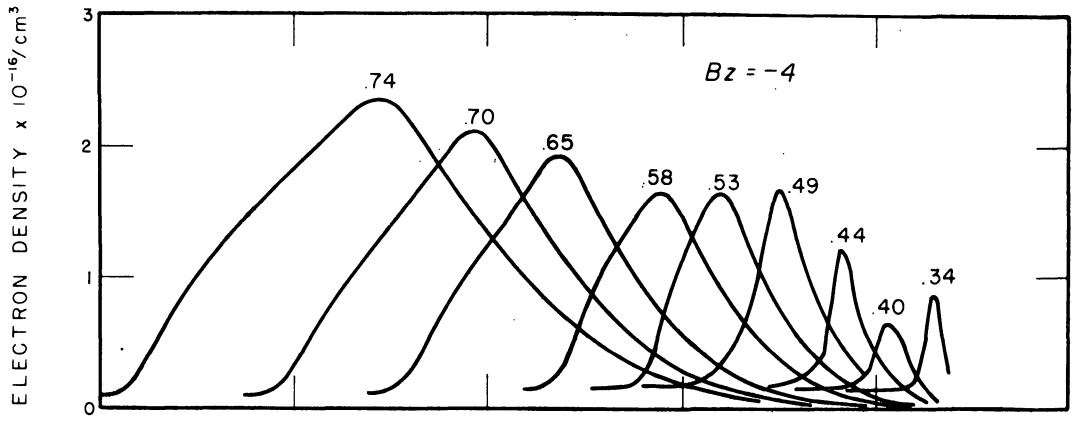
[18] These curves were taken from a more extensive publication (K. Hain and A. C. Kolb) now in preparation. A preliminary report was given at the Nov. 1962 meeting of the American Physical Society, Division of Plasma Physics in Atlantic City, N.J.



Magnetic Field Distribution Calculated From the Full MHD Equations For A θ-Pinch During the First Implosion

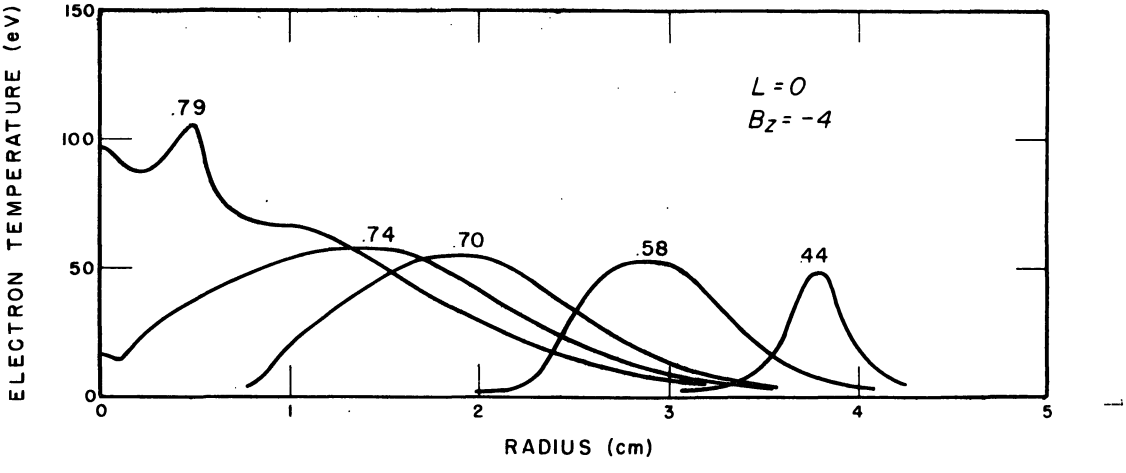
The initial reverse field is -4 kgauss. The indicated times are in μ sec.

FIGURE 7



Electron Density Distribution (See Caption to Figure 7)

FIGURE 8



Electron Temperature Distribution (See Caption to Figure 7)

FIGURE 9

Because of the high current densities (some tens of kiloamperes/cm^2) required to maintain the reversed field inside the diamagnetic plasma column, there is strong ohmic heating of the electrons during the implosion. Because of the relatively high densities and low Alfvén Mach numbers, the viscous shock heating of the ions is negligible (See VII 4). The ions are heated by compression and thermal relaxation with the hot electrons.

Therefore, this is a case where we deal with a resistive shock wave whose structure is determined by the skin effect. At lower densities and longer ion mean-free-paths the viscous effects would be more important as discussed earlier.

At $t \sim 0.79\,\mu$ sec the compression wave reaches the axis and the temperature begins to rise rapidly at the center. The ion temperature is always less than the electron temperature during the initial implosion phase of the discharge because there is not sufficient time for complete thermal equilibration to be reached.

In figure 7 one notices that the magnetic field on the axis is greater than the vacuum field at $0.79\,\mu$ sec. This behavior is typical for dynamic pinches and is due to the large inertial forces which lead to an overcompression of both the plasma and trapped magnetic field. This results in subsequent "bounces" of the plasma column as it undergoes a series of radial oscillations, accompanied by large time variations in the pressures near the axis. This "over-shoot" also leads to a

transient in the temperature near the axis because of the high compression ratio as the shock wave approaches the axis.

At later times (not shown) the temperatures continue to rise with the increasing magnetic pressure, and T_e approaches T_i by thermal relaxation. Since resistivity decreases with electron temperature, the ohmic heating rate tends to decrease and compressional heating becomes relatively more important. Eventually, however, the reverse trapped magnetic field must decay and this has an important influence on the energy balance. The diffusion times can be estimated by the considerations of VI-7, assuming that there are no instabilities (See VIII for a brief discussion of θ-pinch stability). The final temperatures predicted by the theory are in the 10^6 - 10^7 ^oK range at the time of maximum current, depending of course on the initial parameters. Thermal conduction smooths out the temperature gradients.

Such detailed theoretical calculations yield a great deal of information which is invaluable for guidance in the design of experiments and the interpretation of diagnostic measurements. It should be possible to investigate the range of validity of the MHD equations and to study the relative importance of various heating and loss mechanisms by a careful comparison of the consequences of the full nonlinear theory and experiment. The theoretical "information" is now much more abundant than that gleaned from the small amount of reliable experimental data that is available.

VIII. Experimental Observations

1. Apparatus

The magnetic compression of deuterium plasmas in a 0-pinch is under investigation by many laboratories.[19] No attempt will be made to survey all the experimental literature. Instead we will describe some recent results reported at Salzberg obtained with PHAROS, which is a relatively large device of the θ-pinch class.

The relevant parameters of the present experiments are:

Initial plasma radius	4.4 cm
Length	180 cm
Time to maximum current	14 μ sec
Maximum current	$\sim$ 12 megamp
Voltage	18 - 20kV
Initial reverse field	1-4 k gauss
Initial density	10^{15}-10^{16} ions/cm^3
Initial temperature	1-5 ev
Magnetic field	$\sim$ 80 k gauss
Stray inductance	2-3 mμ h

The maximum field depends of course on the bore and length of the solenoid as well as on the stored energy in the capacitor bank. The bank is rated at 2 megajoules but only $\sim$ 1.5 megajoules

[19] See for example the proceedings of the International Conference on Controlled Fusion Research, Salzburg (1961); published in Nuclear Fusion 1962 Supplement. Other references can be found in the several θ-pinch papers.

has been used so far. With full bank energy and some modifications to the coil, fields over 100 k gauss can be generated.

The initial magnetic field is generated by a slow 6 kv capacitor bank normally operated now with $\sim$ 100 k joule. The deuterium is preheated by discharging a small fast bank ($\sim$ 20 kv, 1 k joule, 500 kc/sec) which produces electronic fields of $\sim$ 100-200 v/cm and magnetic fields of a few kilogauss. A combination of ohmic heating and slow compression raises the plasma to a temperature of a few electron volts which is sufficient to produce 70-100% ionization at initial pressures less than 0.1 Torr. The highest degree of initial ionization occurs at the lowest pressures (0.015 Torr).

The initial temperature and density distribution can be measured by standard time-resolved spectroscopic techniques.[20] The temperature is found from the ratio of the Balmer lines to the radiative recombination - bremsstrahlung continuum. The electron density is determined from the absolute continuum intensity (which depends on n_e^2 and weakly on T_e) or from the Stark broadening of the Balmer lines (whose widths are almost proportional to $n_e^{2/3}$).

These relatively low temperatures are sufficient to produce a high enough conductivity initially to form a thin diamagnetic

20 W. H. Lupton, E. A. McLean and D. T. Philips p. 2263; also H. R. Griem p. 1857 in Proceedings 5th Ionization Conference on Ionization Phenomenon in Gases, Munich 1961, North Holland Publishing Co., Amsterdam (1962).

current sheath near the walls when the main capacitor bank is discharged. The initial magnetic field is then trapped by a shell of imploding plasma.

An over-all view of the experiment is shown in figure 10.

2. Shock Velocities and Compression

The implosion velocities of the order 8 cm/μ sec at 0.05 Torr) are measured with a streaking-image camera that has a time-resolution of about 0.1 μ sec in this particular application. A typical streak photograph is shown in figure 11 which exhibits the fast implosion, dynamic bounces discussed earlier, and the subsequent slow compression as the magnetic pressure rises. The measured velocities agree quite well with the theoretical results shown in figures 7, 8, 9.

The plasma can also be observed through the ends of the discharge tube with a framing camera. The photographs shown in figure 12 show that the plasma column is compressed by the magnetic field and maintains a fairly high degree of cylindrical symmetry. The radiation observed photographically is mainly due to bremsstrahlung. The spectral lines of impurities as well as the Balmer lines disappear because of the high ionization rates at high temperatures.

Streak camera photographs taken at different positions along the axis show that with a trapped reverse field there is an axial contraction due to the closed field lines at the ends. Therefore the plasma is confined axially as well as radially



by the magnetic forces so that end losses do not appear to be important in this particular experiment.

3. Temperature

Above 10^6 ^oK, one can measure the electron temperature from the soft X-ray emission with a precision of about 20%. Because of the high densities the electron-electron scattering times are orders of magnitude smaller than the experimental times, so that it is reasonable to assume that the electron gas is nearly Maxwellian. This enables one to compute the intensity of the soft X-rays emission (bremsstrahlung or radiative recombination) from the scattering cross-sections as a function of energy with an average over a Maxwellian velocity distribution for the electrons. This leads to a factor $\exp\left[-h\nu/kT\right]$ in the intensity formula.

The exponential shape has been measured and electron temperatures of about 10^7 ^oK are obtained if the plasma is sufficiently pure. Such temperatures are in the range predicted by the MHD calculations. It appears therefore that the Spitzer ohmic heating rate and the compression can account for the observed electron temperature.

Since the soft X-ray emission does not depend on the ion temperature, the measurement described above yields no information concerning the ion heating rates. Because of the relatively high densities the ions will be heated by electron-ion collisions and by compression and they should also have a fairly high

temperature. However, this is still a conjecture since there are as yet no direct ion temperature measurements in this experiment.

It may be possible to measure the ion temperature indirectly by observing the Doppler broadening of impurity lines in the vacuum ultraviolet or soft X-ray regions of the spectrum. However, experience has shown that such measurements are not always conclusive, especially if there are instabilities of any kind with associated strong electric fields.

It may also be possible to get some information concerning the ion temperature by analyzing the energy of particles which escape from the plasma column. However, the strong magnetic fields can act as a filter and one must be sure that the observed particles are typical of the ions inside the plasma column.

Some idea of the ion energy can be obtained from the measured neutron flux from d-d reactions in the plasma. In the PHAROS experiment 10^8 neutrons are emitted in each pulse at 0.015 Torr, 10^7 at 0.03 Torr and 10^6 at 0.05 Torr. Figure 13 shows typical oscillograms of the X-ray and neutron emission. The ion "temperatures" obtained from such observations are always clouded by the suspicion that instabilities of some kind lead to accelerating electric fields and a non-thermal ion velocity distribution. Neutron intensity measurements alone are not sufficient to prove a thermal origin, except perhaps in some future experiment where the intensity is so high as to leave

little room for reasonable doubt. This day has not yet arrived for the θ-pinch.

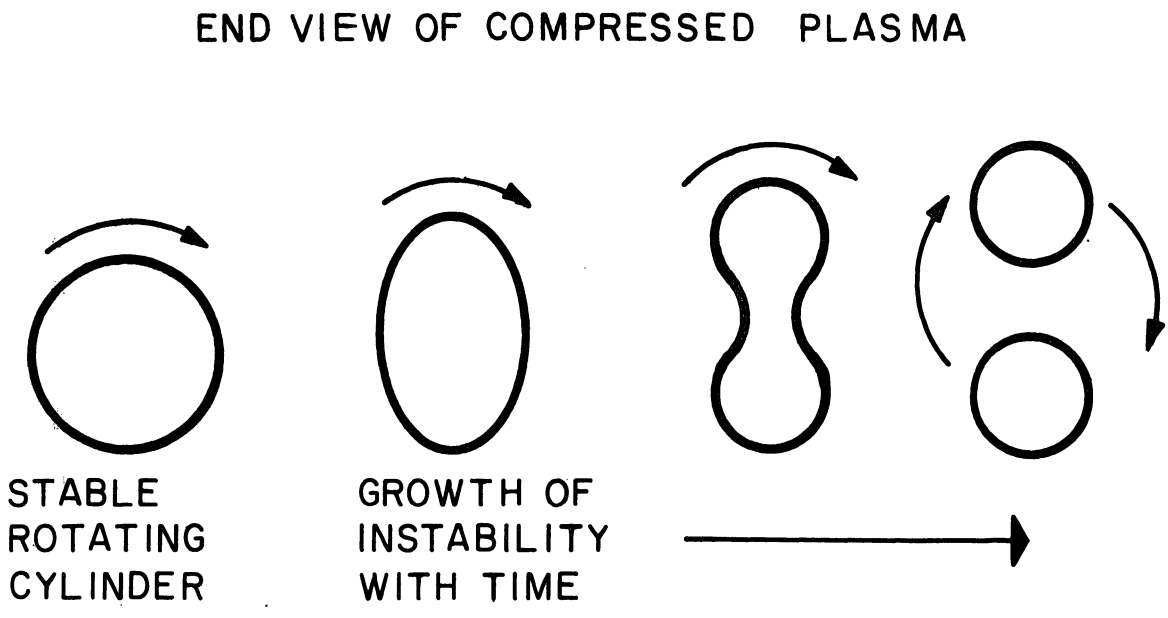
4. Stability

With poor initial preionization the plasma boundary has been observed by several workers to assume a very ragged appearance due to instabilities. It appears, however, that the difficulty can be overcome by having the proper initial conditions.

Another form of instability has been observed[21] in experiments on a smaller scale. The plasma seems to rotate and eventually break up due to the centrifugal forces which drive the so-called flute instabilities. In some cases the plasma separates into two distinct filaments which rotate about one another, (figures 14, 15). At about the same time the end losses are observed to increase rapidly.

It is suspected[21] that the diffusion of the magnetic fields inside the plasma changes the balance of momentum between the ion diamagnetic motion and the drift motion caused by the interaction of the radial space charge electric fields which maintain the electrical neutrality and the axial magnetic field. Calculations based on this model seem to be in qualitative agreement with the measurements.

21 N. Rostoker and A. C. Kolb, Bull. Am. Phys. Soc. 6, 203 (1961), Phys. Rev. 124, 965 (1961). See also ref. 19.



Schematic Of The Rotational Instability Viewed End-on.

FIGURE 15

Because of the relatively large diameter of the PHAROS plasma column, the diffusion times should be longer than in the earlier experiments. This should delay the onset of the rotational instability, and indeed it has not yet been observed during the first half-cycle of the discharge.

However, a new difficulty has arisen: the plasma drifts toward the slot in the coil which feeds the currents, presumably due to small inhomogeneities in the axial field. Attempts are now being made to stabilize this drift by superimposing a multipole field, supplied by an auxiliary capacitor bank.

The drift limits the confinement time to a maximum of about 15 μ sec (depending on the initial density). After the plasma strikes the wall there is a flood of impurities and rapid cooling. However, while the plasma drifts to the wall the axial confinement is maintained and the cross-section of the plasma column does not seem to show gross distortions. However, it should be emphasized that experimental studies of the stability of a θ-pinch are in a very preliminary stage and it is impossible to predict the ultimate possibilities with any degree of certainty. There are all kinds of instabilities connected with finite resistance, density and temperature gradients, electrostatic fields, curved field lines, anisotropic pressure, ad infinitum, that may be important.

The θ-pinch is a useful tool for studying the properties of a high temperature plasma, shock waves, radiation, confine-

ment in a magnetic field and stability. It may also be possible to test certain features of the magnetohydrodynamic theory. However, like all other thermonuclear enterprises, little can be said for the future of such devices in this area.

In conclusion, the experimental situation for large θ-pinches can be summarized as follows: electron temperatures of 10^6 ^oK- 10^7 ^oK can be produced for times up to about 10 μ sec with densities in the range 10^{16} to $\sim 10^{17}/cm^3$, with both radial and axial confinement. The rotational instability which plagued earlier experiments can at least be delayed in time. The magnetohydrodynamic two-fluid theory accounts for the implosion velocities, bounce frequencies and compression ratio. It also seems to predict the initial rate of electron ohmic heating with fair accuracy. Impurities play an important role in the electron heating, and confinement time is presently limited by a slow radial drift of the plasma column as a whole toward the wall of the discharge tube.

"APPLICATION OF SINGULAR EIGENFUNCTION EXPANSION METHODS TO PROBLEMS IN PLASMA OSCILLATIONS"

Roman Zelazny

Case Institute of Technology

In the following lectures we shall investigate the application of singular eigenfunction expansion methods to the study of plasma phenomena. In particular, our major concern shall be with obtaining a solution of the initial value problem for longitudinal plasma oscillations for the case when ions are treated as a positive charge background.

Of course, the latter problem is but a special case of the general investigation of plasma oscillations where one is also concerned with transverse oscillations and coupling between transverse and longitudinal modes. However, because the time in which to present these lectures is severely limited and in virtue of the fact that the notation for the general problem is significantly more involved than is that for the special problem of longitudinal oscillations, it is felt that maximum information can be transferred by restricting the scope of the talk. All essential points of the method of solution can be covered in spite of the limitation we impose and, in fact, should be clearer as a result. If time permits, the modifications in the theory which are necessary for obtaining solutions to the general problem shall be indicated.

The central equation with which we shall work is the collisionless non-relativistic Boltzmann-Vlasov equation. We shall not derive this equation here; rather, we shall obtain a linearized form of the equation and proceed to solve it by taking Fourier transforms with respect to spatial coordinates and then performing an eigenfunction expansion in transform

space. The problem is then reduced to that of obtaining the solution of a system of singular integral equations.

Since they will be very important in the subsequent work, let us start by briefly discussing some properties of Cauchy integrals.

I. Properties of Cauchy Integrals

Definition (Cauchy Integral):

Given a smooth line, L, in the complex plane and a function, $\phi(\tau)$, defined for the points τ on L, then

$$\Phi(z) = \frac{1}{2\pi i} \int_L \frac{\phi(\tau)}{\tau - z} \, d\tau \qquad (1.1)$$

is a Cauchy integral. We have the following theorem:

Theorem:

If L is sufficiently smooth, and we have a function $f(\tau, z)$ continuous with respect to τ and holomorphic[1] with respect to Z, then $\int_L f(\tau, z) d\tau = F(z)$ is also a holomorphic function.

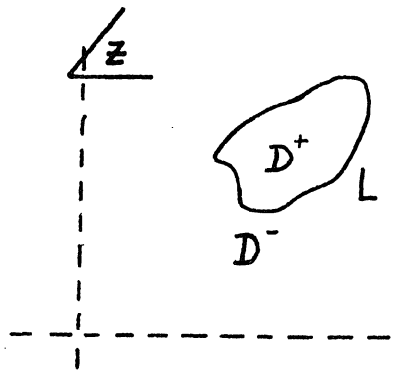
As a consequence, $\Phi(z) = \frac{1}{2\pi i} \int_L \frac{\phi(\tau)}{\tau - z} \, d\tau$ is holomorphic except at Z on L.

Definition:

Suppose L were to separate the complex plane into two

[1] A function is holomorphic in a region, R, if it has derivatives of all orders at every point of R. See N.I. Muskhelishvili, Singular Integral Equations, P. Noordhoff, N.V., Groningen, Holland, (1953) for a thorough discussion of these ideas.

sections, D^+ and D^- . Then the functions $\Phi^+(z)$ for Z within D^+ and $\Phi^-(z)$ for Z within D^- are referred to as the sectionally holomorphic functions given by (1.1) for $\phi(z)$ given on the boundary L .



Definition (Hölder Condition):

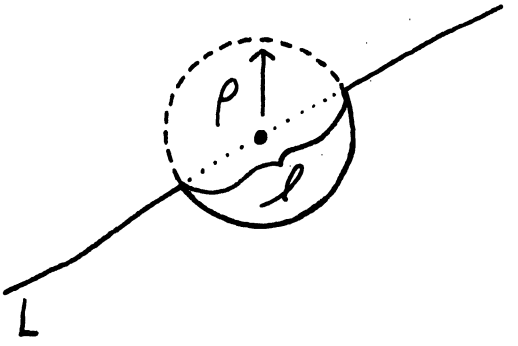
A function, $f(t)$, is in the Hölder class in a given interval iff the following relation is satisfied for any two points t_1 and t_2 within the interval:

$$\left|f(t_1)-f(t_2)\right| < A\left|t_2-t_1\right|^{\lambda} \quad , \tag{1.2}$$

where A is any positive constant and $0 < \lambda \leq 1$.

Definition (Principal Value of Cauchy Integral):

When Z is a point on L we must take (1.1) to be the principal value. We do this by deforming the contour L in the vicinity of Z in the following way:



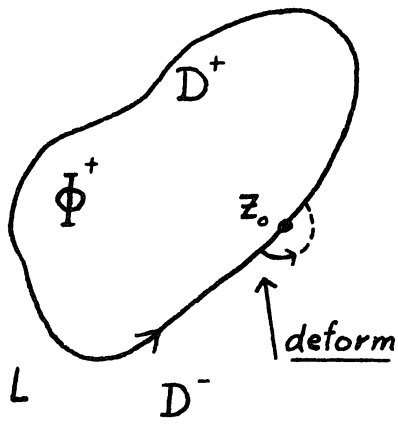
$L \longrightarrow L - \ell$,

where ℓ is the portion of L cut by the circle of radius ρ about Z . Then, the principal value of the integral is defined in the following way:

For Z on L :

$$P\int_L \frac{\phi(\tau)}{\tau - z}\, d\tau \underset{defn}{=} \lim_{\rho \to 0} \int_{L-\ell} \frac{\phi(\tau)}{\tau - z}\, d\tau \qquad (1.3)$$

In light of the previous discussion it is seen that, for Z_o on L , the Cauchy integral will have two values depending on whether the point Z_o is approached from within D^+ or, on the other hand, D^- . This gives rise to the following formulae (the Plemelj formulae):



Assume that the point z on L is approached from within D^+ and let us calculate $\Phi^+(z_0)$, the boundary value of $\Phi(t)$ at $t = z_0$. Deform the contour, L, as shown in the diagram above. Then, by the Cauchy integral formula and the definition of the principal value, we immediately have:

$$\Phi^+(z_0) = \frac{1}{2\pi i}\left\{P\int_L \frac{\phi(\tau)}{\tau - z_0}\,d\tau + i\pi\,\phi(z_0)\right\} = \frac{1}{2}\,\phi(z_0) + \frac{1}{2\pi i}P\int_L \frac{\phi(\tau)}{\tau - z_0}\,d\tau . \tag{1.4a}$$

Analagously,

$$\Phi^-(z_0) = -\frac{1}{2}\,\phi(z_0) + \frac{1}{2\pi i}P\int_L \frac{\phi(\tau)}{\tau - z_0}\,d\tau . \tag{1.4b}$$

We now rewrite the Plemelj formulae in a form in which they are frequently found:

$$\Phi^+(z_0) + \Phi^-(z_0) = \frac{1}{i\pi}P\int_L \frac{\phi(\tau)}{\tau - z_0}\,d\tau \tag{1.5a}$$

and,

$$\Phi^{+}(z_o) - \Phi^{-}(z_o) = \phi(z_o) \tag{1.5b}$$

To conclude these remarks, the following two points should be noted.

Although we have discussed Cauchy integrals for closed boundaries, the notions are easily extended to the case where L is the real axis in the complex z plane. In this case, the principal value of the Cauchy integral means:

$$\mathcal{P}\int_{-\infty}^{+\infty} \frac{\phi(\tau)}{\tau - z} = \lim_{\substack{\epsilon \to 0 \\ N \to \infty}} \left[\int_{-N}^{z-\epsilon} + \int_{z+\epsilon}^{N} \right] \tag{1.6}$$

The Plemelj formulae are defined in an analogous way and the results are the same as for the closed boundary except that

$$\Phi^{\mp}(\infty) = \mp \frac{1}{2} \phi(\infty) \tag{1.7}$$

Secondly, it should be pointed out that if $\phi(\tau)$ belongs to the Hölder class (1.2) then the existence of the Cauchy integral (1.1) is assured. However, this is only a sufficient condition for the existence of the Cauchy integral; in general it is required only that function $\phi(\tau)$ be a distribution in the sense of Schwartz.[2]

[2] L. Schwartz, Theorie des Distributions, Paris (1950/52)

II. The Linearized Vlasov Equation

With these introductions behind us, let us now see how the singular eigenfunction expansion method is applied to the study of plasma oscillations. We start with the Boltzmann-Vlasov equation for electrons moving in a uniform background of positive charge, the latter being due to the ions. The distribution function for electrons will satisfy[3]

$$\frac{\partial F(\underline{r},\underline{v};t)}{\partial t} + \underline{v}\cdot\vec{\nabla}F + \underline{a}\cdot\vec{\nabla}_{\underline{v}}F = 0 \qquad (2.1)$$

$F(\underline{r},\underline{v};t)$ is the probability density for finding an electron at $d\underline{r}$ about $\underline{r}$, having velocity $d\underline{v}$ about $\underline{v}$, at time t.

As we remarked earlier, we shall consider longitudinal plasma oscillations only. In other words, assume that the force acting on the particles is of electrostatic nature only and that the acceleration of the electron is

$$\underline{a} = \frac{e}{m}\underline{E}(\underline{r},t) \quad .$$

[3] Reporter's Note: The Boltzmann-Vlasov equation is an approximation to the full kinetic equation for the electron. However, it should hold for low density plasmas at relatively high temperatures. /For a full discussion of the derivation of this equation, and its relation to more exact kinetic equations, see e.g. R.L. Guernsey, Kinetic Theory of Fully Ionized Gases, University of Michigan Thesis, 1960.

We have, in addition to (2.1), Poisson's equation:

$$\nabla \cdot \underline{E} = 4\pi\rho(\underline{r},t) \quad , \tag{2.2}$$

where

$$\rho(\underline{r},t) \equiv e\int F(\underline{r},\underline{v};t)\,d^3v - \rho_o \tag{2.3}$$

is the excess charge density over background.

Equation (2.1) is a non-linear equation of quite complicated character. Hence, we look for some applicable linearization procedure to simplify the job of its solution. To this end, assume that the equilibrium distribution, $F_{eq} \equiv f_o(v)$, is a function not depending on space or time and that, furthermore,

$$F = f_o(\underline{v}) + f(\underline{r},\underline{v};t), \tag{2.4}$$

where $\left|f(\underline{r},\underline{v};t)\right| << f_o(v)$ for all $\underline{r},\underline{v}$, and t.

Substitute (2.4) into the full Boltzmann-Vlasov equation (2.1) and neglect velocity gradients of $f(\underline{r},\underline{v};t)$ as compared with velocity gradients of $f_o(\underline{v})$. One then has

$$\frac{\partial f}{\partial t} + \underline{v}\cdot\vec{\nabla} f + \frac{e}{m}\,\underline{E}\cdot\vec{\nabla}_v f_o = 0 \tag{2.5}$$

which is referred to as the <u>linearized</u> Boltzmann-Vlasov equation.

Similarly, substitution of (2.4) into (2.3) yields:

$$\rho = e\int f\, d^3v \tag{2.6}$$

in virtue of

$$e\int f_0\, d^3v = \rho_0$$

III. General Solution

We now proceed to the solution of (2.5). First, take the Fourier transform of (2.5) with respect to the spatial coordinate $\underline{r}$. We quickly get:

$$\frac{\partial f}{\partial t} + i\underline{k}\cdot\underline{v} f = -ik\eta(v, \underline{v}_\perp)\int f\, dv\, d^2\underline{v}_\perp \tag{3.1}$$

with

$$\eta \equiv -\frac{4\pi e^2}{m k^2}\frac{\partial f_0}{\partial v} \equiv -\nu_p^2 \frac{\partial f_0}{\partial v} \tag{3.2}$$

Note that for convenience the transform index, k, has been suppressed in the distribution function which we merely write as $f(\underline{v}, t)$. The meaning of v and $\underline{v}_\perp$ is such that $v \equiv \underline{v} \parallel \underline{k}$ and $\underline{v}_\perp \equiv \underline{v} \perp k$. Note that $\underline{v}_\perp$ is a two dimensional vector.

In deriving (3.1) we have used the fact that by Poisson's equation,

$$\phi^T = \frac{4\pi e}{k^2}\int f\, d^3v$$

and that $\vec{E}(r,t) = -\,grad\,\phi(r,t)$ so that $\vec{E}^T = -i\underline{k}\,\phi^T$.

Now, let us search for eigenfunction solutions of (3.1) of the type

$$f \sim e^{-ikvt}\,\phi(v, v) \quad . \tag{3.3}$$

Note that this is really equivalent to performing a separation of variables.

After placing (3.3) into (3.1), one has:

$$(v-\nu)\,\phi(\nu, v, \underline{v}_\perp) = -\eta(v, \underline{v}_\perp)\int \phi\, dv\, d\underline{v}_\perp \quad . \tag{3.4}$$

Clearly, however, is a function only of ν . Call

$$a(\nu) \equiv \int \phi(\nu, v, \underline{v}_\perp)\, dv\, d\underline{v}_\perp \tag{3.5}$$

where $a(\nu)$ is to be determined. We then have

$$(v-\nu)\,\phi = -\eta\, a(\nu) \quad , \tag{3.4'}$$

which has the formal solution:

$$\phi(\nu, v, \underline{v}_\perp) = -\frac{\eta(v, \underline{v}_\perp)\, a(\nu)}{(v-\nu)} + \lambda(\nu, v, \underline{v}_\perp)\,\delta(v-\nu) \tag{3.6}$$

$\lambda(\nu, v, \underline{v}_\perp)$ is subsequently to be related to the initial conditions of the problem.

Equation (3.6) is subject to the following condition on $a(\nu)$ which is obtained according to (3.5):

$$-\int \frac{\eta(v, \underline{v}_\perp)\, a(\nu)}{v-\nu}\, dv\, d\underline{v}_\perp + \int \lambda(\nu, \underline{v}_\perp)\,\delta(v-\nu)\, dv\, d\underline{v}_\perp = a(\nu) \tag{3.7}$$

We see, now, that our problem has been reduced to that of solving a singular integral equation for $a(\nu)$.[4]

This equation can be written in slightly different form:

$$\int \lambda(v, \underline{v}_\perp)\,\delta(v-\nu)\,dv\,d\underline{v}_\perp = a(\nu)\left[1+\int \eta\,\frac{(v,\underline{v}_\perp)}{v-\nu}\,dv\,d\underline{v}_\perp\right] \tag{3.7'}$$

This relationship is a condition on the function a(ν) but also determines the spectrum of eigenvalues $\{\nu\}$. In this regard, we have the following cases:

1. If ν = any real value:

   Integration over v on the l.h.s. of equation (3.7') gives unity. Hence, the pertinent relation for real ν is:

$$\int \lambda(\nu, \underline{v}_\perp)\,d\underline{v}_\perp = a(\nu)\,\lambda(\nu) \quad , \tag{3.8}$$

[4] Reporter's Note: At this point the question was raised by a member of the audience concerning the relation between the eigenfunction expansion method of solution of this paper and an alternate method which involves taking Laplace transforms re. the temporal dependence. It was agreed that an equivalence exists between the two solutions, a fact which has indeed been pointed out by K.M. Case (Ann. of Physics, 4, 349, 1959) and G. Backus (J. Math. Physics 1, 178, 1960). However, the particular analytic form of the solutions differs and each has its own merits for any particular problem.

where

$$\lambda(\nu) \equiv \left[1 + P \int \frac{\eta(v, \underline{v}_\perp)}{v - \nu} \, dv \, d\underline{v}_\perp \right]^{5} \tag{3.9}$$

2. If ν = any complex value:

   Since integration over v on the l.h.s. of equation (3.7') equals zero, the following eigenvalue equation for ν is obtained:

$$\lambda(\nu) = 0 \tag{3.10}$$

3. Also, it is possible that equation (3.10) is satisfied by real valued ν_j. If, for such eigenvalues it is also the case that $\int \eta(\nu_j, \underline{v}_\perp) d\underline{v}_\perp = 0$[6] then the ν_j are legitimate eigenvalues of (3.7') and are to be included in the discrete spectrum.

Therefore, in general, the spectrum of eigenvalues, $\{\nu\}$, consists of a continuous set of real valued quantities (case 1, above) plus a set of discrete complex and real valued quantities (cases 2 and 3). Due to the fact that $\int \frac{\eta(v, \underline{v}_\perp)}{v - \nu} dv \, d\underline{v}_\perp$ is holomorphic there are only a finite number of values ν_j for which (3.10) is satisfied, i.e., the discrete spectrum is finite.

5 Note that $\lambda(\nu)$ and $\lambda(\nu\, v_\perp)$ are, in general, different functions.

6 This arises from the Plemelj formulae and from the requirement that both $\lambda^+(\nu_j) = 0$ and $\lambda^-(\nu_j) = 0$.

If the discrete roots are simple roots, let the associated eigenfunctions be designated by ϕ_j. On the other hand, the j^{th} root may be of multiplicity ρ_j. Then not only is

$$\phi_j(\nu_j, v, \underline{v}_\perp) = -\frac{\eta(v, \underline{v}_\perp) a(\nu_j) e^{-ik\nu_j t}}{v - \nu_j} \tag{3.11}$$

a solution of equation (3.4), but so are the $\rho_j - 1$ linearly independent solutions:

$$\frac{\partial^p}{\partial \nu_j^p}\left[\phi(\nu_j, v, \underline{v}_\perp)\right], \quad p = 1, 2 \cdots, \rho_j - 1 \tag{3.12}$$

However, for convenience of notation, at this point let it be assumed that the roots are simple. Define

$$\phi_j \equiv -\frac{\eta(v, \underline{v}_\perp) a(\nu_j)}{v - \nu_j} \tag{3.13}$$

Thus, in terms of the eigenfunctions associated with $\{\nu\}$, the general solution to (3.4) is given by

$$f(k, \underline{v}, t) = \sum_j \phi_j(\nu_j, v, \underline{v}_\perp) e^{-ik\nu_j t} + \int \phi(\nu, v, \underline{v}_\perp) e^{-ik\nu t} d\nu \,. \tag{3.14}$$

IV. Completeness and the Initial Value Problem

In order for the assertion of (3.13) to be true, it must be shown that the eigenfunctions form a complete set, i.e., that almost any arbitrary function can be expanded in the form (3.13). However, there is a close relation between the demonstration of completeness and the solution of the initial value problem because for almost every initial distribution, $f(k,\underline{v};0)$, one can perform the expansion

$$f(k,\underline{v};0) = \sum \phi_j(\nu_j, v_\perp, \underline{v}_\perp) + \int \phi(\nu, v_\perp, \underline{v}_\perp)\, d\nu \quad (4.1)$$

We now consider the problem of completeness and at the same time provide an explicit calculation of the eigenfunctions $\phi(\nu, v_\perp, \underline{v}_\perp)$ and ϕ_j which solve the initial value problem.

Define

$$g(v_\perp, \underline{v}_\perp) \equiv f(k, \underline{v}; 0) - \sum \phi_j(\nu_j, v_\perp, \underline{v}_\perp) \quad . \quad (4.2)$$

Introducing (3.6) into (4.1) and using (4.2) one obtains

$$\lambda(v_\perp, \underline{v}_\perp) - \eta(v_\perp, \underline{v}_\perp) \int_{-\infty}^{+\infty} \frac{a(\nu)}{v - \nu}\, d\nu = g(v_\perp, \underline{v}_\perp) \quad . \quad (4.3)$$

Integrating (4.3) over $\underline{v}_\perp$ and using (3.8) then

$$a(v)\lambda(v) - \eta(v) \int \frac{a(\nu)}{v-\nu}\, d\nu = G(v) \quad (4.4)$$

where the following quantities have been defined:

$$G(v) \equiv \int g(v_\perp, \underline{v}_\perp)\, d\underline{v}_\perp \quad (4.5a)$$

$$\eta(v) \equiv \int \eta(v, \underline{v}_\perp)\, d\underline{v}_\perp \quad . \tag{4.5b}$$

The problem has thus been reduced to that of a singular integral equation in its dominant form. Using methods described in section I, above, a solution will now be demonstrated.

Define:

$$N(z) \equiv \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{a(\nu)}{\nu - z}\, d\nu \tag{4.6}$$

By (3.13), $f(k, \underline{v}, 0) \sim \int_{-\infty}^{+\infty} \phi(\nu, v, \underline{v}_\perp)\, d\nu$. Because $f(k, v, 0)$ is assumed to be in the Hölder class (or otherwise, a distribution in the sense of Schwartz), then $\int dv\, d\underline{v}_\perp f(k, v, \underline{v}_\perp; 0) \sim \int_{-\infty}^{+\infty} a(\nu)\, d\nu$ is finite. Therefore $N(\pm\infty) \to 0$, and we can write the Plemelj formulae:

$$N^+(\nu) - N^-(\nu) = a(\nu) \tag{4.7a}$$

$$\pi i \left[N^+(v) + N^-(v) \right] = P \int \frac{a(\nu)}{\nu - v}\, d\nu \quad . \tag{4.7b}$$

Also, when the following quantities are introduced,

$$Q(z) \equiv \frac{1}{2\pi i} \int \frac{\eta(\nu)}{\nu - z}\, d\nu \quad , \tag{4.8a}$$

$$M(z) \equiv \frac{1}{2\pi i} \int \frac{G(\nu)}{\nu - z}\, d\nu \quad , \tag{4.8b}$$

we have

$$G(\nu) = M^{+}(\nu) - M^{-}(\nu) \quad , \tag{4.9a}$$

and (from (3.9) and(4.8a))

$$\lambda(\nu) = 1 + \pi i \left[Q^{+} + Q^{-} \right] \tag{4.9b}$$

Using these relationships, the singular integral equation (4.4) can be written in the following form, where all terms are expressed as linear combinations of N(z) and M(z):

$$N^{+}(\nu)\left[1 + 2\pi i\, Q^{+}(\nu)\right] - M^{+}(\nu) = N^{-}(\nu)\left[1 + 2\pi i\, Q^{-}(\nu)\right] - M^{-}(\nu) \tag{4.10}$$

Define:

$$K(z) \equiv N(z)\left[1 + 2\pi i\, Q(z)\right] - M(z) \tag{4.11}$$

so that equation (4.10) can be written as

$$K^{+}(\nu) = K^{-}(\nu) \qquad \text{on the real axis.} \tag{4.12}$$

Because N(z), Q(z), and M(z) are holomorphic, so is K(z). This is a special case of the homogeneous Hilbert problem, and since $K(z) \longrightarrow 0$ as $z \longrightarrow \infty$ then by Liouville's theorem:[7]

[7] See appendix A.2

$$K(Z) \equiv 0,$$

everywhere. Thus, from (4.11), (4.13)

$$N(z) = \frac{M(z)}{1+2\pi i\, Q(z)} \quad . \tag{4.14}$$

N(z) must be everywhere finite, as seen from the definition (4.6). This requirement imposes important conditions on M(z) since M(z) must therefore vanish at the zeros of the denominator of (4.14). From the definitions (4.8) and (4.5b), the denominator is

$$\left[1+2\pi i\, Q(z)\right] = 1+\int \frac{\eta(\nu,\underline{\nu}_\perp)}{\nu - z}\, d\nu\, d\underline{\nu}_\perp \tag{4.15}$$

and (by the previous discussion (3.10)) has zeros of multiplicity ρ_j at the points ν_j ($j = 1,2,\ldots,M$). M(z) must also vanish at these points; in other words (see (3.12))

$$\frac{\partial^{p-1}}{\partial \nu_j^{\,p-1}}\, M(\nu_j) = 0 \qquad \begin{matrix} j = 1,\ldots, M \\ p = 1,\ldots, \rho_j \end{matrix} \tag{4.16}$$

In terms of the definitions (4.5a) and (4.2), equation (4.16) becomes

$$\frac{\partial^{p-1}}{\partial \nu_j^{\,p-1}} \sum_{j=1}^{M} \sum_{\sigma=1}^{\rho_j} \int \frac{\phi^{\sigma}(\nu,\nu,\underline{\nu}_\perp)}{\nu - \nu_j}\, d\nu\, d\underline{\nu}_\perp = \frac{\partial^{\sigma-1}}{\partial \nu_j^{\,\sigma-1}} \int \frac{f(\nu,\underline{\nu}_\perp;0)}{\nu-\nu_j}\, d\nu\, d\underline{\nu}_\perp \tag{4.17}$$

Since the ϕ^{σ}_{j} are related to $a^{\sigma}(\nu_{j})$ through equation (3.13) we see that (4.17) is a set of equations to be used for the solution for the latter quantities. It relates the $a^{\sigma}(\nu_{j})$ explicitly to the initial condition.

Now, everything else quickly follows in a straight-forward manner. Knowing the $a^{\sigma}(\nu_{j})$, one next calculates $g(\nu, \underline{v}_{\perp})$ according to (4.2). Then, $a(\nu)$ follows from (4.5a), (4.7a), and (4.14). Once $a(\nu)$ has been determined, then $\lambda(\nu, \underline{v}_{\perp})$ can be obtained from (4.4).

Finally, knowing $a(\nu)$ and $\lambda(\nu, \underline{v}_{\perp})$, the eigenfunction $\phi(\nu, v, \underline{v}_{\perp})$ is obtained from (3.6). In this way all the elements of the solution (3.14) have been obtained and $f(k,\underline{v};t)$ is uniquely given in terms of the (almost) arbitrary initial condition, $f(k,\underline{v};0)$.

V. Further Remarks

The solution, as has been obtained here, is quite useful in the study of plasma phenomena such as stability, Landau damping, etc.

For example, using the form (3.14):

$$f(k,\underline{v};t) = \sum_j \phi(\nu_j, \nu, \underline{v}_\perp) e^{-ik\nu_j t} + \int \phi(\nu, \nu, \underline{v}_\perp) e^{-ik\nu t} d\nu \quad (3.14)$$

it can be quickly shown that if the plasma is close to equilibrium it is stable (i.e., if $f_o(\underline{v})$ is Maxwellian). Whereas the existence of discrete roots, ν_j, in the upper half plane would give growing oscillations, it can be shown that no such roots exist for Maxwellian f_o. Also, since the contribution of continuous eigenvalues is proportional to $\int_{-\infty}^{+\infty} \frac{a(\nu)}{\nu - \nu} e^{-ik\nu t} d\nu$ and $a(\nu)$ has poles only in the lower half plane, this quantity also provides no growing temporal terms.

However, in the case of damping, it is seen that the continuous eigenfunctions are very important, with the pole lying closest to the real axis being dominant.

We have talked only about the initial value problem of the Boltzmann-Vlasov equation for longitudinal oscillations. However, the B-V with both longitudinal and transverse oscillations, taken together with the complete set of Maxwell equations,

has also been solved. One gets a system of integral equations which may be written in the form:[8]

$$\underline{\underline{\Lambda}}\, a(\nu) - \underline{\underline{\eta}} \int \frac{a(\nu)}{\nu - \nu'} d\nu = \underline{\underline{G}} \quad , \tag{5.1}$$

where

$$\underline{\underline{\Lambda}} = \begin{pmatrix} \lambda\nu & \underline{\Lambda}_\perp \\ \underline{\Lambda}_\perp & \underline{\underline{V}} \end{pmatrix} \quad . \tag{5.2}$$

$\Lambda_\perp$ is due to the isotropic part of the equilibrium distribution function.

Whenever the equilibrium distribution function is anisotropic we may have coupling between the longitudinal and transverse modes of oscillation. Such phenomena might occur, for example, in dense streams of plasmas or plasmas with two different temperatures.[9]

[8] See appendix A.1 for details.

[9] Reporter's Note: There was a question concerning the class of initial distributions for which the methods described above would hold. It was pointed out that a Hölder condition imposed on $f(k,\underline{v};0)$ is sufficient so that $Q(z)$, $M(z)$, and $N(z)$ exist. However, the exposition would be valid for any Schwartz distribution.

Appendix A.1: The Initial Value Problem For Longitudinal and Transversal Plasma Oscillations:

As was indicated in the text, above, the more general solution to the initial value problem which includes the full set of Maxwell equations has also been obtained. The method of solution closely follows, in structure, the solution for the problem for longitudinal plasma oscillations. However, the details are many times more complicated. A short summary of this procedure is given below.

In analogy to (3.3) of the previous treatment, solutions of the following form are sought:

(A1.1a) $$f = e^{-ik\nu t}\,\phi(\nu, v_\parallel, \underline{v}_\perp)$$

(A1.1b) $$\underline{E} = e^{-ik\nu t}\,\underline{e}(\nu)$$

(A1.1c) $$\underline{H} = e^{-ik\nu t}\,\underline{h}(\nu)$$

where $f(k,\underline{v},t)$ has the same meaning as before, $\underline{E}$ is the Fourier transform of the electric field, and $\underline{H}$ is the transform of the magnetic field. One then gets, for the Boltzmann-Vlasov equation:

(A1.2) $$(v_\parallel - \nu)\phi(\nu, v_\parallel, \underline{v}_\perp) + \frac{e}{ikm}\left[\underline{e} + \frac{1}{c}\underline{v} \times \underline{h}\right]\cdot \nabla_v f_0 = 0$$

and for the Maxwell equations in transform space:

(A1.3a) $$i\underline{k} \times \underline{h} + (1/c)ik\nu\underline{e} = (4\pi/c)\underline{j}$$

(A1.3b) $$i\underline{k} \times \underline{e} - (1/c)ik\nu\underline{h} = 0$$

(A1.3c) $$i\underline{k}\cdot\underline{e} = 4\pi g$$

(A1.3d) $$i\underline{k}\cdot\underline{h} = 0$$

In the above equations,

(A1.4a) $$\underline{j}(v) \equiv e\int \underline{v}\,\phi(\nu, v, \underline{v}_\perp)\,d^3\underline{v}$$

(A1.4b) $$g(v) \equiv e\int \phi\, d^3\underline{v}$$

The set of equations (A1.2 - A1.4) can be reduced to the following form (similar to 3.4):

(A1.5) $$(v-\nu)\,\phi(\nu, v, \underline{v}_\perp) + \eta_{\parallel}(v, \underline{v}_\perp)\,\frac{j_{\parallel}(\nu)}{\nu} + \underline{\eta}_\perp(\nu, v, \underline{v}_\perp)\cdot\frac{\underline{j}_\perp(\nu)}{\nu^2 - c^2} + \frac{ik}{4\pi\nu}\,\underline{\eta}_\perp(\nu, v, \underline{v}_\perp)\cdot\underline{e}^\circ_\perp(\nu)\cdot\delta(\nu^2 - c^2) = 0$$

where

(A1.6a) $$\eta_{\parallel}(v, \underline{v}_\perp) \equiv 4\pi e/(ik)^2 m\;\nabla_{v\parallel} f_0$$

and

(A1.6b) $$\underline{\eta}_\perp(\nu, v, \underline{v}_\perp) \equiv \frac{4\pi e}{(ik)^2 m}\left[(\nu - v)\nabla_{v_\perp} f_0 + \underline{v}_\perp \nabla_{v\parallel} f_0\right]$$

$\underline{e}^\circ_\perp(\nu)$ is a function which will be determined from the initial data.

It is convenient to introduce the quantities:

(A1.7a) $$\frac{1}{\nu}\,j_{\parallel}(\nu) \equiv \frac{1}{\nu}\int v\,\phi(\nu, v, \underline{v}_\perp)\,dv\,d\underline{v}_\perp \equiv a_{\parallel}(\nu)$$

(A1.7b) $$\frac{1}{\nu^2 - c^2}\,\underline{j}_\perp(\nu) \equiv \frac{1}{\nu^2 - c^2}\int \underline{v}_\perp\,\phi(\nu, v, \underline{v}_\perp)\,dv\,d\underline{v}_\perp \equiv \underline{a}_\perp(\nu)$$

so that equation (A1.5) may be written as

(A1.8) $$(v-\nu)\phi(\nu,v,\underline{v}_\perp) = -\eta_\parallel a_\parallel(\nu) - \underline{\eta}_\perp(\nu,v,\underline{v}_\perp)\cdot\underline{a}_\perp(\nu)$$
$$-\frac{ik}{4\pi\nu}\underline{\eta}_\perp(\nu,v,\underline{v}_\perp)\cdot\underline{e}_\perp^\circ(\nu)\,\delta(v^2-c^2) \quad .$$

The latter equation may be solved formally, giving:

(A1.9) $$\phi(\nu,v,\underline{v}_\perp) = -\frac{\eta_\parallel a_\parallel}{v-\nu} - \frac{\underline{\eta}_\perp\cdot\underline{a}_\perp}{v-\nu} - \frac{ik}{4\pi\nu}\frac{\underline{\eta}_\perp\cdot\underline{e}_\perp^\circ\,\delta(v^2-c^2)}{v-\nu}$$
$$+\lambda(v,\underline{v}_\perp)\,\delta(v-\nu) \quad ,$$

subject to the conditions (A1.7a and b). (In this way, we have obtained expressions analogous to (3.6) and (3.7) of the text, above.)

One next obtains eigenvalue equations in order to determine the spectrum, $\{\nu_j\}$. Introducing:

(A1.10a) $$\Lambda_\parallel(\nu) \equiv 1 + \int \frac{\eta_\parallel(v,\underline{v}_\perp)}{v-\nu}\,dv\,d\underline{v}_\perp$$

(A1.10b) $$\Lambda_\perp(\nu) \equiv \int \frac{\underline{\eta}_\perp}{v-\nu}\,dv\,d\underline{v}_\perp = \int \frac{v\underline{\eta}_\parallel}{v-\nu}\,dv\,d\underline{v}_\perp$$

and

(A1.10c) $$\underline{\underline{V}}_\perp(\nu) \equiv (v^2-c^2)\,\underline{\underline{I}} + \int \frac{\underline{\eta}_\perp}{v-\nu}\,dv\,d\underline{v}_\perp$$

the following sets of eigenvalue equations are obtained:

If ν is real, then the following conditions must be satisfied:

(A1.11a)
$$\nu\int\lambda(v,\underline{v}_\perp)\,d\underline{v}_\perp = \nu a_\parallel(\nu)\Lambda_\parallel(\nu) + \nu\underline{a}_\perp(\nu)\cdot\Lambda_\perp(\nu) + \frac{ik}{4\pi\nu}\Lambda_\perp(\nu)\cdot\underline{e}_\perp^\circ\,\delta(v^2-c^2)$$

and

$$\int \underline{v}_\perp \lambda(\nu, \underline{v}_\perp) d\underline{v}_\perp = a_\parallel(\nu) \underline{\Lambda}_\perp(\nu) + \underline{a}_\perp(\nu) \underline{\underline{V}}_\perp(\nu)$$

(A1.11b)

$$+ \frac{ik}{4\pi\nu} \left[\underline{\underline{V}}_\perp - \underline{\underline{I}} (\nu^2 - c^2) \right] \underline{e}_\perp^\circ \delta(\nu^2 - c^2)$$

For complex ν , one obtains the condition:

(A1.12)

$$\begin{vmatrix} \Lambda_\parallel(\nu) & \Lambda_\perp(\nu) \\ \underline{\Lambda}_\perp(\nu) & \underline{\underline{V}}_\perp(\nu) \end{vmatrix} = 0$$

If (A1.12) is satisfied by real ν , and in addition

(A1.13a) $$\int \eta_\parallel(\nu_j, \underline{v}_\perp) d\underline{v}_\perp = 0$$

and

(A1.13b) $$\int \underline{v}_\perp \eta_\parallel(\nu_j, v_\perp) d\underline{v}_\perp = 0$$

then these roots, ν_j , are also part of the discrete spectrum.

The eigenfunctions corresponding to discrete ν_j are:

(A1.14) $$\phi_j(\nu_j, v, \underline{v}_\perp) = - \frac{\eta_\parallel(v, \underline{v}_\perp) a_\parallel(\nu_j)}{v - \nu_j} - \frac{\underline{\eta}_\perp(\nu_j, v, \underline{v}_\perp) \cdot \underline{a}_\perp(\nu_j)}{v - \nu_j}$$

(assuming that the roots are simple).

Things follow along in the same way as before, only, of course, their complexity is orders of magnitude greater.[10]

[10] A detailed account of this problem appears in Annals of Physics, September 1962.

Appendix A.2: The Homogeneous Hilbert Problem

The homogeneous Hilbert problem is:

> To find a sectionally holomorphic function $\Phi(z)$ of finite degree at infinity, under the boundary condition
>
> $$\Phi^+(t) = G(t)\Phi^-(t) \quad \text{on L,}$$
>
> where G(t) is a non-vanishing function of the point t on L, satisfying the Hölder condition.[11]

Hence,

(4.12) $K^+(z) = K^-(z)$ on L is a special case. The general solution is somewhat complicated. However, the situation at hand is simplified in virtue of (4.12) and the fact that K^+ and K^- are sectionally holomorphic.

Because of these two properties, by Liouville's theorem, K(z) must be everywhere a constant. Thus, since $K(z) \longrightarrow 0$ as $z \longrightarrow \infty$, $K(z) \equiv 0$.

Liouville's theorem can be stated as follows:[12]

> a function which is analytic for all finite values of z and is bounded everywhere is a constant.

The proof is quickly provided for, by a modification of Cauchy's theorem, the derivative of an everywhere analytic function $f(a)$, is

$$f'(a) = \frac{1}{2\pi i}\oint \frac{f(z)}{(z-a)^2}\,dz \quad .$$

11 Muskhelishvili, p. 87

12 Morse, P. M., and Feshbach, H., Methods of Theoretical Physics, p. 351, McGraw-Hill, New York, 1953.

Let the contour of integration be a circle of radius R centered at a, and let $|f(z)| \leq M$ by hypothesis. Then,

$$|f'(a)| \leq \left(\frac{M}{2\pi R^2}\right) 2\pi R = \frac{M}{R} \to 0 \text{ as } R \longrightarrow \infty \quad .$$ Thus, $f'(a) = 0$ and $f(a)$ is a constant.

FOUR LECTURES ON NEUTRON THERMALIZATION

Noel Corngold

Brookhaven National Laboratory

("A Child's Garden of Thermalization")

Introduction

The subject of neutron thermalization concerns the manner in which the "neutron gas" comes into thermal equilibrium (if it ever does!) with the medium which contains it. Its analysis shows a nice interplay between theories of neutron transport and theories of the solid and liquid state, and is also reminiscent of the classical kinetic theory of gases. More precisely, it resembles the "foreign-gas" problem, where a small number of gas molecules is introduced into a large collection of molecules already in equilibrium at some temperature.

The thermalization problem might be compared with a particularly simple, linearized version of the kinetic theory of gases, were it not for the feature of chemical binding. In all but the simplest models, the atoms of the moderator interact with one another, and the complicated motions which result produce a complex scattering pattern in the laboratory system. Many experimental results in thermalization appear to be quite sensitive to details of binding, and the over-all result is a new and interesting branch of neutron physics.

Four lectures are adequate for little more than a sketch of the field. We shall begin with a short discussion of the "classical" thermalization experiments and the general mathematical apparatus used in their analysis. Then we shall consider some very simple models which will indicate those properties of the scattering operator which control the various experimental results.

Finally we shall discuss some of the many questions that are current and unsolved. In all this, our aim is to give a picture of the field in which terminology and general principles appear, but where fine mathematical detail is absent. The reader interested in pursuing some of these topics should consult the reference section and, in particular, the proceedings of the recent Brookhaven Conference on Neutron Thermalization.

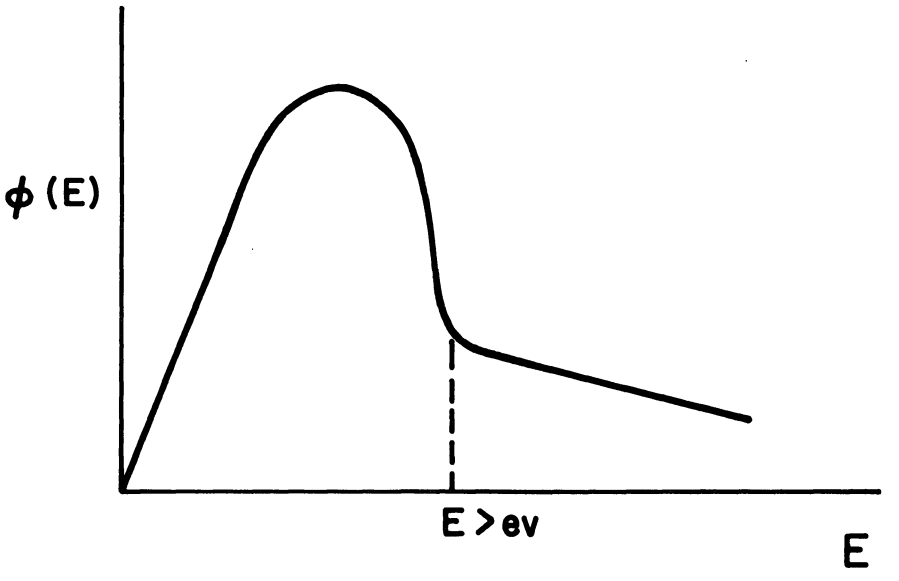
Lecture I

Some Experiments in Thermalization

I. The Infinite Medium, Steady-State Distribution

Imagine that we have a large, homogeneous block of moderating material, and that neutrons of high energy are introduced into the block in a uniform and isotropic manner. The moderating atoms will be supposed to be weak absorbers of neutrons ($\sigma_a << \sigma_s$) though strong enough so that many more neutrons are lost through capture than through leakage. The flux of neutrons in the block will then have the energy distribution sketched in Figure 1.

The infinite medium distribution is described as being "Maxwellian," with "1/E tail." The 1/E behavior, which holds for energies above a few volts, is typical of neutron slowing down theory, There, the average energy loss for a fast neutron (i.e., one whose kinetic energy is much greater than the binding energy of its target) upon collision is proportional to the energy of the neutron, and the cross section for scattering may be taken to be independent of energy. Thus, the number of neutrons in unit volume slowing down past energy E in unit time is proportional to $\Sigma_s \phi(E)E$. If capture is weak in this region, the number slowing down past E in unit time is very close to being the source-strength, S. Thus, $\phi(E) \propto S/\Sigma_s E$.



A Maxwellian Spectrum With 1/E Tail

Figure 1

When the over-all absorption is weak, the portion of (E) near kT has the form of the "Maxwellian,"

$$M(E) = \frac{E}{T^2} \exp(-E/T) \quad . \qquad (1)$$

(We use a system of units in which the Boltzmann constant, k_B is unity.) The presence of the Maxwellian component may be understood if we consider that when capture is small, the neutrons, having long lifetime, come close to being in thermal equilibrium with the moderator. At equilibrium the neutrons assume the canonical distribution of statistical mechanics. The number of neutrons with momenta lying in the element $d^3\vec{p}$ about $\vec{p}$, which we denote by $n(\vec{p})d^3\vec{p}$, is proportional to $\exp(-E/T)d^3\vec{p}$, where E is $\frac{1}{2m}p^2$, the neutron energy. When the neutrons are in a sufficiently large and homogeneous system, $n(\vec{p})$ will depend only upon the magnitude of $\vec{p}$, that is, upon the energy. The number density in energy is then proportional to $|\vec{p}| \exp(-E/T)$, and the flux density, $\phi(E)$, is proportional to $E \exp(-E/T)$.

In the presence of absorption or leakage, the Maxwellian and 1/E component become distorted and less distinct. The Maxwellian shrinks in size, and its maximum, in most cases, is displaced to a larger energy, $E_M > T$, while the slowing-down component is somewhat depressed. The amount of distortion produced by a given number of absorbing atoms is governed by the rate of energy exchange between neutron gas and moderator. One of the goals of thermalization theory is to describe these

distortions quantitatively. This goal has not yet been reached.

Infinite medium spectra have been measured in two ways: by steady-state and by pulsed techniques (1). In the first, a beam is extracted from or near the core of a steady, chain-reacting system and is analyzed into its velocity components by means of a neutron "chopper," or by means of a crystal spectrometer. The second method, which gives greater beam intensity, is somewhat more complicated. It is based on the fact that the steady-state infinite medium spectrum may be obtained by integrating, with respect to time, the time-dependent spectrum produced by a short pulse of fast neutrons. In this case, a pulse of neutrons is injected into the moderator, and the velocity resolution of the extracted beam is carried out by means of time-of-flight. It is clear that the pulsed technique is simplest when the neutron lifetime in the system is short compared with the flight time for thermal and epithermal neutrons. Thus, the earliest results using the pulsed technique have appeared in the study of poisoned H_2O systems (2). Recently, the techniques of data analysis have been improved so as to make measurements of spectra in graphite possible (2a).

2. The Life History Experiment

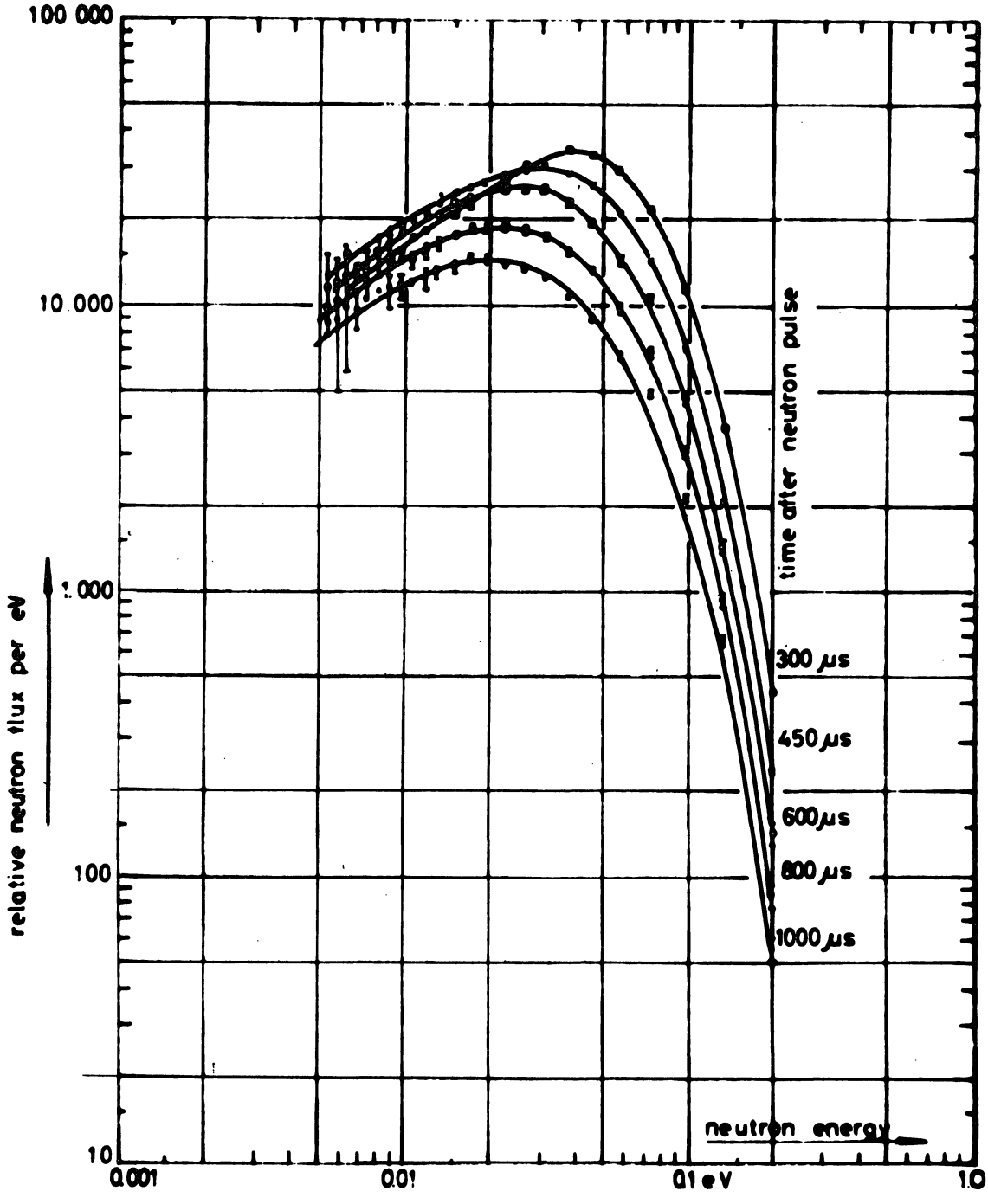
This experiment is aimed at studying in detail the manner in which fast neutrons are thermalized. A burst of high energy neutrons is introduced at t = 0, and ϕ (E,t) can be measured. In the absence of leakage and absorption, ϕ (E,t) tends to a Maxwellian distribution when t becomes very large.

Complete investigations of this sort require intense sources of fast neutrons, and detection apparatus capable of analyzing an emergent beam with respect to its component velocities. M. Poole and his colleagues at Harwell have recently reported the results of such an experiment in graphite (3). Their source is a linear accelerator whose pulsed electron beam produces x-rays which, in turn, liberate neutrons through (γ,n) and (γ,f) reactions. Their detecting equipment consists of a chopper and a time-of-flight system. By "phasing" the accelerator pulse and the opening of the chopper, the experimenters can select various times in the history of the pulse and can analyze the instantaneous spectra by time of flight.

The time history experiment yields a great deal of interesting data, which has only just begun to be analyzed. Figures 2 and 3, which show the time-dependent spectra and the variation of the average energy of the pulse are typical experimental results. We shall next want to know the shape of the final spectrum (if there is one!) the rate at which the pulse "relaxes" into its final form, the behavior of the pulse in the quasi-slowing-down region $E \gtrsim kT$, the dependence of all these upon geometry, poison concentration,... In short, there is much matter here for the neutron physicist.

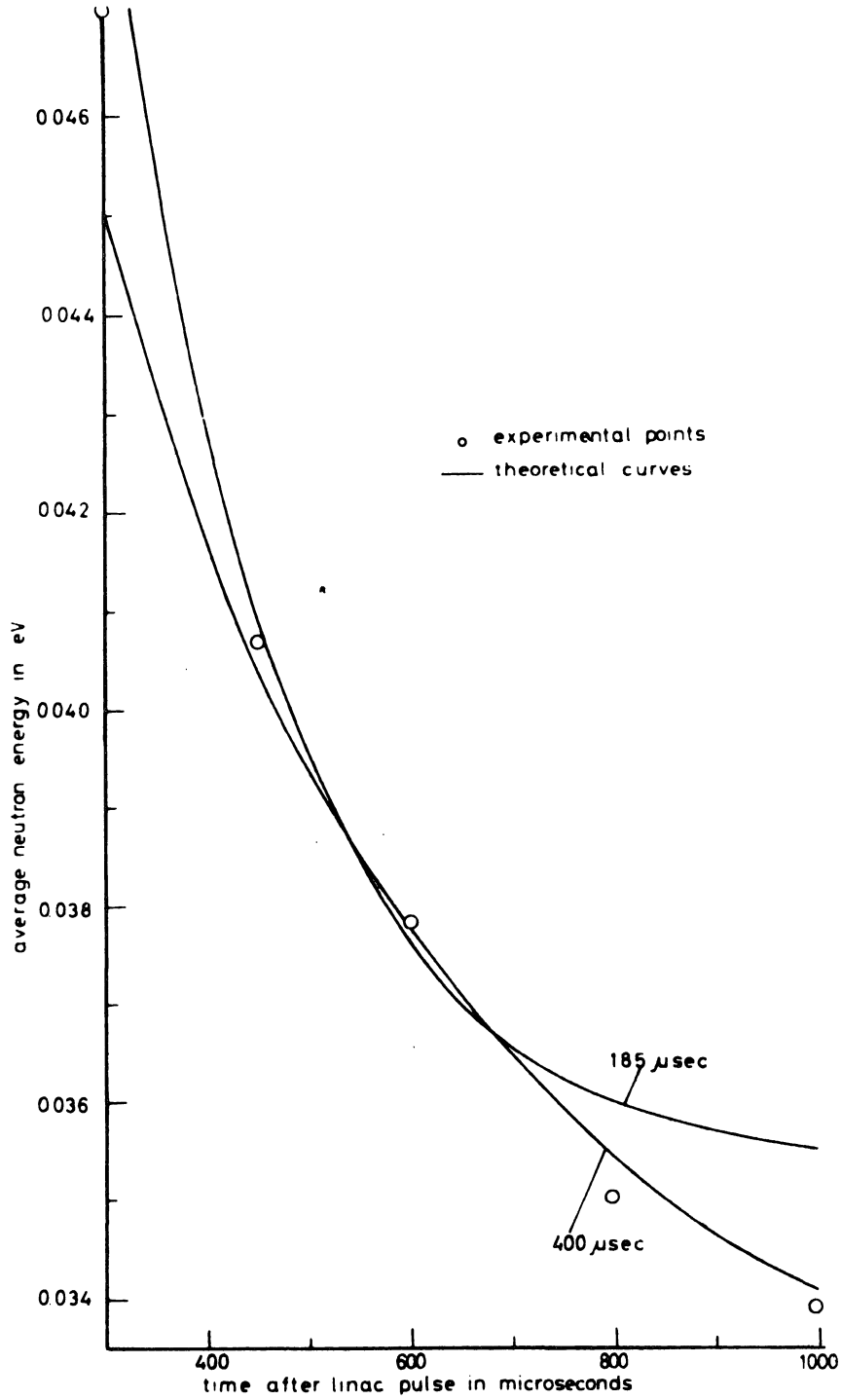
3. The Final, or Asymptotic ($t \to \infty$) Distribution

In this experiment, the nature of the energy spectrum is investigated long after the high-energy pulse is introduced. Once again, a chopper is used in conjunction with time-of-flight



Time Dependent Neutron Spectra in Graphite

Figure 2



Time Behavior of Average Neutron Energy in Graphite

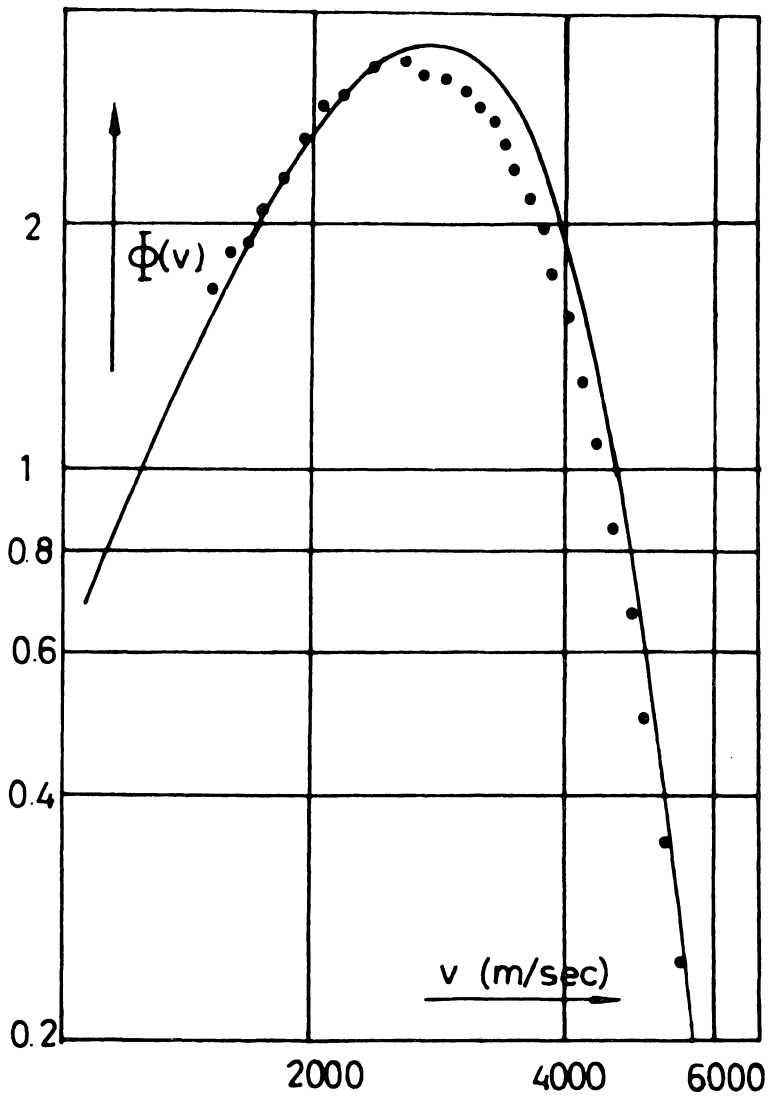
Figure 3

apparatus. The chopper is opened at a suitable time after the pulse, and kept open for a time long enough to secure sufficient neutrons for time-of-flight analysis, but short compared with a typical flight time. In the interior of a system that is weakly poisoned, and sufficiently large, the asymptotic flux will assume the form of the "fundamental decay mode,"

$$\lim_{t \to \infty} \phi(E,\vec{r},t) = e^{-\lambda_o t} \phi_o(E) \qquad (2)$$

The results of such an experiment have been reported by Beckurts (4). The moderating material was light water, and $\phi_o(E)$ was found to be Maxwellian in the interior of the larger blocks. As the blocks were made smaller, the maximum of $\phi_o(E)$ occurred at progressively lower energies. This effect, which is caused by the preferential leakage of fast neutrons, is called "diffusion cooling," and is demonstrated in Figure 4. The neutrons which emerged from the face of the moderating block were also analyzed by Beckurts. These, which are in a sense complementary to the $\phi_o(E)$, had their maximum displaced above the Maxwellian.

These experiments exhibit in a direct manner the interplay of diffusion and inelastic scattering in the determination of the asymptotic spectrum. Their analysis in terms of detailed models of matter has only begun. Related experiments are under way at the General Atomic Laboratory, where R. Beyster and his colleagues are measuring leakage spectra from slabs of light water (5) . It



The Asymptotic Spectrum From the Center of a $(5\ cm)^3$ -Water Cube

Figure 4

is hoped that these experiments will yield the angular variation of the leakage flux. Their interpretation leads one to consider a form of the "multi-velocity Milne problem."

4. Integral Experiments

The integral experiments of thermalization require much less equipment for their performance than do the experiments mentioned above. No attempt is made to measure the distribution-in-energy, so that choppers and time-of-flight equipment are unnecessary. Instead, one concentrates upon the temporal charge of the energy-integrated flux in a pulsed block, or the change in the activation of foils with change of position in a steady state experiment.

a) Temporal Relaxation (λ_o vs. B^2)

In these experiments, a "1/v" counter is placed near the moderating block. When a sufficiently long time has elapsed after initiation of the pulse, the flux has the form of equation (2), and the counting rate decays exponentially. Figure 4 illustrates the decay of a neutron pulse in a block of beryllium, as measured by de Saussure and Silver (6). The linear variation of the experimental points on "semi-log" paper indicates that we are indeed in the asymptotic regime. The slope of the upper curve yields the fundamental decay constant, λ_o.

The variation of λ_o with block size (and shape) will reflect the relative strengths of the diffusive and thermalization processes taking place in the moderator. In a moderator containing a 1/v absorber, diffusion theory predicts that

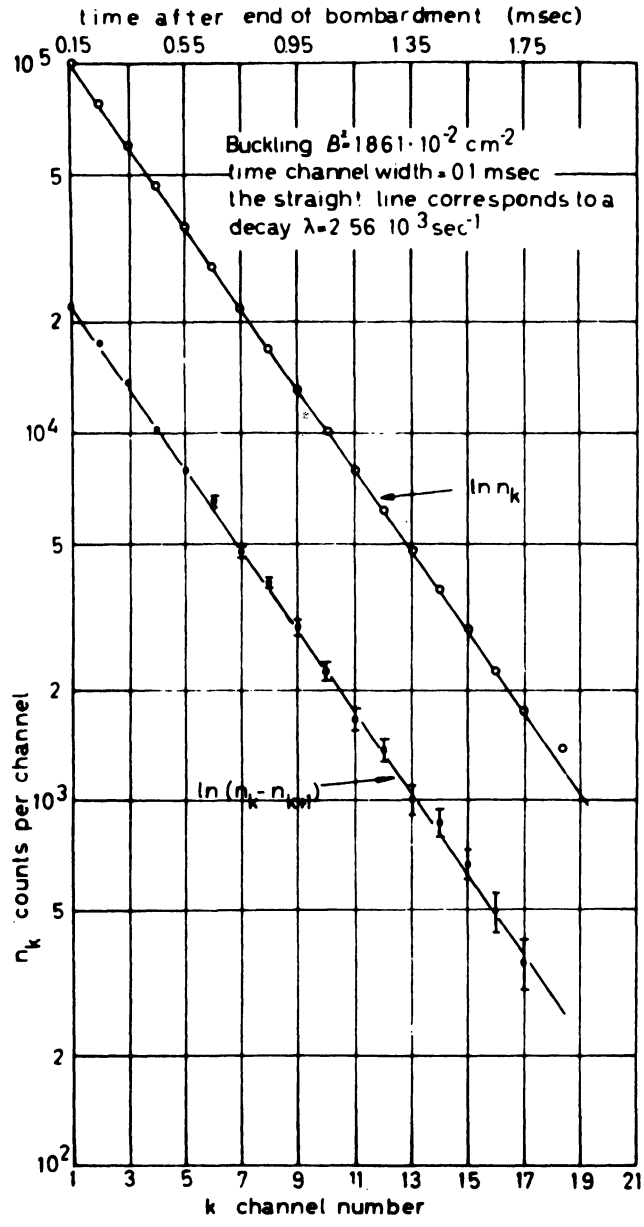
$$\lambda_o = \lambda_a + \langle vD \rangle B^2 \quad , \tag{3}$$

where $\lambda_a = v_o \Sigma_a(v_o)$, B^2 is the buckling of the fundamental spatial mode and $\langle vD \rangle$ is the average of the product of velocity and diffusion coefficient, with respect to the (non-Maxwellian) asymptotic spectrum. The diffusion cooling effect, which is an expression of the distortion of the spectrum, shows itself in the dependence of $\langle vD \rangle$ upon B^2, and the consequent nonlinearity of the λ_o vs. B^2 curve. It is customary, therefore, to rewrite (3) as

$$\lambda_o = \lambda_a + \langle vD \rangle_o B^2 - CB^4 + \ldots \quad , \tag{4}$$

where $\langle vD \rangle_o$ is a Maxwellian average, C is termed the diffusion cooling coefficient, and the dots represent still higher terms in the series. A typical curve of λ_o vs. B^2 may be seen in Figure 6.

To form an idea of the time scale of these processes, let us note that the slowing down time from Mev to ev is roughly 10 microseconds (μ sec) for a neutron in water and roughly 100 μ sec in graphite. In water, approximately 20 μ sec are required for the establishment of an asymptotic mode, while at least 1000 μ sec are needed in graphite. The flight time for a 1-ev neutron is approximately 70 μ sec/meter. Once an asymptotic mode



Decay of the Neutron Flux in a Berryllium Assembly 14 3/8 x 14 3/8 x 16"

Figure 5

has been established, its lifetime is given by equation (4). Some approximate experimental values of λ_o, $\langle vD \rangle_o$, and C follow. It will be seen that the experimental values of C, which characterize the diffusion cooling effect in a crude manner, are far from precise (7).

	λ_a	$\langle vD \rangle_o$	C
H_2O	4850 $\pm$ 50	0.35-0.37 x 10^5	2900-4900
Graphite	88 $\pm$ 1.5	2.1 $\pm$ 0.02 x 10^5	16-38 x 10^5
	sec^{-1}	cm^2/sec	cm^4/sec

b) Spatial Relaxation (κ^2 vs. Σ_a)

Let us imagine a beam of thermal neutrons incident upon the face of a semi-infinite block of material. If we sample the neutron flux, $\phi(E,z)$, throughout the block, we will find in many cases that when the distance from the face is sufficiently large, an asymptotic distribution in energy sets in:

$$\lim_{z \to \infty} \phi(E,z) \longrightarrow e^{-\kappa z} \bar{\phi}_o(E) \qquad (5)$$

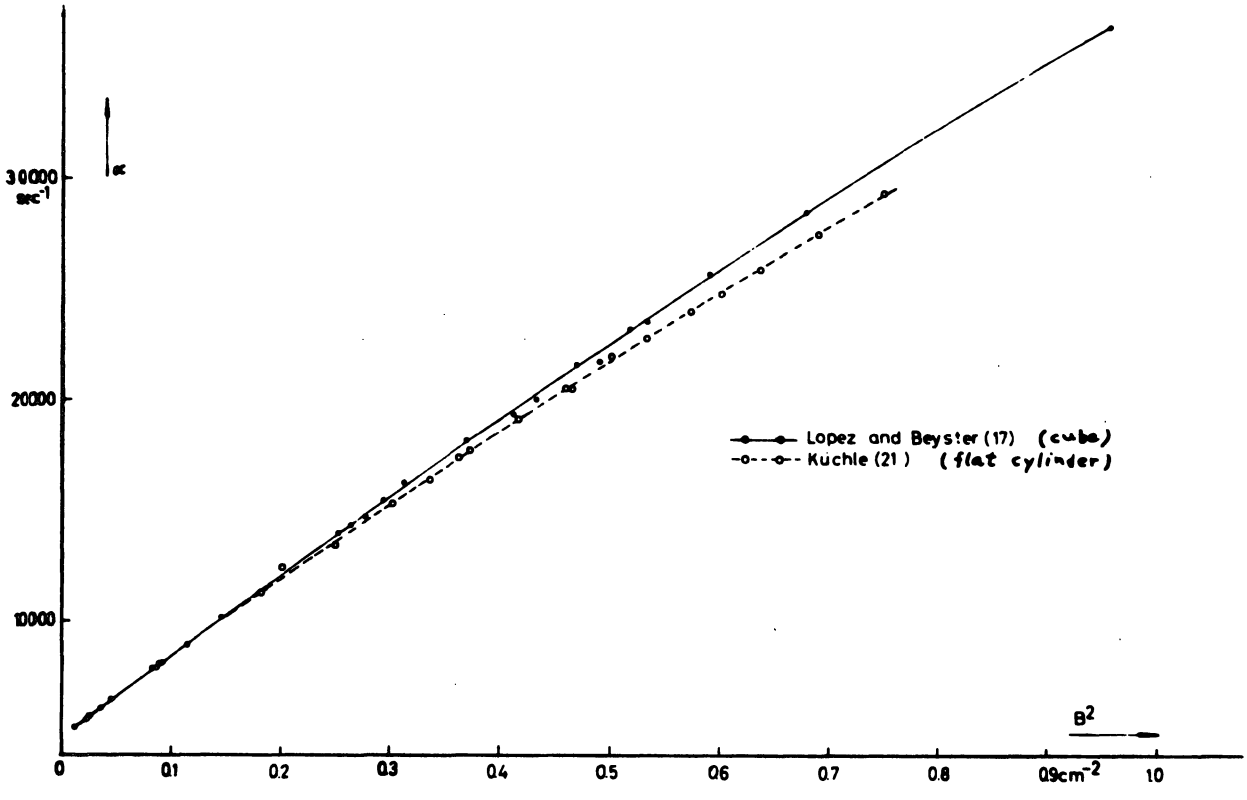
The $\bar{\phi}_o$ occurring in (5) is quite different from the ϕ_o of equation (2). Indeed, $\bar{\phi}_o(E)$ tends to be diffusion heated rather than cooled. The decay constant, κ, is characteristic of the state of the moderator, and is the reciprocal of the thermal diffusion length. Since a determination of $\bar{\phi}(E)$ seems to be difficult to perform, experiments have been limited

to the dependence of κ on material, on temperature, concentration of 1/v absorber, etc. The κ^2 vs. Σ_a measurements tie in nicely with the pulsed neutron experiments, since it is possible to show that the curve of Σ_a vs. κ^2 for 1/v absorber is the analytic continuation of the λ vs. B^2 curve. Thus, either experiment will yield values of $\langle vD \rangle_o$ and $C(\underline{8})$.

Another experiment which illustrates spatial relaxation is the "two-block" experiment. Here, the relaxation of a neutron distribution is studied in a system composed of two contiguous blocks of moderator maintained at different temperatures. Until quite recently ($\underline{9}$), attempts to measure the distribution-in-energy as a function of position have been thwarted by low neutron fluxes. Instead, one can observe the approach to equilibrium indirectly by measuring the dependence of foil activation upon position. Figure 7 shows a typical set of traverses which are analyzed by means of a variant of multigroup diffusion theory ("rethermalization theory") to obtain relaxation lengths. Several experiments, in which graphite and light water serve as moderators have been analyzed by physicists at Hanford, and are discussed in a recent report by R. A. Bennett ($\underline{10}$).

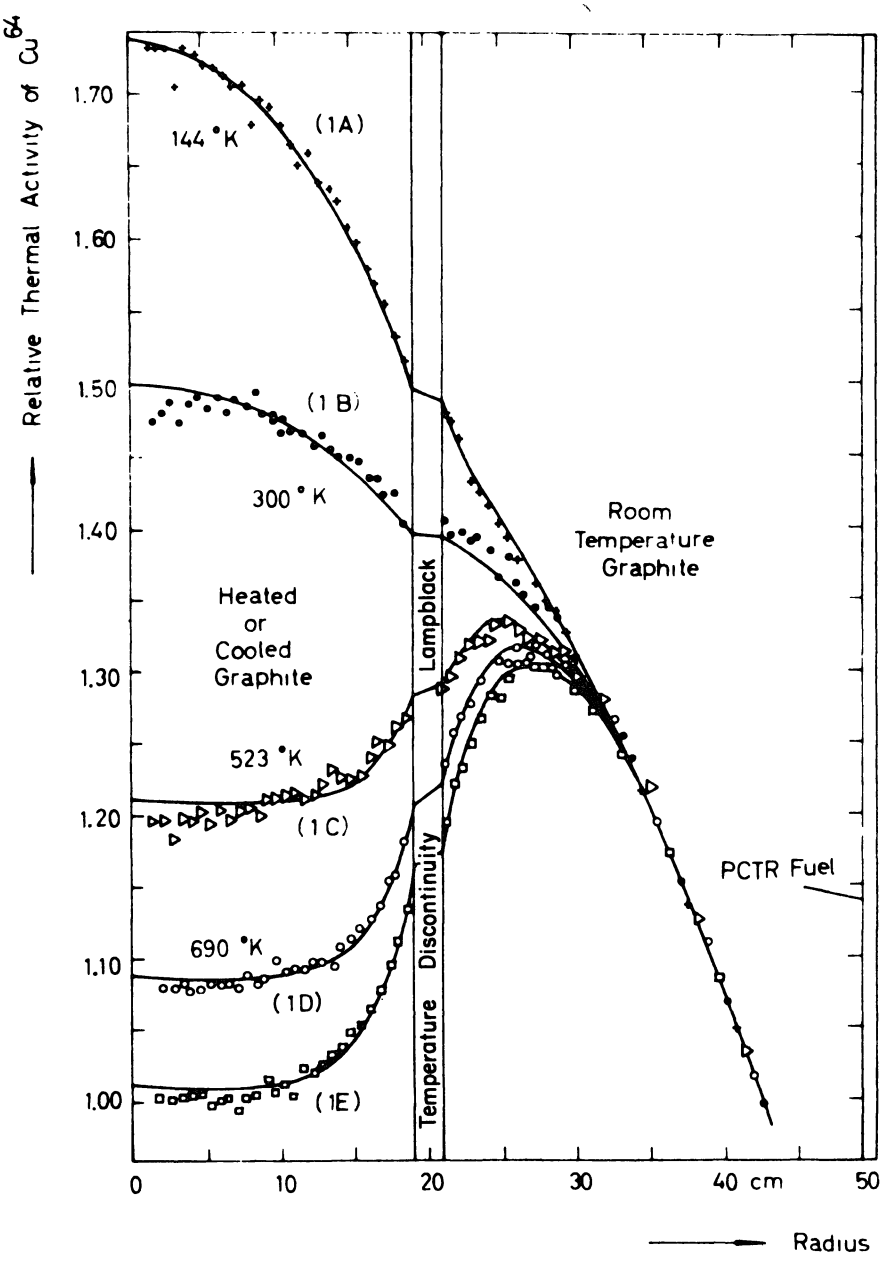
Experiments describing thermalization in lattices are also carried out by the foil traverse technique. In this situation, the neutron distribution varies strongly with position and is far from the asymptotic shape appropriate to either moderator or fuel. Rather special methods leaning strongly upon numerical analysis are used to interpret these experiments. The interested

reader should consult the proceedings of the recent Brookhaven Conference for details.



α VS. B^2 for H_2O at 26,7° C

Figure 6



Radial Traverses of Cu^{64} Activity in Graphite Regions

Figure 7

Lecture II

The Theory of Neutron Thermalization

1. The Boltzmann Equation

The neutron distribution $\phi(E,\vec{r},t,\Omega)$ relevant to a particular experiment in thermalization is a solution of the general Boltzmann equation,

$$\frac{1}{v}\frac{\partial\phi}{\partial t} + \hat{\Omega}\cdot \mathrm{grad}_{\vec{r}}\,\phi + \Sigma_a\phi + \Sigma_s\phi$$

$$= \int d\hat{\Omega}' \int_0^\infty dE'\, \phi(E,\vec{r},t,\hat{\Omega}')\Sigma_s(E',\hat{\Omega}'\to E,\hat{\Omega}) + S \quad , \tag{6}$$

corresponding to particular initial and boundary conditions. The notation and derivation of equation (6) are well-known (11). In solving it for thermal neutron distributions we face a two-horned difficulty. The first is the construction of a scattering kernel, $\Sigma_s(E',\Omega'\to E,\Omega)$, adequate to represent the complicated physical state of the moderator; the second is to solve the transport equation which results. It is important to realize that the "thermal" equation is, in its energy dependence, quite a different animal from the "slowing-down" equation, though it contains the latter as a special case. During slowing-down, neutrons lose energy with each collision. The integral equation is of "Volterra" form and is amenable to numerical methods of solution. In the thermal regime, either gain or loss of energy

may occur during a collision. The integral equation is of "Fredholm" type (though singular!) and is much more demanding of machine capacity, if one resorts to numerical methods. Fortunately, the computing machines presently available are too small to handle the full thermalization problem. There remains a need for a physicist's insight in proposing suitable approximations and simple models, and in seeking general laws in the mountains of "machine output."

The first approximation often made is the assumption that the scattering pattern is isotropic in the laboratory system. Thus, one retains only the first term in a Legendre polynomial expansion of the kernel with respect to the angle of scattering. One finds

$$\Sigma_s(E', \hat{\Omega}' \rightarrow E, \hat{\Omega}) \approx \frac{1}{4\pi} \Sigma_s(E' \rightarrow E) \quad . \tag{7}$$

The validity of this approximation has been established chiefly by "handwaving." The higher terms in the Legendre expansion are proportional to correspondingly higher powers of the ratio of neutron mass to scattering atom mass. One might argue that the ratio is more likely that of neutron mass to effective mass of the scattering atom, the latter being determined by the chemical binding of the scatterer. In the thermal region, these "effective masses" tend to be several times the free atomic masses and, consequently, the higher terms should be small. The approximation is sensible only for the inelastic, incoherent portion of the

scattering, and in the treatment of problems where "last flight" and transport effects are small. Then, it lends to considerable simplification of the mathematics. When one finds it necessary to take anisotropic scattering into account, it may be done by introducing a proper transport cross section into the scattering kernel (matrix) (12).

The replacement of transport theory by diffusion theory represents a second, popular approximation. Here, the implications of the Ansatz are well-known. To remind the reader, we note that the angular flux, $\phi(\vec{r}, \hat{\Omega}, ..)$, is written

$$\phi(\vec{r}, \hat{\Omega}, ..) = \frac{1}{4\pi} \phi_0(\vec{r}) + \frac{\vec{r} \cdot \hat{\Omega}}{|\vec{r}|} \phi_1(\vec{r}) \quad , \tag{8}$$

where $\vec{r}$ is the radius vector appropriate to the geometry under consideration, $\phi_0(\vec{r})$ is the scalar flux, and ϕ_1 is proportional to magnitude of the current. When equation (8) is substituted into the transport equation and we make some additional approximations, among them that the neutron distribution not change too rapidly in a mean collision time, $1/v\Sigma$, we are led to the diffusion equation,

$$-D(E)\nabla^2 \phi(E, t, \vec{r}) + \frac{1}{v} \frac{\partial}{\partial t} \phi + \Sigma_a \phi = S\phi + \mathcal{S} \quad . \tag{9}$$

The energy dependent diffusion coefficient is $D(E) = 1/3\Sigma_{tr}(E)$, and S, the scattering operator, is

$$S\phi = \int_0^\infty dE' \left[\phi(E') \Sigma_s (E', E) - \phi(E) \Sigma_s (E, E') \right] \qquad (10)$$

in the case of isotropic scattering. When we consider systems in which the distribution changes rapidly in time or in space, the approximation will fail. In most cases, however, diffusion theory is a good "first approximation."

2. The Scattering Operator

The scattering operator plays a central role in thermalization. Before entering into its details, we consider some general properties. We shall treat only the isotropic operator; the general operator has quite similar properties.

a) Neutron Conservation

As a consequence of the definition of S,

$$\int_0^\infty dE S \phi(E) = 0 \qquad (11)$$

for any ϕ (E) for which the integral exists. The "physical" meaning of equation (11) is clear. As an example, when the space and time-independent Boltzmann equation is integrated with respect to energy, equation (11) ensures that we are left with "rate of absorption" equals "source strength."

b) Detailed Balance

This important property of the scattering kernel is essentially a sufficient condition for the establishment of statistical equilibrium between the moderator and the neutron

gas. It is grounded upon the transformation properties, under time-reversal and space reflection, of the equations of motion which govern the scattering of the neutron by the atomic nucleus (12a, 13). It requires that the full scattering kernel satisfy

$$M(E)\Sigma_s(E, \hat{\Omega} \to E', \hat{\Omega}') = M(E')\Sigma_s(E', -\hat{\Omega}' \to E, -\hat{\Omega}), \tag{12}$$

while its isotropic part fulfills

$$M(E)\Sigma_s(E, E') = M(E')\Sigma_s(E', E) \ . \tag{13}$$

We show in Appendix I that the general scattering kernel which we consider does indeed satisfy (12).

Equation (13) guarantees that a Maxwellian distribution is a steady state solution to the infinite medium equation. It also implies that our scattering operator is "essentially" symmetric. Thus, if we write $\bar{\phi}(E,\ldots) = \left[M(E)\right]^{-\frac{1}{2}} \phi(E,\ldots)$, there results

$$\left(\frac{1}{v}\frac{\partial}{\partial t} + \ldots\right)\bar{\phi} = \int_0^\infty dE' \bar{\phi}(E') \frac{M(E')\Sigma_s(E', E)}{\sqrt{M(E')M(E)}} - \Sigma_s(E)\bar{\phi}(E) \tag{14}$$

which is an integral equation having symmetric kernel. We shall, however, continue to work with the non-symmetric form of the Boltzmann equation.

The scattering operator adjoint to S will be denoted by S^+; its kernel is the transpose of the scattering kernel in S. From its definition,

$$S^{+}\phi = \int_{o}^{\infty} dE' \Sigma_s(E,E') \left[\phi(E') - \phi(E)\right] \quad , \tag{15}$$

while detailed balance implies the operator identity,

$$MS^{+}\phi = SM\phi \quad , \tag{16}$$

where M is the Maxwellian.

c) Eigenfunctions and Eigenvalues

We noted in our discussion of experiments the role played by separable solutions of the form $e^{-\lambda t}\phi(E)$ and $e^{-\kappa z}\psi(E)$. The search for these solutions leads to eigenvalue equations in which the scattering operator plays an important role. For example, the evolution of a pulse in a very large block of moderator might be expressed by:

$$\phi(E,t) \sim \sum_i a_i \phi_i(E) e^{-\lambda_i t} \quad , \tag{17}$$

where the ϕ_i and λ_i are solutions to the homogeneous integral equation,

$$S\phi_i = -\frac{\lambda_i}{V}\phi_i \quad . \tag{18}$$

In the second case, that of steady state diffusion in a non-capturing halfspace,

$$\phi(E,z) \sim \sum_i b_i \psi_i(E) e^{-k_i z}$$

leads to the equation

$$S\psi_i = -D(E)\kappa_i^2\psi_i \qquad (19)$$

for eigenfunctions and eigenvalues. Note that the different weighting factors on the right hand sides of (18) and (19) cause the eigenfunctions and eigenvalues to be quite different in the two cases. More complex problems, in which one takes absorption and leakage into account further complicate the weight factor.

The calculation of the eigenfunctions and eigenvalues is not an easy job. This becomes apparent when one expands ϕ_i in a complete set of "appropriate" basis functions and seeks to diagonalize the resulting matrix equation. One must calculate many complicated matrix elements in order to obtain the first few eigenvalues with reasonable precision. There are two principal difficulties. The first is that we do not have a set of basis functions as felicitous as, say, the plane waves of quantum mechanics. The second is that the eigenvalue spectrum of the scattering operator is not the usual, discrete point spectrum extending to infinite. Rather, it is likely that all realistic scattering operators have a discrete spectrum with a limit point at some finite value of λ, denoted λ_o, augmented by a continuous spectrum extending from λ_o to infinite. The eigenfunctions corresponding to the continuous spectrum are singular, though integrable, functions.

This unusual eigenvalue spectrum would appear to express the fact that the kernel of the full scattering operator is not square-integrable. Thus,

$$\int_0^\infty dE \int_0^\infty dE' \left| S(E,E') \right|^2$$

diverges, the Fredholm theorems do not apply, and we must face "...the presence of finite accumulation points of the spectrum of eigenvalues, or even of a continuous spectrum, i.e., of eigenvalues filling whole segments of the λ-axis or even the entire λ-axis." [Tricomi, reference (14).] Recently, we have been able to demonstrate precisely this behavior in the case of thermalization by a gas of protons (15).

Finally, the reader should note that equation (16) suggests a very simple relation between the eigenfunctions of S and S^+. If v_λ is an eigenfunction of S^+, corresponding to λ, then Mv_λ is an eigenfunction of S, also corresponding to λ. The eigenvalues are all real.

d) The Scattering Kernel

The cross sections for the scattering of slow neutrons by matter have an appearance that is quite different from those describing the scattering of fast neutrons. The reasons are well known: thermal neutrons have a de Broglie wavelength comparable with the spacing between scattering atoms, and a kinetic energy comparable with typical energy-level spacings in crystals. The first property suggests interference scattering, or scattering from the system "as a whole," rather than from

individual atom. Thus, the "coherent" scattering is anything but a smooth function of angle and energy. Its elastic portion, which is dominant, yields the Bragg pattern of spots when the scattering is due to a single crystal, or the Debye-Scherrer pattern of rings when the scattering sample is a powder of crystallites. The second property suggests that the shape of the kernel will be quite specific and will reflect fine details of the dynamics of the scattering system. It appears, then, that solution of a proper thermal Boltzmann equation will be very difficult if we take account of the full kernel.

The most convenient way of expressing the scattering cross sections is to use the elegant formulation that is due to Von Hove (16) and to Glauber (17). The differential cross section per unit solid angle and unit interval of final neutron energy is divided into its coherent and incoherent parts, and one has:

$$\left(d^2\sigma/d\Omega\, d\epsilon\right)_{coh} = \frac{\langle a\rangle^2 N}{2\pi\hbar}\,\frac{k}{k_o}\int d^3r\,dt\,\exp\left[i(\kappa\cdot r-\omega t)\right]G(\vec{r},t)$$

$$\left(d^2\sigma/d\Omega\, d\epsilon\right)_{inc} = \frac{\{\langle a^2\rangle-\langle a\rangle^2\}N}{2\pi\hbar}\,\frac{k}{k_o}\int d^3r\,dt\,\exp\left[i(\kappa\cdot r-\omega t)\right]G_s(\vec{r},t)\,, \tag{20}$$

where $\epsilon = \hbar\omega = E_o - E$, is the energy transferred by the neutron,

$\hbar\vec{\kappa} = \hbar(\vec{k}_o - \vec{k})$ is the momentum transferred by the neutron, and

N = the number of scattering particles.

The quantity a is the (nuclear) scattering length. It will depend upon nuclear species and upon the spin state of the neutron-nucleus scattering system. Averages of a and a^2 appear in (20). We shall limit ourselves to systems composed of a single type of atom, so that $\langle a \rangle$ and $\langle a^2 \rangle$ refer to spin averages.

Were the scattering amplitude independent of spin, the cross section would be entirely coherent. The incoherent scattering is connected with inability of waves representing neutrons whose spins have been "flipped" during the scattering process to interfere with those whose spin states remain the same. In most cases, the coherent scattering cross section $4\pi \langle a \rangle^2$ is much greater than the incoherent cross section, $4\pi \left\{ \langle a^2 \rangle - \langle a \rangle^2 \right\}$. The proton and the vanadium nucleus are famous exceptions to the rule, while the deuteron occupies an intermediate position. For example,

	Cross Sections in Barns	
	σ_{coh}	σ_{inc}
H	1.8	79.7
V	0.03	4.8
D	5.4	2.2
C	5.5	ess. zero

The key to equation (20) is the pair of functions, $G(\vec{r},t)$ and $G_s(\vec{r},t)$, which contain the dynamical properties of the scattering system. They are complex-valued functions, containing $\hbar$, which possess a simple, physical interpretation when $\hbar \to 0$, and they become real functions. In the classical limit, $G(\vec{r},t)d^3r$

gives the probability of finding a particle in the volume d^3r around $\vec{r}$ when it is known that there was a particle at the origin of coordinates at time t = 0. G(r,t) is, then, a kind of "two-particle, time-displaced, correlation function." It is natural to express it as the sum of two parts: G_s, which gives the probability for finding at r the very same particle which began at the origin, and G_d, which gives the probability of finding a <u>different</u> particle at $\vec{r}$. The reader will note that the incoherent scattering depends only upon G_s, while both G_s and G_d affect the coherent scattering. Further, some reflection will suggest that G_s is <u>much</u> easier to compute than is G_d, since the former is, to a good approximation, a one-particle problem, while G_d involves an N-particle problem.

The relative ease of calculation of G_s leads us to an approximation which has been used extensively in thermalization. It is called the "incoherent approximation" and consists in replacing G by G_s in the expression for the coherent scattering. We then have a single expression,

$$\left(d^2\sigma/d\Omega\, d\epsilon\right)_{total} \approx \frac{\langle a^2\rangle N}{2\pi\hbar}\,\frac{k}{k_o}\int d^3r\, dt\, \exp\left[i(k\cdot r-\omega t)\right]G_s(r,t) \qquad (21)$$

for the scattering kernel. The quantity $\langle a^2\rangle$, when multiplied by 4π, is known as the bound-atom cross section, σ_b. The name "free atom cross section" is given to $\sigma_b/(1+\mu)^2$, where μ is the ratio of neutron mass to scattering atom mass.

It is not at all obvious that the incoherent approximation is a good one, though it has performed well in practice. When the neutron energy is large compared with kT, its wavelength is a small fraction of the atomic spacing, and small values of t (time) dominate the integrations in (20). Then, G_s is a very good approximation to G and the incoherent approximation yields the correct result. On the other hand, for neutron energies below the Bragg cutoff energy for the crystal lattice, the coherent scattering vanishes, and equation (21) if taken literally, gives nonsense. One can find further discussion of the incoherent approximation in references (18) and (19), where it is pointed out that in scattering by crystals the approximation consists in replacing a sum over reciprocal lattice vectors by an integration. In any case, Bragg patterns and the typically saw-toothed behavior of total cross sections no longer appear in the kernel of equation (21); they have been "averaged out" and we deal only with the self-correlation function, G_s.

We shall introduce some additional notation before proceeding to a discussion of G_s. The spatial part of the Fourier transform in (20) and (21) is called the intermediate scattering function. It is denoted by many symbols in the literature, among them $\chi(\vec{\kappa},t)$, $\gamma(\vec{\kappa},t)$, and $I(\vec{\kappa},t)$. We choose

$$\int d^3r \exp(i\kappa\cdot r) G(\vec{r},t) = \chi(\vec{\kappa},t),$$

and (22)

$$\frac{1}{(2\pi)^3}\int d^3\kappa \exp(-ir\cdot\kappa)\chi(\vec{\kappa},t) = G(\vec{r},t)$$

The entire integral in (20), (21) is sometimes called the "scattering law," $S(\kappa, \omega)$. Thus,

$$\left(d^2\sigma/d\Omega d\epsilon\right) = \frac{k}{k_o}\left(N\langle a\rangle^2 \bar{S}_{coh}(\vec{\kappa}, \omega) + N\left\{\langle a^2\rangle - \langle a\rangle^2\right\}\bar{S}_{inc}(\vec{\kappa}, \omega)\right) \quad (23)$$

Finally, there is a convenient form of the scattering law, known as $S(\alpha, \beta)$. When the scattering from a sample is independent of the sample's orientation in space, $\bar{S}(\kappa, \omega)$ depends only upon κ^2. Accordingly, we introduce the dimensionless variables $\alpha = (\hbar^2\kappa^2)/2MT$ and $\beta = \hbar\omega/T$, where M is the mass of the scattering atom (not the neutron.) $S(\alpha, \beta)$ is then defined as

$$S(\alpha,\beta) = \left(e^{\frac{1}{2}\beta}/T\right)\bar{S}(\kappa, \omega) \quad . \quad (24)$$

Note that $S(\alpha, \beta) = S(\alpha, -\beta)$, and that the scattering law is a function of but two variables when it describes such an "isotropic" system. For a discussion of the experimental determination of $S(\alpha, \beta)$, see reference (20).

e) The Self-Correlation Function

The self-correlation function can be calculated for a variety of useful systems (21). One usually proceeds from the intermediate function. For example,

i) ideal gas of point particles, at temperature = T

$$\chi(\kappa^2, t) = \exp\left[-\frac{\kappa^2}{2M}\left(Tt^2 - i\hbar t\right)\right] \quad , \quad (25)$$

from which we deduce

ii) free particle (at rest)

$$\chi(\kappa^2,t)=\exp\left[-\frac{\kappa^2}{2M}(-i\hbar t)\right] \quad , \tag{26}$$

iii) crystal, one atom per cell, isotropic approximation

$$\chi(\kappa^2,t)=\exp\left[-\frac{\kappa^2}{2M}\left\{\int_0^\infty \frac{d\omega}{\omega} f(\omega)\left[\frac{z+1}{z-1}-\frac{z}{z-1}e^{i\omega t}-\frac{1}{z-1}e^{-i\omega t}\right]\right\}\right] \quad . \tag{27}$$

Here $f(\omega)\,d\omega$ is the fraction of normal modes lying between ω and $\omega+d\omega$, and $z = \exp(\hbar\omega/T)$. One obtains the Debye and Einstein models by taking

$$f(\omega)=\frac{3\omega^2}{\omega_D^3} \tag{28}$$

$$f(\omega)=\delta(\omega-\omega_E) \quad , \text{ respectively.}$$

The κ^2 dependence of the χ-function in cases i)-iii) is unusually simple. We may write these functions and the corresponding correlation functions in the form: (22)

$$\chi = \exp\left[-\kappa^2\omega^2(t)\right]$$

$$G_s=\left[2\sqrt{\pi}\;\omega(t)\right]^{-3}\exp\left[-\frac{r^2}{[2\omega(t)]^2}\right] \tag{29}$$

The self-correlation function describes a generalized diffusion of the particle from its position at t = 0. The "width" of the Gaussian, $2\omega(t)$, is a very complicated function of the time. In fact, it is not a real function of time, but rather a real, even function of the complex time $\tau = t - \frac{i\hbar}{2T}$. We are concerned with a quantum mechanical "diffusion," described by a correlation function that coincides with its classical counterpart, when $t \gg \hbar/T$. The manner in which $G_s(t)$ depends upon τ is closely connected with the property of detailed balance (23). (See Appendix I).

While Gaussian models have been used almost exclusively in studies of thermalization, one should realize that all isotropic systems do not carry Gaussian correlation functions. Fluids, and models of crystals more realistic than iii), are non-Gaussian in behavior. However, it is always possible to write down the Gaussian part of a particular correlation function, and to show that the general expression for the width function is (24)

$$\omega^2(t) = -\frac{i\hbar t}{2M} + \frac{1}{3}\int_0^t dt'(t-t')\langle \vec{v}(0)\cdot\vec{v}(t')\rangle_T \quad , \qquad (30)$$

where $\langle \vec{v}(0)\cdot\vec{v}(t)\rangle_T$ is the correlation function for velocity, $\vec{v}$.

f) Some Approximations (25)

We have stressed the relative simplicity of calculation with Gaussian correlation functions. These are, however, still too

complicated to allow exact integration. One must fall back upon additional approximations, some of which we mention here.

1) The Phonon Expansion. This expansion is historically the oldest, and physically, the most perspicuous of all. It is used in the analysis of scattering by crystals, and consists in splitting the χ function of equation (27) into two factors, one containing the time-dependent exponentials, and the other the $\frac{z+1}{z-1}$ term. The first exponential is replaced by a series expansion in powers of its argument, and the expansion is truncated after a few terms. There appear, in the series, individual terms behaving as $1, e^{\pm i\omega t}, e^{\pm 2i\omega t}, \ldots$. These, when Fourier-transformed to obtain cross sections, give contributions from processes in which zero, one, two,...quanta of vibrational energy (phonons) are absorbed or emitted during the scattering.

The exponential which is independent of time is often called the Debye-Waller factor. If we write it as exp(-2W), we see that W is a decreasing function of $\langle q^2 \rangle / \lambda^2$, where $\langle q^2 \rangle$ is the mean square deviation of an atom from its equilibrium position in the lattice, and λ^2 is the square of the neutron wavelength (elastic scattering!). Its effect is most pronounced in elastic, coherent scattering, where it expresses the damping or smearing effect of thermal vibrations upon the scattering pattern.

True phonon expansion is rarely carried beyond the two phonon term. Its convergence is poor unless the system is tightly bound and the neutron energy is "suitably thermal."

2) The Mass Expansion. This expansion, which is a modification of the phonon expansion, was first proposed by G. Placzek (26), then discussed in detail in an important paper by A. Sjölander (27). It has proved to be of great value in the analysis of inelastic, incoherent scattering.

The approximation consists in expanding the entire χ function in powers of the width function, the Debye-Waller factor not being given special treatment. The first few terms in the expansion are significantly better than the corresponding terms in the phonon expansion, especially when the energy transfer is not too large (M large, and/or "tight binding"). Higher terms in the expansion may be approximated through an ingenious application of the Central Limit Theorem of statistics.

3) The Time Expansion. This treatment of the scattering was first given by G. Placzek (28); its present form is due, however, to G. C. Wick (29). The idea is to expand the χ function in a power series in time. Thus, for the self-correlation function,

$$\chi_s(\kappa^2, t) = \exp\left(\frac{i\kappa^2 \hbar t}{2M}\right) \sum_{n=0}^{\infty} S_n(\kappa^2) \frac{(i\hbar t)^n}{n!} \quad , \qquad (31)$$

where the $S_n(\kappa^2)$, called "Placzek moments," express the inner nature of the scattering system. That the $S_n(\kappa^2)$ are, indeed, moments of a distribution, appears when one considers moments $\int d\omega \left(\omega - \frac{\kappa^2}{2M}\right)^n \bar{S}(\kappa, \omega)$ of the scattering law. It may be seen that these quantities are proportional to the $S_n(\kappa^2)$. Some values are:

$$S_o = 1$$

$$S_1 = 0$$

$$S_2 = \frac{4}{3}\,\frac{\hbar^2 k^2}{2M}\,\langle K\rangle_T$$

$$S_3 = \frac{1}{3}\,\frac{\hbar^2 k^2}{2M}\,\langle \frac{\hbar^2}{M}\nabla^2 V\rangle_T \quad , \qquad (32)$$

where $\langle K\rangle_T$ is the average kinetic energy of the scattering atom, and V is its effective potential energy. The moments are generally polynomials in $k^2/2M$, and can be calculated systematically.

When the expansion (31) is substituted into the equations which define the cross section, one finds that each term may be integrated. While the resulting series does <u>not</u> converge, it is a useful asymptotic series in the region where the neutron's energy is considerably larger than a "typical" energy level spacing in the scattering system. This is the "multi-phonon" region, and the series may be regarded as giving a neat summation of these multiple processes. Finally, one should note that when the time expansion is truncated, a polynomial in t remains. Since the polynomial will not be an even function of $(t - \frac{i\hbar}{2T})$, the cross sections derived from it will not satisfy the detailed balance condition. However, they will be used only when $E \gg T$, in which region the loss of detailed balance need not concern us.

The three approximations we have mentioned have been used to interpret neutron scattering experiments with the aim of

learning more about the nature of the solid and liquid states. Recently, in studies of thermalization, the reverse has occurred. One begins with simple pictures of the motion of atoms in typical moderators, constructs the appropriate $\mathcal{K}$ function and attempts to calculate those cross sections needed for an analysis of the Boltzmann equation. For an example of such calculations, the reader should consult the general atomic studies of light water (30) and graphite (31), and remarks from other sources about heavy water (32) and beryllium (33). To date, these scattering kernels have been used most successfully in the analysis of steady state spectra in infinite homogeneous media and in lattices. Needless to say, a great deal of numerical calculation is called for.

Lecture III

Some Simple Models

Though we may grant that within the next decade we shall have computing machines capable of calculating a desired scattering law, then solving the Boltzmann equation to present us with a pile of "printout" containing all (nay, more than all) that we want to know about a particular problem, it remains valuable and satisfying to make some general inquiries about thermalization. For example, what features of the scattering kernel (and the moderator) control the thermalization phenomena we described earlier? We shall try to give a rough answer to the question in this lecture.

Consider the problem of the steady-state spectrum in an infinite, homogeneous medium. We described the general features of the spectrum in an earlier lecture. Now, let us ask how the Maxwellian and slowing-down components are altered as we change the amount of absorber for a given moderator, and as we change the binding of the moderator atoms for a fixed amount of absorber.

A very simple way to describe the spectrum when absorption is weak is to write it as a Maxwellian, $M(E)$, for $0 < E < E^*$ and as a slowing down spectrum, $1/E$, for $E > E^*$. One would take $E^* >> T$. If the absorption is $1/v$, we have $\Sigma_a(E) = \Sigma_a(T)\sqrt{T/E}$. Then, the steady state character of the distribution requires that

$$\phi(E) = A\,\frac{1}{E} \qquad E > E^*$$

$$= A\,\frac{2}{\sqrt{\pi}}\left(\Sigma_a(T)/\xi\,\Sigma_s\right)^{-1} M(E) \quad E < E^*. \tag{33}$$

Thus, the amplitude of the Maxwellian component varies inversely with the macroscopic absorption cross section, which appears in the dimensionless ratio ($\Sigma_a/\xi\,\Sigma_s$). The ratio, which we call $\frac{1}{4}\bar{\Delta}$, is one of absorption to slowing-down strength, and we shall meet it again. The absorption of neutrons is a process which hinders the establishment of thermal equilibrium, while the slowing-down strength aids it. The deviation of the steady-state spectrum from Maxwellian is the result of a competition between these processes, and its magnitude is generally a function of $\bar{\Delta}$ or of a quantity closely related to it.

Equations (33) are hardly more than schematic. The "Maxwellian component" of a true spectrum is not a Maxwellian distribution at the moderator temperature, nor is the slowing-down component simply 1/E. We shall first indicate how the slowing-down spectrum deviates from 1/E as E enters the thermalization domain.

We mentioned earlier that the time expansion of the intermediate scattering function yields expressions for the cross sections in the form of series that are useful when $E \gg kT$. For example, the asymptotic series for the total cross section,

$\sigma_s(E)$, is a power series in (T/E), with coefficients that are combinations of the Placzek moments (32). One might expect, therefore, to find a similar series describing the flux, $\phi(E)$. The form of this asymptotic series may be deduced after considerable calculation. For 1/v capture, and unit source strength, we find (34),

$$\phi E = \frac{1}{\xi \Sigma_s E} \sum_{n=0,\frac{1}{2},1,\ldots}^{\infty} (T/E)^n \sum_{p=0}^{[n]} \tau_n^p \left(-\Delta/2\right)^{2(n-p)} \quad (34)$$

where $[n]$ denotes the largest integer contained in n, and Δ is $2\Sigma_a(T)/\mu\Sigma_s$. Note that Δ becomes identical with $\bar{\Delta}$, the parameter introduced earlier, when μ, the ratio of neutron mass to scattering atom mass approaches zero.

The physical properties of the moderator express themselves through the coefficient τ_n^p, which are functions of the Placzek moments. They are determined from the recursion relation

$$b_0(n+1)\tau_n^p + \mu b_1(n+1)\tau_{n-1}^p + \ldots + \mu^p b_p(n+1)\tau_{n-p}^0 + \mu\tau_{n-\frac{1}{2}}^p = 0 \quad (35)$$

which they satisfy. The connection with the Placzek moments, $S_n(x)$, appears through

$$\mu^p b_p(s) = \sum_{l=0}^{p} \mu^l \gamma_l^{p+l} G_{pl}(s) \quad (36)$$

and

$$S_n(x) = \sum_{l=0}^{[n/2]} \gamma_l^n T^{n-l} (x/2M)^l \tag{37}$$

The $G_{pl}(s)$ in (36) depend only upon the mass-ratio, μ. Thus, the key quantities, the γ_l^n, are the coefficients in the polynomial expression for the moments. The γ_l^n are simply related to moments of the frequency distribution function for lattice vibrations. Thus,

$$\begin{aligned} T^n \gamma_0^n &= \delta_{no} \\ T^{n-1} \gamma_1^n &= \langle \omega^{n-1} \rangle \qquad (n \geq 2) \\ T^{n-2} \gamma_2^n &= \frac{1}{2!} \sum_{\substack{n_1, n_2 = 2 \\ (n_1 + n_2 = n)}}^{n-2} \frac{n!}{n_1! n_2!} \langle \omega^{n_1 - 1} \rangle \langle \omega^{n_2 - 1} \rangle \quad (n \geq 4) \end{aligned} \tag{38}$$

etc.

For n even

$$\langle \omega^n \rangle \equiv \int_0^\infty d\omega f(\omega) \omega^n \,, \tag{39}$$

while for n odd

$$\langle \omega^n \rangle \equiv \int_0^\infty d\omega f(\omega) \omega \coth \frac{\hbar\omega}{2T} \omega^{n-1} \tag{40}$$

Since $f(\omega)\omega \coth \frac{\hbar\omega}{2T}$ is the distribution function for the average vibrational energy of a mode, the moments we encounter

in (38) are even moments of either the distribution of frequencies or of vibrational energy. The quantity $\langle \omega^n \rangle$ (n odd) is also related to the Debye-Waller factor.

In Figure 8 we see curves deduced from the series (34) comparing thermalization by a gas of protons with that by the bound proton model proposed by Nelkin et al. (35). It is clear that there is a difference in the predictions of the two models, and that the bound model gives better agreement with the experiments in light water.

The leading terms in the asymptotic expansion are:

$$\xi \Sigma_s E \phi(E) = 1 - \frac{\Delta}{2} \tau_{1/2}^0 (T/E)^{1/2} + (\tau_1' + (\Delta/2)^2 \tau_1^0)(T/E) + \ldots \quad (41)$$

The second term gives the effect of absorption on the slowing-down spectrum. $\tau_{1/2}^0$ depends only upon μ, and the second term is quite independent of the state of the moderator. In the third term, τ_1^0 again depends only upon μ, and $(\Delta/2)^2(T/E)$ is independent of T. The only thermalization term is $\tau_1'(T/E)$, since τ_1' is proportional to γ_1^2, which in turn is $\langle \omega/T \rangle$. In fact, as $\mu \to 0$, τ_1' approaches γ_1^2. Since we are dealing with the average kinetic energy of a linear oscillator, we may introduce an effective temperature, through $\langle K \rangle \equiv \frac{1}{2} T_{eff}$. Thus, (41) becomes, as $\mu \to 0$

$$\xi \Sigma_s E \phi(E) = 1 - \frac{\Delta}{2}(T/E)^{1/2} + \left(2\frac{T_{eff}}{T} + \tfrac{1}{2}(\Delta/2)^2\right)(T/E) + \ldots \quad (42)$$

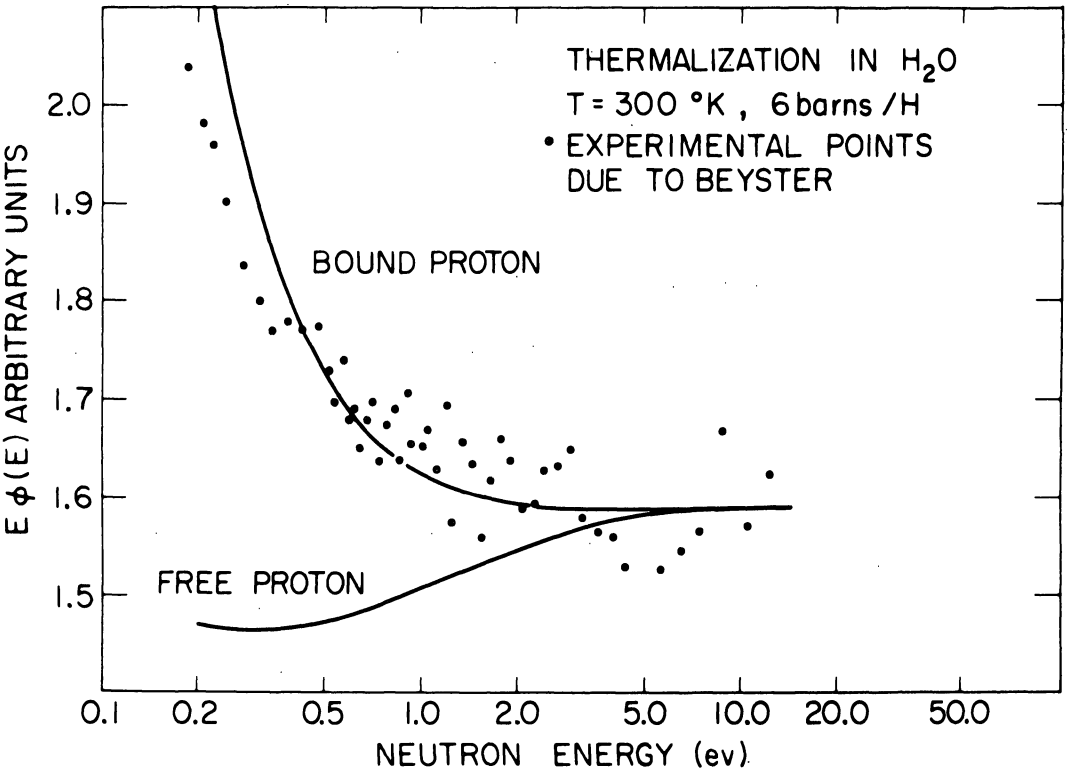

THERMALIZATION IN H_2O
T = 300 °K , 6 barns /H
• EXPERIMENTAL POINTS
DUE TO BEYSTER
BOUND PROTON
FREE PROTON
E φ(E) ARBITRARY UNITS
2.0
1.9
1.8
1.7
1.6
1.5
0.1
0.2
0.5
1.0
2.0
5.0
10.0
20.0
50.0
NEUTRON ENERGY (ev)

Figure 8

The foregoing equations show us that the "asymptotic" part of the thermal flux will be increased as the binding of the scattering atom is increased, since $\langle \omega /T \rangle$ increases with the stiffness of the lattice. Since the area of the "Maxwellian component" will not change--in first approximation--unless the concentration of absorber is changed, it is likely that the peak of the Maxwellian will be shifted to higher energies, as the binding is increased. The shift should be proportional to $\langle \omega /T \rangle$ or T_{eff}/T. We shall soon produce other evidence to support this argument.

The analysis leading to (34) is also applicable to time and space-dependent problems. To give an example, let us recall that the Laplace transform, with respect to time, of the time-dependent Boltzmann equation (infinite medium) is a time-independent equation containing an additional 1/v absorption. Thus, by merely rewriting (34) we have the expansion of $\tilde{\phi}(E,p) \equiv L\,\phi(E,t)$ in powers of the transform variable, p. Since

$$\int_0^\infty dt\; e^{-pt}\,\phi(t,E) \sim \sum_{n=0}^{\infty} \frac{(-p)^n}{n!} \int_0^\infty dt\, t^n\, \phi(E,t)\,, \qquad (43)$$

the coefficient of p^n is proportional to the n^{th} time moment of the time-varying flux seen at $E \gg T$. These results have not yet been fully exploited in the study of the "life-history" experiment.

Let us turn from the distribution for $E \gg T$ to that in the neighborhood of T. It will be convenient to work through the time-dependent problem first, and then make use of the fact that the integral of $\phi(E,t)$ with respect to time gives the steady-state, infinite medium flux. Consider, then,

$$\frac{1}{v}\frac{\partial \phi}{\partial t} + \Sigma_a \phi = \int_0^\infty dE' \left(\phi(E')\Sigma_s(E',E) - \phi(E)\Sigma_s(E,E')\right) + S(E,t) \quad (44)$$

If the cross section is "1/v" it may be eliminated from (44) by writing $\phi = \exp(v\Sigma_a(v)t)\ \phi_1(E,t)$. Further, let us assume that $S(E,t)$ is proportional to $\delta(E-E_o)\,\delta(t)$, where E_o is large. Let us consider equation (44) after enough time has elapsed so that the pulse of neutrons is almost Maxwellian. We shall write ϕ_1 as a Maxwellian with time-dependent temperature and containing a fixed number of neutrons. Such a flux distribution is proportional to $T^{-3/2}Ee^{-E/T}$. Thus, if we write $\phi_1(E,T) = M(E,T_o)\,\psi(E,t)$, and make use of the detailed balance property of $\Sigma_s(E,E')$, we find

$$\frac{1}{v}\frac{\partial \psi}{\partial t} = \int_0^\infty dE' \Sigma_s(E,E')\left(\psi(E',t) - \psi(E,t)\right), \quad (45)$$

with $\psi(E,t) = (T_o/T)^{3/2} \exp\left[-E\left(\frac{1}{T} - \frac{1}{T_o}\right)\right]$. Now, if we call $T(t) = T_o(1+\delta(t))$, and retain only terms of zero and first order in δ, we have $\psi = 1+(\epsilon - \frac{3}{2})\delta$ $(\epsilon = E/T_o = (v/v_o)^2, \tau = v_o t)$

$$\frac{1}{\sqrt{\epsilon}}\left(\epsilon - \frac{3}{2}\right)\frac{\partial \delta}{\partial \tau} = \int_0^\infty d\epsilon' \Sigma_s(\epsilon,\epsilon')(\epsilon'-\epsilon) \quad . \tag{46}$$

Now, multiply by $\epsilon^2 e^{-\epsilon}$ and integrate both sides with respect to ϵ . We may simplify the right-hand side by noting the symmetry of $M(\epsilon)\Sigma_s(\epsilon,\epsilon')$. Thus,

$$\frac{\partial \delta}{\partial \tau} = -\frac{2}{3\sqrt{\pi}}\left[\int_0^\infty d\epsilon \int_0^\infty d\epsilon' M(\epsilon)\Sigma_s(\epsilon,\epsilon')(\epsilon-\epsilon')^2\right]\delta$$

$$= -\frac{2}{3\sqrt{\pi}} M_2 \delta \quad . \tag{47}$$

Equation (47) shows an exponential relaxation of the temperature of the neutron gas to T_o, the moderator temperature. The rate of relaxation is given by the Maxwellian-averaged second moment of energy transfer, which is a good measure of the "width," or capacity for energy exchange, of the kernel. We list some values for the ratio of M_2 (in barns) to the free-atom[1] (i.e., high energy) cross section (in barns), below.

1. This cross section appears in the text in a variety of guises. It will be found as σ_f , σ_{sf} , and $\sigma_s(\infty)$!

The table illustrates nicely the decrease in thermalizing power as one goes from free to bound scattering atom, and from light to heavy mass. It also shows the effect of an increase in temperature upon the effective binding.

Model	M_2/σ_f	
Gas of point-particles	$8\mu/(1+\mu)^{3/2}$	
$\mu = 1$	2.83	
$\mu = 1/12$	0.59	
Bound proton (H_2O), 300°K	2.23	(36)
Graphite, 300°K	0.22	(37)
1200°K	0.53	
∞	0.59	

As a mathematical aside, note that the quantity $2/3\sqrt{\pi}\, M_2$ in (47) must be related to the time-decay eigenvalue, λ_1, mentioned in an earlier lecture. The relationship is easily seen by considering the variational expression for λ_1, which is easily deduced from (9). It is

$$\lambda_1 \leq \frac{(\phi_1^+, S\phi_1)}{(\phi_1^+, \frac{1}{\sqrt{\epsilon}}\phi_1)} \qquad (\phi_1^+, \frac{1}{\sqrt{\epsilon}}\phi_0) = 0 \qquad (48)$$

$$\phi(\epsilon) = T\phi(E)\ ,$$

where S is the scattering operator, and ϕ_0 is the Maxwellian flux. A simple trial function, orthogonal to ϕ_0, is $\phi_1^+ = (\epsilon - \frac{3}{2})$. When it is inserted into the first of

equations (48), the "M_2-formula" appears. It is important to appreciate that the formula is an upper bound to λ_1, and that its quantitative accuracy has not been assessed. We know, for example, that it is within 25% of the correct value in the case of a heavy gas, but wrong by more than a factor of four in the "heavy-crystal" model.

Let us return to problems of the infinite medium steady state flux, by recalling the connection between $\phi(E)$ and the time integral of $\phi(E,t)$. If $\{u_k(e)\}$ is the set of ortho-normal eigenfunctions satisfying

$$Su_k = \frac{\lambda_k}{\sqrt{\epsilon}} u_k \quad \text{and} \quad \left(u_\ell^+, \frac{1}{\sqrt{\epsilon}} u_k\right) = \delta_{\ell k} \,, \tag{49}$$

the series

$$g(\epsilon, \epsilon_s; t) = \sum_k u_k^+(\epsilon_s) u_k(\epsilon) e^{-\lambda_k v_o t} \tag{50}$$

describes the flux distribution induced by a source proportional to $\delta(\epsilon - \epsilon_s)\delta(t)$, in an infinite non-absorbing medium. The steady-state spectrum for 1/v absorption will be

$$\phi(\epsilon; \epsilon_s) = \int_0^\infty dt \, \exp\left[-\Sigma_a(T_o) v_o t\right] g(\epsilon, \epsilon_s; t) \tag{51}$$

$$= \sum_k \frac{u_k^+(\epsilon_s) u_k(\epsilon)}{(\lambda_k + \Sigma_a(T_o)) v_o} \tag{52}$$

$$= \frac{2}{\sqrt{\pi}} \frac{M(\epsilon)}{\Sigma_{ao} v_o} + \frac{u_1^+(\epsilon_s) u_1(\epsilon)}{(\lambda_1 + \Sigma_{ao}) v_o} + \cdots \tag{53}$$

If the absorption is weak enough so that $\lambda_1 >> \Sigma_{ao}$, we may write (53) in a form that should be compared with (33), namely,

$$\phi(\epsilon) = \frac{2}{\sqrt{\pi}} \left(\frac{\Sigma_a(T_o)}{\lambda_1} \right)^{-1} M(\epsilon) + \begin{matrix} \text{non-Maxwellian} \\ \text{component} \end{matrix} \qquad (54)$$

Thus, λ_1 is a more accurate quantity that $\xi \Sigma_s$ for the description of the thermalizing power of a moderator, though when we consider the gas of point particles either parameter appears satisfactory.

We can squeeze a bit more from (53) if we estimate the first moment, or the average energy of the distribution. When the Maxwellian term alone is significant, $\langle E \rangle = 2T_o$, and the average may be used to define an effective temperature. Let us assume that when the absorption is small ($\lambda_1 >> \Sigma_{ao}$), the first two terms of (53) give a good representation of $\phi(\epsilon)$ near the peak of the distribution. Then, if we calculate $\langle \epsilon \rangle$, and compare it with its Maxwellian value, $\langle \epsilon \rangle_o$, we find

$$\frac{\langle \epsilon \rangle}{\langle \epsilon_o \rangle} = 1 + \alpha \frac{\Sigma_a(T_o)}{\lambda_1} \qquad (\alpha > 0), \qquad (55)$$

where α is independent of absorption, independent of source energy when $\epsilon_s >> 1$, and only weakly dependent upon scattering model, Thus, in weakly absorbing systems the effective tempera-

ture of the neutron gas is larger than that of the moderator, and is increased by an increase in absorption, and/or (quite reasonably!) by a decrease in the first relaxation constant.

Earlier, from consideration of the asymptotic solution, we prophesied that the effective temperature would be increased in proportion to $\langle K \rangle$. Now we say that it will increase in inverse proportion to λ_1, or roughly, M_2. Are these ideas at all consistent? To answer, note first that $\langle K \rangle$ and M_2 are respectively proportional to

$$\int_0^\infty d\omega \; \frac{\omega f(\omega)}{\sinh(\hbar\omega/2T)} \cosh(\hbar\omega/2T) \qquad \text{(in general)} \qquad (56)$$

and

$$\int_0^\infty d\omega \; \frac{\omega f(\omega)}{\sinh(\hbar\omega/2T)} \omega^2 K_2(\hbar\omega/2T) \qquad \text{("heavy" crystal),} \qquad (38)$$

where $K_2(y)$ is the modified Bessel function of the second kind. Now, the cosh function increases monotonically, and $\omega^2 K_2$ decreases monotonically as its argument increases, so that the variation of these parameters with stiffness of binding is just what we need. Further, for weak binding (i.e., all significant ω are such that $\hbar\omega < kT$) the two parameters are essentially the same. Finally, one can show that when several suitable models of the moderator are compared, those having the same $\langle K \rangle$ have pretty much the same M_2 (39). In all there is a reasonable consistency in our discussion.

The notion of a Maxwellian distribution of neutrons with temperature shifted slightly from the moderator temperature, T_o, is also useful in understanding thermalization in finite systems. For example, let us ask about the shape of the fundamental mode, and the size of the decay constant in a typical pulsed neutron experiment. If, for convenience, we take the absorption to be "1/v," and the shape, in space, of the fundamental to be $\cos(B\cdot r)$, we have

$$\phi(\epsilon,r,t) = e^{-(\lambda_o+\Sigma_{ao})\tau}\,\phi_o(E)\cos(\vec{B}\cdot\vec{r}) \quad . \tag{57}$$

$\phi_o(E)$ is the fundamental (energy) mode, and it satisfies

$$-\frac{\lambda_o}{\sqrt{\epsilon}}\phi_o = S\phi_o - D(\epsilon)B^2\phi_o \quad . \tag{58}$$

Integration of (58) with respect to ϵ gives

$$\lambda_o = \frac{\int d\epsilon\, D(\epsilon)\, B^2\, \phi_o(\epsilon)}{\int d\epsilon\, \frac{1}{\sqrt{\epsilon}}\,\phi_o(\epsilon)} \tag{59}$$

while multiplication of (58) by ϵ, followed by integration gives

$$\lambda_o = \frac{\int d\epsilon\,(D(\epsilon)B^2 - S)\,\phi_o\,\epsilon}{\int d\epsilon\,\sqrt{\epsilon}\;\phi_o(\epsilon)}$$

Now, if we choose to represent the flux as a Maxwellian with displaced temperature, so that $\phi_o(E) \sim E/T^2 \exp(-E/T)$,

or $\phi_o(\epsilon) \sim \epsilon/\theta^2 \exp(-\epsilon/\theta)$ ($\theta = T/T_o$), equations (59) and (60) enable us to determine θ and λ_o. Since the variation of $D(\epsilon)$ with energy plays an important role, we shall consider those $D(\epsilon)$ having the form $D = D_o \epsilon^{\alpha}$. After a calculation similar to equations (44-47), we find, to lowest order in $\Delta T/T_o$, that

$$\Delta T/T_o = -\frac{D_o B^2}{M_2}(1+2\alpha)\Gamma(2+\alpha) , \tag{60}$$

and

$$\lambda_o = \frac{2}{\sqrt{\pi}} D_o B^2 \Gamma(2+\alpha)\left[1-(1+2\alpha)^2\Gamma(2+\alpha)\frac{D_o B^2}{2M_2}\right]$$

$$= \frac{2}{\sqrt{\pi}} D_o B^2 \Gamma(2+\alpha)\left[1+(1/2+\alpha)\frac{\Delta T}{T}\right] . \tag{61}$$

Equations (60) and (61) are rich in content. First, note that if $\alpha = -\frac{1}{2}$, so that the leakage as well as the volume absorption is 1/v, there is no shift in temperature, and λ_o is simply $D_o B^2$. If α is made greater than $-\frac{1}{2}$, the leakage of fast(er) neutrons is enhanced, and the "temperature" of the neutrons inside is lowered. If α is decreased below $-\frac{1}{2}$ the reverse phenomenon occurs. In both cases, the change in temper-

ature varies inversely with the ability of the moderator to thermalize. Since we are working with trial functions of Maxwellian type, the ubiquitous "M_2" appears. An eigenfunction expansion for $\phi_o(E)$ would introduce λ_1, which is the more precise parameter.

The first term in the expression for the decay constant is the average of $D(\epsilon)B^2$ with respect to a Maxwellian at T_o. The second, or "diffusion cooling" term, expresses the distortion of the spectrum caused by leakage. The reader will note that the diffusion cooling correction to the decay rate is always negative, for, whether α be greater or less than $-\frac{1}{2}$, the use of the undistorted Maxwellian overestimates the leakage.

Finally, we note that similar results may be obtained through the use of a variational expression, similar to equation (48), for λ_o. The trial function may be taken to be a shifted Maxwellian, whose temperature is used as a variational parameter. In this manner, Nelkin (40) first obtained equation (61). We should mention that he was first to introduce M_2 into the study of thermalization, while the use of the distorted Maxwellian originates in the thesis of van Dardel (41).

Synthetic Kernels

The semi-quantitative discussion which we have just given points out the role of energy-transfer moments in thermalization. It serves as a guide to the next step in improving the approximate treatment of the Boltzmann equation, namely, the replace-

ment of the scattering operator, S, by a simpler, "synthetic" operator, $\bar{S}$, which retains only the most important properties of S. The substitution will be valuable only when the "important properties" are few.

A popular and simple device is the replacement of the isotropic scattering operator (10) by a differential operator of second order. The physics of the device is seen more clearly if we work with the adjoint operator, S^+.
Thus, the replacement is

$$S^+ \rightarrow \bar{S}^+ \equiv p(\epsilon)\frac{d^2}{d\epsilon^2} + q(\epsilon)\frac{d}{d\epsilon} + r(\epsilon) \quad , \tag{62}$$

where r, q, and p are three functions to be determined. Our three-fold freedom will be short-lived, however, when we insist that the operator satisfy the conditions of detailed balance and neutron conservation. It is easily seen that we must have $r(\epsilon) \equiv 0$, and

$$M(\epsilon)q(\epsilon) = \frac{d}{d\epsilon}M(\epsilon)p(\epsilon) \tag{63}$$

Thus, only a single arbitrary function remains.

To appreciate the significance of $p(\epsilon)$ and $q(\epsilon)$, note that in general, when S^+ operates upon ϵ it yields $A_1(\epsilon)$, the first moment of the scattering kernel. Since $\bar{S}^+$ gives $q(\epsilon)$ when it operates upon ϵ, we are led to the identification $q(\epsilon) \leftrightarrow A_1(\epsilon)$, where

$$A_n(\epsilon) = \int_0^\infty d\epsilon' \Sigma_s(\epsilon, \epsilon')(\epsilon' - \epsilon)^n . \tag{64}$$

If we now consider $\bar{S}^+$ to act upon ϵ^2, we are led to the second correspondence, $p(\epsilon) \leftrightarrow \frac{1}{2} A_2(\epsilon)$. However, we are not free to use the $A_1(\epsilon)$ and $A_2(\epsilon)$ which would result from the detailed analysis of a scattering model, since they will generally violate (63). It would appear that the best way to proceed is to choose a model for the moderator, calculate its first moment, $A_1(\epsilon)$, take $q(\epsilon) \equiv A_1(\epsilon)$ and determine $p(\epsilon)$ through (63), i.e.,

$$p(\epsilon) = M^{-1}(\epsilon) \int_\epsilon^\infty d\epsilon\, q(\epsilon) M(\epsilon) = \frac{1}{2} \bar{A}_2(\epsilon) \tag{65}$$

A little calculation shows that the synthetic second moment, $\bar{A}_2(\epsilon)$, gives the correct M_2, and that the general behavior of $\bar{A}_2$ is not too objectionable (42). There is one scattering model, the heavy gas, for which $\bar{A}_2$ is precisely A_2. In that case, $q(\epsilon) = 2\mu \Sigma_s (2 - \epsilon)$, and $p(\epsilon) = 2\mu \Sigma_s \epsilon$.

The synthetic kernel technique has not yet been fully exploited. Using it, one deals with an ordinary differential equation of second order, rather than an integral equation. It is then relatively simple to extract series solutions in positive and negative powers of ϵ, and to estimate decay constants and eigenfunctions (42). The asymptotic (high energy) series afford an easy means of checking the accuracy of $\overline{S}$. For example, the heavy crystal series (equation 42) has precisely the same terms, through (T/E), whether we use S or $\overline{S}$. Terms of higher order, which are dependent upon higher moments of energy transfer, will not agree perfectly. Similar conclusions hold for the calculation of decay constants. The lowest modes should be described rather well by the synthetic kernel; the higher (and less important!) modes should be defective because they depend upon higher moments of energy transfer, that is, upon increasingly finer details of the kernel.

Some bibliographical detail is appropriate here. The differential model of the heavy gas moderator was proposed by J. E. Wilkins, Jr. (43) and discussed extensively by Cohen (44), Hurwitz, and Nelkin (45), and many others. A differential operator representation of S similar to equation (62) has been noted in unpublished work by J. Horowitz of Saclay; Schaefer and Allsopp (44a) and Leslie (45a) have recently reported upon their research with the "Horowitz Operator." All treat $p(\epsilon)$ as the "free function." Leslie uses a "best fit" to experiment, while Schaefer and Allsopp use different methods for $\epsilon >> 1$ and $\epsilon << 1$.

In the high energy region, p (ϵ) is chosen to give agreement with the asymptotic series solution to the infinite medium problem; in the low energy region it is determined from knowledge of $A_1(\epsilon)$, as we have suggested. Finally, to emphasize that the subject of synthetic kernels is far from closed, we note a new paper by the Saclay scientists (46) in which an improved version of the Horowitz operator is discussed.

Lecture IV

Eigenvalues and Decay Constants

We have met the time (and space) decay constants several times in our considerations, noting, for example, that the smaller ones can be measured directly and that they aid in the description of spectra distorted from Maxwellian shape. The constants, as we saw earlier, may be thought of as eigenvalues of the scattering operator, with respect to a particular weight-function. In this lecture we shall discuss some aspects of the sequence of time decay constants $\{\lambda_k\} = \lambda_0, \lambda_1, \lambda_2 \ldots$ We are interested in the trend of the $\lambda_k(B^2)$ as $k \longrightarrow \infty$, as $B^2 \longrightarrow \infty$, and as the nature of the moderator is altered. Above all, we should like to know how to calculate them efficiently. Since these questions are the subject of a great deal of contemporary research, our answers must needs be incomplete.

We shall concentrate upon time decay constants, including spatial effects mostly through diffusion theory. The study of spatial decay constants proceeds similarly. To begin, we might inquire about the synthetic kernel of equation (62). There, the quickest way to estimate eigenvalues appears to be through the W.K.B., or phase integral method (47,42). Its application to the heavy gas equation gives eigenvalues accurate to about 1%, and predicts a sequence of λ_k increasing as $k^{3/2}$ for large k and having its limit point at infinity. This behavior is not altered if one includes a leakage term, $D(\epsilon)B^2$, having

reasonable behavior. A similar investigation of the general differential operator (62), with $q(\epsilon) \longleftrightarrow A_1(\epsilon)$, turns up the very same trends. One cannot take these results too seriously because, as we noted earlier, the synthetic kernel does not pretend to give higher moments of energy-transfer---and therefore, higher eigenvalues---accurately.

We shall now consider the first model for thermalization that is analytically tractable, yet not trivial---the proton gas model of Wigner and Wilkins (48). Its simplicity is connected with a special feature of the scattering of neutrons by free particles of mass equal to the neutron mass. In that case, the symmetrized scattering kernel has the form

$$\sqrt{\frac{x'M(x')}{xM(x)}}\ \Sigma_s(x',x) = \begin{cases} a(x')b(x) & x' < x \\ a(x)\,b(x') & x' > x \end{cases} \qquad (66)$$

being the Green's function for a differential equation of second order. (We have introduced the useful and dimensionless velocity variable x ($x^2 \equiv \epsilon = E/T$).) Thus, if one operates upon the Boltzmann equation with the appropriate differential operator, the result is a differential equation in energy. Details of the calculation may be found in reference (48). We merely give the equation which results if we seek a solution to the eigenvalue problem, suggested by $\phi(E,t) \stackrel{?}{=} \phi_n(E)\exp\left[-\bar{\lambda}_n \nu_o t\right]$. With the x-variable, we have $\phi(E) \propto N(x)$, the number

density in velocity, while $N(x) = \sqrt{M(x)}\ \nu(x)$ defines the number density belonging to the symmetrized scattering kernel. The differential equation is:

$$\left\{\frac{d}{dx}\ \frac{1}{P(x)}\ \frac{d}{dx}\left[V(x)-\lambda_n\right]v_n(x)\right\}+$$

$$\left\{\frac{4}{\sqrt{\pi}}+\left[\ \frac{e^{-x^2}}{P^2(x)}\ -\ \frac{x^2}{P(x)}\right](V(x)-\lambda_n)\right\}v_n(x)=0\ ,$$

(67)

where V(x) and P(x) are closely related to the scattering cross section, $\Sigma_s(x)$. We have:

$$V(x)\Sigma_s(\infty)=x\Sigma_s(x)=x\left\{\left(1+\frac{1}{2x^2}\right)erf(x)+\frac{1}{\sqrt{\pi}}\ \frac{e^{-x^2}}{x}\right\}\Sigma_s(\infty)$$

$$P(x)=\sqrt{\pi}\ x\ erf(x)+e^{-x^2}=\sqrt{\pi}\ V(x)-\frac{1}{2x}\ erf(x) \qquad (68)$$

$$erf(x)=\frac{2}{\sqrt{\pi}}\int_0^x dt\ e^{-t^2} \qquad \lambda\Sigma_s(\infty)=\bar{\lambda}\ .$$

We may describe our eigenvalue problem in the following terms: For what values of λ does (67) have solutions which are nowhere singular, which are proportional to x for small x, and proportional to $x\ e^{-\frac{1}{2}x^2}$ for large x? (One can verify that

the solution of the homogeneous integral equation, with which we began, must have this behavior.) Since equation (67) is not particularly transparent, we transform it to one of Schrödinger type and examine the wave-mechanical analog of our thermalization problem. Thus, put

$$\zeta_m(x) = \left[V(x) - \lambda_m\right] \exp\left[-\frac{1}{2}\sqrt{\pi}\int dx\, \frac{erf(x)}{P(x)}\right] v_m(x), \tag{69}$$

whence our equation becomes:

$$\frac{d^2\zeta_n}{dx^2} + \left[E - U(x,\lambda)\right]\zeta_n = 0 \tag{70}$$

with

$$-U(x,\lambda) = \frac{2e^{-x^2}}{P(x)} - x^2 - \frac{3\pi}{4}\,\frac{erf^2(x)}{P^2(x)} + \frac{4}{\sqrt{\pi}}\,\frac{P(x)}{V(x)-\lambda_n}$$

$$E = 0 \tag{71}$$

In quantum mechanical terms, our problem is to find values of the parameter λ, such that the potential $U(x, \lambda)$ will bind a particle having zero energy. ($z(x)$ must be proportional to x as $x \longrightarrow 0$ and to $x^{3/2}\, e^{-\frac{1}{2}x^2}$ as $x \to \infty$.) Though the expression for the potential is complicated, the function itself

is quite smooth, and one can get an idea of its bound states with little effort.

The factor $(V(x) - \lambda)^{-1}$ plays an important role in the analysis. We begin by noting that V(x) is a positive, monotonic increasing function of x whose minimum value is $V(0) = 2/\sqrt{\pi}$ (Figure 9). Thus, the potential for $\lambda < 2/\sqrt{\pi}$ will be quite different in shape than that for $\lambda > 2/\sqrt{\pi}$, when it has a singularity on the x-axis. Figure 10 shows the potential in the former case and Figure 11, the latter. We shall begin our analysis with the case $\lambda = 0$, then imagine λ to increase towards $2/\sqrt{\pi}$. The shape of the potential will change in a smooth manner as λ changes, and the energy levels will change (move) smoothly. As λ increases, the potential well deepens, and the energy levels descend into the well. Every λ for which an energy level coincides with E = 0, is an eigenvalue that we seek. The number that we shall find in the range $0 \leq \lambda \leq 2/\sqrt{\pi}$ depends critically upon the form of $U(x, \lambda)$ as λ approaches $2/\sqrt{\pi}$. Since $V(x) = \frac{2}{\sqrt{\pi}} + \frac{2}{3\sqrt{\pi}} x^2 + \ldots$ the potential, U, will have a $1/x^2$ singularity at x = 0. In this limit, the potential contains an infinite number of levels (49), and we may conclude that an infinite number of eigenvalues, $\bar{\lambda}$, lie between $\bar{\lambda} = 0$ and $\bar{\lambda} = (x\Sigma_s)_{min} = \frac{2}{\sqrt{\pi}} \Sigma_s(\infty)$. Thus, the point $\bar{\lambda} = (x\Sigma_s)_{min}$ is a <u>limit point</u> for the sequence of eigenvalues $\bar{\lambda}_0, \bar{\lambda}_1, \bar{\lambda}_2, \cdots$

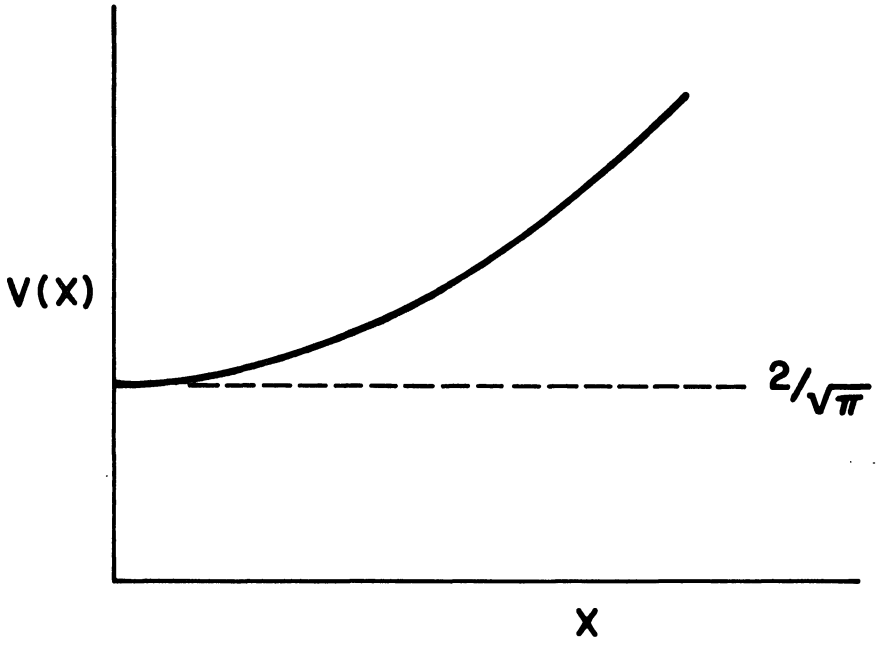
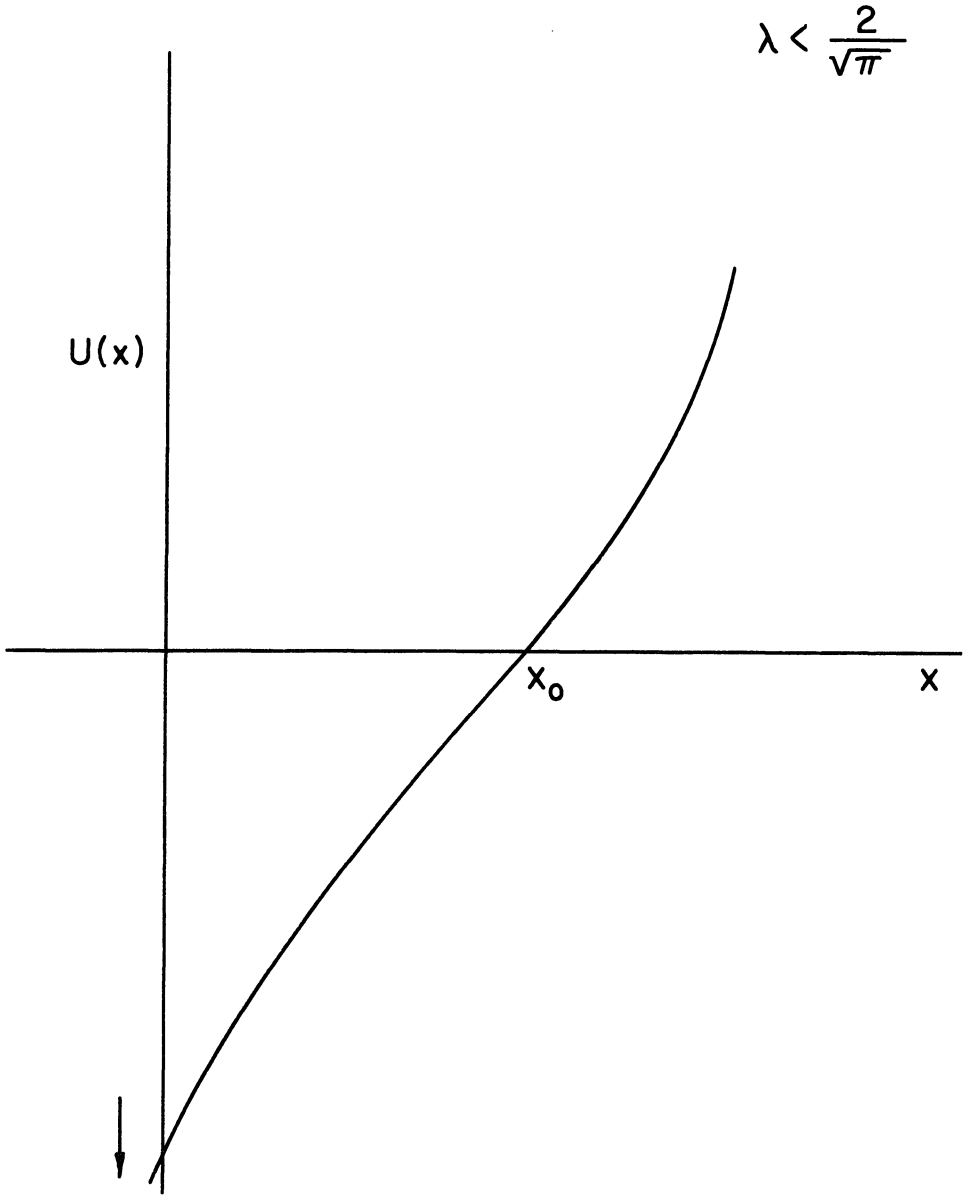
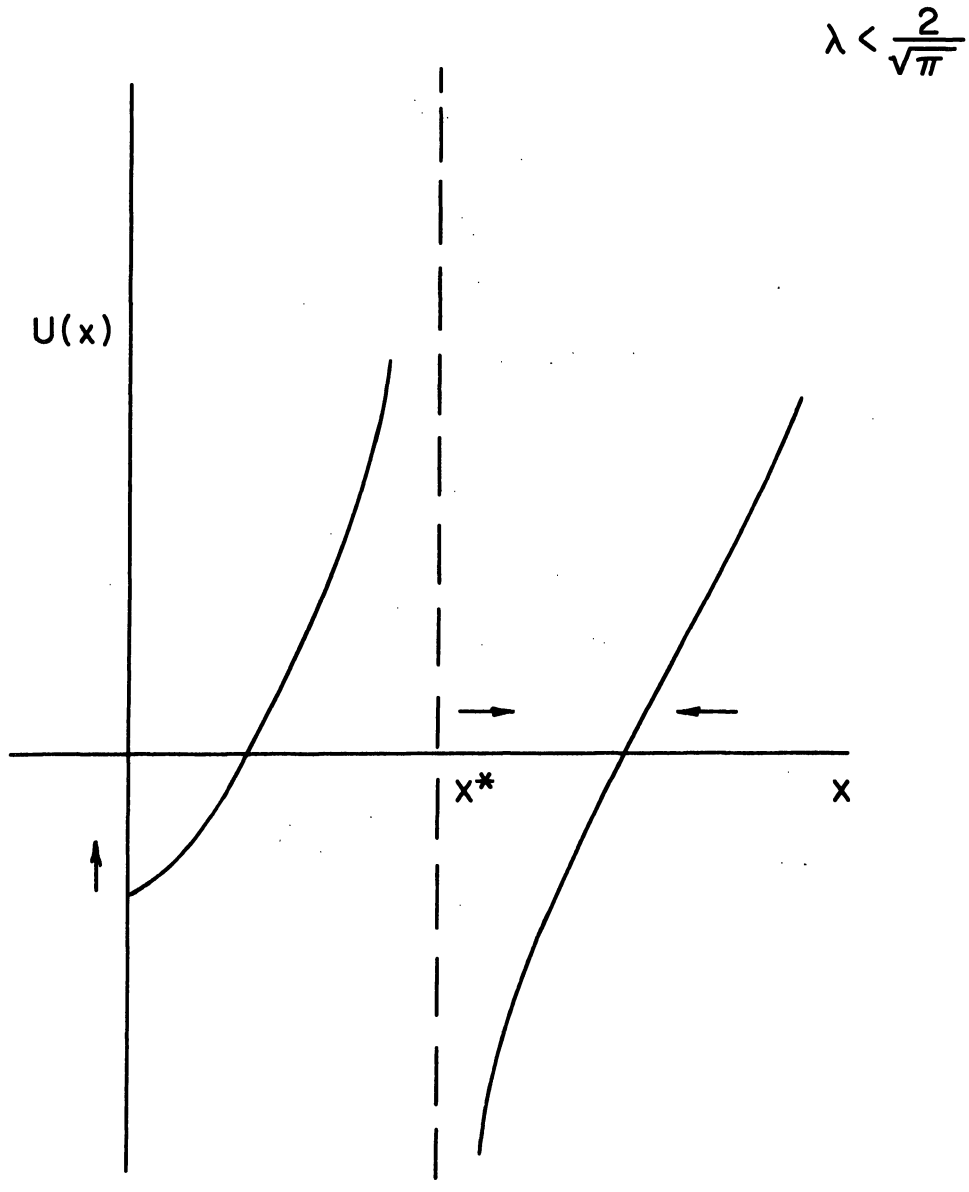


Figure 9



The effective potential, $U(x,\lambda)$, for a value of λ less than $\frac{2}{\sqrt{\pi}}$. x_o is the classical turning point for a particle of zero energy. The arrow indicates the manner in which the potential changes, as λ increases.

Figure 10



The effective potential, U, for a value of λ greater than $\frac{2}{\sqrt{\pi}}$. x* is the point at which the reaction rate, V(x), is equal to λ . The arrows indicate the manner in which the potential changes, as λ increases.

Figure 11

One can make a similar analysis for the region $\lambda > 2/\sqrt{\pi}$; the reader may consult our paper (50) for details. One finds only a finite number of eigenfunctions there, and possibly none at all.

The eigenfunctions belonging to the range $0 \leqslant \lambda \leqslant 2/\sqrt{\pi}$ look like "ordinary" eigenfunctions in that they have zero, one, two,...nodes. However, all of their oscillations occur in a limited range of x, and it follows that they cannot be a complete set for the representation, on $0 \leqslant x < \infty$, of the functions we wish to study. We shall have to augment them by other functions to get a complete set.

To get additional functions, we must abandon one of the requirements placed upon our eigenfunctions---the property of finiteness. If we merely require that our functions be integrable - so that we admit delta-function singularities - we find acceptable eigenfunctions for every $\lambda > 2/\sqrt{\pi}$. Each of these continuum eigenfunctions is singular at that value of x for which $V(x) = \lambda$ (50). The completeness of the augmented set of eigenfunctions has not yet been proved in the case of the proton gas, but it has been shown to be true for a quite broad class of scattering kernels. (60) (Also see Appendix II). It is doubtless a good bet to assume that the set of eigenfunctions consisting of the discrete (ordinary) functions, and the continuum (singular) functions is complete for the representation of all "physical" solutions to the Boltzmann equation. (Compare with the discrete and continuum eigenfunctions in the problem of the hydrogen atom!) Thus, we can say

$$N(x,\tau) = \sum_{k=1}^{\infty} a_k u_k(x) e^{-\lambda_k \tau} + \int_{\lambda^*}^{\infty} d\lambda\, a(\lambda) u(x;\lambda) e^{-\lambda\tau} \qquad (72)$$

with $\lambda^* = (x\,\Sigma_s(x))_{min}$. The discrete eigenfunctions contribute terms having exponential behavior in time, with decay constants lying between 0 and λ^*, while the continuum functions supply a term that is not exponential, but is proportional to $\tau^{-1} \exp\left[-\lambda^* \tau\right]$ as $\tau \longrightarrow \infty$. The convergence of (72) is not likely to be rapid enough to make the series valuable from a computational point of view.

We have suggested that the behavior found in the case of the proton gas model will also be found in a large class of non-trivial models. This conjecture is supported by the following "hand-waving" argument: if we are to have a discrete, or "ordinary" mode, we must have a non-singular distribution of neutrons, $u_k(x)e^{-\lambda_k \tau}$, which is a solution of the Boltzmann equation and which has the property and its form---as a function of x --- is maintained at all times. During the evolution of the distribution, neutrons are being scattered from one velocity to another, with local scattering rate $\sim x\,\Sigma_s(x)$, but the only effect of these scatterings is to give a uniform decrease in the size of the distribution. Since we are dealing with a cooperative effect that treats all parts of the velocity range equally and which is characterized by a single rate constant, λ_k, it

would appear that $\bar{\lambda}_k$ must be smaller than the minimum value of $x\,\Sigma_s(x)$ encountered. To use a pair of homely analogies, remember that a class of school children can proceed no faster than its slowest member, and that a convoy of ships can maintain itself only at speeds slower than that of its slowest member. (53)

We can understand why we failed to observe the limit point behavior in the case of the heavy gas, if we note that $(x\Sigma_s(x))_{min}$ for the point gas of arbitrary mass is proportional to $\sqrt{\mu}$ (51). The decay constants, on the other hand, are proportional to μ, so that the ratio of λ^* to λ_k is proportional to $\mu^{-1/2}$. If we now go to the heavy gas limit ($\mu \longrightarrow 0$) and display our eigenvalues in the form λ/μ, the limit point moves to infinity.

Effects of Diffusion

We shall take diffusion into account by considering the time decay constants for distributions having a spatial variation proportional to $\exp\left[i\vec{B}\cdot\vec{r}\right]$, in a homogeneous and isotropic model of infinite extent. It is known (52) that such a distribution is a good approximation to the truth of the matter in a finite moderator if (a) the size of the moderator is at least several times the maximum scattering mean free path encountered and (b) we wait long enough. We shall consider the dependence of λ upon B^2 when diffusion is governed by diffusion theory and by transport theory. In both cases, we focus our attention upon the proton-gas scatterer to obtain results we believe to be representative of a large class of models.

Consider the Boltzmann equation (with isotropic scattering) in the form in which the scattering operator is split, the total cross section being returned to the left-hand side and the scattering kernel remaining on the right. We seek solutions of the form $\exp\left[i\vec{B}\cdot\vec{r} - \lambda\tau \right] N(x)$. If $B^2 = 0$, the left hand side becomes $(V(x) - \lambda)N(x)$. If $B^2 \neq 0$ and we use diffusion theory, $(V(x) - \lambda)$ becomes $(xd(x)b^2 + V(x) - \lambda)$ where $\left[b \Sigma_s(\infty) \right]^2 = B^2$, and $d(x) = D(x) \Sigma_s(\infty) = \Sigma_s(\infty)/3\Sigma_s(x)$. (Both b and d are dimensionless.) Finally, if we use transport theory, $(V(x) - \lambda)$ is replaced by $\left[\frac{1}{Bx} \tan^{-1} \frac{Bx}{V(x)-\lambda} \right]^{-1}$. Taking the proton gas, we ask how the form of the potential well is altered when these replacements are made.

In the case of diffusion theory, $xd(x)$ is of order x^2 as $x \to 0$ and approaches $\frac{1}{3}x$ as $x \to \infty$. Thus, the addition of $xd(x)b^2$ to $V(x)$ in (71) makes very little difference in the discussion of the eigenvalues; $V(x) + xd(x)b^2 = 2/\sqrt{\pi} + (2/\sqrt{\pi} + d_o b^2)x^2 + \ldots$ for small x, and the effective potential becomes singular as $\lambda \to 2/\sqrt{\pi}$, just as before. The volume of the effective potential becomes infinite, and we may conclude, as before, that an infinite number of discrete eigenfunctions exists in the range $0 \leq \lambda \leq 2/\sqrt{\pi}$. The eigenvalues are changes as B^2 increases. Since λ_o will always increase, the spacing between eigenvalues decreases, and they crowd against the "barrier," $2/\sqrt{\pi}$. See e.g. fig. 12.

The variation of λ with B^2 is quite different in transport theory. As we noted, the effective potential $U(x,\lambda)$ is altered by the replacement

$$\frac{1}{V(x)-\lambda} \longrightarrow \frac{1}{Bx}\tan^{-1}\frac{Bx}{V(x)-\lambda} \quad . \tag{73}$$

Now, recall that as λ approaches $V(0) = 2/\sqrt{\pi}$, the $B^2 = 0$ potential has a $1/x^2$ singularity, and the large volume of the singular potential is responsible for the limit point behavior. In the case $B^2 \neq 0$, the arctangent will always be bounded by $\pi/2$ so that the potential can not be more singular than $1/x$. The latter potential is not strong enough to cause an infinite number of levels to descend into the well, in the language of our earlier argument. In other words, the eigenvalue spectrum of the $1/x$ (Coulomb) potential is bounded from below (i.e., has a ground state), while the spectrum of the $1/x^2$ potential is unbounded. Thus, when $B^2 \neq 0$, there will be a <u>finite</u> number of eigenvalues lying below $2/\sqrt{\pi}$ and possibly one or two with $\lambda > 2/\sqrt{\pi}$. It should not surprise the reader, then, to learn that for sufficiently large B, there are <u>no</u> discrete eigenfunctions at all. We give a proof of this fact for a general kernel, in Appendix III.

It may amuse the reader to learn that the problem we are discussing is quite similar to the propagation of sound waves of small amplitude in a gas of rigid, classical spheres. In that case, one also finds a critical size of propagation vector, $\vec{B}$, such that if $|\vec{B}| > |\vec{B}_{crit}|$, no solution of form

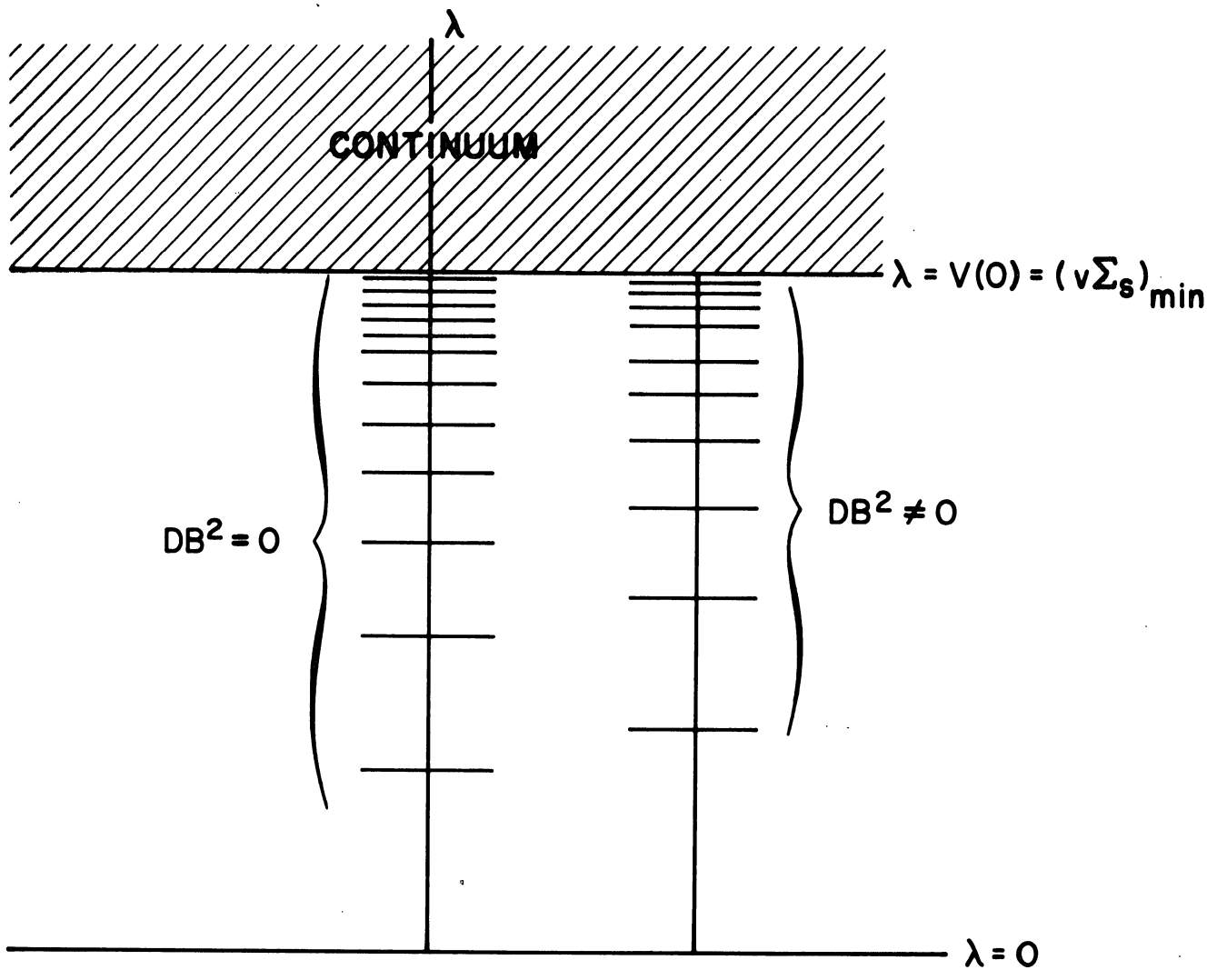


Figure 12

$\exp \left\{ i(\vec{B}\cdot r - \omega t) \right\}$ exists.

While the discussion we have given is based upon the proton gas model, we would conjecture that the same features will be seen in other models. In Appendix II we give a discussion of the general case and the results reinforce our faith in the conjecture. We shall close this section with an observation about the behavior of λ_o, the lowest eigenvalue, as the size of a finite sample of moderator is altered. This point is due to Nelkin (54), and is a modification of arguments given by G. M. Wing (55) and B. Davison (56).

Consider the integral form of the Boltzmann equation (56). We consider the source-free and absorption-free case, and seek solutions of the form $N(\vec{r},x,\tau) = N_\lambda(\vec{r},x)\exp(-\bar{\lambda}\tau)$. The eigenfunction equation is:

$$x N(\vec{r},x) = \frac{1}{4\pi}\int \frac{d\vec{r}'}{|\vec{r}-\vec{r}'|^2} \exp\left\{-\frac{|\vec{r}-\vec{r}'|}{x}\left[x\Sigma_s(x)-\bar{\lambda}\right]\right\} \int_0^\infty dx' x' N(\vec{r}',x')\Sigma_s(x',x) \quad (74)$$

and we seek solutions which are smooth functions of x. Now, the factor 1/x in the argument of the exponential will cause trouble as $x \to 0$ unless $\lim_{x\to 0}\left[x\,\Sigma_s(x) - \lambda\right] > 0$. Since the value of $x\,\Sigma_s(x)$ in this limit appears to be the minimum $x\,\Sigma_s(x)$ in just about all scattering models, (74) tells us the following: if there is to be a (discrete) λ greater than $\left[x\,\Sigma_s(x)\right]_{\min}$ and with it a smooth eigenfunction, the

inner integral must vanish identically for all x. This is impossible when $N(\vec{r},x)$ is the fundamental mode, which is everywhere positive, and is unlikely when $N(\vec{r},x)$ is one of the higher modes, which are oscillating functions of x. Thus, λ_o must be less than $\left[x\Sigma_s \right]_{\min}$.

In (74) the existence of a limiting value for discrete eigenvalues is closely tied to the behavior of "zero-energy neutrons." These neutrons may be thought of as requiring an infinite time to traverse the finite sample of moderator. A similar phenomenon occurs in the study of one-velocity transport theory (55). There, neutrons diffusing in slab geometry can have an infinite transit time if they stream parallel to a slab face. Consequently, the time eigenvalues have a limit point (at $v\Sigma_s$!). On the other hand, when the neutrons diffuse in a sphere, all trajectories lead to escape in a finite time. In that case, one finds a sequence of decay constants extending to infinity and a complete set of discrete eigenfunctions. As a final remark, note that when we discuss neutrons of "zero energy" as agents responsible for these unexpected mathematical phenomena, we are in a domain where the semi-classical transport theory we use has little validity. At present, we are not quite sure as to <u>where</u> we stand!

<u>The Computation of Eigenvalues</u>

Now that we have an idea of the values assumed by the λ_n and how these are altered by changes in sample size, we shall turn to the results of experiment. As we noted in the first

lecture, these have dealt almost entirely with the lowest eigenvalue, λ_o. The first question to ask is: Are all λ_o that have been reported less than $\left[v \Sigma_s(v) \right]_{min}$? Paul Michael and the author have looked into the matter. If one compares the range of values of $\lambda_o(B^2)$ reported in the literature with $(v \Sigma_s)_{min}$, one finds:

Moderator	Range of λ_o	$(v \Sigma_s)_{min}$
Graphite	0-5000 sec^{-1}	2,600 sec^{-1}
Beryllium	0-8000 sec^{-1}	3,800 sec^{-1}
Light Water	0-30000 sec^{-1}	300,000 sec^{-1}

Figure (12a) shows the dependence of the experimental points upon B^2 for a particular moderator (graphite).

It is clear that some of the measurements at large B^2 in crystalline media are not consistent with the theory. The discrepancy could easily be attributed to neutron intensity; the experimenter is unable to wait long enough to ensure that he is measuring the decay constant of a fundamental mode. In some interesting experiments on beryllium, de Saussure and Silver (57) have compared decay constants measured during several time intervals long after the inception of a pulse. Some of their results are shown on Figure 13. While the decay constant for beryllium at 0°C seems to have "settled down" after a waiting time of about $1/\lambda_o$, the decay in beryllium of -100°C has not yet reached its fundamental mode, through relatively few neutrons

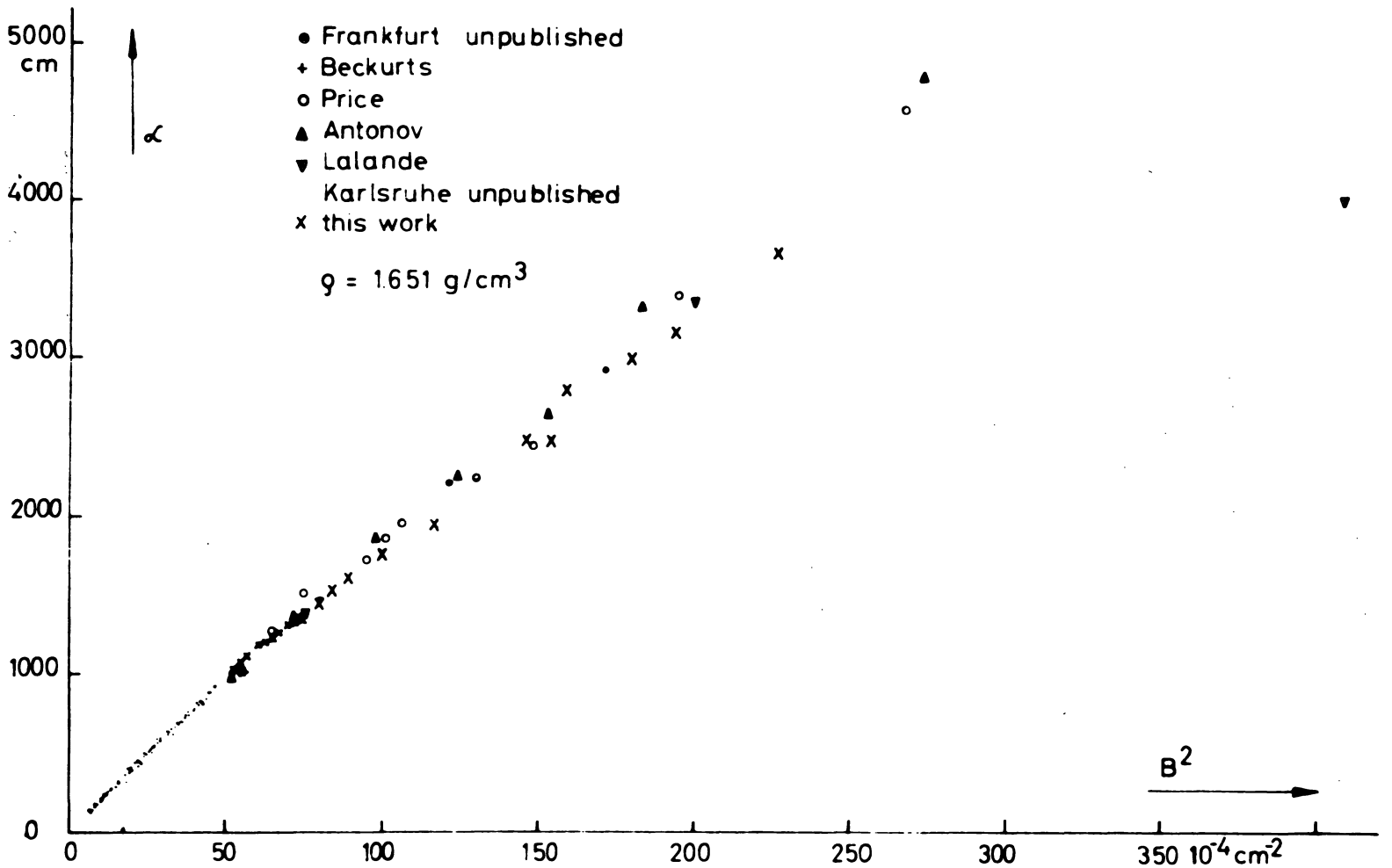


Figure 12a

Decay Constant (α) vs. B^2 (After Klose, et.al.*) in Graphite

* Proc. BNL Conf. III, 935

remain in the block. The difficulty in measuring fundamental modes in small samples of crystalline moderator reflects itself in the wide range of values of diffusion cooling coefficients that have been reported for these substances, and presents a serious challenge to the experimental physicist. Let us ignore experimental difficulties for the moment and ask how one might best calculate the variation of λ_o with B^2 and, in particular, the diffusion cooling coefficient. If we use diffusion theory and neglect effects of absorption, we can write the transport equation as:

$$\frac{\partial}{\partial \tau} N(x,\tau) = \left[S - x D(x) B^2 \right] N(x,\tau) \quad . \tag{75}$$

One approach that will yield an expansion of λ_n in terms of B^2, is the application of conventional perturbation theory, taking the leakage term as the perturbation (58). Thus,

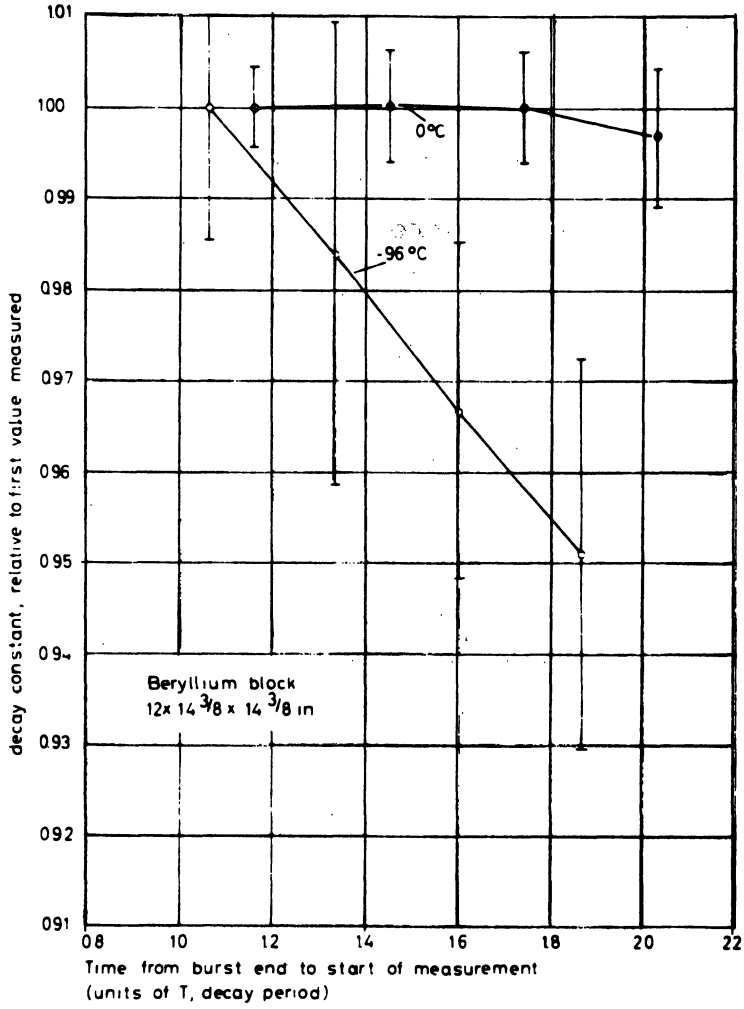
$$\lambda_n = \lambda_n^{(o)} + B^2 \langle n | x D(x) | n \rangle - B^4 \sum_{j \neq n} \frac{\left[\langle n | x D | j \rangle \right]^2}{\lambda_j^{(o)} - \lambda_n^{(o)}} + \cdots \tag{76}$$

where the scalar product is

$$\langle \phi, \psi \rangle \equiv \int_0^\infty dx \left[x^2 e^{-x^2} \right]^{-1} \phi(x) \psi(x) \quad , \tag{77}$$

and the unperturbed eigenfunctions obey

$$S N_k = \lambda_k^{(o)} N_k \quad , \quad \langle N_k , N_\ell \rangle \equiv \langle k | \ell \rangle \delta_{k\ell} \quad . \tag{78}$$



Relative values of decay constant, α, as a function of time elapsed between end of neutron pulse and start of data analysis. Time is in units of $T = 1/\alpha$; α normalized to unity at about $T = 1$. The lines are drawn to connect points obtained from the same run and do not represent the rate of change in α.

Figure 13

(The summation in (76) may be thought of as a summation over discrete eigenfunctions, and an integration over the continuous spectrum). Had we used transport theory in an infinite medium instead of diffusion theory, the perturbing term would be more complicated. There are no corrections to (76) in order B^2, and the transport corrections to B^4 has been examined by Nelkin (59) and found to be small ($\sim$ 10%).

The diffusion cooling coefficient is the coefficient of B^4 in the case n = 0. Thus

$$C = \sum_{j=1}^{\infty} \frac{[\langle 0|xD|j\rangle]^2}{\lambda_j^{(0)}} \tag{79}$$

Our eariler thoughts about time eigenvalues and completeness suggest that equation (79) will converge slowly if at all. Further, there is the difficulty, typical of the kinetic theory, that we do not have a convenient set of unperturbed eigenfunctions with which to work. The solution of the equation $S\psi = -\lambda\psi$ is not much easier than the solution of the full problem. Nevertheless, we can extract <u>some</u> useful information from (79).

To begin with, we can get some simple bounds on C by noting that:

$$C > \frac{\langle 0|xD|1\rangle^2}{\lambda_1^{(0)}} \tag{80}$$

and

$$C < \frac{1}{\lambda_1^{(0)}} \sum_{j=1}^{\infty} \left[\langle 0 | xD | j \rangle \right]^2 = \frac{1}{\lambda_1^{(0)}} \left[\langle 0 | (xD)^2 | 0 \rangle - \left\{ \langle 0 | xD | 0 \rangle \right\}^2 \right] \quad (81)$$

Equation (80) gives a lower bound for C that requires knowledge of the first (infinite medium) eigenfunction and its eigenvalue, while (81) requires knowledge only of the first eigenvalue. Unfortunately, we do not have converged values for these quantities for a single non-trivial model. They should be quite useful, however, as a simple calculation will show. Let us guess that $\lambda_1^{(0)}$ is $0.8(v\Sigma_s)_{min}$. Then, for graphite, we can assume a reasonable energy dependence for the diffusion coefficient and find: $C < 30 \times 10^5$ cm^4/sec as a rough estimate, while the range of C reported is 16 to 38 x 10^5 cm^4/sec. A more refined calculation would be useful, indeed!

We can try to use the lower bound to C by guessing at the first eigenfunction and hoping that the matrix element is not too sensitive to our choice. If we take N_1 proportional to $(1 - \frac{1}{2}\sqrt{\pi}\ x)x^2e^{-x^2}$, which is orthogonal to the Maxwellian, and assume constant diffusion coefficient, we find that the lower and upper bounds are precisely the same! If we try $(1 - \frac{2}{3}x^2)x^2e^{-x^2}$, which is another function orthogonal to the Maxwellian, the ratio of lower to upper bound drops only a little, to about 0.92. Then results are certainly surprising and suggest again that a refined calculation be made.

The second trial function leads to a particularly interesting result. It is:

$$C > \frac{2}{3\pi} \frac{D_o^2}{\lambda_1^{(o)}} \qquad \text{(D(x) constant)} \qquad (82)$$

If we use the trial function to obtain a variational estimate of $\lambda_1^{(o)}$, as well as an estimate for the matrix element, we find $\lambda_1^{(o)} \approx \frac{2}{3\sqrt{\pi}} M_2$ (see equation 48), and $C > D_o^2/\sqrt{\pi}\, M_2$. This last estimate, with "greater than" replaced by "equals," has been in the literature for some time (62). It is precisely the result we obtained with the effective temperature model of lecture III, and we can now get an idea of its accuracy.

The expression on the right hand side of (80) is an underestimate of C, and we reach our final expression after approximating both numerator and denominator. We replace the exact value of $\lambda_1^{(o)}$ by a variational estimate, which is an overestimate and which lowers the value of the fraction by a considerable amount. For example, in the case of graphite, $\frac{2}{3\sqrt{\pi}} M_2$ is several times $(v \Sigma_s)_{min}$, while the correct $\lambda_1^{(o)}$ is almost surely <u>less</u> than $(v \Sigma_s)_{min}$. Next, we approximate the numerator in a manner that is "reasonable," and whose accuracy is difficult to assess. The over-all effect is doubtless to lower what is already a lower bound, and the reader should be wary of using the "M_2-formula" for C in any

but a qualitative manner.

The most accurate calculations of diffusion and diffusion cooling coefficients we possess are due to H. Honeck (63). They have been calculated in a manner totally different from the one we have been discussing. We mentioned in the first lecture that the coefficients also appear in the expansion of the diffusion length in powers of the density of 1/v absorber. Indeed, one can see that the two curves are parts of a single curve (as illustrated in Figure 14). The k^2 vs. Σ_a section is better suited to a machine calculation, and one can compute the diffusion length for a variety of moderator-models, using transport or diffusion theory and a large computing machine. In Figure 14 and 15 one can see Honeck's results and experimental points. The calculations come as close to being exact as one can, at present. The light water calculations based upon Nelkin's model give reasonably good agreement with experiment, while those for graphite show a greater disagreement. The following table gives the results of Honeck's study of some common moderating materials. Both light and heavy water are described by "five-frequency" models of the type introduced by Nelkin (30). The model for graphite is the very subtle one developed by Parks (2a).

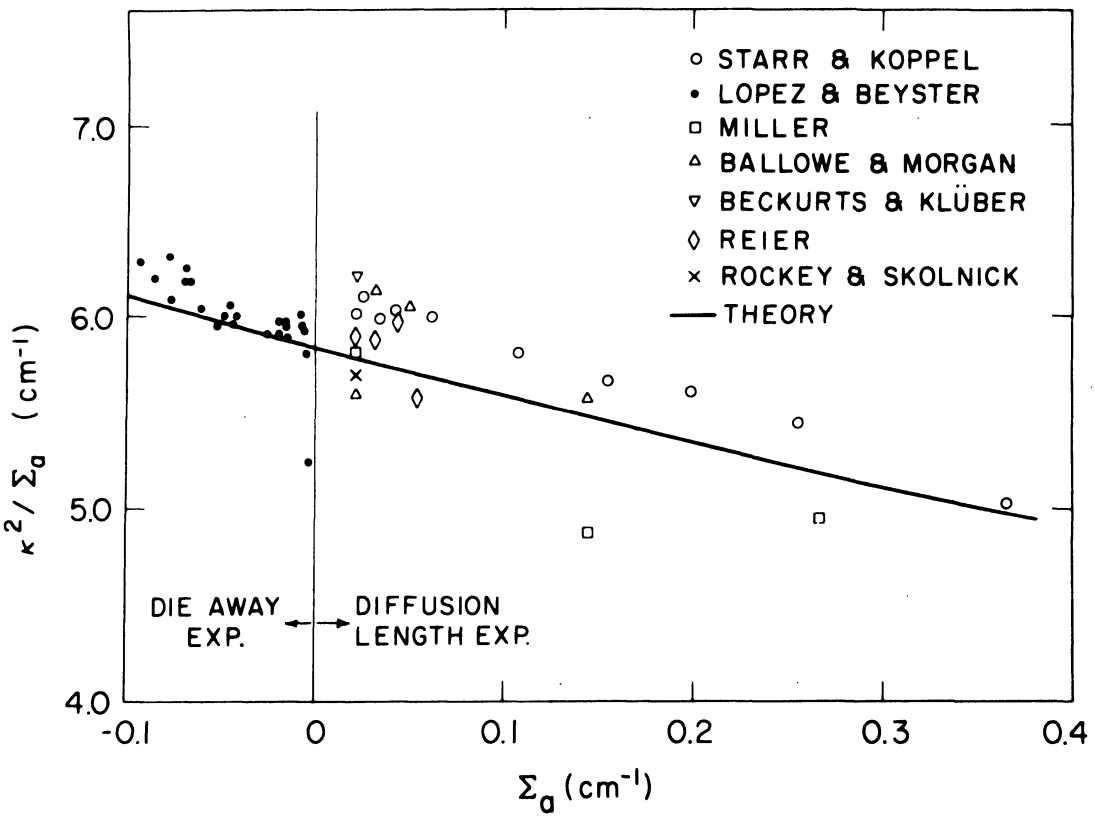
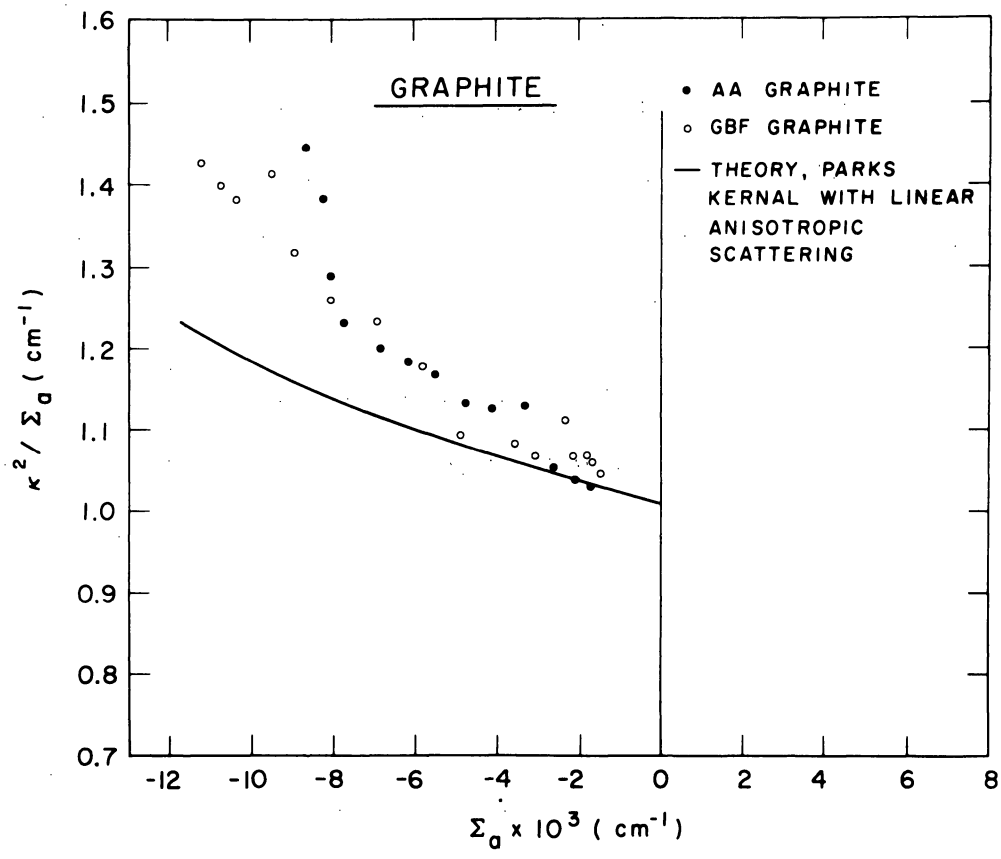

STARR & KOPPEL
LOPEZ & BEYSTER
MILLER
BALLOWE & MORGAN
BECKURTS & KLÜBER
REIER
ROCKEY & SKOLNICK
THEORY
DIE AWAY EXP.
DIFFUSION LENGTH EXP.
7.0
6.0
5.0
4.0
-0.1
0
0.1
0.2
0.3
0.4
κ^2/Σ_a (cm^{-1})
Σ_a (cm^{-1})

Figure 14



Computed and measured diffusion parameters of graphite.

Figure 15

Moderator	C (theory)	C (experiment)
H_2O	3130 cm^4/sec	2900-4900 cm^4/sec
D_2O	5.13×10^5 cm^4/sec	$3.72 \pm 0.5 \times 10^5$ (64)
Graphite	24.6×10^5 cm^4/sec	$16\text{-}38 \times 10^5$

It is clear that we need a tightening of the experimental values. The theoretical values lie well within experimental limits, except in the case of D_2O, where some controversy remains about the analysis of the experimental data.

Systematic studies of higher eigenvalues do not exist at this time. Their measurement is particularly difficult in crystalline moderators where one is not even sure that λ_o can be determined accurately. However, the attempt is being made, and we hope that a few λ_1's will be caught in the next few years.

Mathematical study of the eigenvalues presents special difficulties. As we noted earlier, there is no suitable set of "unperturbed" functions to begin with, which can be combined in a rapidly convergent manner to give the exact eigenfunctions. Further, the limit-point behavior that almost certainly holds for the eigenvalues suggests that variation-iteration methods will converge slowly. Considerable effort has been spent in using, as basis functions in the energy variable, Laguerre polynomials (of first order) multiplied by the Maxwellian (65).

These functions, which, in weak moments, appear "natural" to the thermalization problem, have inappropriate behavior at high energy. They have demonstrated a "natural" rate of convergence that is poor indeed. Another method, suggested by the author, is the use of degenerate, synthetic kernels in a systematic way. It is being studied, at present, at Brookhaven. Finally, an approach based on multi-group theory, which Honeck found successful in analyzing λ_o, has not yet been fruitful in treating higher modes.

We come to the end of these talks with the feeling that we have neglected almost as many topics in thermalization as we have mentioned. In the last lecture, we have directed the reader's attention to questions like a) Do all time-dependent eigenvalues have a limit point at $(v\Sigma)_{min}$? b) How do non-1/v absorption cross sections affect the eigenvalues? c) How do proper, transport-theory boundary conditions affect them?....but only now do we mention that similar problems arise in the study of space-dependent thermalization. In fact, we have barely touched upon that aspect of the subject. We have neglected the velocity-dependent Milne problem, diffusion in a medium containing a temperature gradient, thermalization in lattices...and more. Perhaps these omissions will be partly forgiven if we close by emphasizing that our subject is rich in problems requiring, for their solution, degrees of talent ranging from modest to formidable. If these lectures attract some of the talent to a new and interesting field of physics,

they will have been well worth-while.

APPENDIX I

The Scattering Kernel and Detailed Balance

To illustrate some of the remarks made in text, we consider the scattering from a collection of N identical, spinless nuclei.

$$\frac{d^2\sigma}{d\Omega\, d\epsilon} = \frac{Na^2}{2\pi\hbar}\,\frac{k}{k_o}\int d^3r\,dt\,\exp\left[i(\kappa\cdot r-\omega t)\right]G(\vec{r},t)\,. \tag{I-1}$$

The cross section for scattering from momentum $\vec{p}_o$ to $\vec{p}$ is $\sigma(\vec{p}_o,\vec{p})$ and we have

$$\sigma(\vec{p}_o,\vec{p}) = \frac{1}{m|p|}\,\frac{d^2\sigma}{d\Omega\, d\epsilon} = \frac{Na^2}{2\pi\hbar}\,\frac{1}{m|p_o|}\int d^3r\,dt\ldots \tag{I-2}$$

The space-transform of G(r,t) has been denoted by $\chi(\vec{\kappa},t)$. The general theory (29) gives

$$\chi(\vec{\kappa},t) = \frac{1}{N}\langle\langle \alpha\left|F^{+}(0)F(t)\right|\alpha\rangle\rangle\ , \tag{I-3}$$

where

$$F(t) = \sum_{j=1}^{N}\exp\left[i\kappa\cdot r_j(t)\right]$$

is a quantum-mechanical operator containing the coordinates of the N atoms, and the symbol $\langle\langle\ \ldots\ \rangle\rangle$ represents a quantum-mechanical expectation value taken in state "α" of the

scatterer, followed by a thermal averaging over states. Thus, we may also write

$$\chi(\vec{\kappa},t) = \frac{1}{N} \text{Trace } (F^{+}(0)F(t)e^{-H/T}) \div \text{Trace } e^{-H/T}. \tag{I-4}$$

We wish to show that χ is a real, even function of $\tau = t - \frac{i\hbar}{2T}$, and that this property is closely connected with the property of detailed balance. Consider, then, $\chi_o(t)$, defined through

$$\chi_o(t) \equiv \chi(t + \frac{ih}{2T}) \propto \text{Trace } (F^{+}(0)e^{-H/2T} F(t)e^{-H/2T}). \tag{I-5}$$

We have used, in (I-5), the relation

$$F(t) = \exp\left[\frac{iHt}{\hbar}\right] F(0) \exp\left[-\frac{iHt}{\hbar}\right]. \tag{I-6}$$

The complex conjugate of $\chi_o(t)$ may be written as

$$\chi^{*}_{o}(t) \propto \text{Trace } (F(0)e^{-H/2T} F^{+}(t) e^{-H/2T}), \tag{I-7}$$

if one uses the "cyclic" property of the trace. Now, we shall need two additional assumptions. First, the properties of the scattering system will not depend upon its orientation in space (no single crystals!). We deduce at once that χ_o is a function only of κ^2 and that replacement of $\vec{k}$ by $-\vec{k}$ is of no consequence.

Second, we assume that the properties of the system are unchanged by translation in time, that is, that Trace$(F^+(t_1)\ \ldots F(t_1+t)\ \ldots)$ is independent of t_1. We use the first property after noting that the operator adjoint to F(t) is formed by replacing $\vec{\kappa}$ by $-\vec{\kappa}$. Since this operation, when performed on all of the F(t) which appear, cannot affect the value of the trace, we conclude that (I-7) gives the same result as (I-5). Thus, $\chi_o(t)$ is real.

To prove that $\chi_o(-t) = \chi_o(t)$, we use

$$\chi_o(t) = \chi_o^*(t) \propto \text{Trace}\ (F^+(t)\ e^{-H/2T}\ F(0)\ e^{-H/2T}). \tag{I-8}$$

If we write -t for t in (I-8) and use the time-translation property noted earlier we see that (I-8) is equal to I-5). Thus, we have shown that χ is a real and even function of $\tau = t - \frac{i\hbar}{2T}$.

Now, the property of detailed balance follows easily. Equation (I-2) tells us that

$$|p_o|\ \sigma(\vec{p}_o,\vec{p}) \propto \int_{-\infty}^{\infty} dt\ e^{-i\omega t}\ \chi(\kappa^2,t)$$

or

$$|p_o| e^{-\epsilon/2T}\ \sigma(\vec{p}_o,\vec{p}) \propto \int_{-\infty}^{\infty} dt\ e^{-i\omega t}\ \chi(\kappa^2,\tau+\tfrac{i\hbar}{2T}). \tag{I-9}$$

Since χ is an even function of τ, both sides of the second equation are symmetric in $\vec{p}_o$ and $\vec{p}$, and the detailed balance relation,

$$\left| p_o \right| e^{-E_o/T} \, \sigma(\vec{p}_o, \vec{p}) = \left| p \right| e^{-E/T} \, \sigma(\vec{p}, \vec{p}_o) \quad , \qquad \text{(I-10)}$$

follows at once. ($\epsilon = E_o - E = \hbar\omega$).

APPENDIX II

The Time Decay Constants - In Infinite Media

We remarked in the text that the discrete and continuous behavior of the time-decay constants found in the case of moderation by monatomic hydrogen gas could be expected to hold for more general systems. We shall sketch here the basis for that belief, using the Laplace transform solution of the initial value problem.

Consider the Boltzmann equation in the form

$$\frac{\partial n}{\partial \tau} + \vec{x} \cdot \nabla n + \left|\vec{x}\right| \rho(x) n = \int d\vec{x}' \left|\vec{x}'\right| \rho_s(\vec{x}', \vec{x}) n(\vec{x}', \vec{R}, \tau) + \mathcal{S} \quad \text{(II-1)}$$

We are using dimensionless units, related to the usual $\vec{v}, \vec{r}$, and t by $\vec{v} = v_B \vec{x}$, $\vec{r} = \vec{R} / \Sigma_{sf}$, $t = \tau / v_B$, $\Sigma_s(x', x) = \Sigma_{sf} \rho(x', x)$. Let us take, as a source, a distribution proportional to $\delta(|\vec{x}| - |\vec{x}_o|)\, \delta(\tau)\, \exp\left[iB \cdot R\right]$. When we obtain a description of the evolution of this pulse in the form of a Laplace inversion integral, we shall have, at the same time, its representation in terms of energy modes. To simplify matters, we assume that the scattering kernel is isotropic.

The manner in which (II-1) is analyzed is well known. We remove the space-dependence of $n(\vec{x}, \vec{R}, \tau)$ by setting it propor-

tional to exp $\left[i\vec{B}\cdot\vec{R} \right]$, then obtain the Laplace-transformed equation in the usual way. If λ is the transform variable, we find, in a region where real part of $\lambda > 0$,

$$\left[i\vec{B}\cdot\vec{x} + \lambda + \left|\vec{x}\right| \rho(x) \right] N(\vec{x}, \lambda) = \int d\vec{x}' \left|\vec{x}'\right| \rho(\vec{x}', \vec{x})$$

$$N(\vec{x}') + \frac{1}{4\pi} \delta\left(\left|\vec{x}\right| - \left|\vec{x}_o\right| \right). \qquad \text{(II-2)}$$

Next, we divide by the factor multiplying $N(\vec{x}, \lambda)$ and integrate both sides over orientations of $\vec{x}$. The final equation contains only the magnitudes of $\vec{x}$ and $\vec{B}$ as variables, and is an integral equation,

$$N(x, \lambda) = \frac{1}{2iBx} \ln\left[\frac{\lambda + x\rho + iBx}{\lambda + x\rho - iBx} \right] \left[\int dx' x' \rho(x', x) \right.$$
$$\left. N(x') + \delta(x - x_o) \right]. \qquad \text{(II-3)}$$

We are concerned with the dependence of N upon λ, and it is clear that this is dominated by the logarithm in (II-3). We have assumed that B is real, and it follows that the singularities of the logarithm (and of $N(x, \lambda)$) will be in the left half-plane. For fixed x, the logarithm can be made single-valued if the λ-plane is "cut" between the points $-x\rho + iBx$ and $-x\rho + iBx$. The family of "cuts" obtained by letting x vary from zero to infinity describes an area in the λ-plane bounded by the line of singularities, $\lambda = -\alpha\rho(\alpha) \pm iB\alpha$ $(\infty > \alpha \geq 0)$. When we deform the integration contour to

convert the inversion integral into a "sum of residues," we shall obtain contributions from poles, and then run up against the line of singularities.

Of course, not all of the singularities of $N(x, \lambda)$ will be due to the logarithm. The function will be singular for those λ for which solutions to the homogeneous version of (II-3) exist. These are the "discrete" decay constants, and since we deal with a real kernel which may be made symmetrical, they will be real (and negative). For λ real we may write the homogeneous equation as:

$$N(x,\lambda) = \frac{1}{Bx} \tan^{-1}\left[\frac{Bx}{\lambda + x\rho(x)}\right] \int_0^\infty dx' x' \rho(x';x) N(x') \quad , \tag{II-4}$$

which is a familiar form. Thus, the situation in the λ-plane is as illustrated in Figure (16), and inversion of the Laplace transform gives:

$$N(x,\tau) = \sum_k a_k N_k(x) e^{-\lambda_k \tau} + \frac{1}{2\pi i} \int_{\Gamma(B^2)} d\lambda e^{\lambda\tau} N(x,\lambda) \; . \tag{II-5}$$

The λ_k are functions of B^2, and all lie between zero and $\left[x\rho(x)\right]_{\min}$, while the integration is carried out along the curve $\Gamma(B^2)$, which is the line of singularities mentioned earlier. If the equation for Γ is used to change the complex

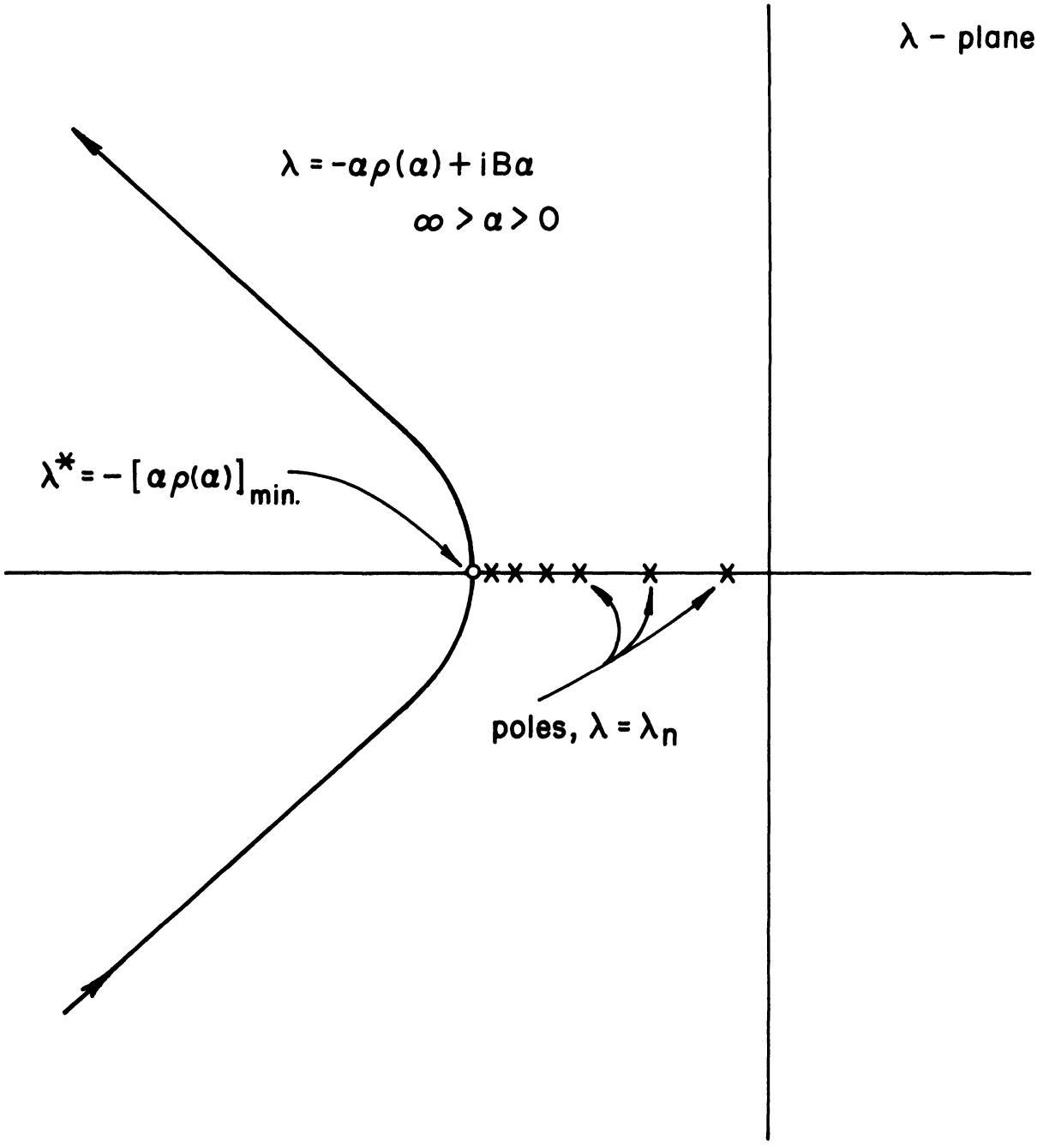


Figure 16

integral into a real integral between $\alpha = 0$ and $\alpha = \infty$, (II-5) will assume an instructive form. For, when $\exp\left[iB\cdot R \right]$ (or $\exp\left[iBz \right]$) is restored we find a superposition of damped "neutron waves." The factor $\exp\left[\lambda\tau\right]$ becomes $\exp\left[-\alpha\rho(\alpha)\tau\right]$. $\exp\left[iB(z\pm\alpha\tau)\right]$ and represents disturbances found previously in connection with the Telegrapher's equation (for neutrons) (11, 61).

There are two special cases of (II-5), worth mentioning. As $B^2 \rightarrow 0$, the curve Γ becomes narrower and approaches the segment of the real axis extending from $-\left[x\rho(x) \right]_{min}$ to . Equation (II-5) then assumes the form of equation (72) of the text. As $B^2 \rightarrow \infty$ a critical B^2 is reached, beyond which no discrete eigenvalues may be found (see Appendix III). Then, only the integral of (II-5) remains.

The connection between (II-5) and the singular, or continuous eigenfunctions mentioned in the text, should be clear. We have really been calculating the Green's function for the time-dependent problem, and (II-5) is its "spectral decomposition." With some manipulation, the equation may be put into the form

$$N(x,\tau) \sim \sum_k N_k^+(x_o) N_k(x) e^{-\lambda_k \tau} + \int_{\lambda^*}^{\infty} d\lambda\, N_\lambda^+(x_o) N_\lambda(x) e^{-\lambda\tau} \qquad \text{(II-6)}$$

$$\left(\lambda^* = \left[x\rho(x)\right]_{min}\right) ,$$

and the discrete and continuous eigenfunctions read off. Thus, existence and completeness appear simultaneously. The reader is invited to try the calculation when the scattering kernel has the simple form, $\rho(x',x) \sim \rho(x')x\rho(x)M(x)$.

APPENDIX III

We wish to show that there exists a critical B, B*, such that when B > B*, no proper solutions of the homogeneous equation (II-4) exist. We shall actually prove an easier version of the theorem, namely, that no energy mode which is everywhere positive exists. Since the fundamental mode is always positive, the theorem refers to it, in particular. The theorem may be proved for <u>any</u> mode, but with additional discussion that is unnecessary here.

The modes which we consider decrease exponentially when x is large. The function $x\rho(x)$ increases linearly when x is large. We shall assume that a solution to the homogeneous equation exists, then multiply the equation by $x\rho(x)$ and integrate. Thus,

$$\int_0^\infty dx\, x\rho(x)n(x) = \frac{1}{B}\int_0^\infty dy\, yn(y)\int_0^\infty dx\, \tan^{-1}\left[\frac{Bx}{x\rho(x)+\lambda}\right]\rho(y,x)\rho(x)\,. \tag{III-1}$$

The left-hand integral will always exist; we shall normalize it to unity. Further, the arc tangent will always be bounded by $\pi/2$. Thus, with the assumption that change of order of integration is allowed (and $n(x) \geq 0$!) we can write

$$1 \leq \frac{\pi}{2|B|}\int_0^\infty dy\, yn(y)g(y)\,,$$

where

$$g(y) \equiv \int_0^\infty dx\, \rho(y,x)\rho(x) \quad . \tag{III-2}$$

We shall show that g(y) is always less than $K\rho(y)$, where K is a positive number independent of B. When this bound is substituted into the first of (III-2) we see that the inequality will lead to nonsense for sufficiently large $|B|$. Then, the positive mode cannot exist.

Consider g(y). It is characteristic of the scattering kernel and is certainly continuous and bounded in the range $\infty > y^* \geq y > 0$. We will want to examine it as $y \to 0$ and as $y \to \infty$. The second limit presents no difficulty since the thermal kernel becomes the slowing-down kernel when y is large. When the slowing-down form is used in (III-2) one sees easily that g(y) is finite as y .

In examining the limit, $y \to 0$, it is convenient to write $\rho(y,x) = \rho(y)\,\pi(y,x)$, where $\pi(y,x)dx$ is the probability that a neutron which is scattered at velocity y will be scattered into the range x, x+dx. $\pi(y,x)$ is never singular for the models we consider, though $\rho(y)$ behaves as 1/y when $y \to 0$. Thus

$$g(y) = \rho(y) \int_0^\infty dx\, \pi(y,x)\rho(x) \tag{III-3}$$

In the limit $y \to 0$, we will consider $\pi(0,x)\,\rho(x)$, which may be singular as $x \to 0$. However, the key point in the argument

is that $\pi(0,x)$ approaches zero as $x \to 0$ in such a way as to ensure the convergence of the integral in (III-3). That is, the probability of elastic scattering vanishes, as the neutron speed vanishes, in all scattering kernels that we consider in thermalization. Thus, $g(y)$ is no more singular than $\rho(y)$, and we may bound the former, $g(y) < K \rho(y)$. The constant, K, is in no way dependent upon B^2. The proof of the theorem follows directly.

References

1. Poole, M. J., et al., Progress in Nuclear Energy, Series I, Vol. II, p. 91, Pergamon Press, New York, 1958.
2. Poole, M. J., J. Nuclear Energy 5, 325 (1957).

2a. Parks, D. E., Beyster, J. R., and Wikner, N. F., Nuclear Sci. and Eng. 13, 306 (1962).

3. Barnard, E., et al., Proc. BNL Conf. on Neutron Thermalization,+ III, 805 (1962).
4. Beckurts, K. H., Z. Naturforschg. 16a, 611 (1961).
5. Young, J. C., et al., Trans. Am. Nuclear Soc. 5, 336 (1962).
6. deSaussure, G., and Silver, E. G., unpublished ORNL reports.
7. Beckurts, K. H., Nuclear Instr. and Methods 11, 144 (1961).
8. For example, see the papers of Starr and Koppel, and of Honeck, in Vols. III and IV of Proc. BNL Conf.
9. Coates, M., unpublished Harwell research (1962).
10. Bennett, R. A., Proc. BNL Conf. III, 838 (1962).
11. Davison, B., Neutron Transport Theory, Oxford (1957); Weinberg, A., and Wigner, E., The Physical Theory of Neutron Chain Reactors, Chicago (1958).
12. Davison, B., op. cit., p. 241.

12a. Coester, F., Phys. Rev. 84, 1259 (1951); Heitler, W., The Quantum Theory of Radiation, Oxford, 1954.

13. Schofield, P., Phys. Rev. Letters 4 (239), 1960.

+ BNL 719 (C-32), Upton, N. Y.

14. Tricomi, F. G., Integral Equations, Interscience, New York (1957).

15. Corngold, N., et al., Proc. BNL Conf. IV, 1103 (1962); Nuclear Sci. and Eng. 15, 13 (1963).

16. Van Hove, L., Phys. Rev. 95, 249 (1954).

17. Glauber, R. J., Phys. Rev. 98, 1692 (1955).

18. Kothari, L. S., and Singwi, K. S., Solid State Physics, 8, 109 (1959), Academic Press, New York.

19. Placzek, G., and Van Hove, L., Nuovo Cimento 1, 233 (1955).

20. Brugger, R. M., Proc. BNL Conf. I, RA1 (1962).

21. Glauber, op cit., Van Hove, op. cit.

22. Vineyard, G., Phys. Rev. 110, 999 (1958).

23. Schofield, op. cit.

24. Rahman, A., et al., Phys. Rev. 126, 986 (1962).

25. Kothari and Singwi, op. cit.

26. Placzek, G., Phys. Rev. 93, 285 (1954); 105, 1240 (1957).

27. Sjölander, A., Arkiv för Fysik 14, 315 (1958).

28. Placzek, G., Phys. Rev. 86, 377 (1952). See also Schofield, P. and Hassit, A., Proc. 2nd Conf. Peaceful Uses Atomic Energy 16, 217 (1958) Geneva.

29. Wick, G. C., Phys. Rev. 94, 1228 (1954).

30. Nelkin, M. S., Phys. Rev. 119, 741 (1960).

31. Parks, op. cit.; also, Kothari and Singwi, op. cit.

32. Honeck, H. C., Trans. Am. Nuclear Soc. 5, 48 (1962).

33. Nelkin, M. S., Nuclear Sci. and Eng. 2, 199 (1957); also Kothari and Singwi, op. cit.

34. Corngold, N., Ann. Phys. 6, 368 (1959); 11, 338 (1960).

35. Nelkin, M. S., ref. 30.

36. Honeck, H. C., private communication.

37. Parks, D., private communication to Purohit, S. N.

38. Purohit, S. N., Proc. BNL Conf. I, 203 (1962); Schofield, P., ibid., I, RB-1 (1962).

39. Purohit, S. N.; Schofield, P., reference 38.

40. Nelkin, M. S., J. Nuclear Energy 8, 48 (1958).

41. von Dardel, G. F., and Sjöstrand, N. G., Phys. Rev. 96, 1245 (1954).

42. Corngold, N., Proc. BNL Conf. IV, 1075 (1962).

43. Wilkins, J. E., Jr., CP 2481 (1944) (unpublished) and Ann. Math. 49, 189 (1948).

44. Cohen, E. R., Nuclear Sci. and Eng. 2, 227 (1957).

44a. Schaefer, G. and Allsopp, K., Proc. BNL Conf. II, 614 (1962).

45. Hurwitz, H., Jr., et al., Nuclear Sci. and Eng. 1, 280 (1956).

45a. Leslie, D. C., ibid., II, 592 (1962).

46. Cadhilac, M., et al., ibid., II, 439 (1962).

47. Morse, P. M., and Feshbach, H., Methods of Theoretical Physics, McGraw-Hill, New York (1953).

48. Wigner, E., and Wilkins, J. E., Jr., AECD 2775 (1948) (unpublished).

49. Case, K. M., Phys. Rev. 80, 797 (1950).

50. Corngold, N., et al., reference 15.

51. Weinberg, A., and Wigner, E., op. cit., p. 336.

52. See reference 11.

53. A related argument is given by de Saussure in Nuclear Sci. and Eng. 12, 433 (1962).

54. Nelkin, M. S., Proc. BNL Conf. IV, RF-1 (1962).

55. Wing, G. M., Proc. 11th Symp. in Appl. Math., p. 140, Amer. Math. Soc., Providence (1961).

56. Davison, B., reference 11, Appendix A.

57. Silver, E. G., Proc. BNL Conf. III, 981 (1962).

58. Takahashi, H., Proc. BNL Conf. IV, 1299 (1962).

59. Nelkin, M. S., Nuclear Sci. and Eng. 7, 210 (1960).

60. Koppel, J. U., private communication.

61. Daitch, P., and Ebeoglu, D. B., Proc. BNL Conf., IV, 1132 (1962).

62. Nelkin, M. S., reference 40; Singwi, K. S., Arkiv för Fysik, 15, 147 (1960); Purohit, S. N., Nuclear Sci. and Eng. 9, 157 (1961). Earlier versions appear in von Dardel, reference 41 and Beckurts, K. H., Nuclear Sci. and Eng. 2, 516 (1957).

63. Honeck, H. C., Proc. BNL Conf. IV, 1186 (1962).

64. Ganguly, N. K., and Waltner, A. W., Trans. Am. Nuclear Soc. 4, 282 (1961).

65. Kazarnovskii, M. V., et al., Proc. 2nd Int. Conf. on Peaceful Uses of Atomic Energy, 16, 279 (1958); Kottwitz, D., Nuclear Sci. and Eng. 7, 345 (1960), and authors listed in reference 62.

THE ELECTRON-NEUTRON INTERACTION

Leslie L. Foldy

Case Institute of Technology

The Electron-Neutron Interaction

Introduction

The study of the electron-neutron interaction started thirty years ago in the same year (1932) in which the neutron was discovered. P. I. Dee investigated the possibility of forces between these particles by looking for recoil electrons and ion pairs produced by neutrons passing through a cloud chamber. He placed an upper limit on the cross-section which was actually several hundred times the figure currently considered to be accurate. In the next fifteen years the problem was not investigated.

The first indication that the interaction was other than that due to magnetic moments came in 1947 from experiments by Havens, Rabi, and Rainwater who measured the total neutron cross-sections of liquid lead and bismuth and inferred the electron-neutron cross-section. (These measurements were redone by Melkonia, Rustad and Havens to a high degree of precision.) Other important determinations of the cross-section were made by Fermi and Marshall, by Hammermesh, Ringo and Wattenberg, and by Hughes, Harvey, Goldberg, and Stafne. These determinations will be discussed below.

The present exposition will be divided into three parts: a discussion of the experiments, a phenomological description of the interaction, and a survey of some aspects of the meson

theoretic structure of the nucleon in its relation to the interaction. Although our discussion is primarily intended to cover the electron-neutron interaction, we shall find that it will naturally generalize to include the proton.

I. Scattering of Slow Neutrons by Atoms

Before discussing the experiments mentioned, it will be necessary to give a brief discussion of the elastic scattering of a slow neutron by an atom as a consequence of its interaction with the nucleus and the atomic electrons. For simplicity, the spin portion of both forces will be ignored.

The force we are looking for is a weak, short-range, spin and velocity independent interaction; this excludes the contribution of the ordinary magnetic dipole interaction known to be present. If the proposed interaction exists it will be describable by a potential depending on the e-n separation. (Henceforth, we will use the small letters "e", "n", and "p" to stand for electron, neutron, and proton respectively. N will stand for nucleon, π for pi meson.) The interaction Hamiltonian for the n-atom system will then be

$$\mathcal{H} = \sum_{i=1}^{Z} V(|\vec{r} - \vec{r}_i|) + U(|\vec{r}|) \quad , \qquad \text{(I-1)}$$

where $\vec{r}$ is the neutron coordinate, $\vec{r}_i$ the coordinate of the i^{th} atomic electron and $U(|\vec{r}|)$ represents the interaction of the neutron with the nucleus, the latter assumed to be infinitely heavy and located at the origin. This latter assumption is for convenience and will be relaxed. The summation is over the Z atomic electrons.

The total Hamiltonian of the system is

$$H = H_a + p^2/2m + \mathcal{H} , \qquad \text{(I-2)}$$

where H_a is the Hamiltonian of the isolated atom and $p^2/2m$ is the kinetic energy of the bombarding neutron.

We now must solve the time-independent Schrodinger equation for the scattering

$$H \Psi(\vec{r}_1,..,\vec{r}_Z,\vec{r}) = E\Psi(\vec{r}_1, \vec{r}_2,..,\vec{r}_Z,\vec{r}). \qquad \text{(I-3)}$$

As noted earlier we only want to look at elastic scattering and therefore (in Born approximation) will only need the ground-state wave-function for the atom, $\psi_o(\vec{r}_1, \vec{r}_2,..,\vec{r}_Z)$. If Eo is the ground state energy of the atom,

$$H_a \psi_o = E_o \psi_o . \qquad \text{(I-4)}$$

The energy of the incident neutron will be written in terms of the neutron wave number k,

$$E = p^2/2m = \hbar^2k^2/2m. \qquad \text{(I-5)}$$

The initial wave number of the neutron is $\vec{k}_i$; the final wave number is $\vec{k}_f$, and for elastic scattering from an infinitely heavy target $\left|\vec{k}_i\right| = \left|\vec{k}_f\right| = k$.

We will calculate the elastic scattering amplitude in the Born approximation. In a quantization volume Ω we can write Ψ_i the initial wave function of the unperturbed system and Ψ_f the final wave function as

$$\Psi_i = \frac{1}{\sqrt{\Omega}} e^{i\vec{k}\cdot\vec{r}} \psi_o(\vec{r}..\vec{r}_z) ,$$

and

$$\Psi_f = \frac{1}{\sqrt{\Omega}} e^{i\vec{k}_f \cdot \vec{r}} \psi_o(\vec{r}_1 \ldots \vec{r}_z) . \tag{I-6}$$

The transition matrix element in the first Born approximation is given by

$$T_{fi} = \langle \Psi_f | \mathcal{H} | \Psi_i \rangle$$

$$= \frac{1}{\Omega} \int d\vec{r}_1 \int d\vec{r}_2 \ldots \int d\vec{r}_z \int d\vec{r}\, e^{i\vec{q}\cdot\vec{r}} \psi_o^* \mathcal{H} \psi_o , \tag{I-7}$$

where $\vec{q} = \vec{k}_i - \vec{k}_f$ is the wave number of the momentum transfer.

The wavelength of low energy neutrons ($\lesssim$ 20 kev) $2\pi/k$ will be long compared to the range of $U(|\vec{r}|)$. Therefore in this energy interval

$$\frac{1}{\Omega} \int d\vec{r}\, e^{i\vec{q}\cdot\vec{r}}\, U(|\vec{r}|) \simeq \frac{1}{\Omega} \int d\vec{r}\, U(r) . \tag{I-8}$$

The factor $\psi_{o*}\ \psi_o$ in this term of (I-7) has already been integrated to one. Introducing a_n the nuclear scattering amplitude for low energy neutrons

$$\frac{1}{\Omega} \int d\vec{r}\, U(r) = -\frac{2\pi\hbar^2}{m\Omega}\ a_n . \tag{I-9}$$

a_n is independent of k as long as (I-8) holds. At somewhat higher energies it will be a slowly varying function of k provided that there are no resonances nearby.

The term involving the interaction of the neutron with the i^{th} electron in (I-7) is

$$\frac{1}{\Omega}\int d\vec{r}_1 \ldots \int d\vec{r}_{i-1} \int d\vec{r}\; e^{i\vec{q}\cdot\vec{r}}\, \psi_o^* V(|\vec{r}-\vec{r}_i|)\, \psi_o \,. \quad \text{(I-10)}$$

Defining

$$\rho(\vec{r}_i) = \int d\vec{r}_1 \ldots \int d\vec{r}_{i-1} \int d\vec{r}_{i+1} \ldots \int d\vec{r}_z\, \psi_o^* \psi_o \quad , \quad \text{(I-11)}$$

equation (I-10) may be written

$$\frac{1}{\Omega}\int d\vec{r}_i \int d\vec{r}\; e^{i\vec{q}\cdot\vec{r}}\, \rho(\vec{r}_i)\, V(|\vec{r}-\vec{r}_i|) \quad . \quad \text{(I-12)}$$

$\rho(\vec{r}_i)$ is the probability of finding the i^{th} electron in the volume element $d\vec{r}_i$ at $\vec{r}_i$. If we now assume that the range of V is small compared to atomic dimensions, we may regard $\rho(\vec{r}_i)$ as essentially constant over the region where V is non-zero in (I-12), evaluated at the point $\vec{r}_i = \vec{r}$. In that case equation (I-12) becomes

$$\frac{1}{\Omega}\int d\vec{r}_i\; e^{i\vec{q}\cdot\vec{r}}\; \rho_i(\vec{r}) \int d\vec{r}_i\, V(|\vec{r}-\vec{r}|) \quad , \quad \text{(I-13)}$$

where the subscript on ρ merely reminds us this is the term for the i^{th} electron.

By analogy with (I-9),

$$\frac{1}{\Omega}\int d\vec{r}_i\, V(|\vec{r}-\vec{r}_i|) = -\frac{2\pi\hbar^2}{m\Omega}\, a_e \quad , \qquad \text{(I-14)}$$

where a_e is the scattering amplitude for low energy neutrons by a fixed electron. Defining a "form factor" $\mathcal{F}_i(\vec{q})$ for the charge distribution of the i^{th} electron

$$\mathcal{F}_i(\vec{q}) = \int d\vec{r}\, e^{i\vec{q}\cdot\vec{r}}\, \rho_i(\vec{r}) \quad , \qquad \text{(I-15)}$$

equation (I-13) becomes

$$-\frac{2\pi\hbar^2}{m\Omega}\, \mathcal{F}_i(\vec{q})\, a_e \quad . \qquad \text{(I-16)}$$

Adding up all such terms, (I-7) becomes

$$T_{fi} = -\frac{2\pi\hbar^2}{m\Omega}\left[a_n + Z\, a_e\, \mathcal{F}(\vec{q})\right] = -\frac{2\pi\hbar^2}{m\Omega}\, a_a \quad , \qquad \text{(I-17)}$$

where $\mathcal{F}(q)$ is the "form factor" for the atom

$$\mathcal{F}(\vec{q}) = \sum_{i=1}^{Z} \mathcal{F}_i(\vec{q})$$

and a_a is the coherent scattering length of the atom.

The differential cross-section is

$$\frac{d\sigma}{d\Omega} = \left| a_a \right|^2 = \left| a_n + z\, a_e\, 7\,(\vec{q}) \right|^2 \quad . \qquad \text{(I-18)}$$

If the possibility of neutron capture is neglected a_n is real. Further, if the electron distribution in the atom is spherically symmetric 7 is real and a function of q (= | q |) only. Equation (I-18) then can be written

$$\frac{d\sigma}{d\Omega} = a_n^2 + 2z\, a_n a_e\, 7\,(q) + z^2 a_e^2 7^2 (q) \quad . \qquad \text{(I-19)}$$

Apart from a factor e, 7 (q) is the Fourier transform of the average electron charge density in the atom. For small q, 7(q) = 1, while in general as q increases from zero, 7(q) decreases monotonically until values of q are reached where 7(q) becomes sensitive to the details of the electron charge distribution. For elastic scattering

$$q = 2k \sin \theta/2,$$

where θ is the angle between $\vec{k}_f$ and $\vec{k}_i$, and so for low energies (I-19) is independent of angle and the total cross-section is

$$\sigma_{tot}\ (k \to 0) = 4\pi \left(a_n + z\, a_e\right)^2 \quad . \qquad \text{(I-20)}$$

For energies such that the neutron wave length ($\sim 1/k$) is small compared to atomic dimensions but still large compared to nuclear dimensions, a_n will still be given by the zero energy form (I-9) while the interference from scattering at various parts of the electron cloud will be nearly complete and the total cross-section will drop to the value

$$\sigma_{tot} = 4\pi a_n^2 \quad . \tag{I-21}$$

Since we shall need a_n in order to evaluate a_e as can be seen in equation (I-19), this is a valuable result.

The forward scattering amplitude (q=o) is always given by

$$a_a = a_n + Z\, a_e .$$

This means that the index of refraction η for neutrons incident on a bulk sample composed of the type of atoms described is

$$\eta^2 = 1 + \frac{\lambda^2}{\pi} N (a_n + Z\, a_e), \tag{I-22}$$

where N is the number of atoms per unit volume and λ is the wavelength of the incident neutrons.

Equations (I-19), 20, and 22 with (I-21) as an important auxiliary relation provide us with the necessary connection between the e-n scattering amplitude, which we want to determine, and experimentally observable quantities.

We have performed our calculations on the basis of an infinitely heavy atom: the results of course are applicable in the center-of-mass system for an atom of finite mass. This will render the distribution of scattered neutrons non-isotropic in the laboratory system even if it is isotropic in the center-of-mass system. If the atoms have a velocity in the lab system as in the case of scattering from a gas, some care is required to separate the effects of atomic motion from the effects of the atomic form factor in converting the observed laboratory distribution to the center-of-mass system where our formulas are applicable. In scattering from solids or liquids there will also be diffraction and interference effects resulting from scattering from different atoms which may effect the determination of the cross-section of a free atom. Neither of the above effects influences the index of refraction.

The interaction of the neutron spin with the spin of the nucleus has been ignored so far. If this is included a_n would become the coherent nuclear scattering amplitude. Actually our evaluation of a_n is not satisfactory even with this modification since the Born approximation is not valid for the small energy neutron-nuclear scattering. However, because nothing in the analysis depends on the way in which a_n is evaluated, the empirical value may be used directly.

Finally, we have ignored the contributions due to scattering by the nucleus and subsequent rescattering by the electrons. This is valid so long as a_n is small compared to atomic dimensions which holds in all cases of present interest.

II. Experiments on the Electron-Neutron Interaction

Experimental determinations of the e-n cross-section are in general of three types: two of these are based on equation (I-19) for the differential cross-section and the other on the expression for the index of refraction in equation (I-22).

In the first two types of experiments, equation (I-19) is used with the term in a_e^2 dropped as small compared to a_n^2, that is, no attempt is made to measure the electron-neutron scattering amplitude directly (incoherently). Instead the experiments rely on the determination of the interference of the "known" amplitude a_n with the unknown amplitude a_e; this requires a good knowledge of a_n. Since a_n is not well known, measurements of the differential cross-section either at several neutron energies or several angles are required to yield a reliable value for a_e. Experiments using the angular dependence were performed by Fermi and Marshall[5] and then more accurately by Hammermesh, Ringo and Wattenberg.[6]

A second type of experiment looks at the energy dependence of the total cross-section based on the expressions in (I-20) and 21. Such experiments have been performed by Havens, Rainwater and Rabi.

The third type of investigation utilizing equation (I-22) for the index of refraction has been done by Hughes, Harvey, Goldberg and Stafne, who measured the relative index of refraction of two materials where the nuclear contribution was similar

for each.

The results of these experiments are usually expressed in terms of a parameter Vo which is defined as the depth of a square potential well which has a radius equal to the classical electron radius e^2/mc^2 and which yields the same scattering as the actual e-n interaction. Using (I-14),

$$\frac{4\pi}{3} V_o \left(e^2/mc^2\right)^3 = \int V(r)\, d\vec{r} = -\frac{2\pi\hbar^2}{M} a_e \,. \qquad \text{(II-1)}$$

The results of the precision experiments which have been performed to date are presented in Table I. The grand average was computed using a weighting according to the percent experimental errors quoted.

Table I

Melkonian,Rustad, & Havens		Vo=-(4165±265)ev.
Hammermesh, Ringo, & Wattenberg	krypton	Vo=-4500 ev.
Crouch, Krohn, & Ringo	xenon	Vo=-(3000)ev.
Mean of the two results above		Vo=-(2900±800)ev.
Hughes, Harvey, Goldberg & Stafne		Vo=-(3860±370)ev.
Grand Average		Vo=-(4050±200)ev.

This value of Vo corresponds to $a_e = 1.46\text{x}10^{-16}$cm.

These experiments seem to indicate clearly the existance of a weak, central interaction between the electron and neutron exclusive of magnetic interactions. The weakness of this force

can be appreciated by noting that for nuclear forces

$V_o \sim -10$ Mev.

III. Character of the Interaction

It has been established in the experiments described above that there does seem to exist a weak, short-range force between the electron and neutron aside from the magnetic dipole-dipole interaction. The next question to explore is that of the origin or type of the force; whether it is some aspect of a familiar interaction or a basically new kind of interaction. We shall ignore the second possibility, at least at first, on the grounds of economy and attempt to associate the findings of the experiments with one of the known forces.

The strong interactions are too strong: as previously noted strong interactions correspond to a Vo of about 10 Mev., a factor of 10^3 too large. The Fermi interaction in β -decay is an attractive possibility but the Vo corresponding to this force is 10^3 times too small. The other simple possibility is that the force is electromagnetic in origin and the coupling in this case is of the right order of magnitude. Indeed the initial investigators expected an electromagnetic interaction based on simple meson-theoretic considerations. In such a picture one expects the neutron to appear as a proton continuously emitting and reabsorbing virtual π^- mesons, a structure which would have electric multipole moments and consequently be capable of interacting with an external change.

We can make a crude estimate of the strength of the interaction based on this simple picture. If the neutron can be

regarded as existing as a proton and a π^- meson essentially all the time, then it will produce a field proportional to the charge e divided by the size of the cloud of $e/(h/m_\pi c)$ where the size of the cloud has been taken to be the Compton wavelength of the meson. The interaction energy for an electron near this configuration is $e^2/(h/m_\pi c)$. Using (II-1) this yields

$$V_o \sim 10^5 \text{ev}.$$

Although this is about 10 times too big, it is encouragingly close in view of the fact that a more accurate version of the calculation would tend to reduce Vo somewhat, since one expects the π^- meson to be separated from the proton only a portion of the time.

We wish to develop a scheme to describe the e-n interaction phenomonologically, that is, given a charge and current distribution characterizing the neutron (assuming the electron is a point distribution), write down the Hamiltonian for the system and the scattering cross-section in terms of these distributions without yet inquiring into the structural dynamics of the neutron which produces them. This procedure will reduce the problem to one of determining the charge and current distributions of the neutron from one theory or another without reference to the electron (provided we neglect external effects such as polarization of the charge and current distributions).

Suppose we take a preliminary look at the neutron regarded as a spherically symmetric charge distribution with center at r

and a charge density $\rho(\vec{b})$ where $\vec{b}$ is measured from $\vec{r}$. In an external electrostatic field ϕ the potential energy of the "neutron" located at the point r is

$$V(\vec{r}) = \int d\vec{b} \quad \rho(\vec{b}) \quad \phi(\vec{r}+\vec{b}). \tag{III-1}$$

If ϕ varies smoothly over the volume of the neutron we may expand it in a Taylor series about $\vec{r}$.

$$\phi(\vec{r}+\vec{b}) = \phi(\vec{r}) + \vec{b}\cdot\nabla\phi + \frac{1}{2}\sum_{i,j} b_i b_j \frac{\partial^2\phi(\vec{r})}{\partial r_i \partial r_j} + \ldots \tag{III-2}$$

from which

$$V(\vec{r}) = \int d\vec{b}\,\rho(\vec{b})\,\phi(\vec{r}) + \nabla\phi(r) - \int \vec{b}\,\rho(\vec{b})\,d\vec{b}$$

$$+ \frac{1}{2}\sum_{i,j}\int \rho(\vec{b})\, b_i b_j \frac{\partial^2\phi}{\partial x_i \partial x_j}\, d\vec{b} + \ldots \tag{III-3}$$

Since the neutron has a net charge of zero, the first term is zero. The second term is zero for $\rho(\vec{b})$ spherically symmetric so that the lowest order non-vanishing term arises.

$$V(r) = \frac{1}{6}\nabla^2\phi(r)\int b^2\rho(\vec{b})\,d\vec{b} = \frac{e}{6}\nabla^2\phi\langle b^2\rangle . \tag{III-4}$$

The quantity $\langle b^2\rangle$ is the second radial moment of the neutron charge distribution. If

$$\phi = - e/r$$

then

$$\nabla^2 \phi = 4\pi e \, \delta(\vec{r}) \ ,$$

and $\langle b^2 \rangle$ can be determined from the experimental value of Vo given in Table I through the relation in (II-1) yielding

$$\langle b^2 \rangle = (3.5 \times 10^{-14} \text{ cm.})^2 ,$$

an estimate of the size of the neutron for the model we have chosen.

It should be noted that the field the neutron was in could have been thought to be produced by any source of an electrostatic field as for example a μ^+ meson. Henceforth we will, for the most part think of the neutron as being in some externally produced electromagnetic field without regard for the source of the field.

The foregoing remarks can form a framework for the more precise procedure we wish to undertake. The structure of the neutron will be represented as the most general charge and current distributions consistant with our qualitative knowledge of the electromagnetic properties of the neutron and basic invariance requirements. The interaction of these charge and current distributions with an external electromagnetic field will then be expressible in terms of quantities which reflect the characteristics of the distributions. Comparison of the predicted scattering with the experimental results may then

specify (although not necessarily unambiguously) the distributions. It is then the task of the various structural theories to reproduce the type of distributions required.

IV. Phenomonology of Electron-Nucleon Scattering: Non-relativistic

We shall now discuss the specification of the interaction of a spin 1/2 particle possessed of a given charge and current distribution with an external electromagnetic field described by the potentials A and ϕ. At first we will restrict our attention to the non-relativistic case since the results are somewhat easier to visualize and will parallel the results of the relativistic calculation for the most part. Note that we will include the proton as well as the neutron in our considerations. First we will write down the Hamiltonian for a non-relativistic point charge in an external electromagnetic field. With $\hbar$=c=1, it is

$$H = \frac{1}{2m}\left[\vec{p} - e\vec{A}(\vec{r},t)\right]^2 + e\,\phi(\vec{r},t) \quad , \tag{IV-1}$$

where e is the charge of the particle and m is its mass.

H can be written in three parts:

$$H = H_o + \mathcal{N} + \mathcal{N}' \quad , \tag{IV-2}$$

where

$H_o = p^2/2m$;

(linear in the field);

$$\mathcal{N} = e\phi - \frac{e}{m}\left[\vec{p}\cdot\vec{A}(\vec{r},t) + \vec{A}(\vec{r},t)\cdot\vec{p}\right] \tag{IV-3}$$

and

$$\mathcal{H}' = e^2/2m\, \vec{A}^2(\vec{r},t) \qquad \text{(quadratic in the field).}$$

If we limit our attention to weak fields $\mathcal{H}'$ will be small and can be ignored. In addition this represents the assumption that the structure of the nucleon is not radically affected by the external field. Note that $\mathcal{H}$ can be rewritten in the more general form

$$\mathcal{H} = \int d\vec{r}' \left[\rho(\vec{r}')\, \phi(\vec{r}',t) - \vec{j}(r',t) \cdot \vec{A}(r',t) \right] . \qquad \text{(IV-4)}$$

For a point charge with no magnetic moment, (IV-4) will reduce to (IV-3) with

$$\rho(r') = e\,\delta(\vec{r} - \vec{r}')$$
$$\vec{j}_c(r') = e/2m \left[\vec{p} \cdot \delta(\vec{r} - \vec{r}') + \delta(\vec{r} - \vec{r}')\, \vec{p} \right] \qquad \text{(IV-5)}$$

j_c is the convection current due to the motion of the particle.

If the particle also has a spin angular momentum and an associated magnetic dipole moment this must be included as a magnetization current,

$$\vec{j}_\mu = \nabla \times \vec{M} \quad , \qquad \text{(IV-6)}$$

where $\vec{M}$ is the magnetization producing this current. For a point magnetic moment with spin

$$M = \mu \vec{\sigma} \delta(\vec{r}-\vec{r}'), \qquad \text{(IV-7)}$$

where μ is the dipole strength carried by the particle. Replacing $\vec{j}$ in (IV-4) by $\vec{j}_c + \vec{j}_\mu$, the Pauli Hamiltonian is obtained,

$$\mathcal{H} = e\phi - \frac{e}{2m}\left[\vec{p}\cdot\vec{A}(\vec{r},t) + \vec{A}(\vec{r},t)\cdot\vec{p}\right] - \mu\vec{\sigma}\cdot(\nabla \times \vec{A}(\vec{r},t)) \; . \qquad \text{(IV-8)}$$

We can generalize $\mathcal{H}$ to the case of extended sources by replacing the δ-functions in ρ, $\vec{j}_c$, and $\vec{j}_\mu$ by finite spherically symmetric distribution functions.

$$\rho(\vec{r}') = e\, f_e(|\vec{r}-\vec{r}'|)$$
$$\vec{A}\cdot\vec{j}_c(r') = e/2m\left[\vec{p}\, f_e(|\vec{r}-\vec{r}'|)\cdot\vec{A} + f_e(|\vec{r}-\vec{r}'|)\,\vec{A}\cdot\vec{p}\right]$$
$$\vec{j}_\mu(\vec{r}') = \mu\vec{\sigma} f_m(|\vec{r}-\vec{r}'|) \; . \qquad \text{(IV-8)}$$

Note that the distribution function in j_μ has not been chosen to be the same as that in ρ and j_c. There is no necessity at this point to make any connection between f_e and f_m.

Note that (IV-9)

$$\int f_m(|\vec{r}-\vec{r}'|)\, d\vec{r}' = 1$$

$$\int f_e(|\vec{r}-\vec{r}'|)\, dr' = \begin{cases} \pm 1 & \text{for particles of charge } \pm e \\ 0 & \text{for particles of charge } 0. \end{cases}$$

The generalizations just made are natural enough but the question of just how general $\mathcal{H}$ has become might be raised. It can be shown that we have included above all possible cases under the following assumptions:

1) If the particle undergoes a translation in space the charge and current distribution associated with it undergo the same translation;

2) ρ is a scalar and $\vec{j}$ is a vector under rotation;

3) ρ is invarient and $\vec{j}$ changes sign under inversion in space or time;

4) the operators for ρ and $\vec{j}$ are functions of $\vec{r}$ and spin and of no higher order than first in the momentum.

The most general form of interaction Hamiltonian in the non-relativistic limit is then

$$\mathcal{H} = \int dr' \left\{ e f_e(\vec{r}-\vec{r}')\,\phi(\vec{r}',t) - \frac{e}{2m}\left[\vec{p}\cdot f_e(\vec{r}-\vec{r}')\vec{A}(\vec{r}',t) + \vec{A}(\vec{r}',t)\cdot f_e(\vec{r}-\vec{r}')p - \mu\vec{\sigma}\, f_m(\vec{r}-\vec{r}')\cdot(\nabla\times\vec{A}(\vec{r}',t))\right\}\right., \tag{IV-10}$$

with $\vec{p}$ and ∇ operating on $\vec{r}$ (not $\vec{r}'$). If the particle has spin 1/2 it has no quadrupole moment and $f(\vec{r}-\vec{r}') = f(\,|\vec{r}-\vec{r}'|\,)$. If the particle has spin 1, f_e is not necessarily spherically symmetric.

The first order perturbation theory gives for the scattering matrix element for the transition $i \rightarrow f$

$$S_{fi} = -\int_{-\infty}^{\infty} dt \left[\Phi_f, \mathcal{H}\Phi_i\right] \tag{IV-11}$$

for a nucleon scattered by an external electromagnetic field. The wave functions in this approximation are plane waves,

$$\Phi_i = \frac{1}{\sqrt{\Omega}} u_i e^{i(\vec{p}_i \cdot \vec{r} - \epsilon_i t)}$$

$$\Phi_f = \frac{1}{\sqrt{\Omega}} u_f e^{i(\vec{p}_f \cdot \vec{r} - \epsilon_f t)} \qquad \text{(IV-12)}$$

where Ω is the normalization volume, u_i and u_f the initial and final neutron spinors, $\vec{p}_i$ and $\vec{p}_f$ the initial and final momenta, and E_i and E_f the corresponding energies. Using these wave functions, (IV-11) becomes

$$S_{fi} = -\frac{1}{\Omega}\left\{ e\langle u_f/u_i\rangle F_e(q)\,\phi_{\vec{q},q_o} - e\langle u_f/u_i\rangle F_e(q)\left(\frac{\vec{p}_i + \vec{p}_f}{2m}\right)\cdot\vec{A}_{\vec{q},q_o} - \langle u_f|\vec{q}\times\vec{\sigma}|u_i\rangle F_m(q)\cdot\vec{A}_{\vec{q},q_o}\right\} , \qquad \text{(IV-13)}$$

where $\vec{q}$ and q_o are the momentum and energy transfers, defined as

$$\vec{q} = \vec{p}_f - \vec{p}_i,$$

$$q_o = E_f - E_i. \qquad \text{(IV-14)}$$

The following definitions have also been made.

$$F_e(q) = \int d\vec{\rho}\; e^{i\vec{q}\cdot\vec{\rho}} f_e(\rho)$$

$$F_m(q) = \int d\vec{\rho}\; e^{i\vec{q}\cdot\vec{\rho}} f_m(\rho)$$

$$\phi\,\vec{q}\,q_o = \int_{-\infty}^{\infty} dt \int d\vec{r}'\; e^{-i\vec{q}\cdot\vec{r}' + iq_o t}\,\phi(\vec{r}', t)$$

$$\vec{A}\,\vec{q}\,q_o = \int_{-\infty}^{\infty} dt \int d\vec{r}'\; e^{-i\vec{q}\cdot\vec{r}' + iq_o t} A(\vec{r}', t) \quad . \qquad \text{(IV-15)}$$

Several conclusions follow from these results, some of which will persist in the relativistic treatment.

1) A transition $i \rightarrow f$ with momentum transfer $\vec{q}$ and energy transfer q_o is possible only if the external fields have a $\vec{q}$ space Fourier component and a q_o time Fourier component. This will be true for the relativistic treatment. The occurance of the transition is then independent of the source of the fields and in the case of electron-neutron scattering we can "replace" the electron by equivalent fields with the same Fourier components as are produced by the electron: these fields are the Møller potentials.

2) A spin-flip transition can only occur in a magnetic field and then only if the scattered particle possesses a magnetic moment. This will not be the case for a Dirac particle.

3) If the scattered particle is neutral ($eFe(q) = 0$), then an electrostatic field has no effect on it. This also is not the case for a Dirac particle.

The calculation we did originally on the scattering of neutrons by atoms included only the electrostatic part of the e-n interaction. (This is the only part of (IV-13) which satisfies our assumptions made at that time about the form of the interaction.) We may compute the e-n cross-section in this approximation from (IV-13) by setting $\vec{A} = 0$. Then

$$\frac{d\sigma}{d\omega} = \left| S_{fi} \right|^2 = \frac{e^2 m^2}{4\pi^2} F_e^2(q) \left| \phi_f \right|^2 \quad .$$

If

$$\phi = \frac{Ze}{r} \quad ,$$

then

$$\phi_{\vec{q}} = \frac{4\pi Ze}{q^2}$$

and

$$\frac{d\sigma}{d\omega} = \left(\frac{Ze^2}{4E_o}\right)^2 \csc\frac{\theta}{2} F_e^2(q) \, .$$

This is the Rutherford cross-section for the Coulomb scattering of two point charges, modulated by the form factor $F_e^2(q)$ arising from the charge distribution of the neutron. If we had allowed the charge Ze to be distributed as well, the cross-section would have been

$$\frac{d\sigma}{d\omega} = \left(\frac{Ze^2}{4E_o}\right)^2 \csc\frac{\theta}{2} F_e^2(q) F^2(q) \quad ,$$

where F(q) is the form factor of the charge Ze. This feature, whereby the form factors occur multiplicatively in the cross-sections, will also be true in the relativistic calculations.

The factor $F_e^2(q)$ can be evaluated in the limit of small q as

$$F_e(q) = \int d\vec{\rho}\, e^{-i\vec{q}\cdot\vec{\rho}} f_e(\vec{\rho}) = \int d\vec{\rho}\left[1 + i\vec{q}\cdot\vec{\rho} - \frac{(\vec{q}\cdot\vec{\rho})^2}{2} + \ldots\right] f_e(\vec{\rho})$$

$$\simeq -\frac{1}{6} q^2 \int d\rho\, \rho^2 f_e(\rho) = -\frac{1}{6} q^2 \langle b^2 \rangle \quad ,$$

where $\langle b^2 \rangle$ is again the second radial moment of the neutron. Higher order terms in the expansion of $F_e(q)$ would be successively higher order (even) moments of the charge distribution of the neutron. If now we connect $\langle b^2 \rangle$ to the

definition of a_e using (III-4) and (I-14), we can regain the result in (I-18) by setting $a_n = 0$ in that equation since we have disregarded the presence of the nucleus in the calculation above.

We have now connected our phenomonological study of the neutron's structure and its effect on the low energy e-n interaction back to the experimental studies which have been made, establishing that these experiments may indeed be explainable on the basis of the neutron possessing some sort of charge (and current) structure. This is essentially all that can be done in the nonrelativistic limit. We must now repeat the calculations for a Dirac particle scattering by an external electromagnetic field. This is essential, not only for purposes of interpreting the recent high energy electron-nucleon scattering experiments such as those done by Hofstadter at Stanford and Wilson at Cornell, but also because the motion of a Dirac particle in an external field exhibits some kinematical peculiarities which will modify the concepts of the origin of the charge and current distributions which we have developed in the nonrelativistic calculation.

V. Kinematics of a Dirac Particle

The Hamiltonian for a free Dirac particle may be written

$$H_o = \beta m + \vec{\alpha} \cdot \vec{p}, \qquad \text{(V-1)}$$

where β and $\vec{\alpha}$ are the Dirac matrices. There exist four linearly independent plane wave solutions to H_o for a given momentum $\vec{p}$. These correspond to an energy E where

$$E = \lambda\sqrt{m^2+p^2}, \qquad \lambda = \frac{\beta m + \vec{\alpha}\cdot\vec{p}}{\sqrt{m^2+p^2}} = \pm 1 . \qquad \text{(V-2)}$$

Two solutions belong to $\lambda = +1$ and two to $\lambda = -1$. The remaining degeneracy cannot be resolved by the Dirac spin operator

$$\vec{\sigma} = \frac{1}{2i}\left[\vec{\alpha} \times \vec{\alpha}\right]$$

since it is not a constant of the motion; Dirac particles cannot stay in states where σ_z is a constant as can Pauli particles. However, there does exist an operator $\vec{\Sigma}$ which is a constant of the motion and which has components parallel or antiparallel to $\vec{p}$, Σ_p which are constants of the motion. For $\lambda = 1$ one solution corresponds to $\Sigma_p = +1$ and the other to $\Sigma_p = -1$ (with a similar situation for $\lambda = -1$.) (For $\Sigma_p = +1$ the spin of the particle is parallel to $\vec{p}$, etc.) Thus all four states are uniquely specified by giving λ and Σ_p.

It is instructive to connect the Dirac Hamiltonian (V-1) to the nonrelativistic Hamiltonian. This may be done by applying

the proper unitary transformation to H to get

$$H_o' = e^{iS} H_o e^{-iS} \quad , \tag{V-3}$$

where

$$S = \frac{i\beta \vec{\alpha}\cdot\vec{p}}{2p} \tan^{-1} \frac{p}{m} \quad . \tag{V-4}$$

This yields for Ho'

$$H_o' = \beta\sqrt{m^2 + p^2} = \beta\left[m + \frac{p^2}{2m} + \ldots\right] \quad , \tag{V-5}$$

where we have expanded in powers of p/m. (V-5) corresponds to the usual non-relativistic Hamiltonian if we ignore the rest energy of the particle and stay at low energies where higher order terms are neglible. If a certain representation for the Dirac matrices is used,

$$\beta = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}, \quad \vec{\alpha} = \begin{pmatrix} 0 & \vec{\sigma} \\ \vec{\sigma} & 0 \end{pmatrix} \quad , \tag{V-6}$$

the four solutions to the Dirac equation take the form

$$\lambda = +1: \begin{pmatrix} \phi_1 \\ 0 \end{pmatrix}, \begin{pmatrix} \phi_2 \\ 0 \end{pmatrix}; \quad \lambda = -1: \begin{pmatrix} 0 \\ \chi_1 \end{pmatrix}, \begin{pmatrix} 0 \\ \chi_2 \end{pmatrix} \quad . \tag{V-7}$$

The matrices in(V-6)are actually 4 x 4 where I is a 2 x 2 unit matrix, etc.: in(V-7)the matrices are 4 x 1 with ϕ_1 a 2 x 1 matrix, etc. Using (V-5) with only the term quadradic in p,

(V-6) and (V-7) yields the usual two component Pauli theory.

Let us now return to equation (V-1). The velocity operator $\dot{\vec{r}}$ has the value

$$\dot{\vec{r}} = -i\left[H_o, \vec{r}\right] = \vec{\alpha} \quad . \tag{V-8}$$

If we compute the value of a component of $\vec{\alpha}$ we find it can only be $\pm c$ (the velocity of light), a most peculiar result. There is, however, another operator, $\vec{R}$, which we can construct which "moves" in a manner consistent with our notion of a position operator, namely $\dot{\vec{R}} \propto \vec{p}$. This operator may be arrived at in the following way.

When the position operator $\vec{r}$ is transformed according to (V-3) its form is very complicated:

$$\vec{r}' = e^{iS}\vec{r}e^{-iS} = \vec{r} - \frac{i\hbar c\beta\vec{\alpha}}{2E} + \frac{ic^3\beta(\vec{\alpha}\cdot\vec{p})\vec{p}}{2E^2(E+mc^2)} - \frac{\vec{\sigma}\times\vec{p}}{2E(E+mc^2)} \tag{V-9}$$

This operator does not represent the position of the particle in the nonrelativistic limit so it is not so surprising that $\dot{\vec{r}}$ in the original representation did not correspond to what we should think of as a velocity. The first term in (V-9) is the position operator of the particle in the low-energy limit. Then applying the F-W transformation (V-3) in reverse to this operator should give us an operator in the original representation which does correspond to what we like to think of as the velocity of the particle. Letting

$$\vec{R} = e^{-is}\vec{r}\,e^{is} \quad , \tag{V-10}$$

where $\vec{R}$ is now in the original representation where H_o has the form (V-1) we get

$$\dot{\vec{R}} = \lambda c^2 \vec{p}/E \quad . \tag{V-11}$$

$\vec{R}$ will be called the mean position operator and from (V-11) it has a time development in keeping with our usual notion of a velocity.

We may interpret the relation between $\vec{r}$ and $\vec{R}$ in the following way. $\vec{r}$ measures the instantaneous position of the particle which oscillates or "dances" around the position $\vec{R}$ with a velocity c (the Zitterbewegung). The extent of the region in which the particle performs these gyrations is h/mc. This is one of the kinematic peculiarities previously mentioned. Forces which are exerted on the particle will act at $\vec{r}$. Thus in an electromagnetic field a point Dirac charge with no intrinsic magnetic moment will appear to have an extended charge distribution spread over a region of dimensions h/mc and, because of the circulation of charge, a magnetic moment. The magnitude of this moment (called the normal Dirac moment) is e/2mc. The effective spreading of the charge distribution of the particle due to the Zitterbewegung gives

rise to the "Darwin" term in the low-energy Hamiltonian which must be included if one is interested in terms of order $1/c^2$. The motion of a Dirac particle carrying a point charge e in an electromagnetic field is described by

$$H = \beta m + e\phi(\vec{r},t) + \vec{\alpha}\cdot\left[\vec{p} - e\vec{A}(\vec{r},t)\right] . \tag{V-12}$$

Thus a point Dirac charge described by (V-12) has an effective finite charge distribution and a magnetic moment which interacts with the field, although the Hamiltonian gives no indication that this would be the case. Pauli pointed out that a Dirac particle could also have an intrinsic point magnetic moment (before dancing) as do the proton and neutron. The form of the Hamiltonian for this addition is

$$H = \beta m + e\phi(\vec{r},t) + \vec{\alpha}\cdot\left[\vec{p} - e\vec{A}(\vec{r},t)\right] - \mu_p\left[\beta\vec{\sigma}\cdot\vec{B}(\vec{r},t) - i\beta\vec{\alpha}\cdot\vec{E}(\vec{r},t)\right] , \tag{V-13}$$

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

and $\vec{E} = -\nabla\phi - \partial\vec{A}/\partial t$.

In the non-relativistic limit, the scattering produced by the Hamiltonian in (V-13) will be that of an extended charge with a magnetic moment $\mu_p + e/2m$ where μ_p is the "anomalous" Pauli moment. The first term in the square brackets multiplying μ_p is the interaction of the Pauli moment with the external magnetic field. The second term is necessary to keep H Lorentz

covarient and may be interpreted as due to the fact that a relativistic particle with an intrinsic magnetic moment which is in motion will develop an electric moment at right angles to both the particle momentum and the intrinsic magnetic moment. This electric moment will then interact with the external E field.

The non-relativistic limit of (V-13) can be used to calculate the scattering of a point neutron from an electrostatic field. There will be scattering as a result of the induced electric moment described above. If the electrostatic field is regarded as being produced by a point charge -e and the results are expressed in terms of V it can be shown that

$$V = \pi \mu_p \left(\frac{e}{M}\right)^2 \left[\frac{4\pi}{3}\left(\frac{e^2}{mc^2}\right)^3\right]^{-1}$$

$$= -4080 \text{ ev},$$

which is almost exactly the strength required by the low energy experiments previously described.

Although initially gratifying, this is a most embarrassing result. Apparently the e-n interaction can be almost exactly explained on the basis of the anomalous Pauli moment of the neutron, or more properly by the electric moment produced by the motion of the anomalous Pauli moment. Of course the experimental errors on the measured value of V will permit a nominal addition to the result deriving solely from the Pauli moment from some intrinsic structure of the neutron, but our estimate

of the contribution which might be expected from a simple separation of charge led to a V_o about one hundred times too large by itself. The Pauli moment must arise from some sort of charge structure, but calculations in weak coupling meson theory which reproduce the anomalous moment also yield a charge separation implying a V many times too big. Also the recent electron scattering experiments conducted on H^2 again suggest that the neutron has no intrinsic charge distribution at all but only a Pauli moment. (It should perhaps be remarked that "intrinsic" as used here will always refer to something attached to the particle and not properties due to the relativistic kinematics.)

VI. Phenomenology of the Electron-Nucleon Interaction: Relativistic

We now wish to write down the Hamiltonian for a Dirac particle with a given (intrinsic) charge and current distribution interacting with an external electromagnetic field in analogy with equation (IV-10) for the non-relativistic case.

The Dirac equation for a free particle can be written in a manifestly Lorentz covarient form:

$$\gamma_\mu \frac{\partial \psi}{\partial x_\mu} + m\psi = 0 \quad , \tag{VI-1}$$

where the γ matrices are obtained from the Dirac matrices upon multiplication of the latter from the left with $i\beta$, and $x = x_1, x_2, x_3, x_4 = x, y, z, i\,t$. Now we should like to add to (VI-1) the most general form of interaction term encompassing the possible structure of the particle which is in first order in A_μ $(= A_x, A_y, A_t, i\phi)$, the external field, independent of the momentum of the particle and gauge invarient. One can show that the Hamiltonian will then become

$$\gamma_\mu \frac{\partial \psi}{\partial x_\mu} + m\psi - i\sum_{n=0}^{\infty}\left[\epsilon^n \gamma_\mu \Box^n A_\mu + \frac{1}{2}\mu_n \gamma_\mu \gamma_\nu \Box^n\left(\frac{\partial A_\mu}{\partial x_\nu} - \frac{\partial A_\nu}{\partial x_\mu}\right)\right]\psi = 0 \quad , \tag{VI-2}$$

where the coefficients ϵ_n and μ_n characterize the charge and current distribution of the particle. If A_μ is static, this is just an expansion in charge and current moments of the particle.

We can write (VI-2) in terms of form factors if A_μ corresponds to a Møller potential.

$$A_\mu \propto e^{i(\vec{q}\cdot\vec{r} - q_o t)}.$$

The effect of the d'Alembertian $\Box$ on this potential is to multiply it by $\vec{q}^2 - q_o^2 = q^2$, the square of the invarient four-momentum transfer. Note that the form factors thus obtained will depend on q^2 rather than $\vec{q}^2$ as was true in the non-relativistic case. The first sum in (VI-2) will then contain something of the form

$$\sum_n \epsilon_n (q^2)^n \longrightarrow F_1(q^2)$$

and the other sum will define a similar function $F_2(q)$. These will be called the form factors of the particle in question. $F_1(q^2)$ can be considered roughly as the Fourier transform of the intrinsic charge distribution of the particle and $F_2(q^2)$ as the transform of the intrinsic current distribution. The actual scattering of such a particle by an electrostatic field will involve <u>both</u> form factors <u>at all energies</u>; the same is true for a static magnetic field.

We have now succeeded in summarizing the information about the intrinsic charge and current distributions of a Dirac particle in an external field in terms of two functions $F_1(q^2)$ and $F_2(q^2)$, where q^2 is the four-momentum transfer associated with a Fourier component of the external field. We have already in effect discussed these functions in the non-relativistic region for the neutron and should now like to consider the recent high energy experiments which have been done.

VII. High Energy Electron-Nucleon Scattering

The proton and neutron taken together are describable by specifying four form factors as functions of $q^2 (= \vec{q}^2 - q_0^2)$, two for each nucleon. These are for the proton $F_1^p(q^2)$ and $F_2^p(q^2)$, and for the neutron $F_1^n(q^2)$ and $F_2^n(q^2)$. We shall always write these with the electric charge e factored out of F_1 and the anomalous magnetic moment factored out of F_2. Then for $q^2 = 0$

$$F_1^p(0) = 1 \quad , \qquad F_2^p(0) = 1$$

$$F_1^n(0) = 0 \quad , \qquad F_2^n(0) = 1.$$

Also from the low energy experiments on electron-neutron scattering

$$\left.\frac{dF_1^n}{d(q^2)}\right]_{q^2 = 0} = 0 \quad .$$

Total knowledge of the nucleon charge and current distributions implies knowing F_1 and F_2 for all values of q^2. Information for large q^2 has been obtained in the experiments at Stanford and Cornell as previously mentioned. The scattering in these experiments can be characterized by the Feynman diagram in Figure 1 representing the scattering of an electron by a nucleon via the exchange of one virtual photon carrying a momentum $\vec{q}$ and an energy q_o.

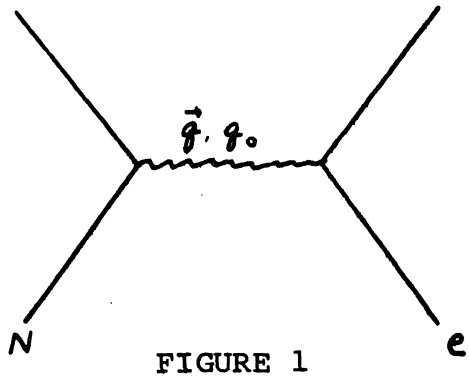


FIGURE 1

The field effective at the position of the nucleon is then the Møller potential previously mentioned. As remarked before, if one assumes that the electron has no structure we can forget about it and consider only the nucleon-photon vertex, where the photon carries a momentum $\vec{q}$ and energy q_o and comes from any source. This vertex will actually appear as shown in Figure 2 where the shaded area denotes that the interaction takes place over an extended region because of the finite charge and current distributions of the nucleon.

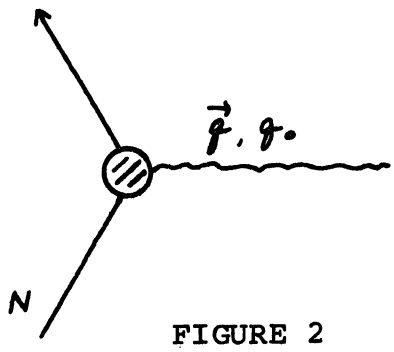


FIGURE 2

The characteristics of this vertex are of course completely determined by F_1 and F_2 and the value of q^2 associated with the photon.

The cross-section for the process shown in Figure 2 has been calculated by Rosenbluth and has the form

$$\frac{d\sigma}{d\omega} = \left(\frac{e^2}{4E_o^2}\,\frac{1}{\sin^4(\theta/2)}\right)\left(\frac{\cos^2(\theta/2)}{1+\frac{2E_o}{M}\sin^2(\theta/2)}\right)\left\{e^2F_1(q^2) - \frac{q^2}{4M^2}\left[2(eF_1(q^2)+2\mu MF_2(q^2))^2\tan^2\theta/2+(2\mu MF_2(q^2))^2\right]\right\},$$

where E_o is the incident energy of the nucleon, M its mass and the magnitude of the anamolous magnetic moment. The first factor in (VII-1) is $1/e^2$ times the Rutherford cross-section. The first term inside the curly brackets corresponds to scattering off the charge distribution of the particle, the term in $\tan^{2\theta}/2$ is the scattering off the total magnetic moment of the particle (normal plus anamolous) and the last term is the scattering off the electric moment induced by the anamolous Pauli moment. Vacuum polarization has been neglected in the derivation of this formula. If we set $F_1 = 1$, $F_2 = 0$ in (VII-1) we have scattering off a point charge with a normal Dirac moment. If $F_1 = 1$, $F_2 = 1$, the scattering is off a point charge with an intrinsic point moment.

If F_1 and F_2 can be obtained for the proton and neutron as a function of q^2 by experiment, we will have obtained a complete phenomonological description of the interaction of nucleons with electromagnetic fields. Observations on e-N scattering for two values of E_o and θ at fixed q^2 will provide sufficient information to determine F_1 and F_2 at that value of q^2.

The proton form factors can in principle be determined with relative ease by scattering high energy electrons off hydrogen. The neutron interaction must be studied indirectly: the most likely source of information is electrodisintegration of the deuteron. However, the theoretical interpretation of the data must take into account many corrections. For instance, the bombarding electron may scatter off the meson associated with the exchange current. Extensive refinements are being made on the theoretical treatment of the deuteron problem and results for the neutron can be obtained with difficulty. The current information on the form factors for the proton and neutron are presented in Figures 3 and 4 respectively.

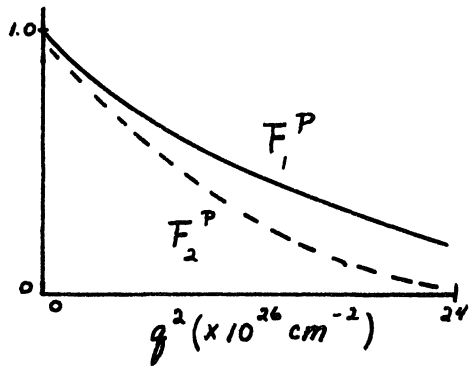


FIGURE 3

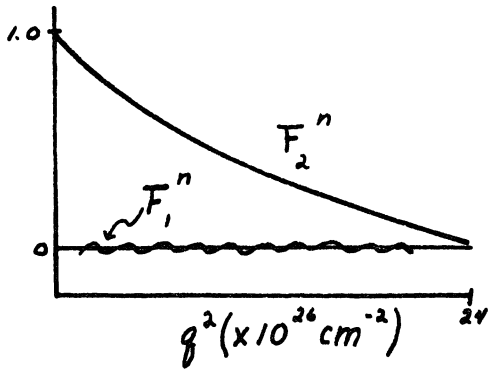


FIGURE 4

The curve for F_1^p seems to be leveling off at large q values, indicating an increasingly dense distribution of charge as one penetrates deeper into the proton; F_2^p seems to be approaching zero at large q indicating that the currents in the proton which are producing the intrinsic magnetic moment are peripheral. F_1^n seems to be zero at all q^2 or at least very small, indicating no intrinsic charge separation in the neutron. The q dependence of F_2^n seems to be commensurate with that for F_2^p.

We have developed a phenomonology for the interaction of nucleons with an external electromagnetic field and found that the interactions can be characterized by two form factors for a nucleon. We should now like to consider the prediction of these form factors from theory, e.g. to investigate the dynamical structure of protons and neutrons.

VIII. Theory of Nucleon Structure

The most general form of the matrix element describing the scattering of a spin 1/2 particle by the absorption of a single photon is dictated to a large extent by requirements of Lorentz covariance and gauge invariance. If j_μ is the four vector current operator for the nucleon which is scattered from a momentum state p_i to p_f:

$$\langle p_f | j_\mu | p_i \rangle = \frac{1}{\sqrt{4 E_i E_f}} \langle u_f | F_1(q^2) \gamma_\mu + i F_2(q^2) \sigma_{\mu\nu} q_\nu | u_i \rangle \quad \text{(VIII-1)}$$

where

$$\sigma_{\mu\nu} = \frac{1}{2} i (\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu) \ .$$

This matrix element is associated with the diagram in Figure 5.

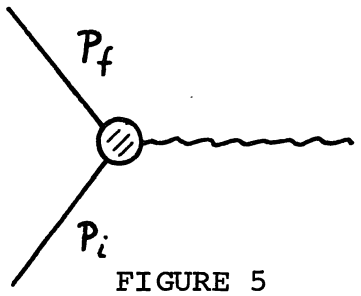


FIGURE 5

The problem of predicting the functions F_1 and F_2 is equivalent to determining what processes are contained in the extended vertex. For instance, the nucleon is capable of emitting a virtual π meson. Then the photon could either be absorbed by the meson or by the nucleon in the presence of the meson. The graph in Figure 5 can be decomposed into a number of graphs to symbolically represent the various contributions to the

scattering. This is done in Figure 6.

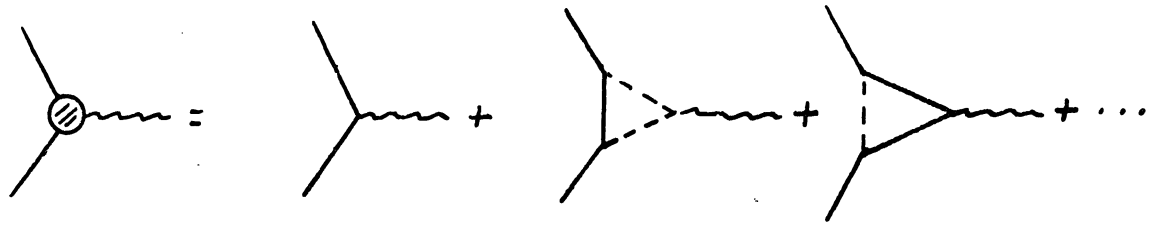


FIGURE 6

The meson is represented by a dotted line. Clearly as more than one meson can be emitted at a time (and even different types of mesons) the actual structure of the vertex in question and hence the prediction of F_1 and F_2 is a very complicated problem. (Note that for the neutron the first graph on the RHS of Figure 6 would be absent.)

A possible line of attack has been investigated using weak coupling perturbation theory and pseudo-scalar mesons. In this theory the ratio of the anamolous magnetic moments of the neutron and proton is independent of the value of the coupling constant. The predicted ratio is about 7 instead of 1, so that little faith can be placed in these results. This value is obtained by considering contributions only from the three graphs in Figure 6. If the last graph is also neglected the ratio obtained is not too bad. Since both of the graphs with one meson processes are of the same order, it is not clear why one should be suppressed.

Another type of calculation is based on the Chew-Low static model for the pion-nucleon system in which one treats the nucleon as in infinitely heavy extended source of mesons. Using this

picture one can obtain results for N-π and N-e scattering which are reasonable, but the assumptions about the meson source preclude a relativistic treatment and an ambiguity exists even in the non-relativistic limit as to what electromagnetic properties to assign to the meson source. Hence it must be concluded that this method of calculation is inadequate to the problem.

The technique of dispersion relations began to be used in 1957-58 in investigations on N-π scattering and has also been applied to the N-e problem. This sort of calculation can be motivated in the following way. Consider the diagram in Figure 7 for N-e scattering.

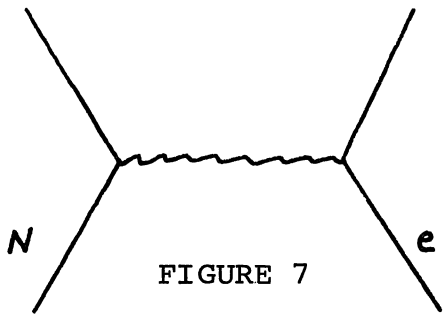


FIGURE 7

This is drawn according to the Feynman convention with time increasing from the bottom to the top of the page. It represents a nucleon and an electron scattering from initial momentum states to final momentum states through the exchange of a virtual photon. We can now think of "reorienting" this graph by redrawing it as shown in Figure 8 with the same convention for the direction of increasing time.

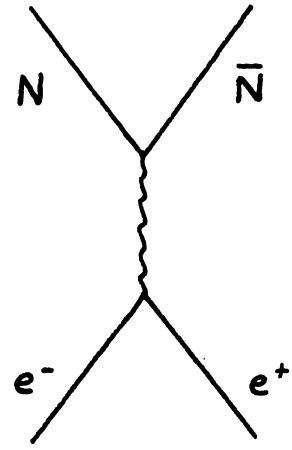


FIGURE 8

This diagram represents an entirely new process; the annihilation of an electron-positron pair with the emission of a photon which subsequently produces a nucleon-antinucleon pair.

The S matrix elements for these two diagrams are identical except for the values of q^2 which one puts into the nucleon form factors F_1 and F_2: in Figure 7, $q^2 > 0$ and in Figure 8, $q^2 < 0$. We have then effectively extended our definitions of F_1 and F_2 to negative values of q^2. They are now defined partly on the $+q^2$ axis and partly on the $-q^2$ axis. This suggests that we extend their definitions into the complex q^2 plane whereby we can utilize the theory of analytic functions.

The analyticity of $F_1(q^2)$ and $F_2(q^2)$ has been investigated in the complex q^2 plane with the result that these functions seem to be analytic at every point except for a cut extending from $q^2 = -4m_\pi^2$ to $q^2 = -\infty$. (m_π is the pion rest mass.) This result has been proved to every order in perturbation theory, but has not as yet been proven for the functions themselves. (That is, even though every term of a series is analytic at

a point the series itself need not be analytic at that point.)

This procedure yields a dispersion relation for either F in the form

$$F(q^2) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\mathrm{Im}\, F(m^2)}{m^2 - q^2 - i\epsilon}\, d(m^2) \quad ,$$

an unsubtracted dispersion relation which requires $F(q^2) \longrightarrow 0$ as $q^2 \longrightarrow \infty$ or

$$F(q^2) = F(0) + \frac{q^2}{\pi} \int_{-\infty}^{\infty} \frac{\mathrm{Im}\, F(m^2)}{m^2(m^2 - q^2 - i\epsilon)}\, d(m^2) \quad ,$$

a subtracted dispersion relation where $F(q^2)/q^2 \longrightarrow 0$ as $q^2 \longrightarrow \infty$.

These dispersion relations (whichever form is pertinent) can only be of use if we know Im $F(q^2)$ for all q^2. This quantity can be expressed as contributions of the following kind. Consider any system into which a virtual photon with $q^2 = m^2$ can go with energy and momentum conservation such that this system can in turn annihilate and produce a nucleon-antinucleon pair. Then each such system contributes to the sum an amount which depends on the two amplitudes involved, i.e. the amplitude for the photon production and the amplitude for subsequent annihilation and nucleon-antinucleon production. Note that for $m^2 > -4m_\pi{}^2$ there are no contributions.

The process involving a $\pi^+ - \pi^-$ pair in the intermediate state is shown in Figure 9.

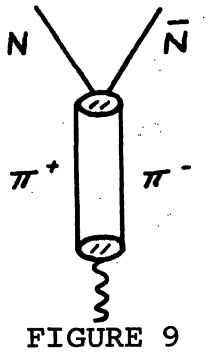


FIGURE 9

The program indicated for a full evaluation of ImF(m^2) is clearly difficult considering the number of processes possible as m^2 gets more and more negative, but several approximate approaches suggest themselves. For q^2 small one might expect the major contributions to the integral to come from small m^2, that is from intermediate states of small mass. The first approximation then is to represent ImF(m^2) by the two pion contribution shown in Figure 9.

From the above discussion it is evident that we need the matrix element for the process $\gamma \longrightarrow \pi^+ + \pi^-$. This is a difficult problem in itself. However, the pion has only one form factor, $F_\pi(q^2)$, which must be determined. We can now repeat the above discussion for $F_\pi(q^2)$ and in the same way arrive at the problem of evaluating the graph in Figure 10.

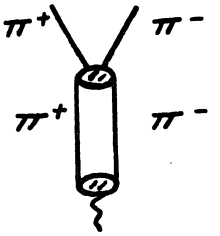


FIGURE 10

This diagram contains the photon-pion vertex again and also $\pi - \pi$ scattering, which we do not now know how to treat. This problem has been investigated by Frazer and Fulco whose work indicates the necessity for a strong T = 1, J = 1^- resonance in $\pi - \pi$ scattering. Among the variety of resonances (or new particles) which have recently been discovered, two seem to have the right quantum numbers to play a role, the ρ with T = 1, J = 1^- and a mass of $\sim$ 780 Mev and the ω with T = 0, J = 1^- and a mass of $\sim$ 790 Mev. The ρ and ω suggest that the isoscalar and isovector parts in the theory can (accidentally) cancel as they should to give $F_1^n(q^2) \simeq 0$.

References

A list of references pertaining to the work referred to in these lectures (except for calculations using dispersion relations) can be found in the review article "Neutron-Electron Interaction" by L. L. Foldy, Rev. Mod. Phys. 30, 471 (1958). References on the use of dispersion relations and discussion of the more recent high energy work on the electron-nucleon interaction can be found in Electromagnetic Structure of Nucleons by S. D. Drell and F. Zachariasen (Oxford University Press, London, 1961). The discussion of the kinematics of a Dirac particle can be found in a much more complete form in the article by L. L. Foldy in Quantum Theory III: Radiation and High Energy Physics, edited by D. R. Bates (Academic Press, New York, 1962).

MESONS AND BARYONS

Suraj N. Gupta

Wayne State University

Introduction

It is indeed a pleasure for me to present these lectures on the theory of mesons and baryons. In order to develop a subject like this in a meaningful way, I think it would be best for me to summarize some well-known ideas in the beginning.

Strongly interacting particles can be divided into two groups according to their statistics. Those which obey the Bose-Einstein statistics are called mesons and those which obey the Fermi-Dirac statistics are called baryons. The mesons and baryons which have been experimentally observed are shown in Table 1. Notice that the antimesons* are listed with the mesons, while the antibaryons are listed separately. We will comment on this difference later.

Table 1

Known Mesons and Baryons

Mesons	Baryons
π^+, π^0, π^- } Pions	p, n } Nucleons
K^0, $\tilde{K}^0$, K^+, K^- } K Mesons	Λ, Σ^+, Σ^0, Σ^-, Ξ^0, Ξ^- } Hyperons
	Plus an antibaryon for each baryon

* We denote the antiparticle of A as $\tilde{A}$. Note that $\tilde{\pi}^+ = \pi^-$, $\tilde{\pi}^0 = \pi^0$, $\tilde{K}^+ = K^-$.

Conservation of charge and baryon number are two well known conservation laws which govern the interactions of these mesons and baryons. Conservation of charge states that the total charge of a system of particles does not change. Conservation of baryon number states that, in a system of particles, B (the number of baryons) minus $\tilde{B}$ (the number of antibaryons) does not change. There are also conservation laws for energy and momentum, but we need not explicitly consider them here.

Conservations of charge and baryon number alone are not sufficient to account for the observed properties of the interactions of mesons and baryons. For example, the following two reactions conserve both charge and baryon number:

$$\pi^- + p \longrightarrow \Lambda + K^\circ \qquad (1)$$

$$\pi^- + p \longrightarrow n + K^\circ \; . \qquad (2)$$

Reaction (1) is found to have a large cross section, while reaction (2) is not observed. This problem was solved by Gell-Mann[1] and Nishijima[2] by introducing the idea of strangeness. They assigned a new quantum number, called strangeness, to each particle, and postulated that strangeness is conserved for strongly interacting particles. We then find that in reaction (1) strangeness is conserved while in reaction (2) it is not.

A mathematical formulation of these ideas was given by d'Espagnat and Prentki[3]. They utilized the concept of the three-dimensional charge space called the isospace and obtained a relation of the form

$$Q = I_3 + \tfrac{1}{2}U, \tag{3}$$

where Q is the charge in units of e, I_3 is the third component of the isospin, and U is the hypercharge. The name hypercharge for the quantity U, which is closely related to the strangeness, was first used by Schwinger[4].

It is well known that the invariance of the Lagrangian density L under rotations in the isospace gives the conservation of I_3. D'Espagnet and Prentki have also tried to derive the conservation of hypercharge by requiring that L be invariant under inversion (reversal of the three axes) in the isospace. But such an inversion only ensures the conservation of parity in the isospace, and gives us

$$\Delta U = 4n, \tag{4}$$

where n is a positive, negative, or zero integer. To overcome this difficulty they imposed the restriction that U can take only the values

$$U = 0, \pm 1, \tag{5}$$

which, together with (4), leads to the conservation of U in the usual type of interactions. However, the restriction (5) is

not only rather artificial but also looks suspicious, because it rules out the existence of any new charged mesons or baryons, although several claims for the possible existence of such particles have been made from time to time.

There have been many attempts to extend this theory. These attempts usually extend the number of dimensions in the isospace to four or more in order to replace the discrete three-dimensional inversion by a continuous four-dimensional rotation. The invariance of L under all rotations in this four-dimensional space gives the conservation of U without limiting its values. However, the number of particles introduced in the theory is much larger than has been observed. One hesitates to accept a theory with many unverified particles.

We shall, therefore, follow a different approach in which we return a three-dimensional isospace but replace the discrete inversion by a suitable continuous transformation. We shall thus obtain the conservation of hypercharge without any artificial restriction and without introducing too many new particles.

I. Three-Dimensional Spinors

Before formulating the theory of mesons and baryons, we will discuss the mathematical theory of three-dimensional spinors by simplifying the original work of Cartan[5].

Let a be a vector with complex components and such that

$$a_1^2 + a_2^2 + a_3^2 = 0. \tag{6}$$

Since a satisfied equation (6) we can construct only two independent quantities from its components. Let us choose them as

$$u = (-a_1 + ia_2)^{\frac{1}{2}},$$

$$v = (a_1 + ia_2)^{\frac{1}{2}},$$

$$uv = a_3, \tag{7}$$

where the third relation fixes the relative signs of u and v. The components of a transform under an infinitesimal rotation as

$$a'_i = a_i + \omega i_k \, a_k \ , \tag{8}$$

where the ω_{ik} are real infinitesimal quantities with the antisymmetrical property $\omega_{ik} = -\omega_{ki}$. Applying equation (8) to equations (7), we see that u and v transform under an infinitesimal rotation as

$$u' = u + \tfrac{1}{2}i\omega_{12}u + \tfrac{1}{2}(i\omega_{23} + \omega_{31})v$$

$$v' = v - \tfrac{1}{2}i\omega_{12}v + \tfrac{1}{2}(i\omega_{23} - \omega_{31})u. \qquad (9)$$

If we write

$$\psi = \begin{pmatrix} u \\ v \end{pmatrix} \qquad (10)$$

then equations (9) can be written in a more compact form as

$$\psi' = \left[1 + \frac{i}{2}(\omega_{23}\tau_1 + \omega_{31}\tau_2 + \omega_{12}\tau_3) \right] \psi \qquad (11)$$

where

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \qquad (12)$$

The two-component quantity ψ is called an isospinor.

Consider the transformation of ψ under inversion. The components of $\vec{a}$ transform under inversion as

$$a_i' = a_i \ ,$$

and therefore, according to (7), ψ transforms either as

$$u' = iu \ , \quad v' = iv \ , \quad \psi' = i\psi \ , \qquad (13)$$

or as

$$u' = -iu, \quad v' = -iv, \quad \psi' = -i\psi \qquad (14)$$

The two preceding transformation relations refer to isospinors of the first and second kind, respectively.

The following two theorems relating to these isospinors are particularly useful:

Theorem 1

Let $\psi = \begin{pmatrix} u \\ v \end{pmatrix}$ be an isospinor of the first kind.

Then $\phi = \begin{pmatrix} -v^* \\ u^* \end{pmatrix}$ is an isospinor of the second kind.

Proof

We see from the complex conjugates of equations (9) that ϕ transforms as an isospinor under infinitesimal rotations:

$$\phi' = \left[1 + \frac{i}{2} \left(\omega_{23} \tau_1 + \omega_{31} \tau_2 + \omega_{12} \tau_3 \right) \right] \phi$$

Further, under inversion, u and v transform as

$u' = iu,$

$v' = iv,$

so that u* and v* transform as

$u^{*\prime} = -iu^*,$

$v^{*\prime} = -iv^*,$

and ϕ transforms as an isospinor of the second kind,

$\phi' = -i\phi.$

Theorem 2

If ψ is an isospinor of the first kind and ϕ is an isospinor of the second kind, then

A) $\psi^* \psi$ and $\phi^* \phi$ are isoscalars,

B) $\phi^* \psi$ and $\psi^* \phi$ are isopseudoscalars,

C) $\phi^* \tau_i \psi$ and $\psi^* \tau_i \phi$ are isovectors,

D) $\phi^* \tau_i \phi$ and $\psi^* \tau_i \psi$ are isopseudovectors.

In Theorem 2, we have used the following definitions:

An isoscalar (isopseudoscalar) is a one component quantity in the isospace which does not (does) change sign under inversion.

An isospinor of the first (second) kind is a two component quantity in isospace which is multiplied by i(-i) under inversion.

An isovector (isopseudovector) is a three component quantity in isospace which does (does not) change sign under inversion.

The proof of the above theorem easily follows from the transformation relations for isospinors under infinitesimal rotation and inversion.

We can construct quantities with any number of isospace components. For example, a four-component quantity is obtained by first multiplying a spinor, $\psi = \begin{pmatrix} u \\ v \end{pmatrix}$, and a vector,

$U = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix}$, to get a six-component quantity

$$\psi_i = \begin{pmatrix} u U_i \\ v U_i \end{pmatrix}, \quad i = 1, 2, 3,$$

and then imposing the condition, $\tau_i \psi_i = 0$, to eliminate two of the six components of ψ_i. A five component quantity is obtained by taking a symmetrical tensor, U_{ik} and imposing the condition $U_{ii} = 0$, and so on.

II. Isocovariance And Isocharge

For any isocovariant quantity ψ, we can express the transformation relation for an infinitesimal rotation as

$$\psi' = \left[1 + i\left(\omega_{23} I_1 + \omega_{31} I_2 + \omega_{12} I_3\right)\right]\psi , \tag{15}$$

where ψ is written as a column matrix, and, in general, can have any number of components, while I_1, I_2, and I_3 are the Hermitian isospin matrices.

We can illustrate this general transformation by several examples. An isoscalar or isopseudoscalar, ψ, transforms as

$$\psi' = \psi , \tag{16}$$

and the isospin matrices are

$$I_1 = I_2 = I_3 = 0.$$

An isospinor of the first or second kind, $\psi = \begin{pmatrix} u \\ v \end{pmatrix}$, transforms as

$$\psi' = \left[1 + \frac{i}{2}\left(\omega_{23} \tau_1 + \omega_{31} \tau_2 + \omega_{12} \tau_3\right)\right]\psi , \tag{17}$$

and the isospin matrices are

$$I_i = \tfrac{1}{2} \tau_i .$$

An isovector or isopseudovector, $\psi = \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix}$, transforms as

$$U'_i = U_i + \omega_{ik} U_k \quad , \tag{18}$$

and the isospin matrices are

$$I_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

$$I_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}$$

$$I_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad .$$

For the isovector and isopseudovector, it is more useful to choose the components of ψ in such a way that I_3 is diagonal. Thus, if we take

$$\psi = \begin{pmatrix} \psi^+ \\ \psi^0 \\ \psi^- \end{pmatrix}$$

where

$$\psi^{\pm} = \frac{1}{\sqrt{2}} (U_1 \mp i U_2) \quad ,$$

$$\psi^0 = U_3 \quad ;$$

then I_3 is diagonal and given by

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Table 2 lists the diagonal I_3 isospin matrices for isocovariants with one to five components.

Table 2

Diagonal I_3 Isospin Matrices

Isocovariant	I_3
One Component	0
Two Components	$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$
Three Components	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$
Four Components	$\begin{pmatrix} 3/2 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -3/2 \end{pmatrix}$
Five Components	$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$

Consider an infinitesmal rotation is isospace through an angle $\delta\ \phi$ about the X_3 axis. For this rotation, the transformation coefficients ω_{ik} are

$$\omega_{23} = \omega_{31} = 0$$

$$\omega_{12} = \delta\phi,$$

and the isocovariant ψ transforms as

$$\psi' = (1 + i\,\delta\phi\, I_3)\,\psi . \tag{19}$$

Since $\delta\phi$ is infinitesimal, we can write (19) as

$$\psi' = e^{i\,\delta\phi\, I_3}\,\psi , \tag{20}$$

and for a finite angle of rotation ϕ, equation (20) becomes

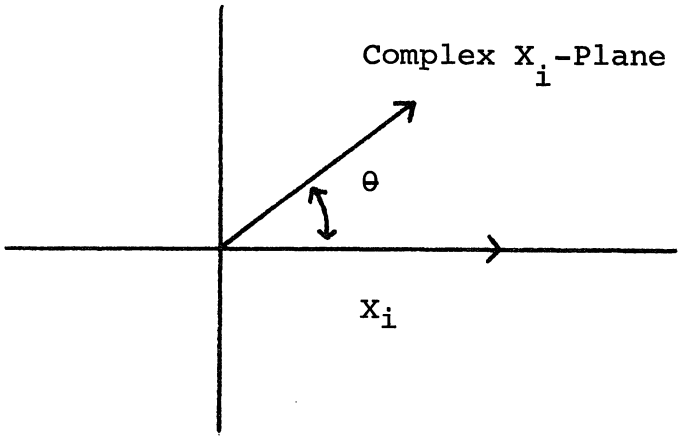
$$\psi' = e^{i\,\phi\, I_3}\,\psi . \tag{21}$$

Let us denote the r^{th} diagonal element of I_3 as Z_r, and call Z_r the isocharge associated with the r^{th} component of ψ.

III. Inversion and Hypercharge

Inversion is usually regarded as a discontinuous transformation. However, we can associate a complex plane with each of the three axes and rotate any axis through an angle θ in its complex plane as shown in Figure 1. Inversion is then obtained by the simultaneous rotation of all three axes through an angle $\theta = \pi$ in their complex planes.

Figure 1



Under a rotation of all three axes in their complex planes through an angle θ, the unit vectors transform as

$$\begin{aligned} \vec{i}' &= \vec{i}\, e^{i\theta} \\ \vec{j}' &= \vec{j}\, e^{i\theta} \\ \vec{k}' &= \vec{k}\, e^{i\theta} \end{aligned} \quad , \qquad (22)$$

At first one would think that if A is an isoscalar and B_i is an isovector, then they would transform under the rotation (22) only as

$$A' = A \ , \tag{23}$$

$$B_i' = e^{-i\theta} B_i \tag{24}$$

But, it follows from (24) that the isoscalars B_i^2 , $B_i^* B_i$ and B_i^{*2} transform as

$$B_i'^2 = e^{-2i\theta} B_i^2 \ , \ B_i^{*\prime} B_i' = B_i^* B_i \ , \ B_i^{*\prime 2} = e^{2i\theta} B_i^{*2} \ , \tag{25}$$

and by multiplying the above relations with each other we can obtain further transformation relations for the isoscalars.

We can, however, obtain the same result more easily in the following way: Let the isoscalar A transform as

$$A' = e^{i\theta Y} A \ , \tag{26}$$

where Y is a constant which is to be determined from the condition that $e^{i\theta Y} = 1$ for $\theta = \pi$. This gives us

$$e^{i\pi Y} = 1 \ , \tag{27}$$

or

$$Y = 0, \pm 2, \pm 4, \ldots . \tag{28}$$

Thus, we obtain infinite types of isoscalars with different transformation properties under (22), which would be indistin-

guishable from each other if we confine ourselves to the usual discontinuous inversion.

In general, any isocovariant ψ transforms as

$$\psi' = e^{i\theta y} \psi \quad , \tag{29}$$

and, since a rotation through an angle π corresponds to inversion, we get the following restrictions on y:

$e^{i\pi y} = 1$, for isoscalars and isopseudovectors, (30a)

$e^{i\pi y} = -1$, for isopseudoscalars and isovectors, (30b)

$e^{i\pi y} = i$, for isospinors of the first kind, (30c)

$e^{i\pi y} = -i$, for isospinors of the second kind. (30d)

Equations (30) imply that

$y = 0, \pm 2, \pm 4, \ldots$, for isoscalars and isopseudovectors, (31a)

$y = \pm 1, \pm 3, \pm 5, \ldots$, for isopseudoscalars and isovectors, (31b)

$y = \frac{1}{2}, \frac{1}{2} \pm 2, \frac{1}{2} \pm 4, \ldots$, for isospinors of the first kind, (31c)

$y = -\frac{1}{2}, -\frac{1}{2} \pm 2, -\frac{1}{2} \pm 4, \ldots$, for isospinors of the second kind. (31d)

We call y the hypercharge associated with the isocovariant ψ.

When I_3 is diagonal, we can also define Q as *

$$Q = I_3 + y, \tag{32}$$

where the r^{th} diagonal element of Q is called the charge asso-

* Note that y in equation (32) and U in equation (3) differ by a factor of two.

ciated with the r^{th} component of the isocovariant ψ .

The values of isocharge for any isocovariant ψ are uniformly distributed about zero, and therefore its average value vanishes. Thus, the hypercharge equals the average charge associated with the components of ψ .

IV. Charge Conservation

We shall now make use of the following well-known theorem from field theory:

> Let a system of fields be described by the isocovariant field operators, $\psi(1), \ldots, \psi(n)$, such that each field operator is a one column matrix. Let $\psi_r^{(s)}$, the r^{th} element of $\psi^{(s)}$, contain the annihilation operators for particles with charge $Q_r^{(s)}$ and the creation operators for particles with charge $-Q_r^{(s)}$. If the Lagrangian density, L, of this system is invariant under the transformation of all the field operators as
>
> $$\psi_r^{(s)\prime} = e^{i\alpha Q_r^{(s)}} \psi_r^{(s)} \quad , \tag{33}$$
>
> with α an arbitrary constant, then the total charge of the system is conserved.

We postulate that L is invariant under arbitrary rotations in isospace as well as under simultaneous rotations of the isospace axes in their complex planes through any arbitrary angle. It then follows that L is invariant under the transformations

$$\psi^{(s)\prime} = e^{i\phi I_3^{(s)}} \psi^{(s)} \tag{34}$$

and

$$\psi^{(s)'} = e^{i\phi y^{(s)}} \psi^{(s)} \quad . \tag{35}$$

Since L must be invariant under the product of transformations (34) and (35), we also obtain invariance of L under the transformation

$$\psi^{(s)'} = e^{i\phi (I_3^{(s)} + y^{(s)})} \psi^{(s)} = e^{i\phi Q^{(s)}} \psi^{(s)} \quad , \tag{36}$$

which, in view of the above field-theoretical theorem, leads to charge conservation.

Transformation (35) contains inversion as a special case, and hence L is invariant under inversion. However, invariance under inversion does not imply invariance under the transformation (35). Thus, the invariance of L under isospace rotations and inversion is not sufficient to imply charge conservation in general.

We have derived charge conservation in strong interactions by considering only the structure of the theory of strong interactions. It has not been necessary to invoke the gauge invariance of the electromagnetic field.

V. Multiplets Of Elementary Particles

Elementary particles with the same values of baryon number, isospin, and hypercharge are said to belong to the same multiplet. Particles belonging to the same multiplet must have the same parity and spin. Their masses should also be the same. However, they are found to be only approximately equal. For example, there is a 4.6 Mev mass difference between the charged and neutral pions. Proposals have been made to account for the mass differences on the basis of electromagnetic interactions[6], but we will not comment on the merits of these proposals.

If we assume that corresponding to every set of values of isospin, isocharge, and hypercharge there exists one meson, one baryon and one antibaryon, and if we restrict the charge of the mesons and baryons in any multiplet to $Q = 0, \pm 1$, we obtain the mesons and baryons shown in Table 3.

Notice that for every baryon there is an antibaryon with opposite isocharge and hypercharge. We are free initially to choose any one of the particles as a baryon or antibaryon. However, the remaining assignments follow from the conservation of baryons. It should also be noted that since the number of mesons is not conserved, it is impossible to divide them into mesons and antimesons in a unique way.

Table 3

Meson and Baryon Multiplets with $Q = 0, \pm 1$

Isocovariance	Iso-spin	Iso-charge	Hyper-charge	Charge	Meson	Baryon	Anti-baryon
Isoscalar	0	0	0	0	ρ^0	Λ	$\tilde{\Lambda}$
Isopseudo-scalar	0	0	1,-1	1,-1	σ^+, σ^-	x^+, y^-	$\tilde{y}^-, \tilde{x}^+$
Isospinor first kind	½	½,-½	½	1,0	K^+, K^0	p, n	$\tilde{\Xi}^-, \tilde{\Xi}^0$
Isospinor second kind	½	½,-½	-½	0,-1	$\tilde{K}^0, K^-$	Ξ^0, Ξ^-	$\tilde{n}, \tilde{p}$
Isopseudo-vector	1	1,0,-1	0	1,0,-1	π^+, π^0, π^-	$\Sigma^+, \Sigma^0, \Sigma^-$	$\tilde{\Sigma}^-, \tilde{\Sigma}^0, \tilde{\Sigma}^+$
Isovector*	1	1,0,-1	-	-	-	-	-

* The smallest value of the hypercharge (1, -1) gives doubly charged particles for the isovector. Thus, there are no isovector multiplets with $Q = 0, \pm 1$.

Without any mathematical inconsistency, we could have defined Q in a more general way as

$$Q = I_3 + cy,$$

where c is any constant. However, it is easy to see that when c = 1, the number of particles in multiplets with $Q = 0, \pm 1$ is maximum. Thus, nature seems to prefer multiplets with $Q = 0, \pm 1$.

VI. Interactions of Mesons and Baryons

We denote the field operator of a particle by the symbol for the particle, as shown in Table 4. For instance, p denotes the operator for the proton field, which contains the annihilation operators for protons and the creation operators for antiprotons.

Table 4

Isocovariance	Field Operator: Meson Field Operator	Field Operator: Baryon Field Operator
Isoscalar	ρ^0	Λ
Isopseudoscalar	σ^+, σ^-	X^+, Y^-
Isospinor first kind	$K = \begin{pmatrix} K^+ \\ K^0 \end{pmatrix}$	$N = \begin{pmatrix} p \\ n \end{pmatrix}$
Isospinor second kind	$K' = \begin{pmatrix} -K^{0*} \\ K^{+*} \end{pmatrix} = \begin{pmatrix} -\bar{K}^0 \\ K^- \end{pmatrix}$	$\Xi = \begin{pmatrix} \Xi^0 \\ \Xi^- \end{pmatrix}$
Isopseudovector	π_i $\pi_1 = \frac{1}{\sqrt{2}}(\pi^+ + \pi^-)$ $\pi_2 = \frac{i}{\sqrt{2}}(\pi^+ - \pi^-)$ $\pi_3 = \pi^0$	Σ_i $\Sigma_1 = \frac{1}{\sqrt{2}}(\Sigma^+ + \Sigma^-)$ $\Sigma_2 = \frac{i}{\sqrt{2}}(\Sigma^+ - \Sigma^-)$ $\Sigma_3 = \Sigma^0$

In writing the interaction terms, we will assume that all mesons have spin zero and all baryons have spin ½. This may

not be sufficiently general, and some particles may have higher spins. However, the theory can be easily adapted to include higher spins.

Table 5 lists all the meson-baryon interaction terms which:

1. are invariant under the transformations leading to conservation of isocharge and hypercharge;
2. are invariant under Lorentz transformations;
3. conserve baryon number;
4. are invariant under charge conjugation; and
5. are linear in meson fields and bilinear in baryon fields.

Table 5

Meson	Interaction Terms
π	$:\Big[i g_1 \bar{N} \gamma_5 \tau_i N \pi_i + i g_2 \bar{\Xi} \gamma_5 \tau_i \Xi \pi_i + g_3 \epsilon_{ijk} \bar{\Sigma}_j \gamma_5 \Sigma_k \pi_i + g_4 (\bar{\Lambda} n \Sigma_i + \bar{\Sigma}_i n \Lambda) \pi_i \Big]:$
K	$:\Big[g_5 (\bar{N} n \Lambda K + K^* \bar{\Lambda} n N) + g_6 (\bar{\Xi} n \Lambda K' + K'^* \bar{\Lambda} n \Xi) + g_7 (\bar{N} n \tau_i \Sigma_i K + K^* \bar{\Sigma}_i n \tau_i N) + g_8 (\bar{\Xi} n \tau_i \Sigma_i K' + K'^* \bar{\Sigma}_i n \tau_i \Xi) + g_9 (\bar{N} n x^+ K' + K'^* \bar{x}^+ n N) + g_{10} (\bar{\Xi} n Y^- K + K^* \bar{Y}^- n \Xi) \Big]:$
ρ°	$:\Big[g_{11} \bar{N} n N \rho^\circ + g_{12} \bar{\Xi} n \Xi \rho^\circ + g_{13} \bar{\Lambda} n \Lambda \rho^\circ + g_{14} \bar{\Sigma}_i n \Sigma_i \rho^\circ + g_{15} \bar{x}^+ n x^+ \rho^\circ + g_{16} \bar{Y}^- n Y^- \rho^\circ \Big]:$
σ	$:\Big[g_{17} (\bar{N} n \Xi \sigma^+ + \bar{\Xi} n N \sigma^-) + g_{18} (\bar{x}^+ n \Lambda \sigma^+ + \bar{\Lambda} n x^+ \sigma^-) + g_{19} (\bar{\Lambda} n Y^- \sigma^+ + \bar{Y}^- n \Lambda \sigma^-) \Big]:$

In Table 5, the symbol : : denotes the ordered product[7], and ϵ_{ijk} is the completely antisymmetric tensor with $\epsilon_{123} = 1$. Further, $\bar{N} = N^* \gamma_4$, etc., where an asterisk denotes the Hermitian conjugate.

The interaction terms also involve the operator η, which is chosen in such a way that the interaction is Hermitian as well as invariant under space inversion. It is well known that the behavior of a field operator under space inversion is given by

$$\psi' = P\psi \qquad \text{for a meson field,} \qquad (37)$$

and

$$\psi' = P\gamma_4\psi \qquad \text{for a baryon field,} \qquad (38)$$

where $P = \pm 1$ denotes the parity of the field operator. It follows that if P_1, P_2, and P_3 are the parities of the three field operators coupled to each other, then

$$\eta = 1 \qquad \text{if } P_1P_2P_3 = 1, \qquad (39)$$

and

$$\eta = i\gamma_5 \qquad \text{if } P_1P_2P_3 = -1. \qquad (40)$$

One might ask whether the parity of a particle is an observable quantity. The answer is that what we observe is the effect of parity on the interaction of particles, and therefore in general $P_1P_2P_3$ is an observable quantity. If in an interaction term it is required by theory that $P_1 = P_2$, then

$$P_1 P_2 P_3 = P_1^2 P_3 = P_3 ,$$

and thus P_3 becomes an observable quantity. However, if the theory does not require any definite relationship between P_1, P_2, and P_3, these parities separately are not observable quantities. The interaction terms for mesons and baryons show that the parity of pions is observable, while the parity of K mesons can be fixed only by convention.

VII. Experimental Evidence for the Mesons and Baryons

All the particles listed in Table 3 have been observed except ρ^0, σ^+, σ^-, X^+ and Y^-, and there is some evidence for the existence of these particles.

The ρ^0 has the following properties: It has zero isospin and is neutral. We would expect it to have zero spin, since the mesons which have been observed, the π and K, have zero spin. However, a higher spin for the ρ^0 is not definitely ruled out. The ρ^0 interacts strongly with every baryon and, thus, should be readily produced. Since the ρ^0 has not been observed, it should be extremely unstable and presumably decay strongly into pions. This, of course, is possible only if the ρ^0 has a mass greater than that of two pions.

Several resonances have recently been reported in pion final states, and one of these could be the ρ^0. Abashian, et al[8] have reported a resonance with the quantum numbers of a scalar, neutral meson. Maglić, et al[9] have reported a resonance with the quantum numbers of a vector, neutral meson. Moreover, Pevsner, et al[10] have reported a neutral pion resonance, which has been classified as pseudoscalar by several authors.[11] Besides these direct experimental results for the existence of the ρ^0, the nucleon-nucleon interaction gives some evidence of the ρ^0.[12]

The σ^+ and σ^- could correspond to the heavy mesons of mass about 1500 m_e and strangeness $\pm$ 2, which are needed for the interpretation of the observations of Fry, Schneps, and Swami[13], and others[14]. The Y^- could correspond to the heavy baryon of mass about 3200 m_e and strangeness - 3 observed by Eisenberg[15], and the relationship between the X^+ and Y^- baryons could be regarded as similar to that between the p and Ξ^- baryons, which ensures that an equal number of positively and negatively charged baryons exist in nature. The above experimental evidence refers to anomalous events observed during the study of cosmic-ray stars, and further experimental work in high-energy laboratories is needed to settle the question of the possible existence of these particles.

VIII. Quantitative Calculations

For strong interactions the coupling constant is greater than one. For instance, for pion-nucleon interactions,

$$g^2/4\pi\hbar c \approx 14 .$$

In perturbation theory an expansion is carried out in powers of the coupling constant, and therefore the higher order terms in the series should get progressively larger. One wonders if perturbation theory has any meaning at all for strong interactions.

It is interesting to observe that in spite of the large coupling constants, it is sometimes possible to obtain meaningful quantitative results by applying the perturbation theory. For instance, it is well known that in the case of nuclear forces the range of the force is roughly proportional to the inverse of the mass of the particles exchanged between the nucleons. Specifically, the range is:

$$R \approx \frac{\hbar}{mc} \tag{41}$$

Thus, if m_π is the pion mass, the nuclear force is approximately given by one pion exchange for distances of nucleon separation greater than about $\hbar/m_\pi c$, by one and two pion exchange for distances greater than about $\hbar/2m_\pi c$, and so on.

One pion exchange has been used by Moravcsik et al.[16], and one and two pion exchange has been used by Breit et al.[16] for proton-proton scattering up to 310 Mev in the laboratory system. Roughly speaking, their calculations show that, except for s and p waves, we can reasonably well account for all higher partial waves in proton-proton scattering by using one and two pion exchange interactions. It is, of course, natural to find that s and p waves cannot be treated in this way, because for such waves the interaction at very small distances cannot be ignored. A full discussion of the nucleon-nucleon interaction would be outside the scope of these lectures, and I have therefore confined myself to these brief remarks.

REFERENCES

1. M. Gell-Mann, Phys. Rev., 92, 833 (1953).

2. T. Nakano and K. Nishijima, Prog. Theor. Phys., 10, 581 (1953).

3. B. d'Espagnat et J. Prentki, Nuc. Phys., 1, 33 (1956).

4. J. Schwinger, Ann. Phys., 2, 407 (1957).

5. E. Cartan, Lecons sur la Théorie des Spineurs, Hermann, Paris, 1938.

6. For example, R. P. Feynman and G. P. Speisman, Phys. Rev., 94, 500 (1951), consider the n-p mass difference.

7. S. N. Gupta, Phys. Rev., 107, 1722 (1957).

8. A. Abashian, et al, Phys. Rev. Letters, 7, 35 (1961).

9. B. Maglić, et al, Phys. Rev. Letters, 7, 178 (1961); and Phys. Rev., 125, 687 (1962).

10. A. Pevsner, et al, Phys. Rev. Letters, 7, 421 (1961).

11. L. Brown and P. Singer, Phys. Rev. Letters, 8, 155 and 460 (1962); M. Gell-Mann, D. Sharp, and W. Wagner, Phys. Rev. Letters, 8, 261 (1962); A. Rosenfeld, et al, Phys. Rev. Letters, 8, 117 and 293 (1962); and G. Shaw and D. Y. Wong, Phys. Rev. Letters 8, 336 (1962).

12. S. N. Gupta, Phys. Rev., 111, 1436 and 1698 (1958); and Phys. Rev. Letters, 2, 124 (1959). G. Breit, Proc. Natl. Acad. Sci. 46, 746 (1960); Phys. Rev. 120, 287 (1960).

13. W. Fry, J. Schneps, and M. Swami, Phys. Rev., 97, 1189 (1955); and Nuovo Cimento, 2, 346 (1958).

14. T. Yamanuchi and M. F. Kaplon, Phys. Rev. Letters, 3, 283 (1959); and T. Yamanuchi, Phys. Rev. Letters, 3, 480 (1959).

15. Y. Eisenberg, Phys. Rev., 96, 541 (1954); and Nuovo Cimento, Suppl. 4, 484 (1956).

16. H. P. Noyes, Proceedings of the Rutherford Jubilee International Conference, 1961, page 65, Heywood, London, 1961; and G. Breit, ibid, pages 17 and 159.