Theory Division
21 May, 1962

SELECTED TOPICS ON SCATTERING THEORY

PART I

RELATIVISTIC KINEMATICS AND PRECESSION OF POLARIZATION

Lectures given by :
R. Hagedorn

G E N E V A

45/TH.167

SELECTED TOPICS ON SCATTERING THEORY

PART I

RELATIVISTIC KINEMATICS AND PRECESSION OF POLARIZATION

TABLE OF CONTENTS

SELECTED TOPICS ON SCATTERING THEORY

PART I

RELATIVISTIC KINEMATICS AND PRECESSION OF POLARIZATION

LECTURE 1

1) Lorentz transformations and invariants

(a) The invariant line element

It is one of the most important facts of physics that the velocity of light, $C = 2.99776 \ldots \times 10^{10}$ cm/sec, is the same in all inertial systems. This has the consequence that if light is supposed to be the fastest means of communication, all measurements involving distances must be influenced by this fact. Indeed, this influence is expressed by the Lorentz transformations.

Let K and K' be two reference systems moving with constant velocity with respect to each other. We call an "event" or a world point the set

$$P \equiv \left\{ xyzt \right\} \tag{I.1}$$

of space-time co-ordinates.

Let the directions of the axes of K and K' be parallel and such that the x and x' axes coincide and are parallel to the relative velocity [see Fig. I.1].

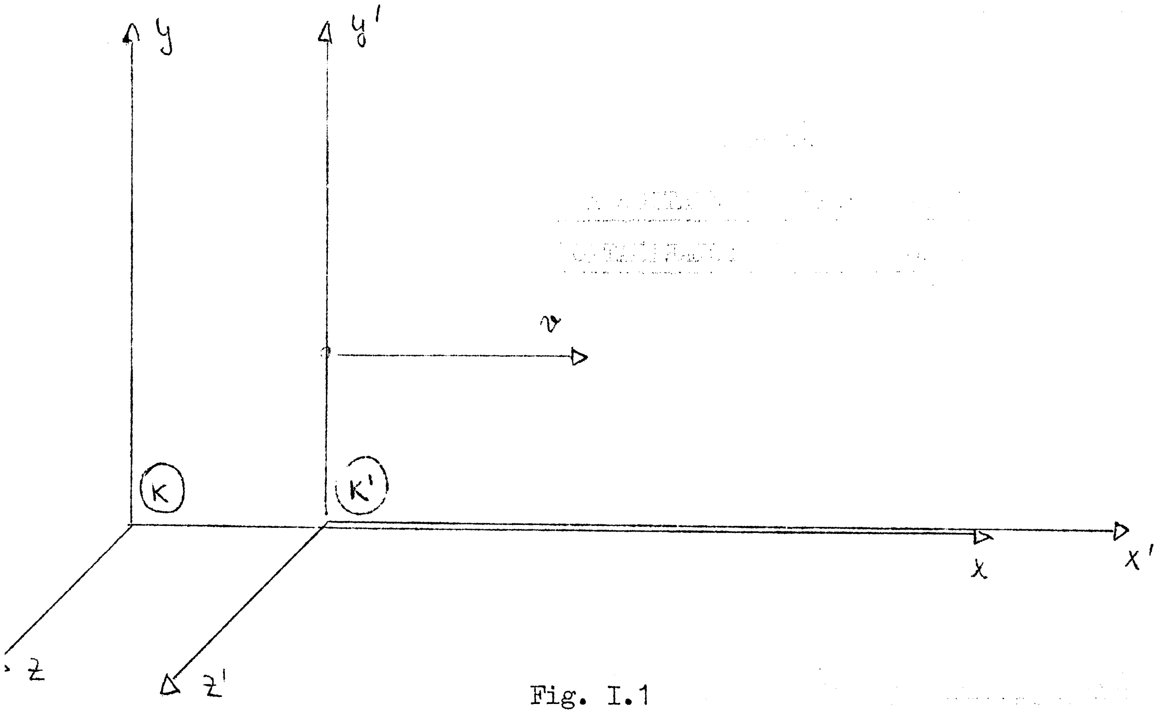


Fig. I.1

Consider two particular events P_1 and P_2 in the frame K, where

$P_1 \equiv (x_1 y_1 z_1 t_1)$ is sending out a light signal at time t_1 from the space point $x_1 y_1 z_1$ and

$P_2 \equiv (x_2 y_2 z_2 t_2)$ is receiving this signal at time t_2 in the space point $x_2 y_2 z_2$.

The distance between the space points is

$$d = \sqrt{(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2}$$

but since C is a fixed constant for all inertial systems, it is also given by

$$d = C(t_2-t_1) ,$$

hence

$$(x_2-x_1)^2 + (y_2-y_1)^2 + (z_2-z_1)^2 = C^2(t_2-t_1)^2 \qquad (I.2)$$

But we could have written down the same in system K' :

$$(x'_2-x'_1)^2 + (y'_2-y'_1)^2 + (z'_2-z'_1)^2 = c^2(t'_2-t'_1)^2 \qquad (I.2')$$

with the same constant C.

We now introduce $\tau \equiv ict$ and go over to infinitesimal distances. We call the square of the distance between any two events which are very near to each other

$$ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2 = -(dx^2 + dy^2 + dz^2 + d\tau^2)$$

and conclude from (I.2) and (I.2') that

$$ds = 0 \qquad \text{implies} \qquad ds' = 0 \qquad (I.3)$$

As ds and ds' are of the same order, it follows that $ds = a\,ds'$ and since K and K' are on an equal footing, $ds' = a\,ds$, hence $a = \pm 1$ and only $a = +1$ remains for reasons of continuity. By integrating between any two events we see that this is an invariant quantity with respect to the co-ordinate transformations from K to K'.

$$\underbrace{\int_{P_1}^{P_2} ds = \sqrt{-(\Delta x^2 + \Delta y^2 + \Delta z^2 + \Delta\tau^2)}}_{K\text{ - system}} = \underbrace{\int_{P'_1}^{P'_2} ds' = \sqrt{-(\Delta x'^2 + \Delta y'^2 + \Delta z'^2 + \Delta\tau'^2)}}_{K'\text{ - system}} \qquad (I.4)$$

The word invariant denotes one of the central ideas of the theory of special relativity and invariants are a very convenient tool to do calculations. We call $ds^2 = -dx^2-dy^2-dz^2 + c^2dt^2$ the invariant line element.

(b) Space-like and time-like distances; future and past; the light cone

We now ask two questions.

Given two events P_1 and P_2 with distance $\{\Delta x, \Delta y, \Delta z, \Delta\tau\}$, is there a system in which these two events appear at the same time ?

We have then $\Delta\tau' = 0$, but from (I.4) it follows that

$$\Delta s^2 = -(\Delta x^2 + \Delta y^2 + \Delta z^2 + \Delta\tau^2) = -(\Delta x'^2 + \Delta y'^2 + \Delta z'^2) \leq 0$$

thus

a system in which two events happen at the same time can be found if and only if $\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \leq 0$ [space-like distance]	(I.5)

On the other hand, we ask : can two events appear to happen at the same place in some system K' ?

We have then

$$\Delta s^2 = -(\Delta x^2 + \Delta y^2 + \Delta z^2 + \Delta\tau^2) = -\Delta\tau'^2 = +c^2\Delta t'^2 \geq 0$$

Therefore

a system in which two events happen at the same place can be found if and only if $\Delta s^2 = c^2\Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2 \geq 0$ [time-like distance]	(I.6)

Since Δs^2 is invariant, either the one or the other happens, no matter what the system is. We have to add the case $\Delta s^2 = 0$ which applies to the distance between two events connected by a light signal [see (I.2) and (I.2')].

Consider now all possible events with respect to a given one. Put the co-ordinate origin into the given event 0 : (x=y=z=t=0) and draw two co-ordinates (x,t) only [Fig. I.2]

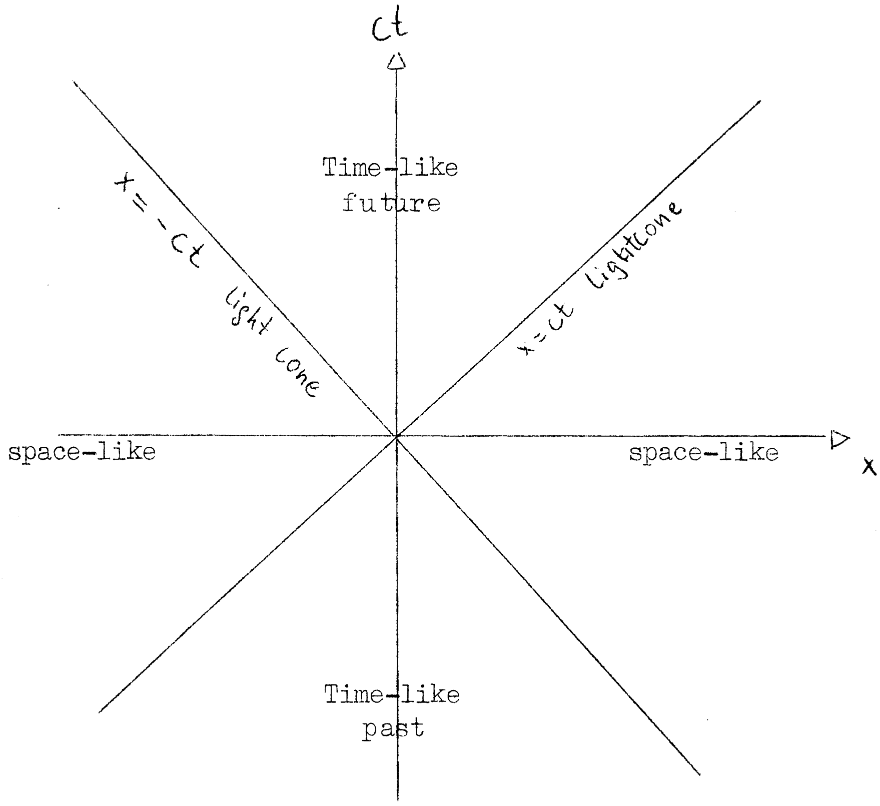


Fig. I.2

Time-like and space-like distances,
past and future, the light cone.

The distance from the origin is given by the invariant

$$s^2 = c^2t^2 - x^2 - y^2 - z^2$$

(i) $s^2 = 0$ connects all those events with the origin which can be reached by a light signal, hence the cone $s^2 = 0$ is called the "light cone" and these events are said to be o n the light cone.

(ii) $s^2 > 0$ if $s > 0$, the event is i n the forward light cone if $s < 0$, it is i n the backward light cone. Clearly an event in the forward light cone is later as 0, in the backward light cone it is earlier as 0.

Since s^2 is an invariant, one cannot transform an event with $s < 0$ into one with $s > 0$, since all these transformations form a connected continuous group.

Therefore the forward light cone contains the absolute future, and the backward cone the absolute past. Only events in the backward cone can have an influence on 0 and 0 can have an influence only on events in the forward cone. The space-like events cannot interact with 0.

This is how causality is expressed on this level of argument.

/Remark : it finds its most direct application in quantum field theory where the measurement of the field at the origin and at a world point $P = \{xyzt\}$ do not interfere, if P lies space-like. Therefore the two operators corresponding to the fields must commute :

$$\left[A(P_1), A(P_2) \right] = 0 \quad \text{if } P_1 \text{ and } P_2 \text{ lie space-like to each other.} \tag{I.7}$$

The dispersion relations are derived from this requirement./

Problem :

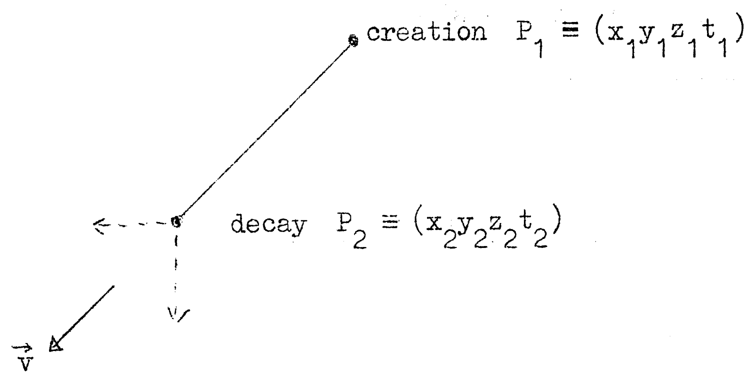
1) Define the *p r o p e r t i m e* of a moving body to be the time shown by a clock which moves with that body. Use the invariance of Δs^2 to establish the lifetime of a particle measured by a clock in the lab. system.

   a) if the particle moves with constant velocity

   b) if the particle moves arbitrarily.

2) Is the light quantum a stable particle ?

Solution 1)



In the lab. system we find the distance between creation and decay

$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 - \Delta y^2 - \Delta z^2$ but in the particles system

$\Delta s^2 = c^2 \Delta t'^2$ since there it is at rest

hence

$$\Delta t' = \Delta t \sqrt{1 - \frac{\Delta x^2 + \Delta y^2 + \Delta z^2}{c^2 \Delta t^2}} = \Delta t \sqrt{1 - \beta^2} \; ; \quad \beta \equiv \frac{v}{c}$$

a) $t_2' - t_1' = (t_2 - t_1) \sqrt{1 - \beta^2}$

b) $t_2' - t_1' = \int_{t_1}^{t_2} dt \sqrt{1 - \beta(t)^2}$

2658

Solution 2)

Suppose the γ-quantum were unstable, then a lifetime should be definable and the only invariant way to define a lifetime is to measure it in the rest system of the particle (i.e., by its proper time). Such a rest system does not exist by supposition (constancy of c). Even formally

$$\text{lifetime} = t'_2 - t'_1 = \underbrace{(t_2 - t_1)}_{\substack{\text{observed lifetime} \\ \text{in the lab.} = \infty}} \cdot \underbrace{\sqrt{1-\beta^2}}_{=0}$$

Thus $t'_2 - t'_1 = 0 \cdot \infty$ = undetermined.

Therefore the question is senseless [for all particles with $m = 0$!]

(c) The Lorentz transformation

We now derive the transformation formula from K to K' where K' moves with constant velocity v as indicated in Fig. I.1. The invariance of Δs^2 requires [$\tau = ict$]

$$\Delta s^2 = -(\Delta x^2 + \Delta y^2 + \Delta z^2 + \Delta \tau^2) = -(\Delta x'^2 + \Delta y'^2 + \Delta z'^2 + \Delta \tau'^2) .$$

If we exclude translations in the $xyz\tau$-space, the only transformations leaving Δs^2 invariant are rotations. We are not interested in space rotations and in fact Fig. I.1 singles out the rotation in the $x\tau$-plane, since only x and the time are involved, y and z staying constant.

Let α be the angle of rotation. The transformation must be of the form

$$\begin{aligned} x &= x' \cos\alpha - \tau' \sin\alpha \\ \tau &= x' \sin\alpha + \tau' \cos\alpha \end{aligned} \qquad [y = y' \text{ and } z = z' \text{ will not be mentioned}]$$

(I.8)

We now determine α by considering an example : we are in K and observe the origin of K' moving with velocity v in our positive x-direction. Its motion in our system is described by x and τ , in the system K' by $x' = 0$ and τ'. Hence

$$x = -\tau' \sin\alpha$$

$$\tau = \tau' \cos\alpha$$

$$\frac{x}{\tau} = \frac{v}{ic} = -\operatorname{tg}\alpha \equiv -i\beta$$

Now

$$\cos\alpha = \frac{1}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{1}{\sqrt{1-\beta^2}} \equiv \gamma$$

$$\sin\alpha = \frac{\operatorname{tg}\alpha}{\sqrt{1+\operatorname{tg}^2\alpha}} = \frac{i\beta}{\sqrt{1-\beta^2}} = i\beta\gamma$$

hence

$$x = x'\gamma - i\cdot ict'\beta\gamma = \gamma(x'+\beta ct')$$

$$ct = -ix'i\beta\gamma - i\cdot ict'\gamma = \gamma(ct'+\beta x') \qquad (I.9)$$

Lorentz transformation

Compare Fig. I.1, which gives the meaning of this transformation.

Problem :

3) Use (I.9) and the fact that any vector can be decomposed into a component in a given direction and another component perpendicular to it, in order to establish the most general form of the Lorentz transformation between two inertial systems K and K' where K' moves with

$$\vec{\beta} = \frac{\vec{v}}{c}$$

with respect to K.

4) Discuss this general transformation

a) specialize it to retain (I.9);

b) solve it for the primed co-ordinates and verify the solution;

c) go to the non-relativistic limit;

d) derive the Lorentz contraction and time dilatation from the general formula.

Solution 3)

We decompose x into a component parallel to $\vec{\beta}$ and one perpendicular to $\vec{\beta}$:

$$\vec{x}' = \vec{x}'_{\parallel} + \vec{x}'_{\perp} = \underbrace{\vec{\beta}\cdot \frac{\vec{\beta}\cdot\vec{x}'}{\beta^2}}_{\vec{x}'_{\parallel}} + \underbrace{\left(\vec{x}' - \vec{\beta}\,\frac{\vec{\beta}\vec{x}'}{\beta^2}\right)}_{\vec{x}'_{\perp}}$$

To $\vec{x}'_{\parallel}$ we apply (I.9) whereas $\vec{x}'_{\perp}$ remains untransformed

$$\vec{x}_{\parallel} = \gamma(\vec{x}'_{\parallel} + c\vec{\beta}t')$$

$$ct = \gamma(ct' + \vec{\beta}\vec{x}'_{\parallel}) = \gamma(ct' + \vec{\beta}\vec{x}')$$

$$\vec{x}_{\perp} = \vec{x}'_{\perp}$$

hence

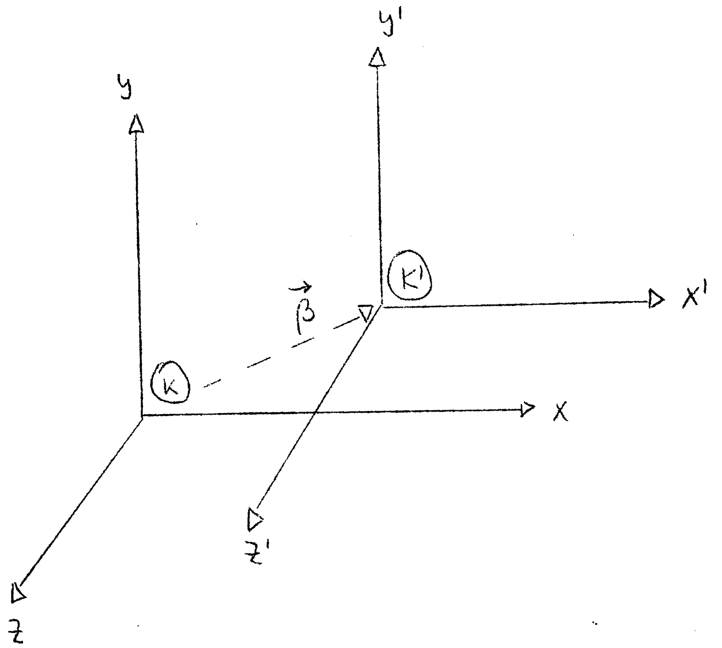
$$\vec{x} = \vec{x}_{\parallel} + \vec{x}_{\perp} = \gamma\left(\vec{\beta}\frac{\vec{\beta}\vec{x}'}{\beta^2} + c\,\vec{\beta}t'\right) + \vec{x}' - \vec{\beta}\frac{\vec{\beta}\vec{x}'}{\beta^2}$$

$$\vec{x} = \vec{x}' + \vec{\beta}\left[\vec{\beta}\vec{x}'\,\frac{\gamma-1}{\beta^2} + \gamma ct'\right]$$

finally with

$$\beta^2 = \frac{\gamma^2-1}{\gamma^2}$$

$$\boxed{\begin{aligned} \vec{x} &= \vec{x}' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\,\vec{\beta}\vec{x}' + ct'\right] \\ ct &= \gamma\left[ct' + \vec{\beta}\vec{x}'\right] \end{aligned}} \qquad \text{(I.10)}$$



Solution 4)

a) $\vec{\beta} \parallel x$ gives

$$y = y'$$

$$z = z'$$

$$ct = \gamma(ct' + \beta x')$$

$$x = x' + \beta\gamma\left(\frac{\gamma}{\gamma+1}\beta x' + ct'\right)$$

$$x'\left(1 + \frac{\beta^2\gamma^2}{\gamma+1}\right) = \gamma x' \; ; \quad (\beta^2\gamma^2 = \gamma^2 - 1 = (\gamma+1)(\gamma-1))$$

b) The K system moves with $-\vec{\beta}$ as seen from K'. Hence we must interchange the primed co-ordinates with the unprimed ones and reverse the sign of β :

$$\vec{x}' = \vec{x} + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\vec{\beta}\vec{x} - ct\right]$$

$$ct' = \gamma\left[ct - \vec{\beta}\vec{x}\right]$$

Verification :

$$\vec{x}' = \vec{x}' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\vec{\beta}\vec{x}' + ct'\right] +$$

$$+\vec{\beta}\gamma\left\{\frac{\gamma}{\gamma+1}\vec{\beta}\left[\vec{x}' + \vec{\beta}\gamma\left(\frac{\gamma}{\gamma+1}\vec{\beta}\vec{x}' + ct'\right)\right] - (ct' + \vec{\beta}\vec{x}')\right\}$$

after some rearrangement :

$$= \vec{x}' + \vec{\beta}\vec{x}'\underbrace{\left[2\vec{\beta}\frac{\gamma^2}{\gamma+1} - \vec{\beta}\gamma^2 + \vec{\beta}\frac{\beta^2\gamma^4}{(\gamma+1)^2}\right]}_{\vec{\beta}\gamma^2\underbrace{\left[2(\gamma+1) - (\gamma+1)^2 + \underbrace{\beta^2\gamma^2}_{\gamma^2-1}\right]}_{2\gamma+2-\gamma^2-2\gamma-1+\gamma^2-1 \,=\, 0}}$$

$$+ ct^2\underbrace{\left[1 + \underbrace{\frac{\beta^2\gamma^2}{\gamma+1}}_{\gamma-1} - \gamma\right]}_{=\,0}\vec{\beta}\gamma \, . \qquad \text{q.e.d.}$$

c) the non-relativistic limit is ($\gamma = 1$; $\beta^2 = 0$)

$$\vec{x} = \vec{x}' + \vec{v}t' \qquad \vec{v} = \vec{\beta} c$$

$$t = t' + \frac{\vec{v}\vec{x}'}{c^2}$$

which is not yet the Galilei transformation $(t = t')$.

d)
$$\vec{dx} = \vec{dx}' + \vec{\beta}\gamma \left[\frac{\gamma}{\gamma+1} \vec{\beta}\vec{dx}' + cdt' \right]$$

$$cdt = \gamma \left[cdt' + \vec{\beta}\vec{dx}' \right]$$

If we wish to know how in K appears a length element $\vec{dx}'$ which is at rest in K', we must measure the position of its end points in K simultaneously, hence for this case

$$cdt = 0 \quad \text{or} \quad cdt' = -\vec{\beta}\vec{dx}'$$

Inserted into the first equation

$$\vec{dx}\Big|_{dt=0} = \vec{dx}' + \vec{\beta}\gamma \left[\frac{\gamma}{\gamma+1} - 1 \right] \vec{\beta}\vec{dx}'$$

$$\underbrace{\vec{dx}\Big|_{dt=0} = \vec{dx}' - \vec{\beta} \cdot \frac{\gamma}{\gamma+1} \cdot (\vec{\beta} \cdot \vec{dx}')}_{\text{generalised Lorentz contraction}} ;$$

if $\vec{\beta} \parallel \vec{dx}'$ we find $dx'\left(1 - \frac{\beta^2\gamma}{\gamma+1}\right) = \frac{dx'}{\gamma}$.

If we wish to know how a clock at rest in the K' system appears in the K-system, we must put $\vec{dx}' = 0$, hence

$$\vec{dx} = \vec{\beta}\gamma c dt'$$

$$c dt = \gamma c dt' \dashrightarrow dt = \gamma dt' \quad : \text{ time dilatation.}$$

Introducing this in the first equation gives

$$\vec{dx} = \vec{\beta} c dt = \vec{v} dt$$

which only says that the clock moves in K with velocity v.

(d) <u>The transformation of velocities</u>

With formula (I.10) from problem 3 we obtain

$$\vec{dx} = \vec{dx}' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\,\vec{\beta}\vec{dx}' + c dt'\right]$$

$$c dt = \gamma\left[c dt' + \vec{\beta}\vec{dx}'\right]$$

and by dividing

$$\frac{1}{c}\frac{\vec{dx}}{dt} = \frac{\vec{v}}{c} = \frac{\vec{dx}' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}(\vec{\beta}\vec{dx}) + c dt'\right]}{\gamma\left[c dt' + (\vec{\beta}\vec{dx}')\right]}$$

or

$$\vec{v} = \frac{\vec{v}' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}(\vec{\beta}\vec{v}') + c\right]}{\gamma\left[1 + \frac{(\vec{\beta}\vec{v}')}{c}\right]} = \frac{\vec{v}' + \vec{V}\gamma\left[\frac{\gamma}{\gamma+1}\frac{(\vec{V}\vec{v}')}{c^2} + 1\right]}{\gamma\left[1 + \frac{(\vec{V}\vec{v}')}{c^2}\right]} \qquad \text{(I.11)}$$

where $\vec{v}$ is the velocity of a point (e.g., a particle) in system K (e.g., the laboratory system), $\vec{v}'$ its velocity in K' (e.g., the centre-of-mass system of a reaction) and $\vec{V} = c\vec{\beta}$ is the velocity of the system K' as seen from K (the velocity of the CM-system seen from the lab. system). One sees that the velocities add in a very complicated way. Assume them to be parallel, then no change in the direction takes place and we have

$$v = \frac{v' + \frac{\beta^2\gamma}{\gamma+1} v' + V\gamma}{\gamma\left(1 + \frac{Vv'}{c^2}\right)}$$

$$\frac{\beta^2\gamma^2}{\gamma+1} = \gamma - 1$$

gives

$$v = \frac{V + v'}{1 + \frac{Vv'}{c^2}}$$

thus

$$\frac{v}{c} = \frac{\frac{V}{c} + \frac{v'}{c}}{1 + \frac{V}{c} \cdot \frac{v'}{c}}$$

$v = 0$ if $V = -v'$ and $v \leq c$ always.

Choose the co-ordinates such that $\vec{\beta}$ points in the positive x-direction and that v' lies in the $x'y'$ and xy planes.

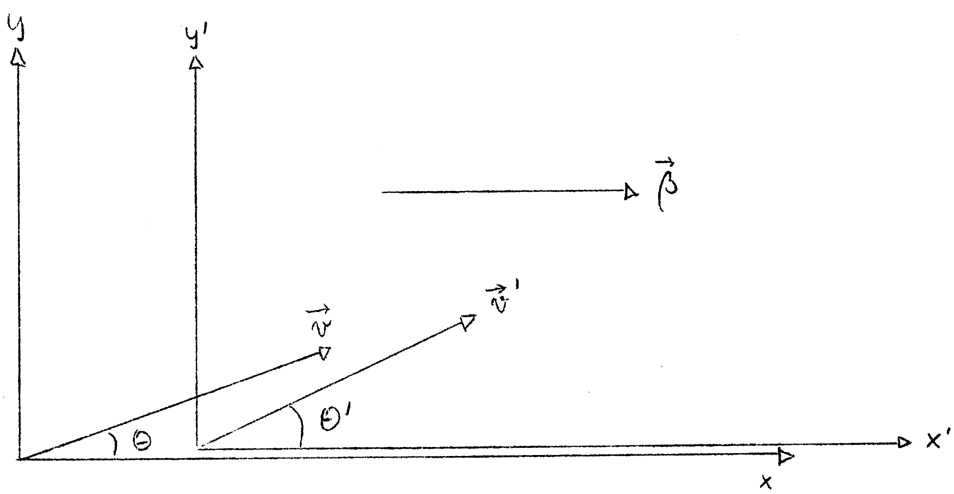


Fig. I.3

Transformation of angles

Then (I.11) reads

$$v_x = v\cos\Theta = \frac{v'\cos\Theta' + \frac{\beta^2\gamma}{\gamma+1}v'\cos\Theta' + \beta\gamma c}{\gamma\left[1 + \frac{\beta}{c}v'\cos\Theta'\right]} = \frac{\gamma\left[v'\cos\Theta' + \beta c\right]}{\gamma\left[1 + \frac{\beta}{c}v'\cos\Theta'\right]}$$

$$v_y = v\sin\Theta = \frac{v'\sin\Theta'}{\gamma\left[1 + \frac{\beta}{c}v'\cos\Theta'\right]}$$

$$\boxed{\mathrm{tg}\Theta = \frac{v'\sin\Theta'}{\gamma\left[v'\cos\Theta' + \beta c\right]}} \quad , \qquad (I.12)$$

which gives the transformation of the angles of a velocity.

LECTURE 2

(e) Four-vectors and invariants

All quantities consisting of a set of 4 numbers which transform under a Lorentz transformation exactly as the components of

$$ds = \left\{ cdt,\ dx,\ dy,\ dz \right\}$$

according to (I.10), are called four vectors.

We required that the Lorentz transformation should leave the line element

$$ds^2 = -dx^2 - dy^2 - dz^2 + c^2dt^2$$

invariant. Consequently, if four vectors transform like ds, then the scalar product of the four vector P with itself

$$P \equiv \left\{ P_t, P_x, P_y, P_z \right\}$$

$$P^2 \equiv -P_x^2 - P_y^2 - P_z^2 + P_t^2$$

is an invariant and then if P and Q are four vectors :

$$(P+Q)^2 = P^2 + 2PQ + Q^2 = P'^2 + 2P'Q' + Q'^2 \quad ,$$

hence

$$PQ = -p_x q_x - p_y q_y - p_z q_z + p_t q_t = \text{invariant} \tag{I.13}$$

Some of the most important four vectors are :

-- the four-dimensional radius vector

$$X = \left\{ ct, \vec{x} \right\} \quad ,$$

-- the energy-momentum vector (four-momentum)

$$P = \left\{ \frac{E}{c}, \vec{p} \right\} \quad ; \quad P^2 = m^2 c^2 \quad \text{(time-like)}$$

-- the four velocity $\left[\vec{\beta} = \frac{\vec{v}}{c} \; ; \; \gamma = (1-\beta^2)^{-\frac{1}{2}} \right]$

$$V = c \, \frac{dX}{ds} = \frac{dX}{d\tau} = \left\{ c\gamma, \; c\gamma\vec{\beta} \right\} \quad ; \quad V^2 = c^2 \; \text{(time-like)}; \; \tau = \text{proper time}$$

-- the current vector

$$J = \left\{ \varrho c, \varrho\vec{v} \right\} \quad ; \quad J^2 = \frac{\varrho^2 c^2}{\gamma^2} = \varrho_o^2 c^2 \quad \text{(time-like)}$$

(ϱ_o = density in the rest system of the volume element considered).

Some words about transformation properties seem to be necessary since it is relatively easy to commit errors.

The transformation properties of scalars = invariants, vectors, tensors, etc., have their origin in the geometric nature of these objects. In addition to that these objects may be functions of the point to which they are attached : they may be fields.

Let us consider a constant scalar field Θ (e.g., a temperature distribution). It is constant, i.e., the same throughout the whole space and therefore the same in every co-ordinate system.

$$\Theta' = \Theta$$

If we have a constant vector field, $\vec{v}$, then a rotation of the co-ordinate system will change the components v_i (though the object $\vec{v}$ itself remains the same) :

$$v_i' = \alpha_{ik} v_k$$

(we always sum over double subscripts) but the scalar product of two such vectors $(\vec{v}\vec{w}) = \Theta$ is invariant again :

$$v_i' w_i' = \alpha_{ik} \alpha_{ij} v_k w_j = \delta_{kj} v_k w_i = v_k w_k$$

is geometrically obvious. Similar considerations apply to a constant tensor field. Generally, we may say that for any constant field the equations

$$\begin{aligned} F' &= S(T) \cdot F \\ x' &= T \cdot x \end{aligned} \qquad \text{(I.14)}$$

($x \equiv \{x_1, x_2, \ldots, x_n\}$ is a n-dimensional radius vector)

express that under a co-ordinate transformation T the constant field F transforms with a matrix $S(T)$. The matrix $S(T)$ may contain one single element $S(T) = 1$ if F is a scalar, 4 elements if F is a spinor, 9 if F is a vector, etc. [These different kinds of S are called the one-, two-, three- etc., dimensional representation of the transformation group T.]

Examples worth studying are the matrices occuring in the Dirac theory. [See, e.g., Jauch and Rohrlich, p.425, or J. Schweber, p.70].

But in general the situation is more complicated in so far as F will be a function of the co-ordinates and not be constant all over the space.

As we already see in the simple case of the scalar, one c a n n o t say that there is no change under a co-ordinate transformation, since

$$\Theta(x) \neq \Theta(x')$$

in general.

2658

Invariance means in fact here that there is a function $\Theta'(x)$ defined in such a way that

$$\Theta'(x') = \Theta(x)$$

It is not easy to explain the meaning of this in words because such an explanation grows so long that it makes the intuitively obvious thing unobvious. The reader is urged to make this equation clear to himself by discussing the simple example where $\Theta(x)$ is, e.g., a temperature distribution in space and T is a simple rotation.

We may use $x' = Tx$ or $x = T^{-1}x'$ to rewrite the equation

$\Theta'(x') = \Theta(T^{-1}x')$ where now x' appears as variable on both sides and may be called x again :

$\Theta'(x) = \Theta(T^{-1}x)$ is then an equivalent definition of Θ'.

Thus the equation

$$\Theta'(Tx) = \Theta(x) \qquad \text{or} \qquad \Theta'(x) = \Theta(T^{-1}x)$$

expresses what is really meant if one says that $\Theta(x)$ is a scalar function. We turn to the general case :

let $\mathfrak{F}(P)$ be a quantity (scalar, spinor, vector, tensor) attached to the point P in our n-space. $\mathfrak{F}$ is defined as a physical quantity, not as a set of numbers [e.g., $\mathfrak{F}(P)$ may be an electric field defined in such an abstract way *): it does not refer to a particular system of co-ordinates but will, if co-ordinates are specified, have a representation (components) which depends on the system chosen. In quantum mechanics one has analogously the abstract operators and their representatives in the systems chosen : the matrices. One must clearly distinguish these two things.]

*) E.g., by an actual experimental set up (distribution of material loaded conductors in space).

If we now introduce two different co-ordinate systems K and K' then we can represent $\mathfrak{F}$ (P) as functions of the co-ordinates x and x' respectively

$$P \equiv \{x_1 \ldots x_n\} \qquad \text{in} \quad K$$

$$P \equiv \{x'_1 \ldots x'_n\} \qquad \text{in} \quad K'$$

and

$$\mathfrak{F}(P) \equiv F(x) \qquad \text{in} \quad K$$

$$\mathfrak{F}(P) \equiv F'(x') \qquad \text{in} \quad K'$$

F' will of course be different from F, if it will describe the same physical situation at the point P.

Now we are for the moment only interested in what happens at P and we may then for a moment replace the field

$\mathfrak{F}(P) \rightarrow \mathfrak{F}_o$, namely by that constant field $\mathfrak{F}_o$, which has the value $\mathfrak{F}_o = \mathfrak{F}(P)$ everywhere. It will appear as the constant field $F(x) \equiv F_o$ in K, and as another constant field $F'(x') \equiv F'_o$ in K'. We know, however, how this constant field appears, if seen from different systems K and K', namely

$F'_o = S(T)F_o$, where S(T) is a representation matrix of the transformation T.

This is true for a constant field $\mathfrak{F}_o$, which has everywhere the value which $\mathfrak{F}(P)$ had at the particular point P.

But in the whole argument no other point than P was used and we may well write the arguments x' and x in the last equation : as far as the point P is concerned, we have

$$\mathfrak{F}(P) \equiv F(x) \quad \text{in} \quad K \quad \ldots$$

$$\mathfrak{F}(P) \equiv S(T)F(x) \quad \text{in} \quad K' \quad \ldots$$

but also,

$$\mathfrak{F}(P) \equiv F'(x') \quad \text{in} \quad K'$$

and the two descriptions in K' must be equal, hence

$$\left.\begin{aligned} &F'(x') = S(T)F(x) = S(T)F(T^{-1}x') \quad \text{or} \\ &F(x) = S^{-1}(T)F'(x') = S^{-1}(T)F'(Tx) = S(T^{-1})F'(Tx) \quad \text{and} \\ &x' = Tx \end{aligned}\right\} \quad (I.15)$$

give the full description of the transformation properties of the quantity F. Equation (I.15) reduces to (I.14) for a scalar or invariant function.

From these considerations it should be clear that an "invariant function" is in general not a constant function.

In this discussion we choose as an example for the scalar field a temperature distribution and not a mass distribution. A mass distribution, more generally a density distribution as such is not invariant. It becomes invariant only after multiplication with a volume element and the whole

$$dm = \rho(x_1 \ldots x_n)\, dx_1 \ldots dx_n$$

is an invariant quantity. The volume element has, however, its own particular transformation properties : it multiplies with a Jacobian determinant. We shall discuss this in paragraph 4, where the transformation of cross-sections is treated. Our example of a temperature distribution avoids these additional complications (why ?).

LECTURE 3

2) Choice of a system of units

We have

$$\left.\begin{aligned} x^2 &= -x_1^2 - x_2^2 - x_3^2 + c^2 t^2 \\ P^2 &= m^2 c^2 \end{aligned}\right\} \qquad (I.16)$$

In these, and in many other formulae, the velocity of light c appears explicitly. It seems therefore convenient to introduce such units that c has the numerical value 1.

On the other hand, in elementary particle physics, one has to do with quantum mechanics at the same time and the de Broglie relation between the four momentum of a particle and its wave vector is obtained from Einstein's equation

$$E = \hbar\omega$$

by extending it to become a relation between four vectors :

$$P = \hbar K$$

$$P = \left\{\frac{E}{c}, \vec{p}\right\} \; ; \quad K = \left\{\frac{\omega}{c}, \vec{k}\right\} \qquad (I.17)$$

by which the wave vector $\vec{k}$ is defined.

This suggests to choose such units that Planck's constant, $\hbar$, has the numerical value 1. We shall achieve both simultaneously and shall adopt from now on $\hbar = c = 1$. This does not yet fix our system of units completely. Let us briefly look into this.

We have to choose three basic units, namely for mass, length, time. Let us choose the proton mass M to be the mass unit.

If

(M), $(\hbar)$, (c) denote the dimensionless numbers indicating the numerical value of these constants in a given system of units, then the elementary-particle-system of units is defined by :

$$(M)_o = (\hbar)_o = (c)_o = 1.$$

The corresponding unit mass, unit length and unit time shall be denoted by m_o, l_o, t_o. Then

$$M = 1 \cdot m_o = (M)_{cgs} \cdot g$$

$$\hbar = 1 \cdot \frac{m_o l_o^2}{t_o} = (\hbar)_{cgs} \cdot \frac{gcm^2}{sec}$$

$$c = 1 \cdot \frac{l_o}{t_o} = (c)_{cgs} \cdot \frac{cm}{sec}$$

We find by solving for m_o, l_o, t_o :

$$m_o = M = (M)_{cgs}\, g \quad = \text{proton mass} = 1.672 \cdot 10^{-24}\, g$$

$$l_o = \frac{\hbar}{MC} = \left(\frac{\hbar}{MC}\right)_{cgs} cm \quad = \text{proton Compton wave length} = 0.211 \cdot 10^{-13}\, cm$$

$$t_o = \frac{\hbar}{MC^2} = \left(\frac{\hbar}{MC^2}\right)_{cgs} sec =$$ the time in which light travels one proton Compton wavelength $= 0.07 \cdot 10^{-23}$ sec

This has the consequence that now mass, length and time are numerically measured in multiples of the proton mass :

let μ be the mass of a particle, then

$$\mu = (\mu)\cdot M = (\mu)\cdot m_c$$

The Compton wavelength of this particle is then

$$\bar{\lambda}_\mu = \frac{\hbar}{\mu c} = \frac{\hbar}{(\mu)MC} = \frac{1}{(\mu)}\cdot l_o$$

and therefore the numerical value of the Compton wavelength becomes

$$(\bar{\lambda}_\mu) = \frac{1}{(\mu)}\,.$$

Similarly, a time is attached to it :

$$t_\mu = \frac{\hbar}{(\mu)MC^2} = \frac{1}{(\mu)}\cdot t_o$$

$$(t_\mu) = \frac{1}{(\mu)}$$

and any given length and time can be expressed by choosing the appropriate value of (μ).

Frequently this is expressed by saying :

"we put $\hbar = c = 1$. Then only one dimension, namely the mass, remains and everything is measured in terms of powers of M or M^{-1}".

This might be a matter of taste, but I personally prefer to keep all dimensions, mass, length, time, distinct and choose the units such that the n u m e r i c a l values become equal to one, but not the physical quantities. In calculations, however, one practically works with equations written as if the mass were the only dimension.

Example

If your age is y years then, expressed in elementary particle units, it will be just the reciprocal of that mass, whose Compton wavelength equals y light years and that mass is measured in units of M_{proton}.

3) Some practical examples for the use of invariants

Some of the following examples are - apart from their usefulness in a high energy laboratory - chosen to illustrate that the concept of "invariants" is not only of theoretical interest, but leads to sometimes surprisingly easy calculations of quantities, which otherwise would be found only after lengthy algebra.

If, e.g., one asks for the centre-of-momentum energy of a system of particles, then the straightforward - but tedious - way would be to make a Lorentz transformation of all four momenta to the centre-of-momentum and add all the transformed energies. Or, if one asks what is the energy of a certain particle as seen from the rest system of another one, then one would transform the four momentum of the particle in question to that system.

If one happens to know, or to have at hand the most general form of the Lorentz transformation - namely Eq. (I.10) - then these Lorentz transformations are not too difficult to carry out; and we shall do this later on in an example. But this formula is not too easy to remember and just when one needs it, it is not available. For the above questions - and similar others - one needs not the Lorentz transformation, however. The point is this :

if a question is of such a nature that its answer will always be the same, no matter in which Lorentz system one starts, then it must be possible to formulate the answer entirely with the help of those invariants which one can build with the available four vectors. One then finds the answer in a particular Lorentz system which one can choose freely and in such a way that the answer is there obvious or most easy. One looks then how the invariants appear in this particular system, expresses the answer to the problem by these invariants and one has found at the same time already the general answer. This looks sometimes like hocus-pocus as you will see in the following examples. It is worthwhile to devote some thinking to this method of calculation until one has completely understood that there is really no jugglery or guesswork in it and that it is absolutely safe.

Let us turn to the examples.

i) Centre-of-momentum (CM) energy and the velocity of the CM.

Suppose we have in a certain Lorentz system - we call it laboratory system, but it may be any one - two particles with four momenta p_1 and p_2 and with masses m_1 and m_2; see Fig. I.4

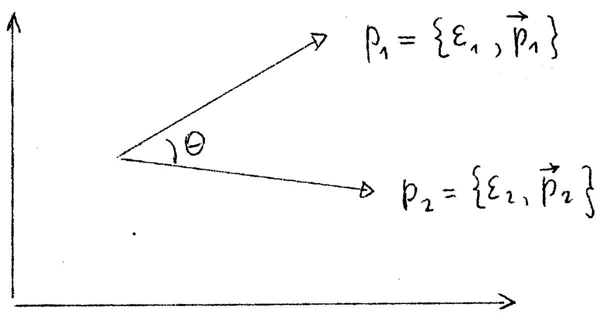


Fig. I.4
The two particles kinematics

/Each of the two momenta p_1 and p_2 may again stay for the total momentum of a whole system of particles./

What is the CM-energy E ? This question must have an answer which is independent on the Lorentz system, in which p_1 and p_2 are given. It must be possible to give this answer in terms of the three invariants

$$p_1^2 = m_1^2 \quad \text{and} \quad p_2^2 = m_2^2 \quad \text{and} \quad \left[p_1 p_2 \quad \text{or} \quad (p_1+p_2)^2 \quad \text{or} \quad (p_1-p_2)^2 \right]$$

The answer is obvious in the CM-system itself, namely there /we denote CM-quantities with an asterisk/

$$\vec{p}_1^{\,*} + \vec{p}_2^{\,*} = 0, \qquad \text{hence} \quad p_1^* + p_2^* = \left\{ \varepsilon_1^* + \varepsilon_2^* , \vec{0} \right\} \text{ and } E^* = \varepsilon_1^* + \varepsilon_2^* .$$

Hence

$$E^{*2} = (\varepsilon_1^* + \varepsilon_2^*)^2 = (p_1^* + p_2^*)^2 = (p_1 + p_2)^2$$

since $(p_1+p_2)^2$ is invariant. We may define the total mass M of the system by the square of its total four momentum

$$P^2 = (p_1+p_2)^2 = M^2 = E^{*2} = (\varepsilon_1 + \varepsilon_2)^2 - (\vec{p}_1+\vec{p}_2)^2 = \text{invariant} \tag{I.18}$$

I.e., kinematically our two particles p_1 and p_2 are equivalent to one single particle with four momentum P and mass $M = E_{CM}$. Therefore each p_1 and p_2 can in turn be considered as representing a system of particles.

Furthermore, any four momentum can be written

$$\vec{p} = m \vec{v} \gamma \qquad \varepsilon = m\gamma \qquad \gamma = \frac{1}{\sqrt{1-v^2}} \tag{I.19}$$

This is one of the most useful formulae. It should be known by heart.

Hence, for the total four momentum of the two-particle system :

$$\vec{P} = M \cdot \vec{\beta}\gamma$$

$$E = M\gamma$$

therefore

$$\vec{\beta}_{CM} = \frac{\vec{P}}{E} = \frac{(\vec{p}_1+\vec{p}_2)}{(\varepsilon_1+\varepsilon_2)}$$ is the velocity of the CM seen from the lab., and

$$\gamma_{CM} = \frac{1}{\sqrt{1-\beta^2}} = \frac{E}{M} = \frac{\varepsilon_1+\varepsilon_2}{\sqrt{(\varepsilon_1+\varepsilon_2)^2 - (\vec{p}_1+\vec{p}_2)^2}}$$ is the corresponding γ

(I.20)

For practical calculations one has to express everything either in

$\varepsilon_1, \varepsilon_2$ and $\cos\theta$

or in

$|\vec{p}_1|, |\vec{p}_2|$ and $\cos\theta$

using

$$\varepsilon = \sqrt{m^2+\vec{p}^2}$$

Problem :

5) In nuclei the kinetic energy of the bound nucleons goes up to the order of 20 MeV. Illustrate formulae (I.18) and (I.19) by calculating the effect of this motion when it is parallel or antiparallel to an incoming beam of 25 GeV (kin. energy) protons. (Put $M \cong 1$ GeV.)

a) which is the difference in the CM-energy ?

b) which energies must incoming protons have to produce the same CM-energies on nucleons at rest ?

c) which is the difference in the $\vec{\beta}$ and γ of the CM-system ?

Solution 5)

a) We express everything by $\varepsilon_1 \varepsilon_2$ and $\cos\theta$ and obtain

$$E^{*2} = 2\,\varepsilon_1 \varepsilon_2 + (m_1^2 + m_2^2) \mp 2\sqrt{(\varepsilon_1^2 - m_1^2)(\varepsilon_2^2 - m_2^2)}$$

where the $-$ sign is valid for parallel motion and the $+$ sign for antiparallel motion. We have then

$$m_1 = m_2 = 1$$

$$\varepsilon_1 = 1 + 0.02 \qquad \varepsilon_1^2 - 1 = 1 + 0.04 - 1 = 0.04$$

$$\varepsilon_2 = 1 + 25 \qquad \varepsilon_2^2 - 1 = 1 + 50 + 625 - 1 = 675$$

$$\sqrt{(\varepsilon_1^2 - 1)(\varepsilon_2^2 - 1)} = \sqrt{4 \cdot 6.75} = \sqrt{26} = 5.1$$

$$E^{*2} = 2(1 + 26.5 \pm 5.1) = \begin{cases} 65.2 & \text{antiparallel} \\ 44.8 & \text{parallel} \end{cases}$$

$$E^{*} = \begin{cases} 8.08 & \text{for antiparallel motion} \\ 7.35 & \text{for nucleon at rest} \\ 6.70 & \text{for parallel motion} \end{cases}$$

b) If the energy of the incoming proton which produces these CM-energies on a nucleon at rest is denoted by ε', then

$$E^{*2} = 2(\varepsilon' + 1) = \begin{cases} 65.2 \\ 54.0 \\ 44.8 \end{cases}$$

$$\varepsilon' = \begin{cases} 31.6 & \text{for antiparallel motion} \\ 26 & \text{for nucleon at rest} \\ 21.4 & \text{for parallel motion} \end{cases}$$

That means that the 20 MeV nucleon motion is equivalent to about 5 GeV difference in primary energy !

c) From (I.20)

$$\beta_{\mp} = \frac{p_1 \mp p_2}{\varepsilon_1 + \varepsilon_2} = \frac{26 \mp 0.2}{27} = \begin{cases} 0.956 & \text{for antiparallel motion} \\ 0.97 & \text{for parallel motion} \end{cases}$$

$$\gamma_{\mp} = \frac{\varepsilon_1 + \varepsilon_2}{E_{cm}^{(\mp)}} = \frac{27}{\begin{Bmatrix} 8.08 \\ 6.70 \end{Bmatrix}} = \begin{cases} 3.34 & \text{for antiparallel motion} \\ 4.03 & \text{for parallel motion} \end{cases}$$

The reader should discuss the colliding-beam machine in this way !

Problem :

6) Suppose that a group of A nucleons (at rest) as a whole would interact with an incoming proton of 25 GeV kin. energy.

a) which energy would be available for the production of particles and for kinetic energy ? (Put $M \cong 1$ GeV.)

b) how do β_{CM} and γ_{CM} depend on A ?

Solution 6)

a) From the formula derived in problem 5), putting $\varepsilon_1 = m_1$, we obtain

$$E^{*2} = 2\,\varepsilon_1 \varepsilon_2 + m_1^2 + m_2^2$$

where

$\varepsilon_1 = m_1 = A$

$\varepsilon_2 = 26$

$m_1 = A$

$m_2 = 1$

The available energy E (nucleon conservation !) is

$$E = E^* - (A+1) = \sqrt{52A + A^2 + 1} - (A+1)$$

We give below a numerical example

A	1	2	5	10	20	40	100	∞
E	5.35	7.4	10.9	13.9	16.9	19.6	22.2	25

The result for ∞ is obvious.

b) For β we have with $\vec{p}_1 = 0$ and $\varepsilon_1 = A$

$$\beta(A) = \frac{26}{26 + A}$$

which as long as A remains small compared to 26, does not change very much, whereas

$$\gamma(A) = \frac{1}{\sqrt{1-\beta^2(A)}}$$

depends on A much more critically in the neighbourhood of $\beta \approx 1$.

---

ii) The energy, momentum and velocity of one particle seen from the rest system of another one.

Suppose in Fig. I.4 we sit on particle 1, moving with it, which would be for us the energy of particle 2 ?

The answer to this question must always be the same, no matter in which Lorentz system we start. It must be therefore expressible by invariants and the only invariants are again

$$p_1^2 = m_1^2 \; ; \quad p_2^2 = m_2^2 \; ; \quad p_1 p_2 \quad \text{or} \quad (p_1+p_2)^2 \quad \text{or} \quad (p_1-p_2)^2$$

We call the wanted energy E_{21}. It is the energy of particle 2 if we look at it in the rest system of 1 :

$$E_{21} = \varepsilon_2$$

in the system where $\vec{p}_1 = 0$.

We only need to write this in an invariant form, namely by expressing it by the three invariants.

In this particular system the last invariant is

$$p_1 p_2 = \varepsilon_1 \varepsilon_2 = m_1 \varepsilon_2 \quad ,$$

hence

$$E_{21} = \varepsilon_2 = \frac{p_1 p_2}{m_1} \quad .$$

Since the right hand side is obviously invariant, we have already the general result. Further from

$$|\vec{p}_{21}|^2 = E_{21}^2 - m_2^2 = \frac{(p_1 p_2)^2 - m_1^2 m_2^2}{m_1^2}$$

and from (I.19) it follows :

if p_1 and p_2 are the momentum four vectors of any two particles in any Lorentz system, then

$$E_{21} = \frac{p_1 p_2}{m_1}$$

$$|\vec{p}_{21}|^2 = \frac{(p_1 p_2)^2 - m_1^2 m_2^2}{m_1^2} \quad ; \quad \left[p_1 p_2 \equiv \varepsilon_1 \varepsilon_2 - \vec{p}_1 \vec{p}_2 \right] \qquad (I.21)$$

$$v_{21}^2 = \frac{|\vec{p}_{21}|^2}{E_{21}^2} = \frac{(p_1 p_2)^2 - m_1^2 m_2^2}{(p_1 p_2)^2}$$

are the energy (E_{21}) and momentum $(|\vec{p}_{21}|)$ of particle 2 seen from particle 1 and v_{21} is the relative velocity (symmetric in 1 and 2).

All these expressions are invariant and can be evaluated in any Lorentz system.

iii) The energy, momentum and velocity of a particle seen from the centre-of-momentum system.

We solve this problem immediately by observing that we need only consider what these quantities are as seen from a fictitious particle M, namely the "centre-of-momentum-particle", whose four momentum is

$$P = p_1 + p_2$$

and we need only apply formulae (I.21) with p_1 replaced by P and p_2 by the four momentum of that particle whose energy, momentum and velocity we wish to know.

Let us denote centre-of-momentum quantities with an asterisk, then from (I.21)

$$\mathcal{E}^*_i = \frac{Pp_i}{M} \;;\quad |\vec{p}^{\,*}_i|^2 = \frac{(Pp_i)^2 - M^2 m_i^2}{M^2} \;;\quad v_i^{*2} = \frac{(Pp_i)^2 - M^2 m_i^2}{(Pp_i)^2}$$

using explicitly $P = p_1+p_2$ and

$$p_1p_2 = \tfrac{1}{2}\left[(p_1+p_2)^2 - p_1^2-p_2^2\right] = \tfrac{1}{2}(M^2-m_1^2-m_2^2)$$

we obtain almost immediately the result :

$$\mathcal{E}^*_1 = \frac{M^2+(m_1^2-m_2^2)}{2M} \;;\quad \mathcal{E}^*_2 = \frac{M^2-(m_1^2-m_2^2)}{2M} \;;\quad \mathcal{E}^*_1 + \mathcal{E}^*_2 = M$$

$$|\vec{p}^{\,*}|^2 = |\vec{p}^{\,*}_1|^2 = |\vec{p}^{\,*}_2|^2 = \frac{M^4-2M^2(m_1^2+m_2^2)+(m_1^2-m_2^2)^2}{4M^2} =$$

$$= \frac{\left[M^2-(m_1+m_2)^2\right]\left[M^2-(m_1-m_2)^2\right]}{4M^2} \qquad \text{(I.22)}$$

$$v_i^{*2} = \left[\frac{|\vec{p}^{\,*}|}{\mathcal{E}^*_i}\right]^2$$

where $\mathcal{E}^*_i$ and v^*_i are energy and velocity respectively of particle i, as seen from their common centre-of-momentum system and $M^2 = P^2 = (p_1+p_2)^2$ is the total mass squared.

All these expressions are obviously invariant as they answer a question which must lead to the same statement independent of the frame of reference in which p_1 and p_2 are given.

Eq. (I.22) gives at the same time the energies, the momentum and the velocities of two particles m_1 and m_2 into which a particle of mass M decays.

Problem :

7) Given a decay process

$$M \longrightarrow m_1 + m_2$$

what are the energy and momentum of particle 2 seen from particle 1 ?

Solution 7)

We use the fact that

$$M^2 = (p_1+p_2)^2 = m_1^2+m_2^2 + 2p_1p_2 \quad ,$$

hence

$$p_1p_2 = \tfrac{1}{2}(M^2 - m_1^2 - m_2^2) \ .$$

This inserted into (I.21) gives immediately the answer.

LECTURE 4

4) The Lorentz transformation to the rest system of an arbitrary particle (or to the centre-of-momentum system)

So far we only have calculated invariant quantities. This could be done without invoking the Lorentz transformation explicitly. We only made use of the invariants p_1^2, p_2^2, $(p_1+p_2)^2$ and of the trick to calculate things in that system where the quantities had the simplest form.

If we wish to know directions of momenta, then we can no longer proceed that way. But we still can write down immediately everything we wish by means of the general Lorentz transformation (I.10) [now applied to four momenta $p \equiv \{\varepsilon, \vec{p}\}$ and with $c = 1$]:

$$\left.\begin{aligned} \vec{p} &= \vec{p}\,' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\vec{\beta}\vec{p}\,' + \varepsilon'\right] \\ \varepsilon &= \gamma\left[\varepsilon' + \vec{\beta}\vec{p}\,'\right] , \end{aligned}\right\} \qquad (I.23)$$

and of the relations between energy, velocity and momentum (I.19) :

$$\vec{\beta} = \frac{\vec{P}}{E} \; ; \quad \gamma = \frac{E}{M} , \qquad (I.24)$$

which give the velocity $\vec{\beta}$ and $\gamma = 1/\sqrt{1-\beta^2}$ for any particle (or system of particles) with (total) four momentum $P = \{E, \vec{P}\}$, $P^2 = M^2$.

We solve (I.23) for quantities with prime by interchanging them with those without prime and changing $\vec{\beta}$ to $-\vec{\beta}$. Then the complete set of formulae is

$\vec{p}\,' = \vec{p} + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\vec{\beta}\vec{p} - \varepsilon\right]$	$\vec{p} = \vec{p}\,' + \vec{\beta}\gamma\left[\frac{\gamma}{\gamma+1}\vec{\beta}\vec{p}\,' + \varepsilon'\right]$
$\varepsilon' = \gamma\left[\varepsilon - \vec{p}\vec{\beta}\right]$	$\varepsilon = \gamma\left[\varepsilon' + \vec{\beta}\vec{p}\,'\right]$
$\vec{\beta} = \frac{\vec{P}}{E}$; $\gamma = \frac{E}{M}$; $P = \{E, \vec{P}\}$; $P^2 = M^2$	

(I.25)

According to the meaning of the quantities in (I.25) these equations describe the following situation :

let P and p be the four momenta of any two particles (or systems of particles) in a certain reference frame K,

let further K' be the rest system of P [i.e., there $\vec{P}\,' = 0$],

then (I.25) gives the transformation of the four momentum p from K to K' and vice versa.

Two examples :

(α) let there be only one particle and transform to its rest-system [this example is trivial and only checks formulae (I.25)]

$$P = p \;;\quad \vec{\beta} = \frac{\vec{p}}{\varepsilon} \;;\quad \gamma = \frac{\varepsilon}{m}$$

$$\vec{p}\,' = \vec{p} + \frac{\vec{p}}{m}\left[\frac{\varepsilon}{\varepsilon+m}\,\frac{\vec{p}^{\,2}}{\varepsilon} - \varepsilon\right] = 0$$

$$\varepsilon' = \frac{\varepsilon}{m}\left[\varepsilon - \frac{\vec{p}^{\,2}}{\varepsilon}\right] = m$$

as it ought to be.

(β) let there be two particles p_1 and p_2 and put

$$P = (p_1+p_2) = (\varepsilon_1+\varepsilon_2, \vec{p}_1+\vec{p}_2) \equiv (E,\vec{P})$$

we obtain then the transformation to the CM-system :
[replacing here the prime by an asterisk]

$$\vec{p}_1^{\,*} = \vec{p}_1 + \frac{\vec{P}}{M}\left[\frac{\vec{P}\vec{p}_1}{E+M} - \varepsilon_1\right] \; ; \quad \varepsilon_1^* = \frac{1}{M}\left[E\varepsilon_1 - \vec{P}\vec{p}_1\right] = \frac{Pp_1}{M} \qquad \text{(see p.35)}$$

$$\vec{p}_2^{\,*} = \vec{p}_2 + \frac{\vec{P}}{M}\left[\frac{\vec{P}\vec{p}_2}{E+M} - \varepsilon_2\right] \; ; \quad \varepsilon_2^* = \frac{Pp_2}{M} \; . \qquad \text{(see p.35)}$$

And, as it should be

$$\vec{p}_1^{\,*}+\vec{p}_2^{\,*} = \vec{p}_1+\vec{p}_2 + \frac{\vec{P}}{M}\left[\frac{\vec{P}(\vec{p}_1+\vec{p}_2)}{E+M} - \varepsilon_1 - \varepsilon_2\right] = \vec{P} + \frac{\vec{P}}{M}\left[\frac{\vec{P}^2}{E+M} - E\right] = 0$$

Similarly, one may transform to the rest system of particle 1 or 2 and obtain the motion of the other particle seen in that system. One may, of course, do analogous transformations if more than two particles are involved by considering groups of particles as being represented by their common four momentum $P = p_1+p_2+p_3+ \dots$.

5) The transformation of differential cross-sections Jacobian determinants

We shall see later how a cross-section might be defined in an invariant way. Presently, let us discuss the following problem :

given a differential cross-section in one system of co-ordinates, what is the corresponding differential cross-section in another one ?

We first recall a few formulae which are derived in every book on integral calculus.

If between two sets of co-ordinates $x_1 \dots x_n$ and $y_1 \dots y_n$ in an n-dimensional space a transformation

$$\left.\begin{aligned} x_i &= x_i(y_1 \dots y_n) \\ \text{with the inverse} \qquad y_k &= y_k(x_1 \dots x_n) \end{aligned}\right\} \qquad (I.26)$$

is defined, then for integrations in this space

$$\int_{R_x} \dots \int f\left[x_1 \dots x_n\right] dx_1 \dots dx_n = \int_{R_y} \dots \int f\left[x_1(y_1 \dots y_n) \dots x_n(y_1 \dots y_n)\right] \times \frac{\partial(x_1 \dots x_n)}{\partial(y_1 \dots y_n)} \, dy_1 \dots dy_n \qquad (I.27)$$

where R_x is a certain boundary expressed by equations in the variables $x_1 \dots x_n$, R_y the same boundary, expressed by equations in the variables $y_1 \dots y_n$ [one obtains it from R_x inserting $x_i = x_i(y_1 \dots y_n)$ in it], and where

$$\frac{\partial(x_1 \dots x_n)}{\partial(y_1 \dots y_n)} \equiv \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_2}{\partial y_1} & \dots & \frac{\partial x_n}{\partial y_1} \\ \vdots & & & \\ \frac{\partial x_1}{\partial y_n} & \dots & \dots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} \qquad (I.28)$$

is the "Jacobian" determinant expressing how the n-dimensional volume element $dy_1 \dots dy_n$ differs from the element $dx_1 \dots dx_n$.

(I.27) is easy to remember : if one formally cancels

$$\frac{\partial(x_1\ldots x_n)}{\partial(y_1\ldots y_n)}\, dy_1\ldots dy_n = dx_1\ldots dx_n$$

and suppresses the arguments $y_1\ldots y_n$ in f on the r.h.s., then both sides are equal. If several transformations $x_i \to y_i \to z_i$ are carried out one after the other, then one obtains the "chain rule"

$$\frac{\partial(x_1\ldots x_n)}{\partial(z_1\ldots z_n)} = \frac{\partial(x_1\ldots x_n)}{\partial(y_1\ldots y_n)} \cdot \frac{\partial(y_1\ldots y_n)}{\partial(z_1\ldots z_n)} \qquad \text{(I.29)}$$

[that means : one can formally cancel], from which follows, with $z_i = x_i$ (identity)

$$\frac{\partial(x_1\ldots x_n)}{\partial(y_1\ldots y_n)} \cdot \frac{\partial(y_1\ldots y_n)}{\partial(x_1\ldots x_n)} = 1 \qquad \text{(I.30)}$$

The Jacobian is 1 if and only if the volume element is preserved. Particular examples are :

-- the rotations and translations
-- the canonical transformations (see any book on classical mechanics).

Let us return again to (I.27). If we take R_x to be that volume element which in the x-co-ordinates becomes $dx_1\ldots dx_n$ then R_y describes the same volume element, however, expressed in the co-ordinates $y_1\ldots y_n$. We have therefore

$$f\left[x_1\ldots x_n\right] dx_1\ldots dx_n = f\left[x_1(y_1\ldots y_n)\ldots\right]\frac{\partial(x_1\ldots x_n)}{\partial(y_1\ldots y_n)}\, dy_1\ldots dy_n \qquad \text{(I.31)}$$

We can look at this in two different ways :

i) take for a moment $f = \text{const}$ (i.e., independent of $x_1 \ldots x_n$). Then

$$dx_1 \ldots dx_n = \frac{\partial(x_1 \ldots x_n)}{\partial(y_1 \ldots y_n)} \, dy_1 \ldots dy_n \qquad (I.32)$$

i.e., the Jacobian gives the ratio between the volume elements $dx_1 \ldots dx_n$ and $dy_1 \ldots dy_n$.

So, if we consider the Jacobian as belonging to the volume element, then (I.31) can be interpreted as defining a new function

$$g(y_1 \ldots y_n) \equiv f\left[x_1(y_1 \ldots y_n), \; x_2(y_1 \ldots y_n) \; \ldots\right] = f(x_1 \ldots x_n) \qquad (I.31.1)$$

or, since symbolically $x = T^{-1}y$ (T being the transformation),

$$g(y) = f(T^{-1}y) \; ,$$

which is the transformation law (I.17) of a scalar function [put $y = x'$ and $g = f'$ to obtain that formula]. This function $g(y)$ has physically the same signification as $f(x)$, as it is related to one and the same volume element, namely to $dx_1 \ldots dx_n$ on the left hand side and to

$$\frac{\partial(x_1 \ldots x_n)}{\partial(y_1 \ldots y_n)} \, dy_1 \ldots dy_n$$

on the right hand side of (I.31).

ii) for physical reasons it can be preferable n o t to relate the quantity on both sides to the s a m e volume element : $dx_1 \ldots dx_n$ as well as $dy_1 \ldots dy_n$ might have convenient geometrical and/or physical interpretations and then one might relate the function to $dx_1 \ldots dx_n$ on the left and to $dy_1 \ldots dy_n$ on the right hand side. These two volume elements are frequently different not only by a numerical factor but even by their physical dimension.

If we then define a new function $h(y)$ by requiring

$$f(x_1 \dots x_n)\, dx_1 \dots dx_n \equiv h(y_1 \dots y_n)\, dy_1 \dots dy_n \quad , \qquad (I.33)$$

it is clear that $f(x_1 \dots x_n)$ and $h(y_1 \dots y_n)$ are no longer equal. One reads off from (I.31) or finds by dividing the equation which defines $h(y)$ by $dy_1 \dots dy_n$ [see (I.32)]

$$h(y_1 \dots y_n) \equiv f(x_1 \dots x_n) \frac{\partial (x_1 \dots x_n)}{\partial (y_1 \dots y_n)} \qquad (I.31.2)$$

[Of course, the x_i on the r.h.s. are to be expressed by the y_i.]

Therefore the transformation law for scalar functions no longer applies here. Indeed, the functions $f(x_1 \dots x_n)$, $g(y_1 \dots y_n)$ and $h(y_1 \dots y_n)$ have the properties of a density. But if one transforms the co-ordinates in such a way that the volume element changes, then the density cannot remain the same.

The practical examples which we are interested in are the transformations of cross-sections and here we shall discuss two cases : the transformation to polar co-ordinates and the transformation from one Lorentz frame to the other.

We may define a differential cross-section *) either by the number of a definite kind of particles [per event] going into the volume element $dp_1 dp_2 dp_3$ in momentum space or by the number going into the solid angle element and having momentum between p and $p+dp$. Call the first quantity S, then

---

*) The quantity which we define here becomes in fact a dimensionless number after multiplication with the volume element. Therefore it is not yet a "cross-section". The difference is, however, trivial in the present context and the transformation formulae with which we end up (I.46) are correct as well for the numbers we define now as for the cross-sections proper. Therefore we may use the word "cross-section" throughout.

$$\frac{\partial^3 S(p_1 p_2 p_3)}{\partial p_1 \partial p_2 \partial p_3} dp_1 dp_2 dp_3 = \frac{\partial^3 S\left[p_1(p\vartheta\varphi) p_2(p\vartheta\varphi) p_3(p\vartheta\varphi)\right]}{\partial p_1 \partial p_2 \partial p_3} \frac{\partial(p_1 p_2 p_3)}{\partial(p\vartheta\varphi)} dp d\vartheta d\varphi \tag{I.34}$$

As everybody knows ;

$$dp_1 dp_2 dp_3 = p^2 \sin\vartheta \, dp d\vartheta d\varphi \quad ,$$

in other words, the Jacobian is

$$\frac{\partial(p_1 p_2 p_3)}{\partial(p\vartheta\varphi)} = p^2 \sin\vartheta \quad , \tag{I.35}$$

One now defines, on the right hand side of (I.34), the differential cross-section with respect to momentum and solid angle, by splitting the Jacobian $p^2 \sin\vartheta$ into p^2 and $\sin\vartheta$. The first factor is absorbed into the new defined cross-section but the factor $\sin\vartheta$ remains in the differential :

$$\underbrace{\frac{\partial^3 S(p_1 p_2 p_3)}{\partial p_1 \partial p_2 \partial p_3} \cdot p^2}_{\frac{\partial^2 \sigma(p,\vartheta\varphi)}{\partial p \partial \Omega}} \underbrace{\sin\vartheta \, dp \, d\vartheta \, d\varphi}_{dp \, d\Omega} \quad ,$$

$$\left.\begin{aligned} \frac{\partial^2 \sigma(p,\vartheta,\varphi)}{\partial p \partial \Omega} &= p^2 \frac{\partial^3 S(p_1, p_2, p_3)}{\partial p_1 \partial p_2 \partial p_3} \\ d\Omega &= \sin\vartheta \, d\vartheta \, d\varphi \end{aligned}\right\} \tag{I.36}$$

This is an intermediate example between the two ways of looking at the transformation of a function discussed above. Here one relates the function to two different volume elements but one does not absorb the w h o l e Jacobian into the new function. How much of it is left in the new volume element is suggested by physical considerations. The solid angle and the magnitude of the momentum are convenient; one therefore relates the cross-section to them.

The next example is the transformation of a cross-section from one Lorentz system, K', to another one, K.

Let us choose the co-ordinates such that the axes of the system K' are parallel to those of K and that their relative motion is along the z-axis. The Lorentz transformation then is [see (I.9), replace x,ct by p_3,E]

$$\left.\begin{aligned} p_3 &= \gamma\left[p_3' + \beta E'\right] & \qquad p_3' &= \gamma\left[p_3 - \beta E\right] \\ p_2 &= p_2' & p_2' &= p_2 \\ p_1 &= p_1' & p_1' &= p_1 \\ E &= \gamma\left[E' + \beta p_3'\right] & E' &= \gamma\left[E - \beta p_3\right] \end{aligned}\right\} \quad (I.37)$$

We introduce polar co-ordinates

$$p_1 = p \sin\vartheta \cos\varphi$$
$$p_2 = p \sin\vartheta \sin\varphi$$
$$p_3 = p \cos\vartheta$$

and find that $p_2 = p_2'$ and $p_1 = p_1'$ imply $\varphi = \varphi'$. Hence

$$\left.\begin{aligned} p\cos\vartheta &= \gamma\left[p'\cos\vartheta' + \beta E'\right] & \qquad p'\cos\vartheta' &= \gamma\left[p\cos\vartheta - \beta E\right] \\ p\sin\vartheta &= p'\sin\vartheta' & p'\sin\vartheta' &= p\sin\vartheta \\ \varphi &= \varphi' & \varphi' &= \varphi \\ E &= \gamma\left[E' + \beta p'\cos\vartheta'\right] & E' &= \gamma\left[E - \beta p\cos\vartheta\right] \end{aligned}\right\} \quad (I.38)$$

For the transformation of the cross-section we require that the number of particles going into the solid angle element $d\Omega$ and having a momentum between p and p+dp be the same as the number going into the corresponding solid angle element $d\Omega'$ and having a corresponding momentum between p' and p'+dp' :

$$\frac{\partial^2\sigma(p,\vartheta,\varphi)}{\partial p\,\partial\Omega}\,dp\,d\Omega = \frac{\partial^2\sigma'(p'\vartheta'\varphi')}{\partial p'\,\partial\Omega'}\,dp'\,d\Omega' \qquad (I.39)$$

This is an example of the second way to look at a transformation : the Jacobian is absorbed into the cross-section [see (I.33) and (I.31.2)], and consequently

$$\frac{\partial^2\sigma(p,\vartheta,\varphi)}{\partial p\,\partial\Omega} = \frac{\partial^2\sigma'(p'\vartheta'\varphi')}{\partial p'\,\partial\Omega'}\,\frac{\partial(p'\Omega')}{\partial(p\Omega)}$$

$$\frac{\partial^2\sigma'(p'\vartheta'\varphi')}{\partial p'\,\partial\Omega'} = \frac{\partial^2\sigma(p,\vartheta,\varphi)}{\partial p\,\partial\Omega}\,\frac{\partial(p\Omega)}{\partial(p'\Omega')} \qquad (I.40)$$

where the primed and unprimed variables are related by (I.38). The only task which remains is to calculate

$$\frac{\partial(p\Omega)}{\partial(p'\Omega')} .$$

We could do this directly, using (I.38), but we shall proceed through another way : for a product of transformations one obtains the product of the Jacobians [see (I.29)]. Therefore, if we transform in five steps, namely

$$p,\Omega \rightarrow p\,\vartheta\,\varphi \rightarrow p_1 p_2 p_3 \rightarrow p_1' p_2' p_3' \rightarrow p'\vartheta'\varphi' \rightarrow p'\Omega'$$

the Jacobian is

$$\frac{\partial(p\Omega)}{\partial(p\vartheta\varphi)}\cdot\frac{\partial(p\vartheta\varphi)}{\partial(p_1p_2p_3)}\cdot\frac{\partial(p_1p_2p_3)}{\partial(p_1'p_2'p_3')}\cdot\frac{\partial(p_1'p_2'p_3')}{\partial(p'\vartheta'\varphi')}\cdot\frac{\partial(p'\vartheta'\varphi')}{\partial(p'\Omega')} = \frac{\partial(p\Omega)}{\partial(p'\Omega')}$$

The Lorentz transformation is contained only in the third factor. We first consider the other ones :

$$\frac{\partial(p\Omega)}{\partial(p\vartheta\varphi)} = \frac{dp\ \sin\vartheta\, d\vartheta\, d\varphi}{dp\ d\vartheta\, d\varphi} = \sin\vartheta$$

$$\frac{\partial(p\vartheta\varphi)}{\partial(p_1p_2p_3)} = \left[\frac{\partial(p_1p_2p_3)}{\partial(p\vartheta\varphi)}\right]^{-1} = \frac{1}{p^2\sin\vartheta}$$

[transformation between Cartesian and polar co-ordinates (I.35)].

Hence the first two and the last two factors give together simply

$$\frac{p'^2}{p^2} = \frac{\sin^2\vartheta}{\sin^2\vartheta'}$$

[see (I.38)]

There remains to calculate the factor in the middle, which can be found immediately from (I.37) : since only p_3 and E are transformed, the Jacobian reduces to

$$\frac{\partial(p_1p_2p_3)}{\partial(p_1'p_2'p_3')} = \frac{\partial p_3}{\partial p_3'} = \gamma + \beta\frac{\partial E'}{\partial p_3'} = \gamma + \beta\frac{p_3'}{E'} = \frac{E}{E'} \quad . \qquad (I.41)$$

Hence, all factors together,

$$\frac{\partial(p\Omega)}{\partial(p'\Omega')} = \frac{E\ \sin^2\vartheta}{E'\sin^2\vartheta'} = \frac{p'^2E}{p^2E'} \quad . \qquad (I.42)$$

This, introduced into (I.40), settles the question of the transformation completely.

We end this discussion by including the cases where the cross-section in one or both systems is expressed in terms of solid angle and energy. For this we need three other Jacobians

$$\frac{\partial(p\Omega)}{\partial(E'\Omega')}\ ,\quad \frac{\partial(E\Omega)}{\partial(p'\Omega')}\ ,\quad \frac{\partial(E\Omega)}{\partial(E'\Omega')}$$

which are found again most easily by applying the "chain rule" (I.29), and using the Jacobian, (I.42), which we already know :

$$\frac{\partial(p\Omega)}{\partial(E'\Omega')} = \frac{\partial(p\Omega)}{\partial(p'\Omega')}\cdot\frac{\partial(p'\Omega')}{\partial(E'\Omega')} = \frac{\partial(p\Omega)}{\partial(p'\Omega')}\cdot\frac{E'}{p'} = \frac{p'E}{p^2} = \frac{E}{p'}\cdot\frac{\sin^2\vartheta}{\sin^2\vartheta'} = \frac{E}{p}\frac{\sin\vartheta}{\sin\vartheta'} \qquad (I.43)$$

from which follows immediately

$$\frac{\partial(E\Omega)}{\partial(p'\Omega')} = \frac{p'^2}{pE'} = \frac{p}{E'}\frac{\sin^2\vartheta}{\sin^2\vartheta'} = \frac{p'}{E'}\frac{\sin\vartheta}{\sin\vartheta'} \qquad (I.44)$$

$$\frac{\partial(E\Omega)}{\partial(E'\Omega')} = \frac{\partial(E\Omega)}{\partial(p\Omega)}\ \frac{\partial(p\Omega)}{\partial(p'\Omega')}\cdot\frac{\partial(p'\Omega')}{\partial(E'\Omega')} = \frac{p}{E}\cdot\frac{\partial(p\Omega)}{\partial(p'\Omega')}\cdot\frac{E'}{p'} =$$

$$= \frac{p'}{p} = \frac{\sin\vartheta}{\sin\vartheta'}\ . \qquad (I.45)$$

Hence, for the cross-sections :

$$\left.\begin{aligned}
\frac{\partial^2\sigma(p\vartheta\varphi)}{\partial p\,\partial\Omega} &= \frac{\partial^2\sigma'(p'\vartheta'\varphi')}{\partial p'\,\partial\Omega'}\cdot\frac{E'\sin^2\vartheta'}{E\sin^2\vartheta}\\
\frac{\partial^2\sigma(p\vartheta\varphi)}{\partial p\,\partial\Omega} &= \frac{\partial^2\sigma'(E'\vartheta'\varphi')}{\partial E'\,\partial\Omega'}\cdot\frac{p}{E}\frac{\sin\vartheta'}{\sin\vartheta}\\
\frac{\partial^2\sigma(E\vartheta\varphi)}{\partial E\,\partial\Omega} &= \frac{\partial^2\sigma'(p'\vartheta'\varphi')}{\partial p'\,\partial\Omega'}\cdot\frac{E'}{p'}\frac{\sin\vartheta'}{\sin\vartheta}\\
\frac{\partial^2\sigma(E\vartheta\varphi)}{\partial E\,\partial\Omega} &= \frac{\partial^2\sigma'(E'\vartheta'\varphi')}{\partial E'\,\partial\Omega'}\cdot\frac{\sin\vartheta'}{\sin\vartheta}
\end{aligned}\right\} \qquad (I.46)$$

where (I.38) gives the relation between the primed and unprimed quantities.

Problem :

8) Prove Eq. (I.41) without explicit use of the Lorentz transformation : define the number of particles going into the four-dimensional volume element in p-space and discuss its invariance properties. Use the δ-function and one integration in order to eliminate all four momenta not belonging to the particle considered (its mass m is given).

Solution 8)

Let $N(p_0 p_1 p_2 p_3) dp_0 dp_1 dp_2 dp_3$ be the number of particles with four momentum $p = \{p_0, \vec{p}\}$ going into the four-momentum element $d^4p \equiv dp_0 dp_1 dp_2 dp_3$.

If we make a Lorentz transformation then this number must be the same

$$N(p)d^4p = N'(p')d^4p'$$

But as the L-transformation is orthogonal, we have

$$\frac{\partial(p_0 p_1 p_2 p_3)}{\partial(p'_0 p'_1 p'_2 p'_3)} \equiv 1$$

or $d^4p = d^4p'$, hence

$$N(p) = N'(p')$$

behaves as a true scalar under L-transformations (under more general transformations it behaves as a density). We now impose the condition that the particles in question have mass m by using the δ-function :

$$N(p)\,\delta(p^2-m^2)d^4p = N'(p')\,\delta(p'^2-m^2)d^4p'$$

is still obviously invariant.

We now integrate both sides from 0 to ∞ over p_o :

$$\int N(p)\, \delta(p^2-m^2)dp_o\, d\vec{p} = \int N'(p')\, \delta(p'^2-m^2)dp'_o\, d\vec{p}' \quad .$$

This is allowed since $p = \{p_o, \vec{p}\}$ is a time-like vector and no Lorentz transformation can ever change the sign of p_o. Therefore if p_o goes from $0 \to \infty$, p'_o does the same. Now

$$\delta(f(x)) = \sum \frac{\delta(x-x_i)}{\left|\frac{df}{dx}\right|_{x_i}}$$

where $f(x_i) = 0$. (Proof ?).

Here this gives with

$$f(p_o) = p_o^2 - \vec{p}^2 - m^2 \quad ; \qquad p_{o,i} = \pm\sqrt{\vec{p}^2+m^2} \equiv \pm E$$

$$\left|\frac{df}{dp_o}\right|_{p_{o,i}} = 2E$$

$$\boxed{\delta(p^2-m^2) = \frac{\delta(p_o+E) + \delta(p_o-E)}{2E} \quad ; \quad E = \sqrt{\vec{p}^2+m^2}} \qquad \text{(I.47)}$$

/A formula which one should know - at least know how to find it !/ With this in the integrals /which go over p_o, p'_o positive/, we find

$$N(E,\vec{p})\, \frac{d\vec{p}}{E} = N'(E'\vec{p}')\, \frac{d\vec{p}'}{E'}$$

Here, e.g., the left hand side means the number of particles of mass m going into the momentum space element (p_i, p_i+dp_i) $i = 1,2,3$.

But from the derivation we know that

$$N(E,\vec{p}) = N'(E'\vec{p}')$$

hence

$$\frac{d\vec{p}}{E} = \frac{d\vec{p}'}{E'} \quad \text{or} \quad \frac{\partial(p_1 p_2 p_3)}{\partial(p_1' p_2' p_3')} = \frac{E}{E'} \qquad \text{q.e.d.} \qquad [\text{see (I.41)}]$$

Introducing another name for $N(E,\vec{p})$ we find our old formula (I.34) by putting

$$\frac{N(E,p_1 p_2 p_3)}{E} \equiv \frac{\partial^3 S(p_1 p_2 p_3)}{\partial p_1 \partial p_2 \partial p_3}$$

then

$$\frac{\partial^3 S(p_1 p_2 p_3)}{\partial p_1 \partial p_2 \partial p_3}\, dp_1 dp_2 dp_3 = \frac{\partial^3 S'(p_1' p_2' p_3')}{\partial p_1' \partial p_2' \partial p_3'}\, dp_1' dp_2' dp_3' \;.$$

We now try to obtain some intuitive feeling for the transformation of a momentum spectrum by discussing the transformation from the CM to the lab. system as a function of the CM-velocity β_{CM} in the model case where the CM spectrum consists of one single peak $p^* = p_o^*$ and is zero otherwise. The angular distribution is isotropic. (When does this model case happen ?). In this discussion we will disregard the Jacobian.

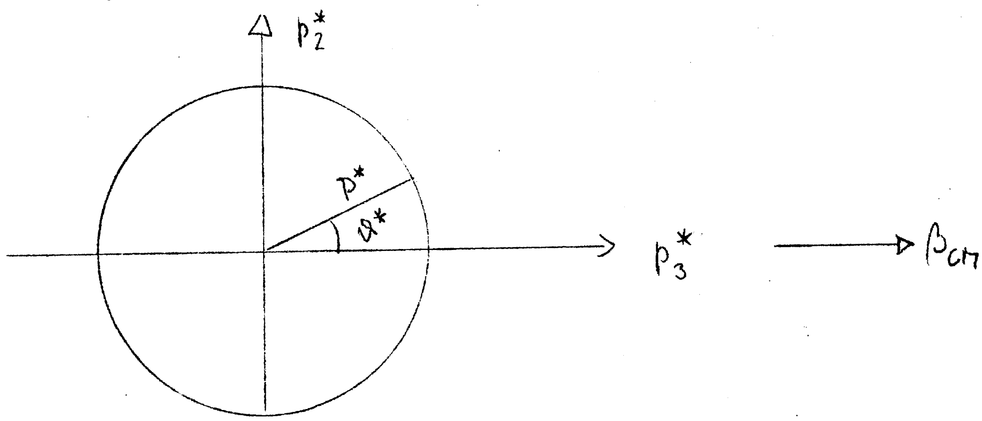


Fig. I.5

In $p^* - \vartheta^*$ polar co-ordinates the spectrum is $\neq 0$ only in a circular ring at $p^* = p_o^*$. The particle has a velocity in the CM

$$v^* = \frac{p_o^*}{E^*} = \frac{p_o^*}{\sqrt{p_o^{*2} + m^2}}$$

We distinguish three cases :

i) $\beta > v^*$

ii) $\beta = v^*$

iii) $\beta < v^*$

i) $\beta > v^*$

All particles go more or less in the forward direction in the lab.; they all have $v_z > 0$ and $\vartheta < \pi/2$, since even those which go backward in CM are bent over by the large β . The circular ring of the $p^* \vartheta^*$ plane is shifted in the $p\,\vartheta$-plane so far that the origin $p = 0$ lies outside. It is no longer a circular ring (why ?).

Since no particle goes backwards in the lab. system there must be a maximum angle $\vartheta_{max} < \frac{\pi}{2}$. At any given $\vartheta < \vartheta_{max}$ one observes two peaks in the lab. spectrum : one at a large momentum coming from particles going forward in CM and one with a low momentum, coming from particles going backward in CM.

For $\vartheta = 0$ these two peaks have the largest separation, giving the maximal and minimal lab. momenta; for $\vartheta = \vartheta_{max}$ they coincide but they are somewhat smeared out.

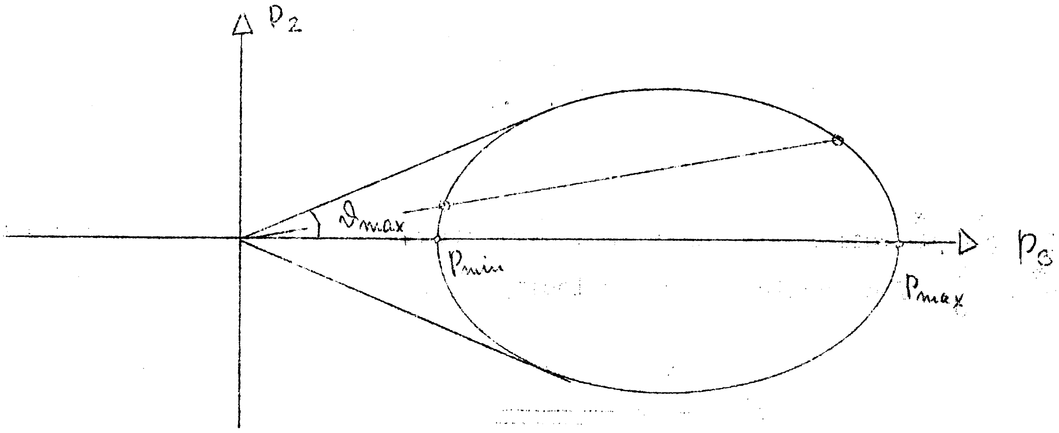


Fig. I.6

ii) $\beta = v^*$

With decreasing β the maximal lab. angle increases and reaches $\pi/2$ for $\beta = v^*$; namely, when the ring reaches the origin. Hence in this case one observes still two separated peaks in the lab. for all $\vartheta < \vartheta_{max} = \pi/2$. One of them has high energy, the other one lies at $p = 0$. If $\vartheta \rightarrow \pi/2$ the high energy peak shifts to zero, both melt together and become smeared out.

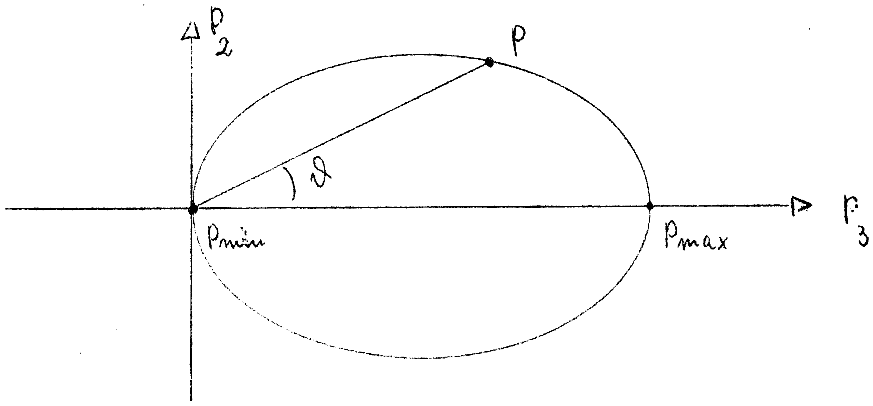


Fig. I.7

iii) $\beta < v^*$

The ring has crossed the origin $p = 0$, which now lies inside the ring. Here also in the lab. system particles can fly backward. There are no longer two peaks and no maximum angle exists.

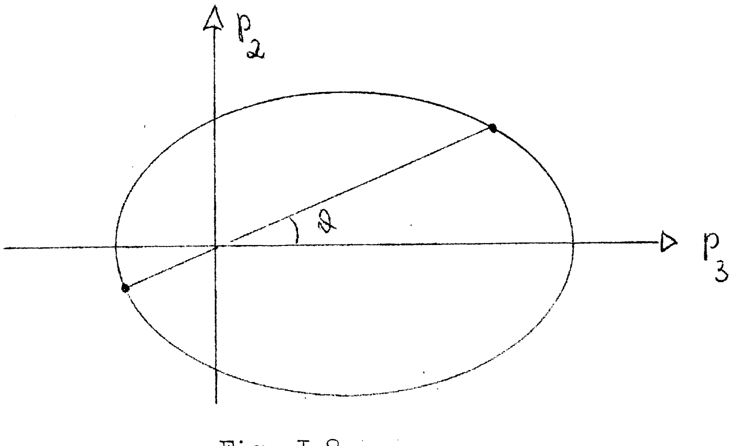


Fig. I.8

This last case happens if the particles considered have zero mass since then $v = c$ is always $> \beta$. We will discuss this later in detail.

For a general spectrum we may always imagine that it is composed of such δ-spectra which, however, are no longer isotropic in the CM system. We may still draw in the CM system ring-shaped regions (not necessarily circular) and we may apply the preceding considerations to each ring. There may, or may not be a certain part of the CM spectrum which in the lab. system appears twice (e.g., one peak in the CM will give two peaks in the lab. if it lies in that part of the CM spectrum; it will give one peak in the lab. if it does not lie there. See Fig. I.9.)

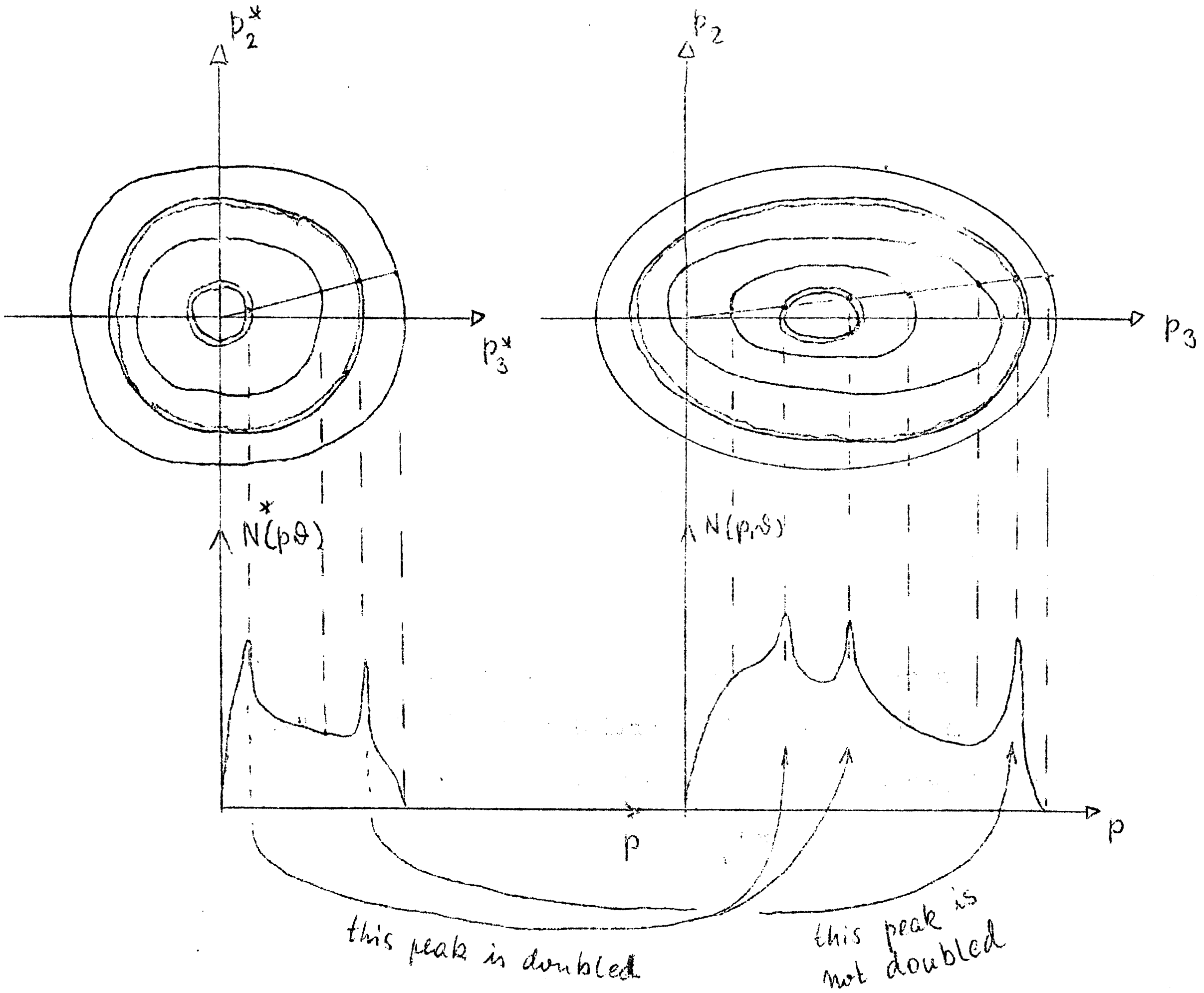


Fig. I.9

The preceding considerations do not give the transformation of the spectra - the full story is contained in formulae (I.44) -- but they give a feeling of what happens and which part of a spectrum goes where. [What we called a "ring" is of course, in space, a "shell".]

We now carry through the preceding qualitative discussion quantitatively assuming a δ-shaped isotropic distribution in the CM (spherical shell). The following questions will be answered :

(a) which form has the shell in the lab. system ?

(b) which are the two momenta for a given angle ϑ in the lab. system ? Which is the maximum angle in the lab. system and the corresponding angle in the CM ?

(c) is there a simple graphical construction for finding the two CM momenta (directions) which will appear under the same angle in the lab. system ?

We use momenta instead of velocities since the components of p transform like the space components of a four vector - which is not true for velocities. Let p^* be the magnitude of the momentum in CM.

(a) The transformation for our problem is (I.37), (I.38). Only p_3 is changed. In fact

$$p_1 = p_1^*$$
$$p_2 = p_2^* \qquad \text{(I.48)}$$
$$p_3 = \gamma\left[p_3^* + \beta E^*\right]$$

E^* is a constant. To any possible value $-p^* \leq p_3^* \leq p^*$ exists also the value $-p_3^*$ with the same p_1^* and p_2^*. If we call $\Delta p_3^* = 2\,|p_3^*|$, the corresponding length in the lab. system becomes $\Delta p_3 = \gamma \Delta p_3^*$ according to (I.48), whereas p_1^* and p_2^* remain unchanged. The spectrum, which in the CM appears to be a spherical shell with radius p^*, will appear "Lorentz enlarged" in the lab. system, namely as a rotational ellipsoid shell with half axes

$$a_1 = p^*$$
$$a_2 = p^*$$
$$a_3 = \gamma p^* .$$

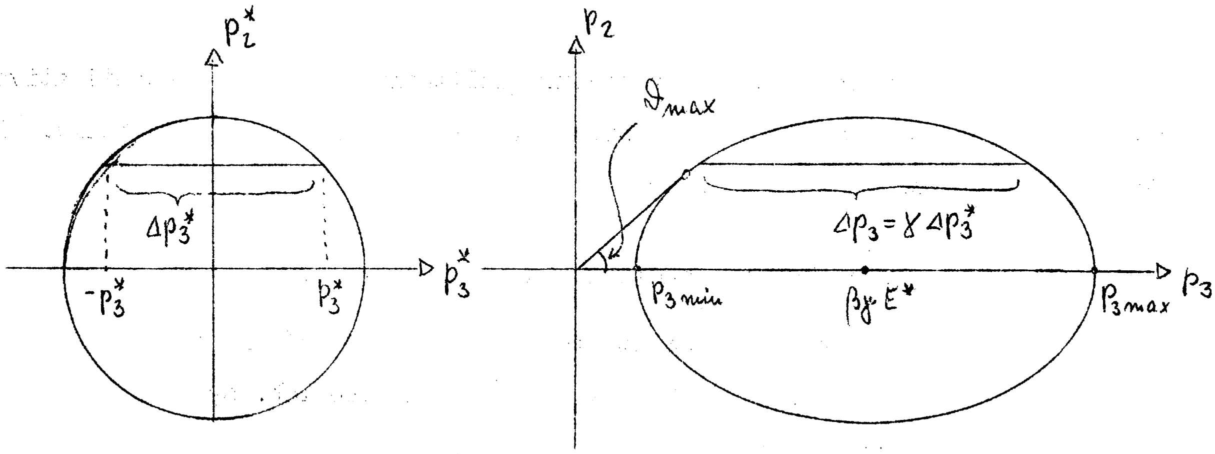


Fig. I.10

The centre of this ellipsoid lies in the middle between

$$p_{3,max} = \gamma\left[p^* + \beta E^*\right]$$

and

$$p_{3,min} = \gamma\left[-p^* + \beta E^*\right]$$

hence at

$$p_{3,centre} = \gamma\beta E^*.$$

Indeed, in the CM system we have $p_1^{*2} + p_2^{*2} + p_3^{*2} = p^{*2}$, hence with (I.48)

$$\frac{p_1^{*2}+p_2^{*2}+p_3^{*2}}{p^{*2}} = \frac{p_1^2}{p^{*2}} + \frac{p_2^2}{p^{*2}} + \frac{(p_3 - \beta\gamma E^*)^2}{\gamma^2 p^{*2}} = 1 \qquad \text{(I.49)}$$

which is the equation of the rotational ellipsoid in the lab. system. This ellipsoid touches the point $p_3 = 0$ if

$$p_{3,min} = \gamma\left[-p^* + \beta E^*\right] = \gamma p^*\left[(\beta/v^*) - 1\right] = 0$$

i.e., for $v^* = \beta$, as we found in the qualitative discussion. For $v^* > \beta$ it shifts over $p_3 = 0$ to the left and then the origin $(p_1=p_2=p_3=0)$ remains inside.

The position of the focal points follows from $\varepsilon^2 = a_3^2 - a_2^2$:

$$\varepsilon^2 = p^{*2}(\gamma^2-1) = \beta^2\gamma^2 p^{*2} .$$

Since ε is the distance of the focal points from the centre these points lie at

$$f_{1,2} = \beta\gamma\,(E^* \pm p^*) .$$

If the particles have zero mass, one focal point lies at $p_3 = 0$.

(b) We can always turn the axes such that $p_1 = 0$. Then

$$p_2 = p_3 \mathrm{tg}\,\vartheta$$

We introduce this into the equation of the ellipse $(p_1=0\ !)$ and obtain

$$p_3^2\,\mathrm{tg}^2\vartheta + \frac{(p_3 - \beta\gamma E^*)^2}{\gamma^2} = p^{*2}$$

This quadratic equation for p_3 has the two solutions

$$p_3^{(\pm)} = \frac{\beta\gamma E^* \pm \sqrt{\beta^2\gamma^2 E^{*2} - (1+\gamma^2\mathrm{tg}^2\vartheta)(\beta^2\gamma^2E^{*2} - \gamma^2 p^{*2})}}{1+\gamma^2\mathrm{tg}^2\vartheta} \qquad \text{(I.50)}$$

Since this equation involves only $\mathrm{tg}^2\vartheta$, it remains true even if the origin $p_1=p_2=p_3=0$ is inside the ellipse :

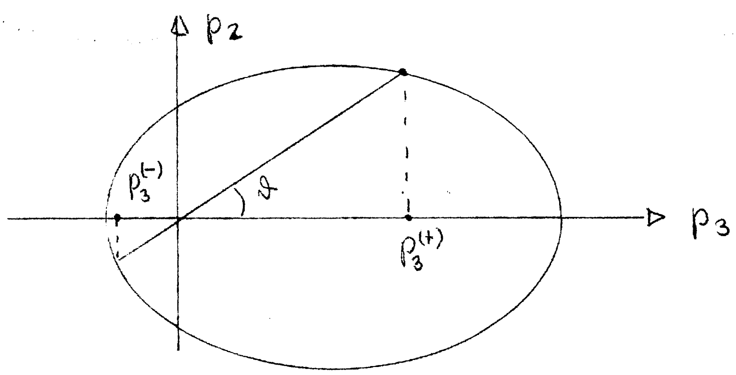


Fig. I.11

The figure shows to what the two roots correspond in that case. The maximal angle implies that the two roots of (I.50) coincide [see Fig.I.10]. Hence the square root must vanish. This gives

$$\mathrm{tg}^2 \vartheta_{max} = \frac{v^{*2}}{\gamma^2(\beta^2 - v^{*2})} \qquad \left(\text{with } v^* = \frac{p^*}{E^*}\right) \qquad (I.51)$$

As we know already from our qualitative discussion, $\mathrm{tg}\,\vartheta_{max} = \infty$ for $v^* = \beta$. [For $v^* > \beta$ there is no real solution, the ellipsoid encloses the origin $p_1 = p_2 = p_3 = 0$.] The corresponding angle $\vartheta^*(\vartheta_{max})$ can be found by the angle transformation formula (I.12), but there is a simpler method :

suppose we attach to the lab. system a fictitious particle of the same mass m; this particle is at rest in the lab. system and has momentum $-\beta\gamma m$ in the CM. We have then the following situation : the maximum angle ϑ_{max} is given by the tangent cone from the origin $p_1 = p_2 = p_3 = 0$ to the ellipsoid and in the CM system we have the corresponding tangent cone from the point $-\beta\gamma m$ to the sphere (see figure) :

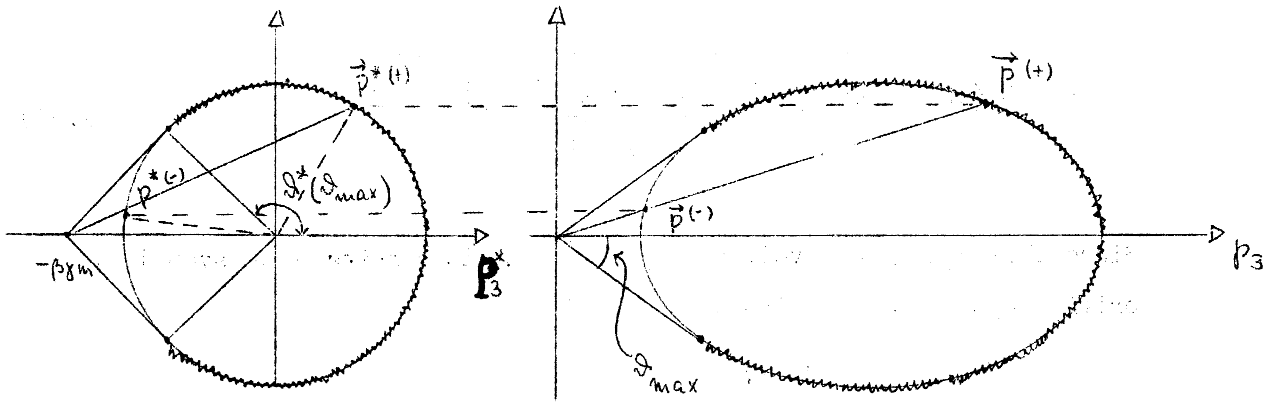


Fig. I.12

Hence :

$$-\cos\vartheta^*(\vartheta_{max}) = \frac{p^*}{\beta\gamma m} .$$

We write $p^* = \frac{v^* . m}{\sqrt{1-v^{*2}}}$ and put $\frac{1}{\sqrt{1-v^{*2}}} \equiv \gamma^*$. We then obtain

$$\cos\vartheta^*(\vartheta_{max}) = \frac{v^*\gamma^*}{\beta\gamma} . \qquad \text{(I.52)}$$

(c) This last remark leads to a very simple graphical construction [see Fig. I.12] :

draw the ellipse and the circle corresponding to lab. and CM spectra. To one angle in the lab. (straight line leaving $p_2=p_3=0$) and two corresponding momenta (the intersections of the straight line with the ellipse) correspond in the CM system t w o angles and two momenta such that for corresponding momenta $p_2 = p_2^*$ always. This method works also for $v^* > \beta$ [see Fig. I.11].

LECTURE 6

I wish to add here an application of the foregoing considerations which results in the surprising statement :

"if an observer looks at (or photographs) a fast moving object ($\beta \approx 1$) which approaches him under a small angle α of observation then, if $\alpha \gtrsim \sqrt{1-\beta^2}$, he sees no longer the frontside a of that object, but he can see the backside c !"

[See the figure below, the object is assumed to be a cube].

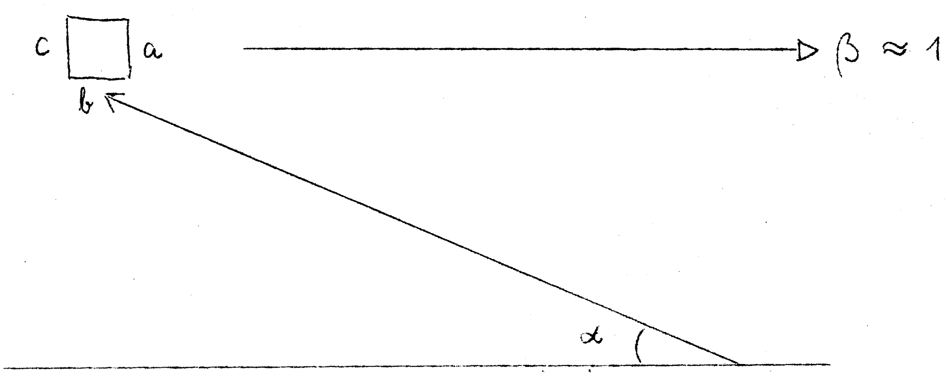


Fig. I.13

Arrangement of the observation of a fast approaching object

If $\beta \approx 1$, the critical angle $\alpha_0 \approx \sqrt{1-\beta^2}$ may be very small. If then an object appears under an angle $\alpha \gtrsim \alpha_0$, one would see practically only the frontside a, if the object would not move; whereas our fast moving object will show us no longer the frontside a but b and the backside c.

To prove this statement we have only to apply our above considerations to the transformation of spectra. Assume the whole surface of the cube covered with light sources which emit an isotropic and monochromatic radiation. Let the object be opaque, so that any radiation which has a component towards the object will be absorbed.

Now consider the radiation of one of these light-points. It has just all the properties we required in our model case for the discussion of how a spectrum transforms. We meet here the case in which the velocity v^* of the particles in the object's frame (CM) is greater than β. Hence the situation is as shown in the next figure :

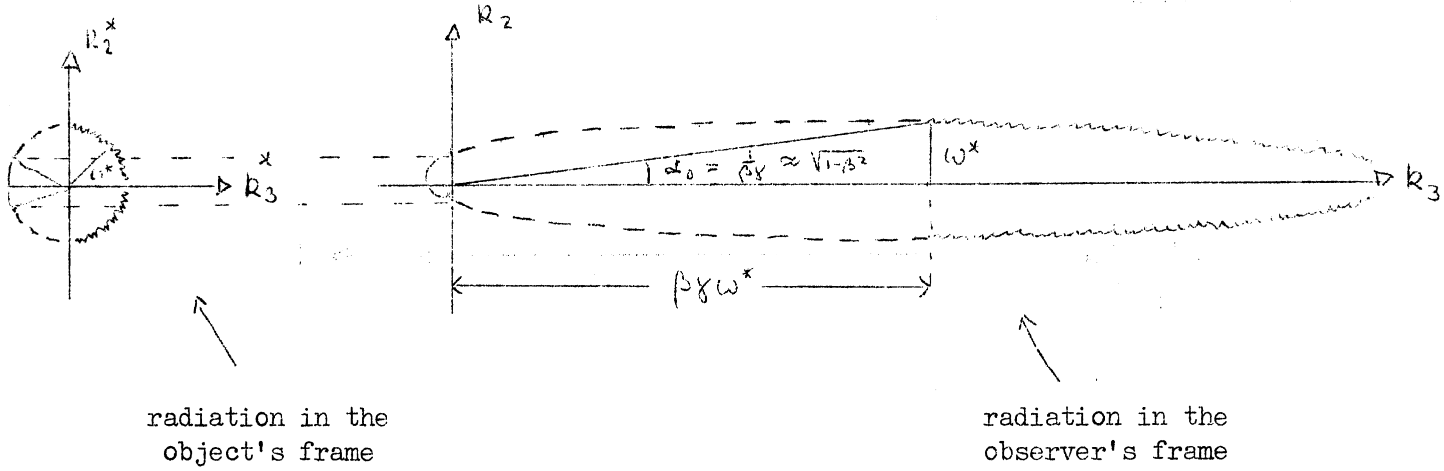


Fig. I.14

We apply our considerations to the case $p^* = E^* = \omega^*$ (the frequency of light in the object's frame).

In the lab. system the angular and momentum distribution is then given by an ellipsoid with centre at $\beta\gamma\omega^*$. Since $v^*=c=1 > \beta$, the ellipse surrounds the origin in the lab. system. All the light which in the object's frame goes forward is emitted within the angle $\alpha_0 = \frac{1}{\beta\gamma} \approx \sqrt{1-\beta^2}$, as follows immediately from the figure (indicated by a wavy line). A part of the light going backwards in the object's frame still goes forward in the lab. system, where it appears at angles between α_0 and $\pi/2$ (indicated by broken lines). Only a small part of the light emitted backwards will go backward in the lab. system too (thin line). Therefore the isotropic and monochromatic radiation as it was in the rest system of the object, appears like coming from a spotlight in the lab. system : it is sharply focussed in the forward direction and, of course, its frequency depends on the angle. Therefore, if we look at the object at a certain angle (let the object be so small or so far away that α practically does not vary over the object), then

-- we can see the radiation coming from the frontside "a" of our cube only as long as $\alpha < \alpha_0 \approx \sqrt{1-\beta^2}$;

-- we can see always the radiation coming from side "b" ;

-- we can even see the radiation coming from the backside "c" as soon as $\alpha > \alpha_0 \approx \sqrt{1-\beta^2}$.

To this one should add the Doppler effect, as can be seen from the figure. What do we then observe ? First, when the cube appears far away, we see its frontside "a" and, shortened by perspective, the side "b"; both radiating ultra-violet. Then, if α grows, the cube seems to turn and if $\alpha = \alpha_0 = \frac{1}{\beta\gamma}$, then we see only side "b", still violet. If α becomes greater than α_0, we no longer see the frontside "a" but now it has turned so far that the back-side "c" becomes visible; the colour becomes less violet. Finally, when $\alpha = \pi/2$ we see practically only the backside, radiating infra-red. The picture remains then nearly unchanged until the cube disappears ($\alpha = \pi$). This quite unexpected behaviour can be explained in many other ways too and these other ways lead to a new surprise : if one l o o k s at an object, or photographs it, then this above-described apparent rotation of the object is the only thing that happens. There will be no Lorentz contraction observed : a fast moving sphere appears as a sphere and not as a pancake. This does not mean that the Lorentz contraction did not exist; however, the Lorentz contraction takes place under the condition that the position of all points of a moving object is determined simultaneously, i.e., at one given time in the lab. system, whereas "seeing" or "photographing" supposes that the light pulses coming from a moving object do not leave it at one given time, but instead arrive at the eye (or shutter) of the observer at the same time. This condition implies that they left the different points of the object at different times against the supposition under which the Lorentz contraction is derived [see p.13]. For further details see : V. Weisskopf, Phys. Today 13, N°.9, 24 (1960).

It is left to the reader to discuss the even more curious case where we assume the object to be covered with a β-radiating material and to be photographed with a β-ray sensitive camera. To make it a true science fiction assume that the β-rays are monochromatic in the object's frame. Suppose then that the velocity of the object is greater than that of the electrons in the object's frame.

6) Variables and co-ordinate systems frequently used in elastic scattering

We shall discuss here some notations and techniques which have become usual in recent work on scattering, in particular in discussions on dispersion relations and the Mandelstam representation.

Consider an elastic scattering event and define the momenta before and after scattering, as shown in Fig. I.15 :

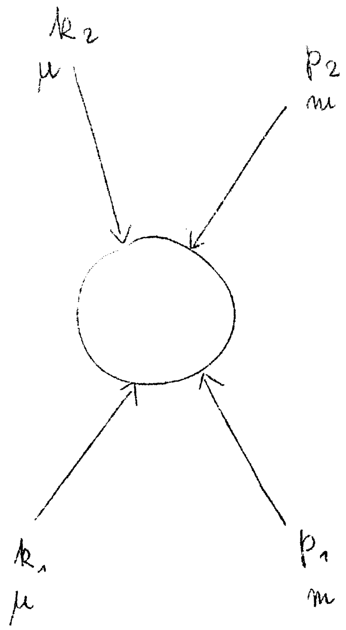


Fig. I.15

We use the convention that all four momenta are ingoing. This has the advantage that one may consider any two of the four to be the incoming particles and the other two as the outgoing ones; the physical momentum of the outgoing particles is then the negative of that one which we use in the present formulation. We shall, however, denote the physical momentum of an outgoing particle by a prime; if, e.g., in Fig. I.15 the particle corresponding to the arrow with k_2 is outgoing, then we call its physical momentum $k_2' = -k_2$.

(a) The independent variables of the scattering process

We shall disregard spin, isospin and excitation of the particles. Then the two pairs of four momenta p_1, k_1 and p_2, k_2 completely determine the initial and final states respectively. The transition amplitude - which describes quantum mechanically the process -- can therefore depend only on these four four-vectors :

$$T_{fi} = T(p_1, k_1, p_2, k_2) \ . \tag{I.53}$$

It seems therefore to depend on 16 variables - namely all the components - but we shall show that these are not independent and, therefore, the number of variables reduces to two only :

* the amplitude T has to be a Lorentz invariant quantity. In fact, its square gives the probability to find a state "f" when the initial state was "i". This probability cannot depend on the Lorentz system of the observer. Therefore T must depend on the invariants which one can construct out of the involved four momenta :

$$p_1^2,\ k_1^2,\ p_2^2,\ k_2^2,\quad p_1k_1 \quad p_1k_2 \quad p_1p_2 \quad k_1k_2 \quad k_1p_2 \quad k_2p_2$$

* these ten invariants are not all useful for the description :

$$p_1^2 = p_2^2 = m^2 \qquad \text{and} \qquad k_1^2 = k_2^2 = \mu^2$$

are fixed parameters which we need not mention as variables.

* the remaining six invariants are indeed variables which can be used to describe the scattering process. They are, however, not independent : four-momentum conservation requires

$$p_1+k_1+p_2+k_2 = 0 \quad .$$

This four-vector equation is equivalent to four simple equations; therefore the number of variables reduces to 6 minus 4, namely 2 independent ones.

It is not arbitrary which two we select : taking, e.g. :

$$p_1k_2 \quad \text{and} \quad p_2k_1$$

we would make an impossible choice. Namely, by multiplying

$$p_1+k_1+p_2+k_2 = 0 \qquad (I.54)$$

by k_1, k_2, p_1, p_2 respectively, we obtain the four equations

$$\begin{aligned} k_1(p_1+p_2) &= -\mu^2 - k_1k_2 \\ k_2(p_1+p_2) &= -\mu^2 - k_1k_2 \\ p_1(k_1+k_2) &= -m^2 - p_1p_2 \\ p_2(k_1+k_2) &= -m^2 - p_1p_2 \quad . \end{aligned} \qquad (I.55)$$

From the first and the second pair of equations, follows

$$\begin{aligned} (k_1-k_2)(p_1+p_2) &= 0 \\ (p_1-p_2)(k_1+k_2) &= 0 \quad . \end{aligned} \qquad (I.56)$$

Adding and subtracting these equations results in :

$$\begin{aligned} k_1p_1 &= k_2p_2 \\ k_1p_2 &= k_2p_1 \quad . \end{aligned} \qquad (I.57)$$

Thus, if we use p_1k_2 as one variable, then k_1p_2 cannot serve as the second one since it is identical with the first one. Similarly one can use either k_1p_1 or k_2p_2, but not both as the variables of the process.

Adding the first pair of (I.55) one obtains

$$(k_1+k_2)(p_1+p_2) = -2\mu^2 - 2k_1k_2$$

while the second pair gives

$$(p_1+p_2)(k_1+k_2) = -2m^2 -2p_1p_2 \quad ,$$

hence

$$\begin{aligned} -\tfrac{1}{2}(k_1+k_2)(p_1+p_2) &= \mu^2 + k_1k_2 = m^2 + p_1p_2 \\ \tfrac{1}{2}(k_1+k_2)^2 = \tfrac{1}{2}(p_1+p_2)^2 &= \mu^2 + k_1k_2 = m^2 + p_1p_2 \end{aligned} \tag{I.58}$$

where $k_1+k_2 = -(p_1+p_2)$ has been used. If we combine this with (I.56), we see then that

$$\begin{aligned} -\tfrac{1}{2}(k_1+k_2)(p_1+p_2) &= - k_1(p_1+p_2) = -k_2(p_1+p_2) \\ &= - p_1(k_1+k_2) = -p_2(k_1+k_2) \quad , \end{aligned}$$

hence

$$-k_1(p_1+p_2) = -k_2(p_1+p_2) = -p_1(k_1+k_2) = -p_2(k_1+k_2) = \mu^2+k_1k_2 = m^2+p_1p_2 \ . \tag{I.59}$$

Therefore k_1k_2 and p_1p_2 cannot serve at the same time as variables.

We can now check explicitly whether we really retain only two variables. We had 6 useful invariants :

$$p_1k_1,\ p_1k_2,\ p_1p_2,\ k_1k_2,\ k_1p_2,\ k_2p_2 \tag{I.60}$$

Assume we select the first one, then p_2k_2 drops out (I.57). Take furthermore the second one, then k_1p_2 drops out (I.57). We are left with

$$p_1k_1,\ p_1k_2 \ \text{ and } \ p_1p_2 \ .$$

However, from (I.59), it follows that

$$p_1p_2 = -m^2 - p_1k_1 - p_1k_2 \quad .$$

2658

Therefore the variable p_1p_2 is a linear combination of the first two and gives nothing new. Our particular choice leads thus to two independent variables, namely

$$p_1k_1 \quad \text{and} \quad p_1k_2 \quad .$$

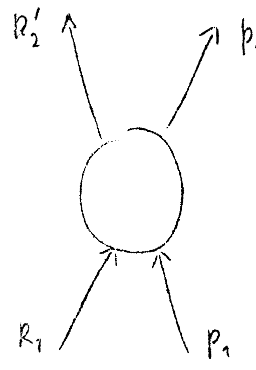
We could of course have taken two other independent invariants; the most general choice consists of any two independent linear combinations of the six invariants (I.60).

Equations (I.55) through (I.59) are sometimes useful in calculations arising in a change of variables.

LECTURE 7

(b) Useful Lorentz systems for the description of the scattering process

Particular Lorentz systems become preferable if, by their use, convenient variables assume simple forms and/or if certain symmetries are exhibited. We therefore expect useful Lorentz systems in the cases where the 3-momentum of one particle or the sum of 3-momenta of two particles vanish :



(α) $\vec{k}_1$ or $\vec{p}_1$ vanishes in the lab. system (if k_1 or p_1 refers to the target particle).

(β) $\vec{k}_1 + \vec{p}_1$ and $\vec{k}_2' + \vec{p}_2'$ vanish in the CM system.

(γ) $\vec{k}_1 + \vec{k}_2'$ or $\vec{p}_1 + \vec{p}_2'$ vanishes in the "Breit"-system.

These are the most natural choices. The CM system exhibits the highest degree of symmetry. Choices in which $\vec{k}_1+\vec{p}'_2$ or $\vec{p}_1+\vec{k}'_2$ vanishes seem not to be useful since the "k" and "p" particle will generally have different masses. If $m=\mu$, however, this reduces to case (γ).

Choices where the difference of two momenta vanishes, are partly impossible and partly not useful :

* impossible if the two momenta belong to equal masses. If, e.g., $\vec{k}_1-\vec{k}'_2 = 0$, then $|\vec{k}_1| = |\vec{k}'_2|$, hence $\omega_1 = \omega'_2$ and $(k_1-k'_2) = (\omega_1-\omega'_2, \vec{k}_1-\vec{k}'_2) = (0,0)$. In other words : if such a system exists, then there is no scattering; or : if there is scattering, then such a system does not exist.

* not useful if the two momenta belong to different masses.

When we say "not useful", we mean that hitherto such Lorentz systems have not shown practical importance. They may, perhaps, do so in particular cases. Disregarding them here, we are left essentially with the three choices : (α), (β) and (γ) mentioned above which we shall discuss now.

(α) <u>The laboratory system</u>

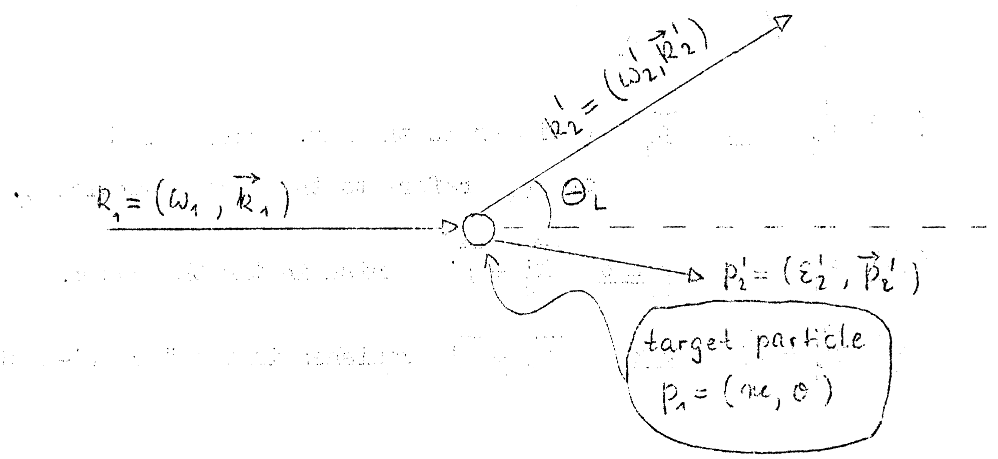


Fig. I.16
The lab. system

This is the system where one of the incoming particles is at rest. Let this be the "p_1"-particle. We then use the notation :

$$p_1 = (m,0) \qquad k_1 = (\omega_1, \vec{k}_1)$$

$$p_2' = (\varepsilon_2', \vec{p}_2') \qquad k_2' = (\omega_2', \vec{k}_2')$$

Useful variables are ω_1, ω_2', ε_2' and $\cos\Theta_L$, where Θ_L is the angle between the direction of the incoming and outgoing "k"-particle. It is easy to express these variables in an invariant form using the procedure explained on p. 27. From the fact that the three-momentum $\vec{p}_1 = 0$, it follows at once that

$$\omega_1^{(L)} = (p_1k_1)\cdot\frac{1}{m} = \text{lab. energy of the incoming particle}$$

$$\omega_2'^{(L)} = (p_1k_2')\cdot\frac{1}{m} = \text{lab. energy of the scattered particle}$$

$$\varepsilon_2'^{(L)} = (p_1p_2')\cdot\frac{1}{m} = \text{lab. energy of the target particle after the scattering.}$$

(I.61)

Since m is an invariant, we have already expressed these variables in an invariant form.

The scattering angle Θ_L will follow from the scalar product of k_1 and k_2', namely

$$k_1k_2' = \omega_1\omega_2' - |\vec{k}_1|\cdot|\vec{k}_2'|\cdot\cos\Theta_L$$

With $|\vec{k}| = \sqrt{\omega^2-\mu^2}$ one finds

$$\cos\Theta_L = \frac{\omega_1\omega_2' - k_1k_2'}{\sqrt{(\omega_1^2-\mu^2)(\omega_2'^2-\mu^2)}} = \frac{(k_1p_1)(k_2'p_1) - m^2(k_1k_2')}{\sqrt{\left[(k_1p_1)^2-m^2\mu^2\right]\left[(k_2'p_1)^2-m^2\mu^2\right]}}$$

2658 (I.62)

Here the ω 's have been expressed invariantly by (I.61); this formula may be transformed into other expressions by means of formulae (I.55) through (I.59).

(β) The centre-of-momentum system

This system exhibits the symmetries of two-body kinematics most clearly.

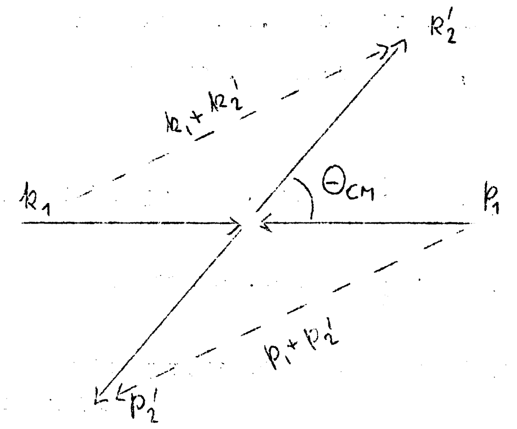


Fig. I.17
The CM system

The CM system is defined by

$$\vec{k}_1+\vec{p}_1 = \vec{k}_2'+\vec{p}_2' = 0 \ .$$

Therefore

$$|\vec{k}_1| = |\vec{p}_1| = K$$

and

$$|\vec{k}_2'| = |\vec{p}_2'| = K' \ .$$

But K and K' are equal for the following reason :

the CM energy is given by [notation : $p \equiv (\varepsilon, \vec{p})$; $k \equiv (\omega, \vec{k})$]

$$E_{CM}^2 = (p_1+k_1)^2 = (p_2'+k_2')^2.$$

Evaluated in the CM system this becomes

$$E_{CM}^2 = (\varepsilon_1+\omega_1)^2 = (\varepsilon_2'+\omega_2')^2$$

or

$$E_{CM}^2 = \left[\sqrt{K^2+m^2} + \sqrt{K^2+\mu^2}\right]^2 = \left[\sqrt{K'^2+m^2} + \sqrt{K'^2+\mu^2}\right]^2$$

hence $K = K'$. We could also have invoked our formula (I.22), p.35, which states that the momentum K is a unique function of the CM energy E_{CM} and of the masses of the particles. [There the notation is different : $K \longleftrightarrow p^*$, $E_{CM} \longleftrightarrow M$; $m \longleftrightarrow m_1$; $\mu \longleftrightarrow m_2$ are the correspondences.]

[This conclusion does not hold if the outgoing particles have masses different from the incoming ones.]

From $K = K'$ follows at once

$$\begin{aligned} \varepsilon_1 = \varepsilon_2' = &= \sqrt{K^2+m^2} \\ \omega_1 = \omega_2' = &= \sqrt{K^2+\mu^2} \end{aligned} \qquad (I.63)$$

Frequently E_{CM}^2 is called s and used as a variable. The second variable cannot be K since it is uniquely related to s by means of Eq. (I.22) :

$$K^2 = \frac{\left[s-(m+\mu)^2\right]\left[s-(m-\mu)^2\right]}{4s}$$ = magnitude of the CM momentum of either particle before and after scattering.

$$s = (p_1+k_1)^2 = (p_2'+k_2')^2 = (p_1+k_1)(p_2'+k_2') = \left[\sqrt{K^2+m^2} + \sqrt{K^2+\mu^2}\right]^2 = \left[\text{CM Energy}\right]^2 \qquad (I.64)$$

2658

Either of these - but not both - may be used as variables. The scattering angle might be conveniently taken to be the second variable : it follows from $k_1 k'_2$, but more conveniently from

$$(k_1-k'_2)^2 = 2\mu^2 - 2k_1k'_2 = 2\left[\mu^2-\omega^2+K^2\cos\theta_{CM}\right]$$

with $\omega^2 = K^2+\mu^2$ we have

$$\begin{cases} t \equiv (k_1-k'_2)^2 = 2K^2(\cos\theta_{CM}-1) = \dfrac{\left[s-(m+\mu)^2\right]\left[s-(m-\mu)^2\right]}{2s}\left[\cos\theta_{CM}-1\right] \\ \cos\theta_{CM} = 1+\dfrac{t}{2K^2} = 1+\dfrac{2ts}{\left[s-(m+\mu)^2\right]\left[s-(m-\mu)^2\right]} \end{cases} \tag{I.65}$$

where we introduced the frequently used notation $t = (k_1-k'_2)^2$.

Note the high symmetry of the process in the CM system : all magnitudes of momenta are equal and the individual energies conserved.

(γ) The Breit system (brick-wall)

This system is also useful as it exhibits symmetries. We apply such a Lorentz transformation that $\vec{k}_1+\vec{k}'_2 = 0$ [see Fig. I.1].
Therefore k_1 and k'_2 will have the form

$$k_1 = (\omega, \vec{k}) \qquad k'_2 = (\omega, -\vec{k}) \tag{I.66}$$

From energy conservation follows then that the energies of the "p" particle before and after the collision, ε_1 and ε'_2 must be equal; hence $|\vec{p}_1| = |\vec{p}'_2|$ and

$$p_1 = (\varepsilon, \vec{p}_1) \qquad \text{with} \quad |\vec{p}_1| = |\vec{p}_2'| = p = \sqrt{\varepsilon^2 - m^2} \qquad \text{(I.67)}$$
$$p_2' = (\varepsilon, \vec{p}_2')$$

From $k_1 - k_2' = p_2' - p_1$ follows

$$k_1 - k_2' = (0, 2\vec{k}) = p_2' - p_1 = (0, \vec{p}_2' - \vec{p}_1) \qquad \text{(I.68)}$$

$2\vec{k}$ is the "three-momentum transfer".

Equations (I.66), (I.67) and (I.68) yield the following picture. Both particles seem to be reflected on a hard wall, the "k" particle perpendicularly

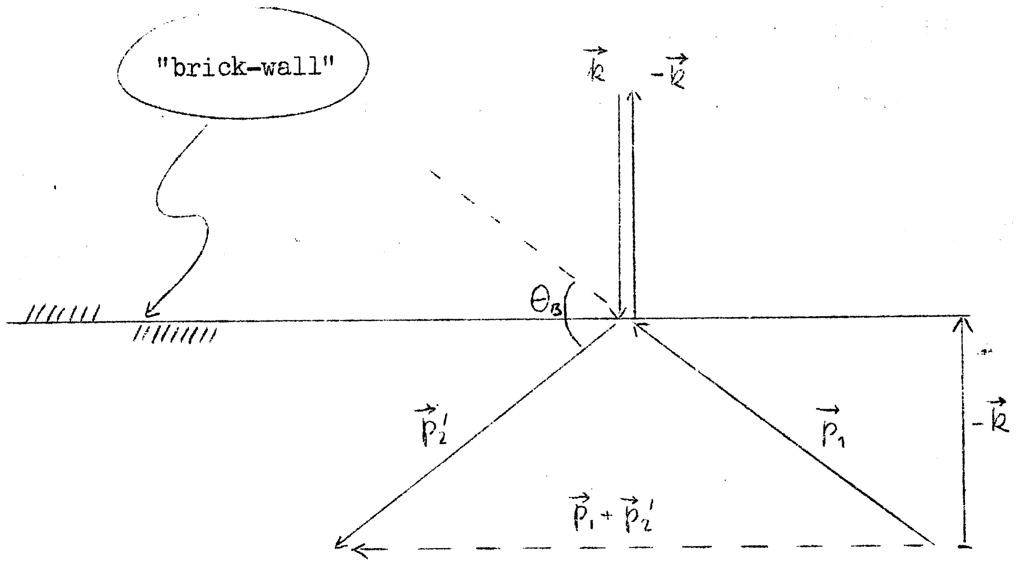


Fig. I.18
The Breit system

Indeed $(p_1 + p_2') \cdot 2k = (p_1 + p_2')(p_2' - p_1) = p^2 - p^2 = 0.$

Whereas in the CM system the energies ε and ω - of the "p" and "k" particles respectively - were not independent; they are so in the present system. Therefore they are convenient as variables. We shall express them invariantly. This is easily done by noting that

$$k_1+k_2' = (2\omega, 0) .$$

Hence

$$\omega_B^2 = \tfrac{1}{4}(k_1+k_2')^2 \qquad (I.69)$$

and

$$\varepsilon_B = \frac{1}{2}\,\frac{(p_1+p_2')\cdot(k_1+k_2')}{\sqrt{(k_1+k_2')^2}} \quad . \qquad (I.70)$$

In this system the variable t is very simple :

$$t = (k_1-k_2')^2 = (0,2\vec{k})^2 = -|2\vec{k}|^2 = \text{square of the 3-momentum transfer.} \qquad (I.71)$$

The scattering angle for the "k" particle is 180° by definition, that of the "p" particle is found from

$$(p_1-p_2')^2 = 2p^2(\cos\theta_B-1) .$$

But $(p_1-p_2')^2 = (k_1-k_2')^2 = t$, hence

$$\left.\begin{array}{l} \left.\begin{array}{l} \cos\theta_B = 1 + \dfrac{t}{2p^2} \\ t = (k_1-k_2')^2 = (p_1-p_2')^2 = 2p^2(\cos\theta_B-1) \end{array}\right\} p^2 = \varepsilon^2-m^2 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [\text{see (I.70)}] . \end{array}\right\} \qquad (I.72)$$

We may express $\omega_B^2 = |\vec{k}|^2+\mu^2$ by means of t :

$$\omega_B^2 = \mu^2 + \frac{t}{4} \qquad (I.73)$$

LECTURE 8

(c) The variables s, t, u

In this section we shall use again, as in (a), the momenta p_2 and k_2 thus leaving open which particles are incoming and which are outgoing.

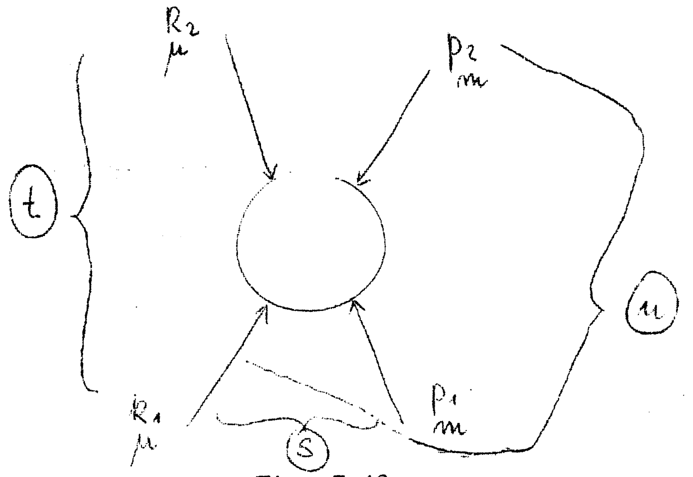


Fig. I.19

Definition of s,t,u

In the foregoing discussion we found already that s and t were useful variables and we shall define a third one, u, though all three are no longer independent of each other :

$$s = (k_1+p_1)^2 = (k_2+p_2)^2 = -(k_1+p_1)(k_2+p_2)$$

$$t = (k_1+k_2)^2 = (p_1+p_2)^2 = -(k_1+k_2)(p_1+p_2) = 2(\mu^2+k_1k_2) = 2(m^2+p_1p_2);$$

$$u = (k_1+p_2)^2 = (p_1+k_2)^2 = -(k_1+p_2)(p_1+k_2) \qquad \text{[see (I.58)]} \tag{I.74}$$

The physical significance of these variables can be expressed in two ways :

i) s is the square of the CM energy if k_1 and p_1 or k_2 and p_2 are incoming

t " " " " " " " if k_1 and k_2 or p_1 and p_2 are incoming

u " " " " " " " if k_1 and p_2 or k_2 and p_1 are incoming

This is a rather artificial description since each variable is defined by another process. The three processes in which s,t,u are the squared CM energies, are called the "s-, t-, u- channel" respectively.
To have an example : let $k_{1,2}$ describe pions, $p_{1,2}$ nucleons. Then the

s-channel means $\pi+N \rightarrow \pi+N$ or $\pi+\bar{N} \rightarrow \pi+\bar{N}$

t-channel means $\pi+\pi \rightarrow N+\bar{N}$ or $N+\bar{N} \rightarrow \pi+\pi$

u-channel means $\pi+N \rightarrow \pi+N$ or $\pi+\bar{N} \rightarrow \pi+\bar{N}$

ii) if we describe the meaning of s,t,u in a definite process, e.g., the "s-channel", then

s is the square of the CM energy;

t is the squared four-momentum transfer. In particular it reduces to the squared three-momentum transfer in the Breit system (I.71);

u has no simple physical meaning since there is no Lorentz system where it reduces to anything obvious. This is a consequence of being the difference of physical momenta of particles with different masses [see remark on p.74, under (b)].

As s,t,u are not independent, we can write down a relation. From (I.74) :

$$s+t+u = 4\mu^2+2m^2+2k_1\underbrace{(p_1+p_2+k_2)}_{=-k_1} = 2\mu^2+2m^2$$

$$s+t+u = 2\mu^2+2m^2 \quad . \tag{I.75}$$

Let us anticipate the notion of the scattering amplitude, namely, that complex function which completely describes the scattering process. It will be a function of two independent invariants, but we may write it as a function of s,t,u if we only keep in mind that one of these variables is redundant. Let then

$$T(s,t,u) = \text{scattering amplitude.} \tag{I.76}$$

One can prove - independently of perturbation theory - that this function is an analytic function of any two of the variables if these are considered to be complex.

There are then certain domains in the complex st- (or su- or tu-) space in which these variables become real and have "physical" values. These regions - as we shall see - are disconnected and belong to different physical processes, namely the three processes described on p. 82 under point i). That T(s,t,u) is an analytic function of any two complex variables out of s,t,u means then that the "physical scattering amplitude" is the boundary value of that general function when s,t,u take on physical values. In other words : the "physical scattering amplitude" is obtained in any channel from the general function simply by specializing to the "physical values" of s,t,u for that channel. As the analytic functions are essentially determined by their singularities, it is important to know the singularities of T(stu). We are far from knowing their structure. In recent work on strong interactions, a conjecture of Mandelstam about these singularities has been widely applied.

Though it has led to very intuitive descriptions of strong interaction processes and has supplied us with a new technique, it remains a conjecture. This does not exclude its value in practical limited calculations. Namely, it may turn out one day when we know more about these things, that Mandelstam's ansatz neglected singularities which in many calculations show only little influence.

We shall not go into the "analytic structure of scattering amplitudes" (which would require a special lecture), but only explain the graphical representation of the variables s,t,u and exhibit their "physical regions".

We remember from elementary geometry a theorem on triangles [Fig. I.20]

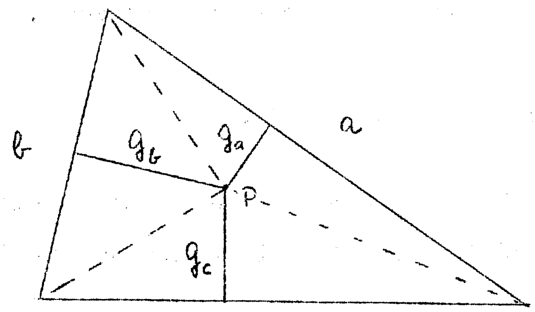


Fig. I.20

If from any point P the three distances g_a, g_b, g_c to the sides a, b, c respectively are taken, then

$$ag_a + bg_b + cg_c = ah_a = bh_b = ch_c = 2F \quad ,$$

namely two times the surface of the triangle. [This is also true for points outside the triangle if proper care of the sign of the distances g_i is taken.] Here h_a, h_b, h_c are the three hights perpendicular on a,b,c respectively. Taking ch_c and dividing by c we have

$$\frac{a}{c} g_a + \frac{b}{c} g_b + g_c = h_c \ .$$

With this we compare (I.75)

$$u + s + t = 2m^2 + 2\mu^2$$

and see that we only need to identify

$$\frac{a}{c} g_a = u \; ; \quad \frac{b}{c} g_b = s \; ; \quad g_c = t \; ; \qquad h_c = 2m^2+2\mu^2$$

to have our relation between s,t,u fulfilled. Therefore any three co-ordinate axes intersecting such that they form a triangle with $h_c = 2m^2+2\mu^2$ can serve to represent s,t,u in a plane. Of course one chooses particular triangles where the representation becomes simple.

The best choice seems to be a=b=c; $h = 2m^2+2\mu^2$ (Fig. I.21a). This is very symmetrical but has one disadvantage : the boundaries of the "physical regions" will be given in the form of equations between s and t. Such curves are easier to draw in a rectangular co-ordinate system.

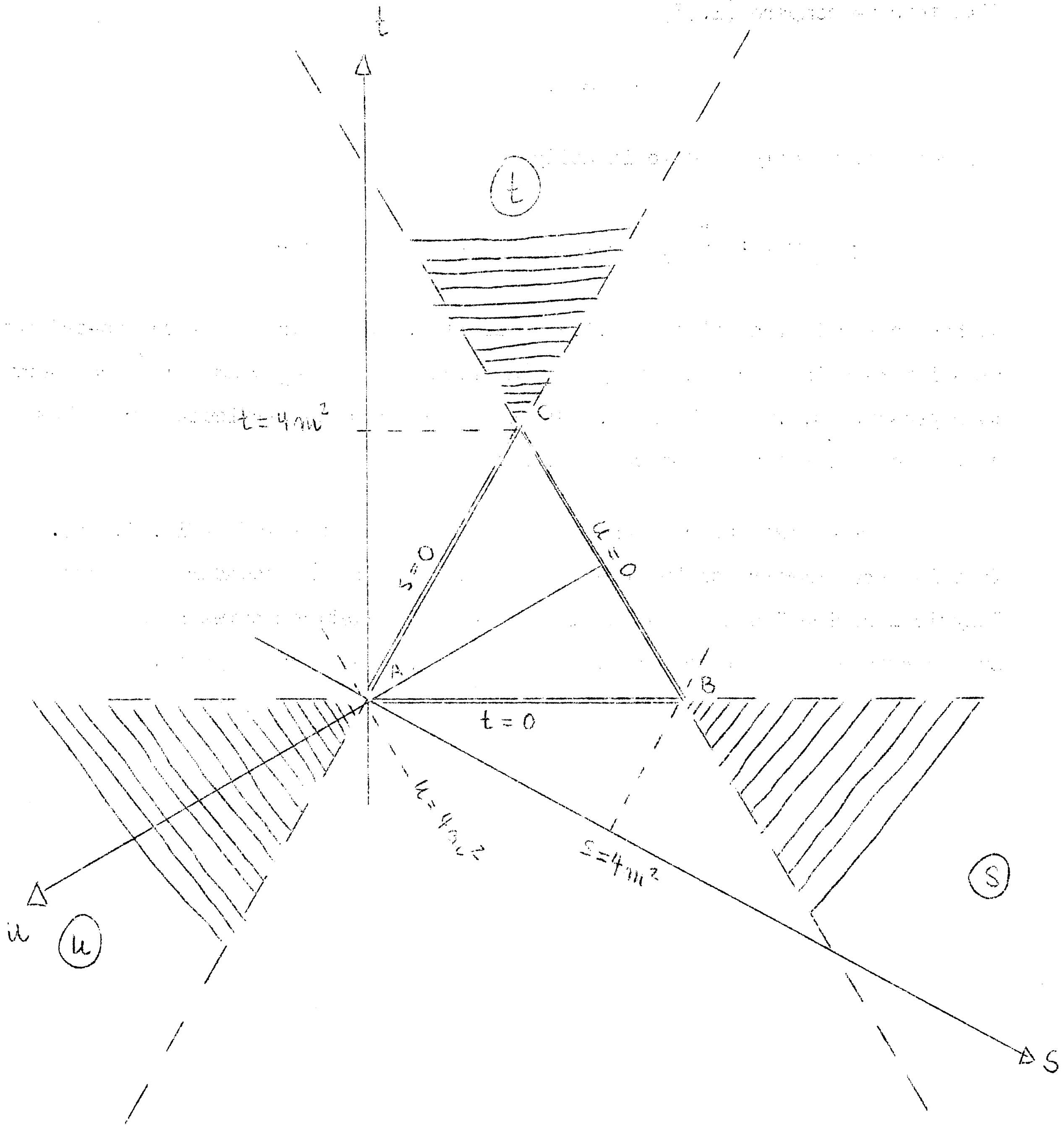


Fig. I.21 a

Physical regions of s,t,u-channels in symmetrical representation for $m = \mu$. Every point in the plane satisfies
$s+t+u = 4m^2$

In the latter case we choose

$$b = c = \frac{a}{\sqrt{2}} = h = 2m^2+2\mu^2 \qquad [Fig. 21.b]$$

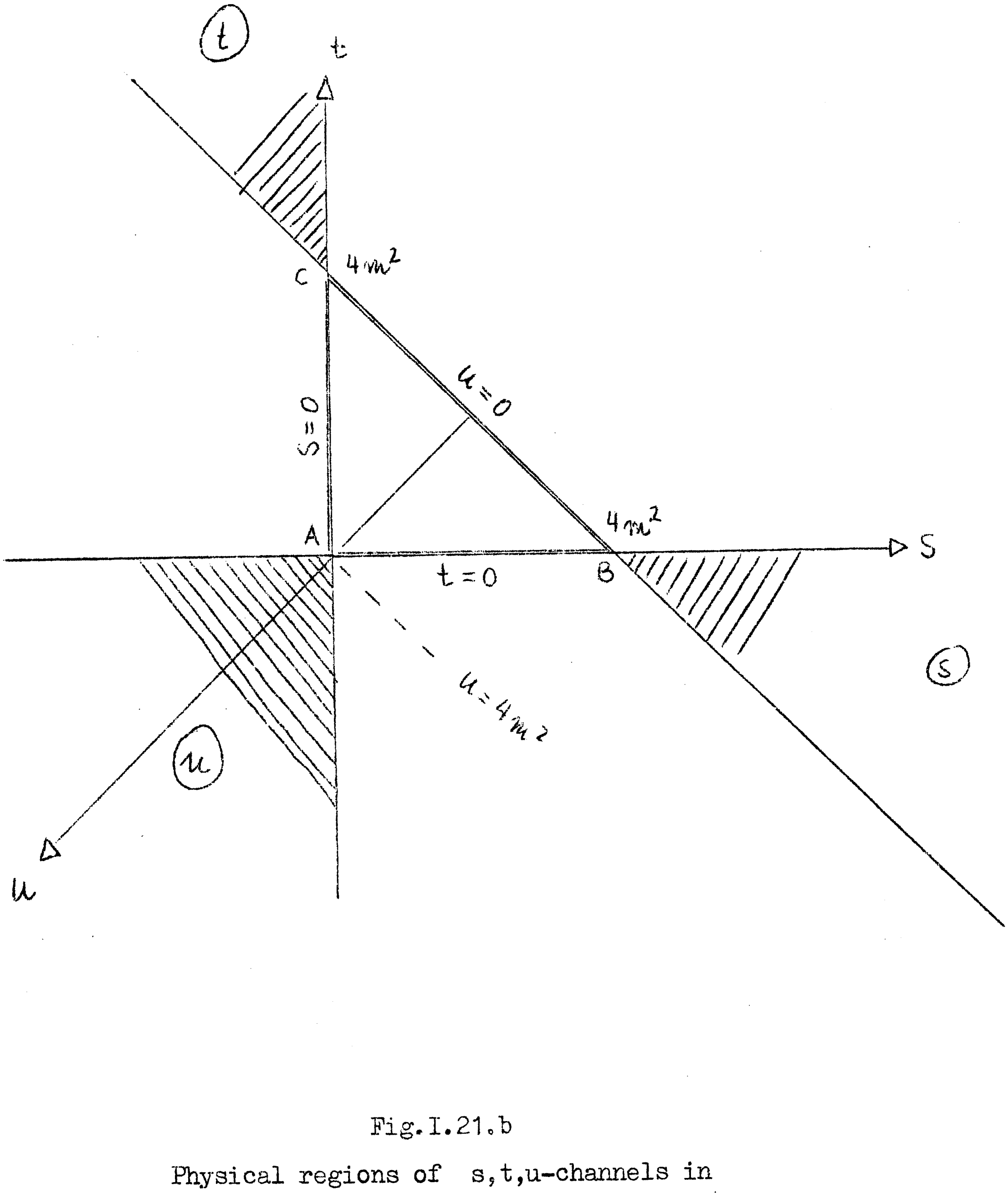


Fig. I.21.b

Physical regions of s,t,u-channels in cartesian s-t-plane for $m = \mu$. Every point in the plane satisfies $s+t+u=4m^2$ (Note that the unit along the u-axis is smaller by a factor $1/\sqrt{2}$ as compared to the s- and t-axes.)

Let us find the "physical regions" of s,t,u in the three possible channels. We draw first three figures : Fig. I.22s, I.22t, I.22u

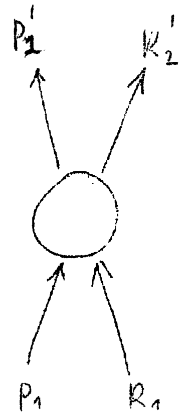


Fig. I.22s
s– channel

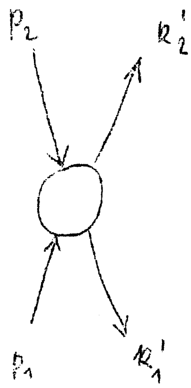


Fig. I.22t
t – channel

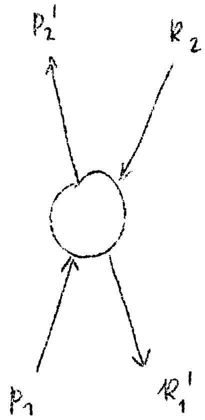


Fig. I.22u
u – channel

We see immediately that even for $m \neq \mu$ there is one symmetry : namely, that t is the momentum transfer in both – the s-channel and the u-channel – whereas s and u are interchanged; we expect therefore that the physical regions in the s and u channels map on each other if s and u are interchanged (this is the famous "crossing-symmetry); it will show up most clearly in the symmetrical representation – /Fig. I.21.a, see also Fig. I.23.a/.

If the masses are equal : $m = \mu$, then there is more symmetry :

going from the s-channel to the t-channel, u keeps its meaning
going from the s-channel to the u-channel, t keeps its meaning
going from the t-channel to the u-channel, s keeps its meaning.

The physical regions are therefore mapped on each other if we :

i) interchange $s \longleftrightarrow t$ and keep u ;
ii) interchange $s \longleftrightarrow u$ and keep t ;
iii) interchange $t \longleftrightarrow u$ and keep s .

These are, in the symmetrical representation /Fig. I.21.a/, the reflections of the whole plane with respect to the three symmetry axes of the basic triangle ABC. In Fig. I.21.b the different scale along the axes makes the figure apparently less symmetric, but one easily translates the physical regions from Fig. I.21.a to Fig. I.21.b.

This symmetry allows us to discuss the s-channel only. We shall restrict the consideration to the most symmetric case $m = \mu$. We have in the s-channel (CM-system)

$$s = (p_1+k_1)^2 = (p_2'+k_2')^2 = (2E)^2 = 4(m^2+K^2)$$
$$t = (p_2'-p_1)^2 = (k_2'-k_1)^2 = 2K^2(\cos\theta_{CM} - 1)$$

with K the momentum of all four particles. Hence the "physical region" in the s-channel is given by

$$4m^2 \leq s$$

$$t_{min} \leq t \leq 0 \quad ; \qquad t_{min} = -4K^2 = 4m^2-s \; .$$

With $s+t+u = 4m^2$ we find $s+t_{min}+u = 4m^2 = s+t_{min}$. Hence the boundary $t_{min} = 4m^2-s$ is identical with the line $u = 0$. The physical region of the s-channel is therefore given by the two conditions :

$$t \leq 0$$
$$u \leq 0 \; .$$

This region is shown in Figs. I.21.a and I.21.b shaded and marked (s) . The corresponding regions for the other two channels follow from the above symmetry considerations. The case $m > \mu$ will by treated in Problem 9 below.

Only in the physical regions can physical measurements yield information about the scattering amplitude. A remarkable circumstance is that the different channels correspond to very different processes :

assume that p and k mean nucleon and pion, respectively. Then

s-channel : π-N scattering (elastic);

t-channel : $N+\bar{N} \longrightarrow 2\pi$ [or $\pi + \pi \longrightarrow N+\bar{N}$];

u-channel : $\bar{\pi}$-N scattering (crossed process with respect to s-channel).

All these processes, different as they are, become "one and the same" if one considers the whole complex stu-plane.

Problem :

9) Discuss the physical regions in the s-, t-, u-channels respectively in the CM system if $m > \mu$.

Solution 9)

We consider first the s-channel. In the general case with two different masses we have

$$s = (k_1+p_1)^2 = (k_2'+p_2')^2 = m^2+\mu^2+2\varepsilon\omega+2K^2 \geq (m+\mu)^2$$
$$t = (p_2'-p_1)^2 = (k_2'-k_1)^2 = 2K^2(\cos\theta - 1) \ .$$

We note first that, whatever t means, the $\cos\theta$ has a physical meaning in all three channels, whereas K^2 might become negative (i.e., t positive, namely in the t-channel).

If we therefore express K^2 – which is an invariant in the sense of p. 27 – a function of s, then we can immediately obtain a relation between s and t which determines the boundary of the physical region. Now K is the magnitude of the momentum of all four particles and this is a unique function of the CM-energy and the masses of the particles involved. We have derived a formula for this, namely (I.22) p. 35. We must replace M^2 by s, m_1 by m, m_2 by μ and $|\vec{p}^*|^2$ by K^2 :

$$K^2 = \frac{[s-(m+\mu)^2][s-(m-\mu)^2]}{4s}$$

Therefore we obtain

$$t = \frac{[s-(m+\mu)^2][s-(m-\mu)^2]}{2s} \cdot (\cos\theta - 1) \qquad \text{(I.77)}$$

$t_{max} = 0$, namely for $\cos\theta = 1$ is one boundary.

The other one is then

$$t_{min} = -\frac{[s-(m+\mu)^2][s-(m-\mu)^2]}{s} \qquad \text{(I.78)}$$

2658

This boundary is a hyperbola. We find its asymptotes by letting

i) $s \to +0\,;\quad t \to -\infty$

ii) $s \to +\infty\,;\quad t \to -\infty \quad \text{as} \quad t \to -s+2(m^2+\mu^2)$

The first one is the line $s = 0$, the second one is the line $u = 0$ (since $u=2m^2+2\mu^2-s-t$). The hyperbola intersects the line $t = 0$ at $s = (m\pm\mu)^2$, as shows Eq. (I.78). We draw this hyperbola in our symmetrical representation [Fig. I.23.a]. Then Eq. (I.78) determines the shaded region marked (s) as the physical region of the s-channel. The physical region for the u-channel follows immediately from crossing symmetry (see remarks on p. 88) : shaded region marked (u) . It remains to find the physical region for the t-channel. Obviously, $t = (p_1+p_2)^2$ implies

$$t \geq 4m^2$$

Here the two incoming particles have both mass m and therefore equal energies E. But this is also the energy of the outgoing particles

$$t = (p_1+p_2)^2 = 4E^2 = (k_1'+k_2')^2 \; .$$

The momentum transfer, $s = (p_1-k_1')^2$, is here between particles of different mass. Therefore

$$s = m^2+\mu^2-2E^2+2\sqrt{(E^2-m^2)(E^2-\mu^2)} \cdot \cos\varphi = -(p_1-k_1')^2 < 0$$

where φ is the angle between p_1 and k_1'. Putting $\cos\varphi = \pm 1$ we obtain the extreme s-values (still < 0)

$$s_{extr} = -4E^2+2(m^2+\mu^2) - \frac{(\mu^2-m^2)^2}{s_{extr}} \; .$$

But $4E^2 = t$, hence

$$t = -s_{extr}+2(m^2+\mu^2) - \frac{(\mu^2-m^2)^2}{s_{extr}} \; ; \quad t \geq 4m^2 \qquad \text{(I.79)}$$

This, in fact, is the same as equation (I.78) for t_{min}. Here, however, it determines the s_{extr} for given $t \geq 4m^2$. Obviously this is the other branch of the hyperbola. From this follows then the physical region in the t-channel.

Figs. I.23.a, I.23.b, show the physical regions shaded and marked with (s) (t) (u) respectively.

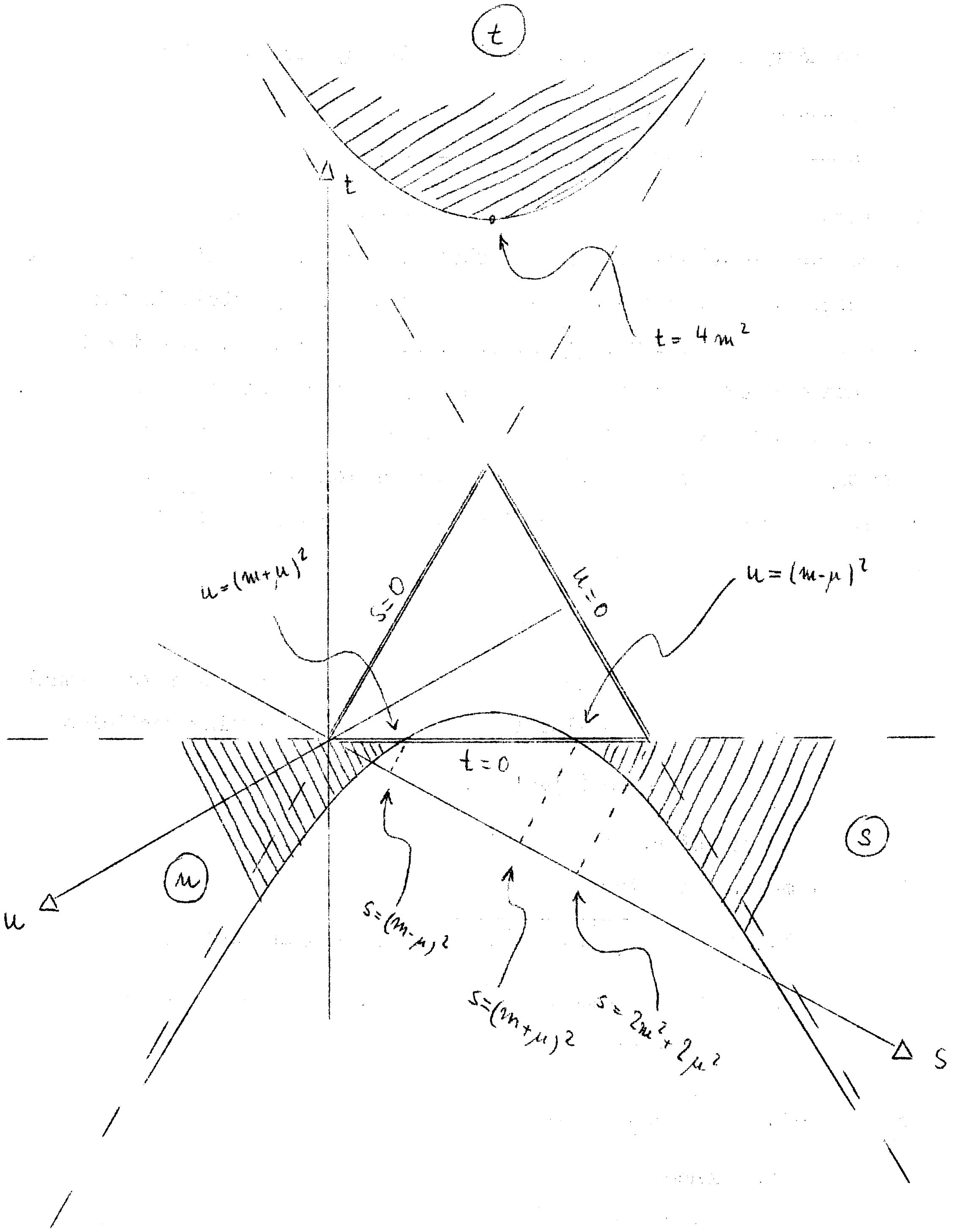


Fig. I.23.a

Physical regions of s,t,u-channels in symmetrical representation $(m > \mu)$

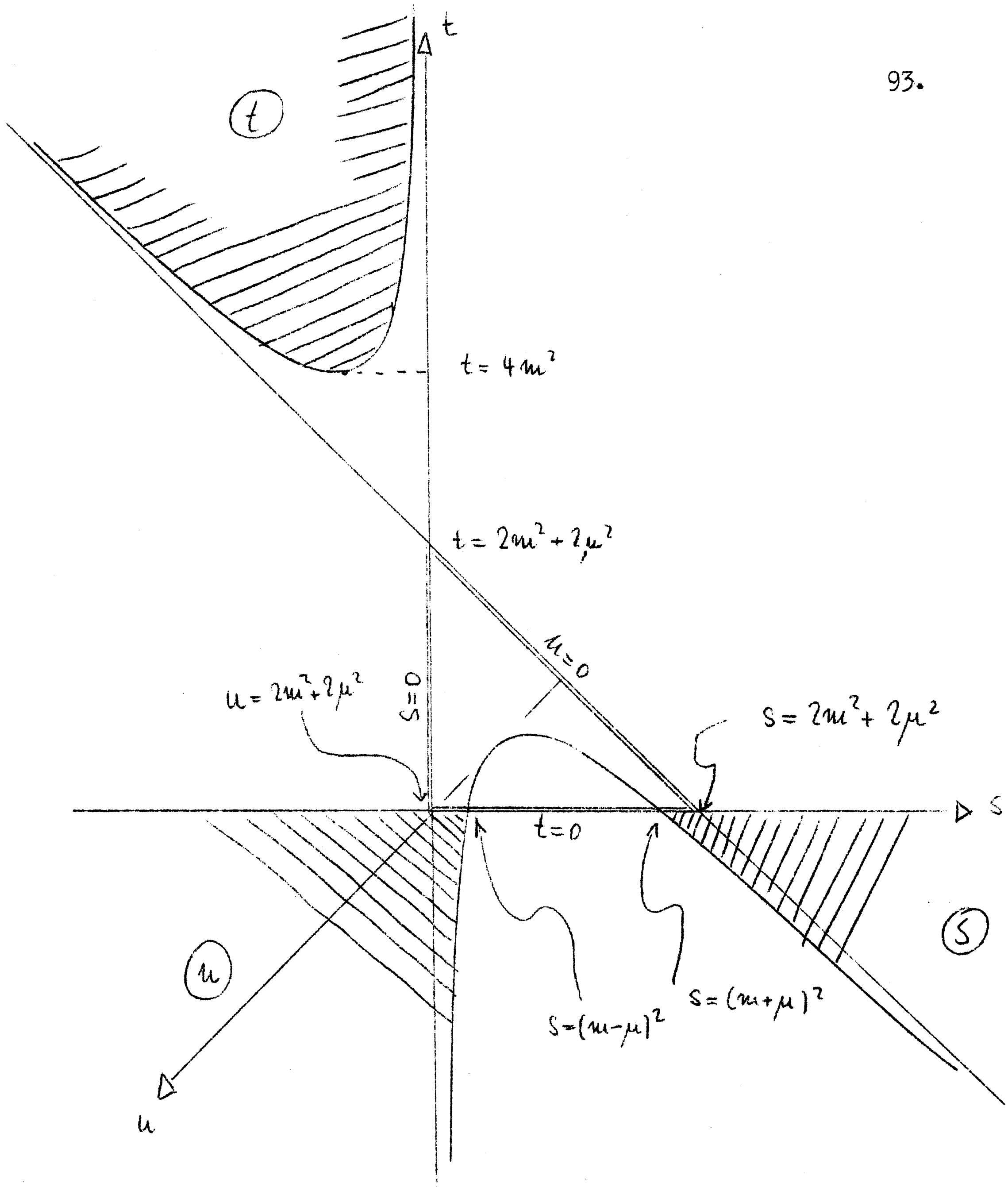


Fig. I.23.b

Physical regions of s, t, u-channels
in cartesian st - plane for $m > \mu$

We could have argued immediately :

Eq. (I.77) gives a relation between s and t and $\cos\theta$. As in all channels $\cos\theta$ has a physical significance, it is always bound between ± 1. One value, $\cos\theta = +1$, leads to $t = 0$ as a boundary. The other one gives the full hyperbola

$$t = -s+2(m^2+\mu^2) - \frac{(\mu^2-m^2)^2}{s}$$

as another boundary. The two branches belong to two situations :

i) $t < 0$; s- (or u-) channel. Then the lower branch is selected and it determines t_{min}.

ii) $t > 0$; this is the t-channel. The upper branch is selected and determines the extreme values of s.

Problem :

10) The dispersion relation for π-N scattering in the model case that both particles are neutral and scalar, takes in the CM system the form

$$\mathrm{Re}\, T(s,t) = g\left[\frac{1}{s+t-m^2-2\mu^2} - \frac{1}{s-m^2}\right] +$$

$$+ \frac{1}{\pi} P \int_{(m+\mu)^2}^{\infty} ds' \mathrm{Im}\, T(s',t) \left[\frac{1}{s'+t-2(m^2+\mu^2)+s} + \frac{1}{s'-s}\right] \qquad (I.80)$$

[For a derivation of this formula see, e.g., Introduction to Field Theory and Dispersion Relations, by R. Hagedorn, CERN 61-1. There it is Eq. (108) on p. 78.]

Derive from this the forward dispersion relation in the lab. system, where the energy ω of the incoming pion serves as variable.

Solution 10)

i) Forward dispersion relation means $t = 0$. This holds in the CM system as well as in the lab. system.

ii) We evaluate in the lab. system ($p_1 = 0$)

$$s = (p_1+k_1)^2 = p_1^2+k_1^2+2p_1k_1 = m^2+\mu^2+2m\,\omega \qquad (I.81)$$

where ω is the pion energy.

iii) Putting $t = 0$ and inserting $s = m^2+\mu^2+2m\omega$; $ds' = 2m\,d\omega'$ gives immediately

$$\mathrm{Re}\ T(\omega) = \frac{g}{2m}\left[\frac{1}{\omega-\frac{\mu^2}{2m}} - \frac{1}{\omega+\frac{\mu^2}{2m}}\right]$$

$$+\frac{1}{\pi}\,P\int_{\mu}^{\infty} d\omega'\ \mathrm{Im}\ T(\omega')\left[\frac{1}{\omega'+\omega}+\frac{1}{\omega'-\omega}\right] \qquad (I.82)$$

LECTURE 9

7) Short considerations on relativistic notation

So far we have avoided to use such terms as "metric tensor", "covariant" and "contravariant" vector components. But sooner or later one may run into trouble with signs if one does not know how to handle these things. They are also used in many books and papers and in the next paragraph we shall need them explicitly. Since some authors prefer the metric $x^2 = t^2 - \vec{x}^2$ and others $x^2 = -t^2 + \vec{x}^2$, we shall confront here both notations.

I wish to stress once more the following point : scalars, vectors, tensor operators and such things are defined abstractly as physical or geometrical quantities. As soon as a system of co-ordinates is introduced, these quantities will be represented by components with respect to the co-ordinate axes. This representation will depend on the choice of the co-ordinates, whereas the abstract quantities are independent : their existence logically precedes the existence of co-ordinates. If in the following we shall find different representations of, e.g., a four vector by covariant and contravariant components, and if these representations are different in different metrics, then they nevertheless represent one and the same physical quantity.

The "contravariant" components of four vectors like x or p are defined in any metric

$$
\begin{aligned}
x &= (x^0, x^1, x^2, x^3) = (t, \vec{x}) \\
p &= (p^0, p^1, p^2, p^3) = (\varepsilon, \vec{p}) \; .
\end{aligned}
\qquad (\text{I.83})
$$

They have therefore as components those with "the right sign". The metric is defined by the "metric tensor" in the two frequently used notations (we use that of the right-hand side) :

$$
g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} -1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix}
\qquad \Bigg| \qquad
g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & & & 0 \\ & -1 & & \\ & & -1 & \\ 0 & & & -1 \end{pmatrix} \; . \qquad (\text{I.84})
$$

Then one obtains the "covariant" components by "lowering" the indices

$$
\begin{aligned}
x_\mu = g_{\mu\nu}\, x^\nu &= (x_0, x_1 x_2 x_3) \\
&= (-t, \vec{x})
\end{aligned}
\qquad \Bigg| \qquad
\begin{aligned}
x_\mu = g_{\mu\nu}\, x^\nu &= (x_0, x_1 x_2 x_3) \\
&= (t, -\vec{x}) \; .
\end{aligned}
\qquad (\text{I.85})
$$

Here, and everywhere, the convention is to sum over any double index appearing once up and once down. These sum indices are dummy indices whose names μ , ν, ρ τ, λ, etc., are irrelevant [see, e.g., (I.88), (I.89)]. In any tensor (vector) any index can be lowered or raised in this way :

$$
\begin{aligned}
a_\mu^{\ \nu} &= a_{\mu\rho}\, g^{\rho\nu} = g_{\mu\rho}\, a^{\rho\nu} \\
a^{\mu\nu} &= g^{\mu\rho}\, a_{\rho\tau}\, g^{\tau\nu} \quad \text{etc.}
\end{aligned}
\qquad (\text{I.86})
$$

Applying this rule to the metric tensor itself, it follows in any metric

$$
g_\mu^{\ \nu} = g_{\mu\tau}\, g^{\tau\nu} = \delta_{\mu\nu} = \begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & 1 \end{pmatrix} \; . \qquad (\text{I.87})
$$

The scalar product of two four vectors, e.g., x and p is defined as the sum over products of covariant with contravariant components :

$$px \equiv p_\mu x^\mu = p^\nu g_{\nu\mu} x^\mu = x_\nu p^\nu \ . \qquad (I.88)$$

In the two metrics this becomes explicitly

$px = p_\mu x^\mu = p^\varrho x_\varrho = -\varepsilon t + \vec{p}\vec{x}$	$px = p_\mu x^\mu = p^\lambda x_\lambda = \varepsilon t - \vec{p}\vec{x}$
$p^2 = p_\mu p^\mu = -m^2$	$p^2 = p_\mu p^\mu = +m^2$.

(I.89)

Similarly invariants can be formed quite generally, e.g.,

$$aBc \equiv a_\mu B^{\mu\lambda} c_\lambda$$

$$BC \equiv B^{\mu\varrho} C_{\varrho\mu} \ .$$

The rule is therefore : a quantity is invariant if and only if each index appears twice, once as an upper and once as a lower one.

Care is needed when differentiating. If F is a scalar, then dF must be likewise a scalar, i.e., a relativistic invariant. Hence

$$dF = \frac{\partial F}{\partial x_\mu} dx_\mu = \frac{\partial F}{\partial x^\nu} dx^\nu = \text{invariant} \ . \qquad (I.90)$$

Therefore with (I.88) we conclude

$$\frac{\partial F}{\partial x_\mu} \equiv \partial^\mu F \quad \text{are contravariant vector components}$$

$$\frac{\partial F}{\partial x^\mu} \equiv \partial_\mu F \quad \text{are covariant vector components.} \qquad (I.91)$$

That is : the components of the gradient have the opposite character to the co-ordinates with respect to which we differentiate. Hence in either metric

$$\frac{\partial}{\partial x^{\mu}} \equiv \partial_{\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \tag{I.92}$$

whereas the contravariant components of the gradient become

$$\frac{\partial}{\partial x_{\mu}} \equiv \partial^{\mu} = \left(-\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \qquad \Bigg| \qquad \frac{\partial}{\partial x_{\mu}} \equiv \partial^{\mu} = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right) . \tag{I.93}$$

The Klein-Gordon operator may be defined as $\Box \equiv \partial_{\mu}\partial^{\mu}$ and becomes

$$\Box = -\frac{\partial^2}{\partial t^2} + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \qquad \Bigg| \qquad \Box = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} . \tag{I.94}$$

A plane wave with positive energy is a physical concept and will be written in either metric

$$\psi_p(x) = \frac{1}{(2\pi)^{3/2}} e^{i(\vec{p}\vec{x} - \varepsilon t)} , \tag{I.95}$$

therefore the invariant notation will read

$$\psi_p(x) = \frac{1}{(2\pi)^{3/2}} e^{+ipx} \qquad \Bigg| \qquad \psi_p(x) = \frac{1}{(2\pi)^{3/2}} e^{-ipx} . \tag{I.96}$$

These plane waves are solutions of the Klein-Gordon equation :

$$\Box\psi = \partial_{\mu}\partial^{\mu}\psi = -p_{\mu}p^{\mu}\psi = m^2\psi \qquad \Bigg| \qquad \Box\psi = \partial_{\mu}\partial^{\mu}\psi = -p_{\mu}p^{\mu}\psi = -m^2\psi . \tag{I.97}$$

<u>Definition</u> A "covariant" equation means not an equation written in covariant components, but an equation which is consistent insofar as the quantities on both sides transform in the same way : namely "covariantly".

That is to say, both sides must be scalars, vectors, tensors, respectively. In particular, any index which is not a sum index, must occur on both sides of the equation and in the same position ! A non-covariant equation is wrong (except perhaps in a particular Lorentz frame). Examples :

$a_\mu b^\mu = a$	is correct
$a_\mu b^\mu = D^{\rho\sigma} a_\rho k_\sigma$	is correct
$a_\mu b^\mu c_\lambda = f^\mu$	is wrong
$a_\mu b^\mu c_\lambda = d_\lambda$	is correct
$a_\mu b^\mu c_\lambda = d^\lambda$	is wrong, but $= g_{\lambda\rho} d^\rho$ is correct.

<u>Problem</u> :

11) The tensor of the electromagnetic field is

$$F = (F^{\mu\nu}) \equiv \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & H_z & -H_y \\ -E_y & -H_z & 0 & H_x \\ -E_z & H_y & -H_x & 0 \end{pmatrix} \qquad (I.98)$$

a) How do the tensors $F_\mu{}^\nu$, $F^\mu{}_\nu$, $F_{\mu\nu}$ look in our notation $(x^2=t^2-\vec{x}^2)$?

b) Write the invariant expression for the trace of a tensor and show that it vanishes for any antisymmetric tensor.

c) Write down the simplest invariant constructed with F.

d) Show that the equations

$$\left.\begin{aligned} &\frac{\partial F^{\mu\nu}}{\partial x^\nu} = j^\mu \\ &\partial^\rho F^{\mu\nu} + \partial^\nu F^{\rho\mu} + \partial^\mu F^{\nu\rho} = 0 \end{aligned}\right\} \qquad (I.99)$$

give the Maxwell equations $[j^\mu = (\rho\ ,\rho\vec{\beta})]$

Solution 11)

a) $F_\mu{}^\nu = g_{\mu\rho}\, F^{\rho\nu}$; $F^\mu{}_\nu = F^{\mu\rho}\, g_{\rho\nu}$; $F_{\mu\nu} = g_{\mu\rho}\, F^{\rho\sigma}\, g_{\sigma\nu}$. Since g is diagonal, we have always

$$|F_\mu{}^\nu| = |F^\mu{}_\nu| = |F_{\mu\nu}| \ .$$

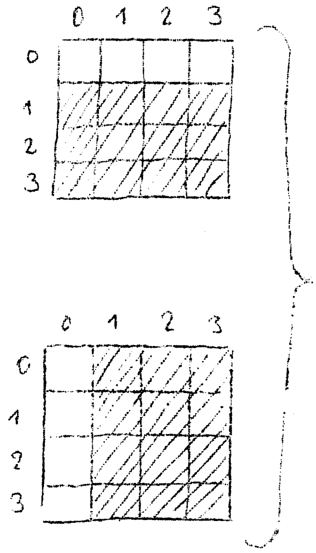
Since $g_{00} = 1$ and $g_{11} = g_{22} = g_{33} = -1$, we have changes of sign as soon as one index of g is 1 or 2 or 3. Hence

$$\begin{aligned} F_\mu{}^\nu &= \begin{cases} F^{\mu\nu} & \mu = 0 \\ -F^{\mu\nu} & \mu = 1,2,3 \end{cases} \\ F^\mu{}_\nu &= \begin{cases} F^{\mu\nu} & \nu = 0 \\ -F^{\mu\nu} & \nu = 1,2,3 \end{cases} \end{aligned} \qquad (I.100)$$

Therefore, if one looks at the matrix F, then the shaded parts will change sign :

$F^{\mu\nu} \rightarrow F_{\mu}{}^{\nu}$

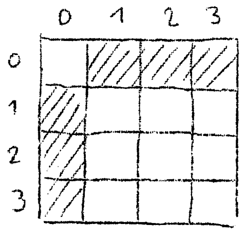
$F^{\mu\nu} \rightarrow F^{\mu}{}_{\nu}$



no longer skew symmetric !

Finally, taking down both indices, both changes occur simultaneously, hence

$F^{\mu\nu} \rightarrow F_{\mu\nu}$



again skew symmetric.

b) The trace of a tensor $T^{\mu\nu}$ must be written, in order to be invariant :

$$\mathrm{Tr}\, T = T^{\mu}{}_{\mu} = T_{\nu}{}^{\nu} = g_{\nu\mu} T^{\mu\nu} \; . \qquad (I.101)$$

If $T^{\mu\nu} = -T^{\nu\mu}$, then since the "names" of the sum-indices are irrelevant :

$$\underbrace{g_{\nu\mu} T^{\mu\nu} = g_{\mu\nu} T^{\nu\mu}}_{\nu \text{ and } \mu \text{ interchanged}} \underbrace{= g_{\nu\mu} T^{\nu\mu}}_{\text{since } g_{\mu\nu} \text{ is symmetric}} \underbrace{= -g_{\nu\mu} T^{\mu\nu}}_{\text{since } T^{\mu\nu} \text{ is antisymmetric}} = 0 \; .$$

c) The simplest invariant, made of F, is

$$F^{\mu\nu} F_{\nu\mu} = -F^{\mu\nu} F_{\mu\nu} \; .$$

This is - apart from signs - the sum of the squares of all components. As follows from a), only the E-components change sign in going from $F^{\mu\nu}$ to $F_{\mu\nu}$. Since each component occurs twice, we obtain

$$F^{\mu\nu} F_{\nu\mu} = -(F^{\mu\nu} F_{\mu\nu}) = -2(-\vec{E}^2 + \vec{H}^2) = 2(\vec{E}^2 - \vec{H}^2) \; . \qquad (I.102)$$

d)

$$\frac{\partial F^{\mu\nu}}{\partial x^{\nu}} = j^{\mu}$$

gives for

$$\mu = 0 \quad : \quad \frac{\partial E^k}{\partial x^k} = \rho \qquad \text{or} \qquad \operatorname{div} \vec{E} = \rho$$

$$\mu = 1 \; (2;3) \; : \; -\frac{\partial E_x}{\partial t} + \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = j_x \longrightarrow \operatorname{curl} \vec{H} = \vec{j} + \frac{\partial \vec{E}}{\partial t}$$

$$\partial^{\rho} F^{\mu\nu} + \partial^{\nu} F^{\rho\mu} + \partial^{\mu} F^{\nu\rho} = 0$$

is identically fulfilled as soon as any two indices are equal, this is due to $F^{\mu\nu} = -F^{\nu\mu}$.

Take therefore

$\rho, \mu, \nu = 0,1,2$ etc. [remember (I.93) !]

$$\frac{\partial H_z}{\partial t} - \frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x} = 0 \qquad \longrightarrow \qquad \operatorname{curl} \vec{E} = -\frac{\partial \vec{H}}{\partial t}$$

$\rho, \mu, \nu = 1,2,3$

$$-\frac{\partial H_x}{\partial x} - \frac{\partial H_y}{\partial y} - \frac{\partial H_z}{\partial z} = 0 \qquad \text{or} \qquad \operatorname{div} \vec{H} = 0$$

8) Precession of the polarization of particles moving in an electromagnetic field *)

(a) Equation of motion of a "polarization four vector"

Although we are leaving with this discussion the pure kinematics, we shall consider this problem here as it is a nice application of some general techniques, namely :

> to find the general invariant (or - in this case - covariant) description of a motion or a process or something else by formulating it in a particular reference system by means of the invariants (or covariants) which can be formed with the given four vectors or tensors.

We used this method extensively in paragraph 3) (p.26), and in paragraph 6) (p.66) [on variables and useful Lorentz systems].

Our problem is the following one :

suppose a beam of polarized particles is given (the polarization may be described by a polarization vector $\vec{s}$, whose length determines the degree of polarization and whose direction coincides with the direction of polarization; see below). Frequently, for technical reasons, this beam has to be guided and deflected by means of lenses and bending fields. What will be the polarization of the beam after such a procedure ?

If we consider a particle with spin σ , then a measurement of its spin component with respect to a given direction $\vec{e}$ (unit vector) will yield one of the $2\sigma+1$ possible eigenvalues, namely $m = \sigma, \sigma-1, \sigma-2, \dots -\sigma$. If we repeat the same experiment very often, i.e., we apply it to a beam, then we will observe a certain frequency distribution $W(\vec{e},m)$ of the m-values. The average over this distribution is

$$\langle \vec{\sigma}.\vec{e} \rangle = \sum_{m=-\sigma}^{\sigma} W(\vec{e},m)\cdot m \qquad (I.103)$$

---

*) See, V. Bargmann, L. Michel and V.L. Telegdi, Phys.Rev.Lett. 2, 435 (1959).

This is the expectation value of the spin component in the direction $\vec{e}$. The probability distribution $w(\vec{e},m)$ can be calculated only from quantum theory. The main point, however, is that the expectation value follows a classical equation of motion. This is the consequence of a general theorem by Ehrenfest (see, e.g., Schiff, Quantum Mechanics) which states that expectation values of quantum mechanical observables follow classical equations of motion. The expectation value $\langle \vec{\sigma}\cdot\vec{e} \rangle$ may be zero for all choices of the direction $\vec{e}$, then the beam is unpolarized. (This does not mean that there might not be an "alignment" of spins - but there is nothing left which could be described by a v e c t o r polarization.) If this is not the case then there exists a certain direction $\vec{e}_o$, in which the expectation value reaches a maximum :

$$\max \langle \vec{\sigma}.\vec{e} \rangle = \langle \vec{\sigma}.\vec{e}_o \rangle = s \; ; \qquad 0 \leq s \leq \sigma \; . \tag{I.104}$$

We call then s the degree of polarization. We may now introduce the polarization vector

$$\vec{s} \equiv \vec{e}_o \cdot s \; . \tag{I.105}$$

This vector has an obvious meaning. As it is defined entirely in terms of expectation values, it must follow a classical equation of motion. We know from classical physics that in the rest system of the particles considered this equation of motion is

$$\frac{d\vec{s}}{dt} = g\mu_o \, \vec{s}\times\vec{H} \qquad ; \quad {}^{*)} \tag{I.106}$$

where $g\mu_o\sigma$ is the magnetic moment of the particles considered. For charged particles, we have

$$g\mu_o\sigma = g\,\frac{e}{2m}\,\sigma \tag{I.106'}$$

but we shall write $g\mu_o$ in order to cover also the case of neutral particles with magnetic moment (neutron, Λ, Ξ^o, Σ^o). In the Dirac theory $g=2$, but in quantum electrodynamics, corrections are obtained such that $g\neq 2$ for electrons and muons.

---

*) For an explicit derivation of this equation, see, e.g., R. Hagedorn, The Density Matrix (Lecture) p.27; yellow report CERN 58-7.

We shall now generalize this equation into the covariant equation of motion of a polarization four vector (axial)

$$\left.\begin{array}{ll} S \equiv (s_0, \vec{s}) \;; & S = (0, \vec{s}) \text{ in the rest system} \\ S^2 = -s^2 & \end{array}\right\} \qquad (I.107)$$

It is not trivial that such a thing exists since $\vec{s}$ is an axial vector and should be described properly by a skew symmetric tensor, whose generalization will be a skew symmetric four tensor (see, e.g., below, the tensor $F^{\mu\nu}$ of the electromagnetic field). Also, if a polarization four vector can at all be defined, it is not obvious that its time component should be zero in the particle's rest system. Eq. (I.107) is therefore an ansatz and we have to try to see whether it leads to consistent equations.

The rate of change of the polarization at any instant t can depend only on the following quantities :

i) on the polarization S at that instant t;

ii) on the electromagnetic field;

iii) on the motion of the particle in that field.

The equation of motion of the polarization four vector, namely the generalization of Eq. (I.106) will then be of the form

$$\frac{dS}{d\tau} = Z \;; \qquad \tau = \text{ proper time of the particle } [\text{see pp. 7 and 14}]$$

where Z is a four vector constructed out of these three quantities. The polarization four vector S is already defined by Eq. (I.107). The electromagnetic field must be written in its relativistic form as a skew symmetric tensor [see problem 11), p. 101]

$$F^{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & H_3 & -H_2 \\ -E_2 & -H_3 & 0 & H_1 \\ -E_3 & H_2 & -H_1 & 0 \end{pmatrix} \qquad \begin{array}{l} F^{OK} = E_K \\ F^{ij} = H_K \end{array} \qquad (I.108)$$

and the motion of the particle may be described by its four velocity [see p.18]

$$V = (\gamma, \gamma\vec{\beta}) \qquad ; \qquad V = (1,0) \text{ in the rest system.} \qquad (I.109)$$

We construct now the four vector Z by generalizing (I.106). We first look for an equation of motion for s_o.

We observe first that

$$SV = s_o v_o - \vec{s}\vec{v} = 0 \qquad (I.110)$$

since this is true in the rest system. Hence

$$\frac{dS}{dt} V = -S \frac{dV}{dt} .$$

But in the rest system $V = (1,0)$, therefore

$$\left(\frac{dS}{dt} V\right)_R = \frac{ds_o}{dt} = -S \frac{dV}{dt} . \qquad (I.111)$$

Thus, in the rest system

$$\left(\frac{dS}{dt}\right)_R \equiv \left(\frac{ds_o}{dt} , \frac{d\vec{s}}{dt}\right)_R = \left(-S \frac{dV}{dt} , g\mu_o \vec{s}\times\vec{H}\right) = Z_R \qquad (I.112)$$

LECTURE 10

We now express these components by covariant expressions.

i) All time derivatives will be replaced by derivatives with respect to the proper time. We write a dot :

$$\cdot \text{ means } \frac{d}{d\tau} = \frac{d}{d(t/\gamma)} = \gamma \frac{d}{dt}$$

In the rest system this does not mean any change.

ii) In the rest system $S_R = (0, \vec{S}_R)$. Hence, if we use the notation

$$SF \equiv S_\mu F^{\mu\nu} = -F^{\nu\mu} S_\mu = -FS$$

we obtain

$$\vec{S}_R \times \vec{H} = \overrightarrow{(SF)}_R \ .$$

Therefore

$$(\dot{S})_R = (-S\dot{V},\ g\mu_0 \overrightarrow{(SF)}_R) \equiv Z_R \qquad (I.113)$$

iii) We generalize Z_R into a four vector. We observe that Z_R is linear homogeneous in S and linear in F. Further, it contains $\dot{V}$.

Z, the generalization of Z_R, will therefore be a four vector :

(α) linear homogeneous in S;

(β) linear in F.

The only non-constant four vectors which can be formed with $S, V, \dot{V}, F$ and which fulfil (α) and (β), are

$$SF \; ; \quad V(S\dot{V}) \; ; \quad V(SFV) \; . \tag{I.114}$$

A product $SF\dot{V}$ is not permitted since $\dot{V}$ = function of F hence $SF\dot{V}$ is not linear in F.

Therefore the general form of $\dot{S} = Z$ is

$$\dot{S} = a\, SF + b\, V(S\dot{V}) + c\, V(SFV) \; . \tag{I.115}$$

We now go to the rest system to find a, b, c.

$$(\dot{S})_R = \left\{ a(SF)^0_R + b\, S\dot{V} + c(SF)^0_R \, , \; a\overrightarrow{(SF)}_R \right\} . \tag{I.116}$$

On the other hand from (I.113) :

$$(\dot{S})_R = \left\{ -S\dot{V}, \; g\mu_0 \overrightarrow{(SF)}_R \right\} \tag{I.117}$$

hence, by comparison, $a = g\mu_0 = -c$; $b = -1$, and therefore with (I.114)

$$\dot{S} = g\mu_0 \left[SF - V(SFV) \right] - V(S\dot{V}). \tag{I.118}$$

This is the u n i q u e generalization of (I.113). Namely, a, b, c have been determined uniquely. A common factor d, say, which in the rest system would reduce to unity, must be equal to one in all systems, since a common factor must be invariant - otherwise it would destroy the four vector character of $\dot{S}$. An additive term Z', say, cannot exist since in the rest system $Z'_R = (0,0)$ and this remains $(0,0)$ in all Lorentz frames.

This is another example of the rule stated on p. 27, which more generally reads :

> if an equation given in a particular Lorentz system can be written in a manifestly covariant form (that is : both sides have the same transformation property !) which in the particular Lorentz frame reduces to the equation given originally, then this covariant form is the unique generalization of the equation given.

Let us return to (I.118).

It should be noted that we tacitly assumed that our particles had

* a constant magnetic moment;
* no electric moment (of any order) and no higher magnetic moments (quadrupole, etc.).

If these two conditions were violated then already our equation in the rest system would look different. It will then still be possible to define the four vector S. Complications arise, however, if the particle is also electrically polarizable. Then no such four vector exists, since already in the rest system only a skew symmetric tensor is sufficient to describe the polarization.

In a h o m o g e n e o u s f i e l d one has the equation of motion for a charged particle (that this gives the usual equations of motion, should be checked by the reader).

$$\dot{V} = -\frac{e}{m} FV \qquad (I.119)$$

In this case

$$\dot{S} = g\mu_0 SF + \left(\frac{e}{m} - g\mu_0\right) V\,(SFV) \qquad (I.120)$$

where the term with the factor $\frac{e}{m}$ vanishes for neutral particles.

Putting in $g\mu_o = g\,\frac{e}{2m}$ gives for charged particles

$$\dot{S} = \frac{e}{2m}\left[g\ SF - (g-2)\ V\ (SFV)\right] \qquad (I.120')$$

We still have to check the consistency of our formulae :

$$\left.\begin{aligned} S^2 &= -s^2 \\ SV &= 0 \end{aligned}\right\} \qquad (I.121)$$

These were the two equations which followed from the definition of the four vector S. Our equation of motion makes sense only if it does not contradict these equations. Eqs. (I.121) have the following consequence

$S\dot{S} = -(s^2)^{\cdot}$ = (change of the degree of polarization)

$\dot{S}V = -S\dot{V}$.

From (I.118), we find

$$S\dot{S} = g\mu_o\left[SFS - SV(SFV)\right] - SV(S\dot{V}) = 0 = -(s^2)^{\cdot}$$

because $SFS \equiv 0$ (antisymmetry of F!) and $SV = 0$. Furthermore, with (I.118)

$$\dot{S}V = g\mu_o\left[SFV - V^2(SFV)\right] - (S\dot{V})V^2 = -S\dot{V}$$

because $V^2 = 1$.

This shows that (I.121) is consistent with the equations of motion. We have thus established the existence of a polarization four vector and found its equation of motion.

The polarization has an obvious meaning only in the rest system of the polarized particle, whereas the equation of motion has been set up in a covariant form mainly in order to apply it in the lab. system, where the electromagnetic fields are most simply described. We therefore must study the Lorentz transformation of the polarization between these two systems.

The relevant transformation formula is obtained from p. 11. If we assume in the figure shown there that K' is the rest system (R) of the polarized beam and K is the lab. system (L), then formula (I.10) gives with $S_R = (0,\vec{s}_R)$ replacing $x' = (ct',\vec{x}')$ and S_L replacing X :

$$S_L = (s_{0L}, \vec{s}_L) = (\gamma \vec{\beta}\vec{s}_R, \ \vec{s}_R + \vec{\beta}\frac{\gamma^2}{\gamma+1}\vec{\beta}\cdot\vec{s}_R) \qquad (I.122)$$

where $\vec{\beta}$ is the velocity of the polarized beam in the lab. system.

We now may ask : what is $\vec{s}_L$?

It is a three vector which indeed has very little to do with what we feel when we hear the word polarization.

* its magnitude depends on $\vec{\beta}$, since $S^2 = s_0^2 - \vec{s}^2 = -s^2$ is invariant. Hence - remember s is the invariant magnitude of polarization -

$$\vec{s}_L^2 = s_{0L}^2 + s^2 ,$$

hence

$$\vec{s}_L^2 = s^2 \left[1 + \gamma^2 \beta^2 \cdot \cos^2(\theta_R) \right] , \qquad (I.123)$$

where θ_R is the angle between $\vec{\beta}$ and the direction of polarization in the rest system.
If $\gamma \gg 1$ and $\cos(\theta_R) \neq 0$, $\vec{s}_L^2$ increases proportionally to γ^2. For massless particles it becomes ∞. In this case we must use another description (see below).

* its direction depends on $\vec{\beta}$ as one sees directly in (I.122) : if $\gamma \gg 1$, the term $\vec{s}_R$ is negligible against the next term which is parallel to $\vec{\beta}$ if $\vec{\beta}\vec{s}_R > 0$ and antiparallel if it was < 0. That is : when $\vec{\beta} \to 1$ the polarization three vector $\vec{s}_L$ becomes parallel or antiparallel to $\vec{\beta}$. The angle is given by

$$\cos^2\Theta_L \equiv \cos^2(\vec{s}_L\vec{\beta}) = \frac{(\vec{s}_L\vec{\beta})^2}{\vec{s}_L^{\,2}\cdot\beta^2} .$$

This is easily calculated from $SV = \gamma(s_o - \vec{s}\vec{\beta})_L = 0$, hence with (I.123) and (I.122)

$$\cos^2\Theta_L = \frac{s_{0L}^2}{\beta^2\vec{s}_L^{\,2}} = \frac{\gamma^2\beta^2 s^2\cos^2(\Theta_R)}{\beta^2 s^2(1+\beta^2\gamma^2\cos^2(\Theta_R))}$$

$$\cos^2\Theta_L = \frac{\gamma^2\cos^2(\Theta_R)}{1+\gamma^2\beta^2\cos^2(\Theta_R)} . \qquad \text{(I.124)}$$

Introducing the helicity h by

helicity = component of the polarization in the direction of flight,

we can write

$$h = \frac{(\vec{s}_L\vec{\beta})}{\beta} = \frac{s_o}{\beta} = \gamma\frac{(\vec{s}_R\vec{\beta})}{\beta} = \gamma\cdot s\cdot\cos\Theta_R .$$

This clearly shows that the polarization of a beam should be considered always in the rest system, because the polarization $\vec{s}$ of a beam depends on the observer in magnitude and direction.

The polarization is completely determined by giving $\vec{s}_R$, i.e., by the magnitude (degree of polarization) s and direction of the polarization in the rest system.

Our proof of the consistency of the equations of motion with $S^2 = -s^2$ has shown, however, that S^2 is not only invariant, but is even a constant of the motion :

$$\frac{d}{d\tau}S^2 = -\frac{d}{d\tau}s^2 = 0 \quad ;$$

that is : the degree of polarization of a beam cannot be changed by passing it through electromagnetic fields whatsoever as long as the inhomogeneities across the beam can be neglected. [Otherwise one must split up the beam into a bundle of sufficiently many smaller beams over whose individual cross-sections the fields are constant, and calculate the changes of directions of polarization for each one separately. In this case the degree of polarization of the beam as a whole can of course change.]

Therefore, the degree of polarization s is irrelevant for the description of the actual state of polarization and only the directions must be given. That needs two angles. In most cases even one single angle yields the relevant information : the angle between the polarization in the rest system and the direction of motion. It describes how transversal (or longitudinal) the polarization is.

We shall now give an equation for the rate of change of the angle θ between the polarization $\vec{s}_R$ and the direction of motion $\vec{\beta}$. To this end, we introduce in the lab. system, at the instant $t = t_o$, two unit vectors, one $\vec{\ell}$ parallel to $\vec{\beta}$ and the other $\vec{n}$ perpendicular to $\vec{\ell}$ such that $\vec{s}_L$ lies in the plane spanned by these two unit vectors [see Fig. I.24]

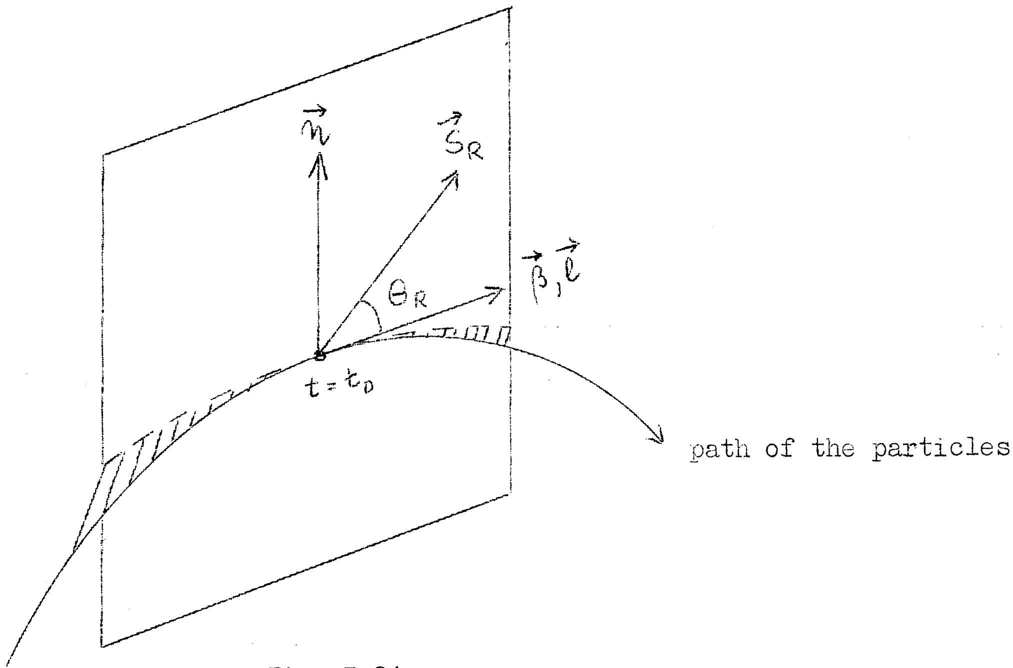


Fig. I.24

The unit vectors $\vec{n}$ and $\vec{\ell}$
(perspective drawing, $\vec{n} \perp \vec{\ell}$)

Then also $\vec{s}_R$ lies in the plane spanned by $\vec{n}$ (the normal direction) and $\vec{\ell}$ (the longitudinal direction). Therefore $\vec{\ell}$ and $\vec{n}$ can serve in both reference frames - the lab. and the rest system - and both are defined in an invariant way as far as these two systems are concerned. Clearly

$$\left.\begin{aligned} \vec{\ell} &= \vec{\beta}/\beta \\ \vec{n}\cdot\vec{\ell} &= 0 \\ \vec{n}^2 &= \vec{\ell}^2 = 1 \end{aligned}\right\} \qquad (I.125)$$

With these unit vectors we can write (I.122) - remember $\theta_R \equiv \sphericalangle(\vec{s}_R,\vec{\beta})$ -

$$S_L = (s_{0L},\vec{s}_L) = s(\beta\gamma\cos\theta_R,\ \vec{\ell}\cos\theta_R + \vec{n}\sin\theta_R + \vec{\ell}\,\frac{\beta^2\gamma^2}{\gamma+1}\cos\theta_R)\ .$$

With $\beta^2\gamma^2 = \gamma^2-1$ we obtain

$$S_L = s(\beta\gamma\cos\theta_R,\ \gamma\vec{\ell}\cos\theta_R + \vec{n}\sin\theta_R) = s\cdot\cos\theta_R\cdot(\beta\gamma,\vec{\ell}\gamma) + s\cdot\sin\theta_R\cdot(0,\vec{n}).$$

Introducing the four vectors

$$\left.\begin{aligned} L &\equiv (\beta\gamma,\vec{\ell}\gamma) \\ N &\equiv (0,\vec{n}) \end{aligned}\right\} \qquad (I.126)$$

we can write

$$S = sL\cos\theta_R + sN\sin\theta_R \quad {}^{*)} \qquad (I.127)$$

---

*) This is also the correct representation in the rest system since there we obtain

$$S_R = s(0,\vec{\ell}\cos\theta_R + \vec{n}\sin\theta_R)\ .$$

Indeed : by applying the Lorentz transformation Lab $\longrightarrow$ Rest system, L and N are transformed into what they should become, namely

$$L_R = (0,\vec{\ell}) \quad \text{and} \quad N_R = (0,\vec{n}).$$

Therefore the angle between S and L is the same as between $\vec{s}_R$ and $\vec{\ell}$.

The particular properties of the four vectors L and N are :

$$\left.\begin{aligned} & L^2 = N^2 = -1 \\ & LN = LV = NV = 0 \quad ; \qquad V = (\gamma, \vec{\beta}\gamma) \\ \text{hence} \quad & \\ & L\dot{L} = N\dot{N} = 0 \\ & L\dot{N} = -N\dot{L} . \end{aligned}\right\} \qquad (I.128)$$

We introduce (I.127) into the equation of motion (I.120) in a homogeneous field :

$$\begin{aligned} \dot{S} &= s\left[\dot{L}\cos\theta_R + \dot{N}\sin\theta_R + \dot{\theta}_R(N\cos\theta_R - L\sin\theta_R)\right] = \\ &= s\left[g\mu_o(LF\cos\theta_R + NF\sin\theta_R) - (g\mu_o - \frac{e}{m})V(LFV\cos\theta_R + NFV\sin\theta_R)\right] . \end{aligned} \qquad (I.129)$$

Using (I.128) we can solve for $\dot{\theta}$ by multiplying by N from the right :

$$L\dot{N}\cos\theta_R - \dot{\theta}_R \cos\theta_R = g\mu_o\, LFN \cos\theta_R .$$

The rest annihilates because $NFN \equiv 0$ (F is antisymmetric). Hence

$$\dot{\theta}_R = \dot{L}N - g\mu_o\, LFN \qquad (I.130)$$

(multiplying by L would give the same).

Now, the term $\dot{L}N$, is easy to calculate : from (I.126) follows

$$\dot{L}N = -L\dot{N} = \gamma \vec{\ell}\,\dot{\vec{n}} = -\gamma \vec{n}\dot{\vec{\ell}} .$$

On the other hand, writing $V = (\gamma, \vec{\ell}\beta\gamma)$

$$N\dot{V} = -n\cdot(\vec{\ell}\,(\beta\gamma)^{\cdot} + \beta\gamma\dot{\vec{\ell}}) = -\beta\gamma\vec{n}\dot{\vec{\ell}} ,$$

hence

$$\dot{L}N = \frac{1}{\beta} N\dot{V} = \frac{e}{m}\frac{1}{\beta} NFV .$$

2658

Introducing this into (I.130) we obtain (AFB = -BFA !)

$$\dot{\Theta}_R = \left[\frac{e}{m\beta} V - g\mu_o L\right] FN. \qquad (I.131)$$

We have with (I.108) and (I.126)

$$FN = F^{\mu\nu} N_\nu = (-\vec{E}\vec{n}, -\vec{n}\times\vec{H})$$

$$\frac{e}{m\beta}V - g\mu_o L = \frac{e}{m}\left(\frac{\gamma}{\beta}, \vec{\ell}\gamma\right) - g\mu_o(\beta\gamma, \vec{\ell}\gamma) = \gamma\left(\frac{e}{m\beta} - \beta g\mu_o, \vec{\ell}\left(\frac{e}{m} - g\mu_o\right)\right).$$

Inserting this into (I.131) results $\left[\text{with } \dot{\Theta}_R = \frac{d\Theta_R}{d\tau} = \gamma\frac{d\Theta_R}{dt}\right]$

$$\frac{d\Theta_R}{dt} = (\vec{E}\cdot\vec{n})\left(g\mu_o\beta - \frac{e}{m\beta}\right) + \left(g\mu_o - \frac{e}{m}\right)\vec{\ell}\cdot\vec{H}\times\vec{n} \qquad (I.132)$$

This is valid for any particle with magnetic moment $g\mu_o\sigma$ and charge e. If the charge is $\neq 0$, then with $g\mu_o = g\frac{e}{2m}$ we obtain

$$\frac{d\Theta_R}{dt} = \frac{e}{2m}\left[(\vec{E}\cdot\vec{n})\frac{(g-2)-g/\gamma^2}{\beta} + (g-2)\cdot\vec{\ell}\cdot\vec{H}\times\vec{n}\right] \qquad (I.133)$$

[Remember : Θ_R is the angle between the polarization $\vec{s}_R$ and $\vec{\beta}$ measured in the rest system, $\vec{E}$ and $\vec{H}$ are h o m o g e n e o u s fields in the lab. system.]

Notice that $\frac{d\Theta_R}{dt}$ is independent of the degree s of polarization ! The case of an electric dipole moment can be described similarly. [See paper quoted on p. 105.] For inhomogeneous fields one must go back to Eq. (I.118)*).

It should be mentioned that $\dot{\Theta}_R$ is invariant in the sense of the remark on p. 27, hence it must be possible to express it by invariants only. Indeed, Eq. (I.131) is an invariant definition. It is useful in some cases to combine these equations with (I.129), we therefore list below the various four vectors appearing on the r.h.s. of (I.129), explicitly written in form of three vectors :

---

*) See also : R.H. Good, Jr., Phys.Rev. 125, 2112 (1962).

$$
\left.
\begin{aligned}
LF &= -FL = \gamma(\vec{E}\cdot\vec{\ell}\,,\ \beta\vec{E} + \vec{\ell}\times\vec{H})\\
NF &= -FN = (\vec{E}\cdot\vec{n},\ \vec{n}\times\vec{H})\\
VF &= -FV = \gamma(\beta\vec{E}\cdot\vec{\ell}\,,\ \vec{E} + \beta\vec{\ell}\times\vec{H}) = \frac{m}{e}\frac{dV}{d\tau} \quad \text{in homogeneous fields}\\
LFV &= -VFL = \vec{E}\cdot\vec{\ell}\\
NFV &= -VFN = \gamma(\vec{E}\cdot\vec{n} + \beta\vec{n}\cdot\vec{\ell}\times\vec{H})\ .
\end{aligned}
\right\}
\quad (I.134)
$$

For the proof of these equations and some examples of applications of Eqs. (I.132)-(I.134), see Bible, The General Epistle of James 1, 22, and Problems 12-15 below.

Problem :

12) Verify Eqs. (I.134). The solution is straightforward and follows from the definitions.

Problem :

13) Give a full discussion of the equations of motion for the case of a homogeneous field such that $\vec{E}\times\vec{\ell} = \vec{H}\times\vec{\ell} = 0$.

Solution 13)

i) We have first to check whether the condition $\vec{E}\times\vec{\ell} = \vec{H}\times\vec{\ell} = 0$ is conserved by the equations of motion of the particle. Indeed, with (I.134) and $\dot{V} = -\frac{e}{m} FV$, we obtain

$$\dot{V} = (\dot{\gamma}, \dot{\vec{\ell}}(\beta\gamma) + \vec{\ell}(\beta\gamma)^{\cdot}) = \frac{e}{m}\gamma\,(\beta\vec{E}\vec{\ell}\,,\ \vec{E} + \underbrace{\beta\vec{\ell}\times\vec{H}}_{=0})\ .$$

Writing $\vec{E} = \vec{\ell}E$ we see that $\dot{\vec{\ell}} = 0$. All higher derivatives of $\vec{\ell}$ also vanish. Hence $\vec{\ell}$ is constant and the condition $\vec{\ell}\times\vec{E} = \vec{\ell}\times\vec{H} = 0$ is conserved.

ii) Now (I.132) tells us immediately

$$\frac{d\Theta_R}{dt} = 0.$$

iii) $\dot{S}$ however is not zero. We take (I.129) with $\dot{\Theta} = 0$

$$\dot{S} = s\left[\dot{L}\cos\Theta_R + \dot{N}\sin\Theta_R\right]$$

$$= s\left(\gamma\frac{e}{m}E\cos\Theta_R,\ \beta\gamma\frac{e}{m}E\vec{\ell}\cos\Theta_R + g\mu_0\,\vec{n}\times\vec{H}\sin\Theta_R\right).$$

Here (I.134) was used to write down explicitly the r.h.s. of Eq. (I.129). Comparing coefficients gives

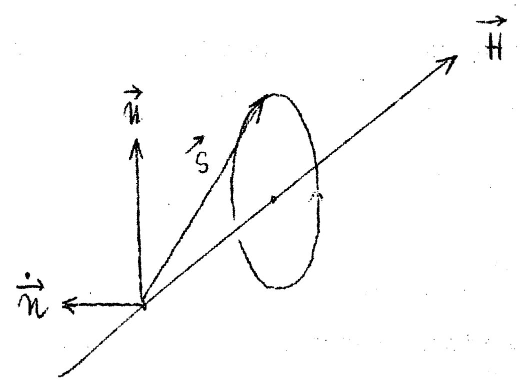
$$\dot{L} = \frac{e}{m}E\,(\gamma, \vec{\beta}\gamma) = \frac{e}{m}EV$$

which gives $V\dot{L} = \frac{e}{m}E = -L\dot{V}$, as it should be, and

$$\dot{N} = g\mu_0\,\vec{n}\times\vec{H}.$$

These equations state that L changes because the particle is accelerated and that N precesses in a left screw and with constant angular frequency around $\vec{H}$:

$$\frac{dN}{dt} = \frac{d\vec{n}}{dt} = \frac{1}{\gamma}\dot{N} = \frac{1}{\gamma}\dot{\vec{n}} = \frac{g\mu_0}{\gamma}\vec{n}\times\vec{H} \quad ; \quad \left|\frac{d\vec{n}}{dt}\right| = \omega = \frac{g\mu_0 H}{\gamma}$$



Problem :

14) Consider the case $\vec{E} = 0$; $\vec{H}(\vec{n}\times\vec{\ell}) = \vec{H}$ in the same way as in the previous problem.

Solution 14)

i) We first check whether these conditions are conserved, using again the equations of motion of the particle (I.134)

$$\dot{V} = (\dot{\gamma}, \dot{\vec{\ell}}\beta\gamma + \vec{\ell}(\beta\gamma)^{\cdot}) = \frac{e}{m}\gamma(0, \beta\vec{n}H) .$$

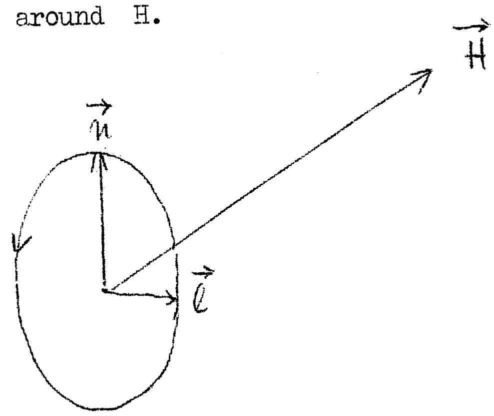
It follows

$$\dot{\vec{\ell}} = \frac{eH}{m}\vec{n}$$

which means that $\vec{\ell}$ and $\vec{n}$ rotate with angular frequency

$$\left|\frac{d\vec{\ell}}{dt}\right| = \frac{eH}{m\gamma} = \omega_o$$

in a left-screw around $\vec{H}$.



Therefore the condition is conserved. This is the well-known behaviour of a charge moving in a constant magnetic field : it goes on circles with angular frequency ω_o (Larmor frequency).

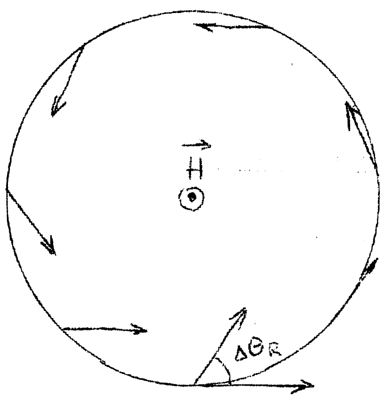
ii) The equation of motion (I.132) for Θ_R gives now

$$\frac{d\Theta_R}{dt} = (g\mu_o - \frac{e}{m}) H = \frac{eH}{m} (\frac{g}{2} - 1) .$$

That is, the angle between the direction of motion and the polarization increases (or decreases) with constant rate if $g \neq 2$. Note that this is independent of the velocity of the particles ! The particles move on a circle and turn around once in the time $T = \frac{2\pi}{\omega_o} = 2\pi \frac{m\gamma}{eH}$. Therefore

$$\Delta \Theta_R = \Theta_R(T) - \Theta_R(0) = T \cdot \frac{d\Theta_R}{dt} = 2\pi\gamma (\frac{g}{2} - 1) .$$

Since after one turn $\vec{\ell}$ has its old position, this $\Delta \Theta_R$ is the change of the direction of polarization per turn.



See above figure ($g > 2$ was assumed).

This fact has been used to measure the g-factor of the μ .
Further examples are found in the reference given on p. 105.

Problem :

15) Regarding Eq. (I.132) in the case of a fictitious particle with spin and with $e \neq 0$, but without magnetic moment $(g=0)$, one sees that

$$\frac{d\theta_R}{dt} = -\frac{e}{m}\left[\frac{\vec{E}\vec{n}}{\beta} + \vec{\ell}\cdot\vec{H}\times\vec{n}\right].$$

Physical intuition tells us, however, that the polarization should not change if there is no magnetic moment. Show that there is no contradiction.

Solution 15)

Since θ_R is the angle between the polarization and the direction of motion and since we expect that the polarization $\vec{s}_R = \text{const}$, we presume that $\frac{d\theta_R}{dt}$ describes the change of the direction of motion with respect to any fixed direction. We calculate everything in the lab. system.

i) From (I.134) we have $\left(' \equiv \frac{d}{dt} = \frac{d}{\gamma\, d\tau}\right)$

$$\dot{V} = \gamma V' = \frac{e\gamma}{m}\left(\beta\vec{E}\vec{\ell}\,,\ \vec{E} + \beta\vec{\ell}\times\vec{H}\right) = \gamma\left(\gamma',\ \beta\gamma\vec{\ell}' + \vec{\ell}(\beta\gamma)'\right),$$

hence, comparing time components :

$$\gamma' = \frac{e}{m}\vec{E}\vec{\ell}\beta.$$

Further

$$(\beta\gamma)' = \frac{\gamma'}{\beta} = \frac{e}{m}\vec{E}\vec{\ell},$$

therefore, comparing space components :

$$\frac{e}{m}\left(\vec{E} + \beta\vec{\ell}\times\vec{H}\right) = \beta\gamma\vec{\ell}' + \frac{e}{m}\vec{\ell}(\vec{E}\vec{\ell})$$

$$\vec{\ell}' = \frac{e}{m\beta\gamma}\left[\vec{E} - \vec{\ell}(\vec{E}\vec{\ell}) + \beta\vec{\ell}\times\vec{H}\right].$$

ii) At a given instant t_o we take in the lab. system a constant four vector $A = S(t_o)$. Then

$$A = L\cos\Theta + N\sin\Theta \quad ; \qquad \Theta(t_o) = \Theta_R$$

$$A' = L'\cos\Theta + N'\sin\Theta + \Theta'(N\cos\Theta - L\sin\Theta) = 0 \quad .$$

Therefore by multiplication by N [see Eqs. (I.126), (I.128)]

$$\Theta' = NL' = -\gamma\,\vec{n}\,\vec{\ell}\,' \quad .$$

Taking the $\vec{\ell}\,'$ just calculated, we find

$$\Theta' = -\frac{e}{m}\left[\frac{\vec{E}\vec{n}}{\beta} + \vec{\ell}\cdot\vec{H}\times\vec{n}\right]$$

which coincides with $\frac{d\Theta_R}{dt}$ as stated in the problem.

(b) <u>The case of mass zero</u>

Our polarization four vector

$$s^{\mu} = (\gamma\vec{\beta}\vec{s}_R,\ \vec{s}_R + \vec{\beta}\,\frac{\gamma^2}{\gamma+1}\,\vec{\beta}\vec{s}_R)$$

does not work since $\gamma \rightarrow \infty$ and since $\vec{s}_R$ is meaningless. On the other hand, on p. 114 I said that "polarization" is defined only in the rest system. Clearly, our whole concept fails to apply to particles with mass zero. And yet particles with mass zero and spin $\neq 0$ exist.

We may solve the problem by analogy with the four velocity : $V^{\mu} = (\gamma, \vec{\beta}\gamma)$ is also diverging if we consider $m \rightarrow 0$. There exists, however, another four vector, which is proportional to V and does not become meaningless for $m \rightarrow 0$, namely the four momentum $P = mV = (m\gamma,\ m\vec{\beta}\gamma) = (\mathcal{E}, \vec{p})$.

Let us multiply S by m and consider the four vector

$$W^{\mu} = mS^{\mu} = (m\gamma\vec{\beta}\,\vec{s}_R,\ m\vec{s}_R + \vec{\beta}\,\frac{m\gamma^2}{\gamma+1}\,\vec{\beta}\,\vec{s}_R) \cdot \qquad (I.135)$$

Let us forget about the meaning of $\vec{s}_R$ and look at $|\vec{s}_R| \cdot \cos(\vec{\beta}\,\vec{s}_R)$ as a constant s. Then

$$W^{\mu} = (m\gamma/\beta s, m\vec{s}_R + \vec{\beta}\frac{m\gamma^2}{\gamma+1}/\beta s).$$

This W^{μ} transforms as a four vector, no matter what origin the constant s has : if we give this expression to somebody and tell him that this were m times the polarization four vector of a particle with four momentum P, then he is able to tell us how this four vector will look in any other Lorentz system. This is sufficient. As long as $m \neq 0$ he can even transform to the rest system and discover the physical significance of the constant s.

$\vec{\beta} \to \vec{\ell}$ But now we allow the mass to go to zero. Then, with $\gamma \to \infty$

$$W^{\mu}(m \to 0) = s(\varepsilon, \vec{\ell}\varepsilon) = s\,P^{\mu}(m = 0).$$

We may again call $|s|$ the degree of polarization and find

$$m = 0 : \qquad \left.\begin{array}{c} W^{\mu} = s\,P^{\mu} \\ \vec{s} = s\vec{\ell} \\ W^{\mu}W_{\mu} = W^{\mu}P_{\mu} = P^{\mu}P_{\mu} = 0\ . \end{array}\right\} \qquad (I.136)$$

I.e., : since W^{μ} and P^{μ} are four vectors, s must be an invariant. The direction of polarization is always parallel $(s > 0)$ or antiparallel $(s < 0)$ to the direction of motion and the magnitude s is the same in all Lorentz frames : the helicity is always $\pm$ s.

There is therefore no need for an equation of motion of the polarization.

Our representation of the polarization of massless particles used here is different from that for polarized light by means of the Stokes parameters [see lecture on Density Matrix, CERN 58-7], but it is also applicable to light quanta.

LECTURE 11

(c) The relation of the polarization four vector to the angular momentum tensor

We defined S as the expectation value of the spin of a beam. It therefore follows classical equations of motion. This suggests to try to establish a connection between S and the angular momentum of a classical system. Corresponding relations will then exist between the operators.

We consider a closed system of N s p i n l e s s classical mass points with co-ordinates and momenta

$$x_i^\mu = (t, \vec{x}_i) \qquad i = 1 \ldots N \; .$$
$$p_i^\mu = (\varepsilon_i, \vec{p}_i)$$

The angular momentum of the system is the skew symmetric tensor

$$M^{\mu\nu} = \sum_i x_i^\mu p_i^\nu - x_i^\nu p_i^\mu \; . \tag{I.137}$$

The time components are *)

$$M^{ok} = t \sum p^k - \sum x^k \varepsilon \; . \tag{I.138}$$

---

*) Notation : Greek indices 0...3; Latin indices 1..3.
The indices i labelling the mass points will be suppressed wherever possible.

Introducing the total four momentum

$$P = \sum p_i = (\sum \varepsilon \ , \sum \vec{p}) \tag{I.139}$$

and going to the rest system $(\vec{P} = 0)$, we obtain

$$(M^{ok})_R = -(\sum x^k \varepsilon)_R \ .$$

From angular momentum conservation, it follows that this quantity is constant in time. We define now in R the four vector (with constant space components !)

$$(X^{\mu})_R = (t, \frac{1}{\sum \varepsilon_i} \sum \vec{x}_i \varepsilon_i)_R \ .$$

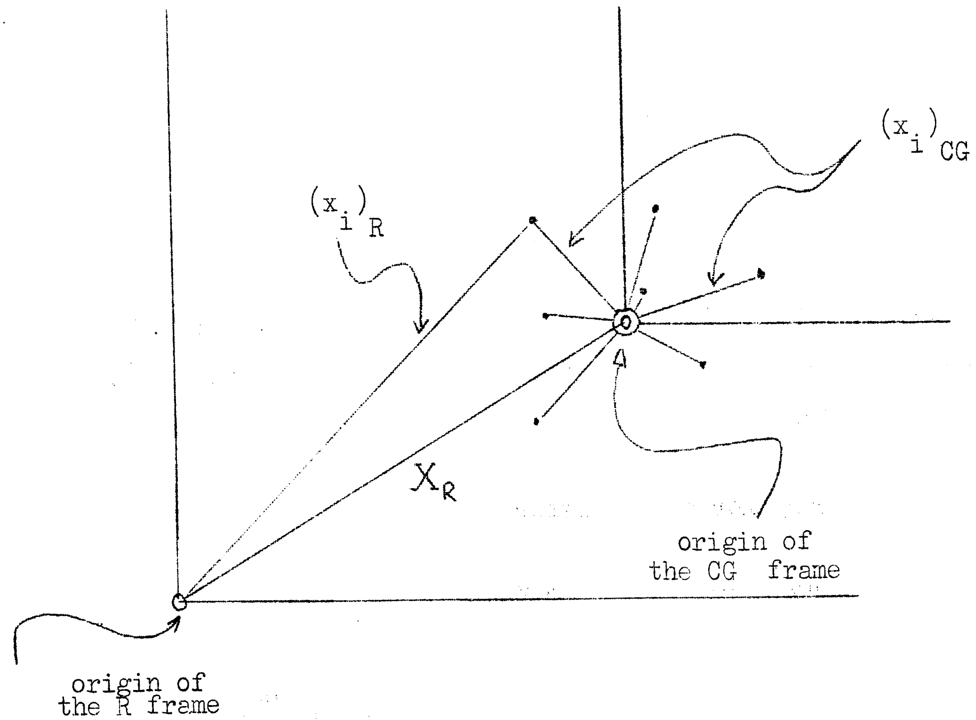
Since R is defined invariantly [starting from any Lorentz system, the condition $\sum \vec{p} = 0$ leads always to the same system R (up to a space rotation and translation)], the four vector $(X^{\mu})_R$ is defined invariantly and may then be transformed to any Lorentz system. We call this general X the centre of gravity (CG)

$$X^{\mu} = L \cdot (X^{\mu})_R = L \cdot (t, \frac{\sum \vec{x}\varepsilon}{\sum \varepsilon})_R \ . \tag{I.140}$$

The notation means that $(X^{\mu})_R$ has to be Lorentz transformed to obtain X^{μ}. This X^{μ} is different from what one would obtain by Lorentz transforming each x_i^{μ} and p_i^{μ} and constructing $(t, \frac{\sum \vec{x}\varepsilon}{\sum \varepsilon})$ in the new Lorentz frame. Therefore we had to define X^{μ} in this complicated way to make sure that it is a four vector. The word "centre of gravity" is chosen because for non-relativistic particles we have

$$\left(\frac{\sum \vec{x}\varepsilon}{\sum \varepsilon}\right)_R \longrightarrow \frac{\sum \vec{x}m}{\sum m} \ .$$

Now in the rest system R (where $\vec{P} = 0$) we may introduce new co-ordinates $x^{\mu}_{i,CG}$ by measuring all distances from the centre of gravity $(X^{\mu})_R$ [see figure]



$$(x_i)_R = X_R + (x_i)_{CG} \tag{I.141}$$

Note that the $(x_i)_{CG}$ have the time components zero. The momenta remain unchanged. Then the angular momentum tensor becomes in R

$$(M^{\mu\nu})_R = \sum (x^\mu p^\nu - x^\nu p^\mu)_R = \sum (x^\mu p^\nu - x^\nu p^\mu)_{CG} + (X^\mu P^\nu - X^\nu P^\mu)_R \tag{I.142}$$

and the general $M^{\mu\nu}$ in any Lorentz system is found by a suitable Lorentz transformation. This $(M^{\mu\nu})_R$ has remarkable properties :

> if we imagine the origin in R shifted to X_R, then the term $X^\mu P^\nu - X^\nu P^\mu$ vanishes since then $\vec{P} = 0$ and $\vec{X} = 0$.

Therefore the term

$$\sum (x^\mu p^\nu - x^\nu p^\mu)_{CG}$$

represents the intrinsic angular momentum of the system. It has no time components :

$$\sum (x^o p^k - x^k \varepsilon)_{CG} = 0$$

since $\sum p^k = 0$ and $\sum x^k \varepsilon = X^k \cdot \sum \varepsilon = 0$.

2658

We define now

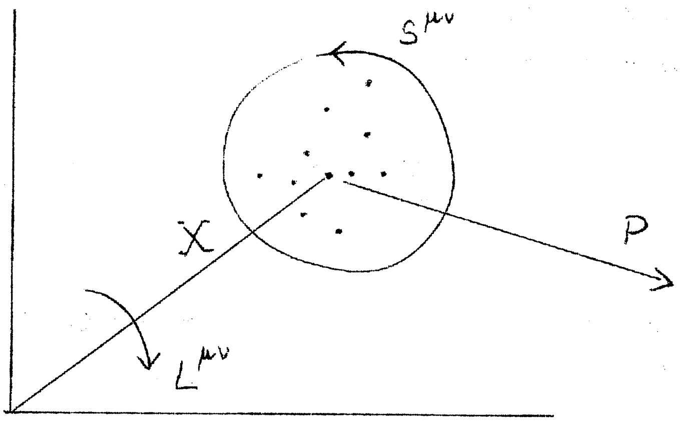
$$\sum (x^\mu p^\nu - x^\nu p^\mu)_{CG} \equiv (S^{\mu\nu})_R$$
$$(X^\mu P^\nu - X^\nu P^\mu)_R \equiv (L^{\mu\nu})_R \qquad (I.143)$$

Only the three components of $(S^{\mu\nu})_R$ are different from zero : $(S^{23}, S^{31}, S^{12})_R$ and they can be considered as the three components of an axial vector $\vec{s}_R$. We shall come back to this below.

Now by an arbitrary Lorentz transformation :

$$M^{\mu\nu} = S^{\mu\nu} + L^{\mu\nu} = S^{\mu\nu} + X^\mu P^\nu - X^\nu P^\mu \quad . \qquad (I.144)$$

We have thus separated in a covariant way the orbital part $L^{\mu\nu}$ from the intrinsic part $S^{\mu\nu}$ of the total angular momentum $M^{\mu\nu}$ [see figure] :



We note one important covariant property of $S^{\mu\nu}$, namely

$$S^{\mu\nu} P_\nu = 0 \; . \qquad (I.145)$$

This is true in the rest system R, because there $S^{ok} = 0$ and $\vec{P} = 0$. But if the four vector $S^{\mu\nu} P_\nu$ is zero in one system, then it is zero in all Lorentz systems (all its components vanish).

Eq. (I.145) is an implicit definition of $S^{\mu\nu}$, but it cannot serve to find $S^{\mu\nu}$ once $M^{\mu\nu}$ is given. On the other hand, since we were able to define $S^{\mu\nu}$ in a covariant way (although in a rather complicated description involving words), we also should be able to define it by means of a covariant formula (remember the remarks on p. 27).

Roughly speaking, we shall do something like this : by multiplying $M^{\mu\nu}$ by P_ν we obtain

$$M^{\mu\nu} P_\nu = L^{\mu\nu} P_\nu$$

since

$$S^{\mu\nu} P_\nu = 0.$$

If we now could undo the multiplication by P_ν , then we would have projected out the $L^{\mu\nu}$ part and then $S^{\mu\nu} = M^{\mu\nu} - L^{\mu\nu}$. Actually it is not quite so simple but we can do it in two steps :

i) we introduce the completely antisymmetric tensor $\varepsilon_{\mu\nu\rho\sigma}$ of rank four

$$\varepsilon_{\mu\nu\rho\sigma} = \begin{cases} +1 & \text{if } (\mu\nu\rho\sigma) \text{ is even} \\ -1 & \text{if } (\mu\nu\rho\sigma) \text{ is odd .} \end{cases} \tag{I.146}$$

"Even" and "odd" mean that $(\mu\nu\rho\sigma)$ is obtained from (0123) by an even or odd number of transpositions respectively. Consequently : $\varepsilon_{\mu\nu\rho\sigma}$ is zero if any two indices are equal; $\varepsilon_{\mu\nu\rho\sigma} = -\varepsilon_{\nu\rho\sigma\mu}$ (cyclic permutation). Furthermore, raising or lowering one index (by means $g^{\mu\nu}$) changes the sign if that index was 1,2,3 ; it does not change the sign if that index was 0.

With this $\varepsilon_{\mu\nu\rho\sigma}$ we define the (pseudo-) four vectors

$$W^\mu = \tfrac{1}{2}\varepsilon^\mu{}_{\nu\rho\sigma} P^\nu M^{\rho\sigma} = \tfrac{1}{2}\varepsilon^\mu{}_{\nu\rho\sigma} P^\nu S^{\rho\sigma}$$
$$S^\mu = \frac{1}{m} W^\mu \quad ; \quad m \equiv \sqrt{p^2} \ . \tag{I.147}$$

The second member of the first equation is true since the orbital part drops out :

$$\varepsilon^{\mu}{}_{\nu\rho\sigma} P^{\nu} (X^{\rho} P^{\sigma} - X^{\sigma} P^{\rho}) = 0;$$

each term vanishes because of the antisymmetry of ε . The physical significance of W^{μ} is easily found by going to the rest system where $\vec{P} = 0$ and $P^0 = m$:

$$(W^0)_R = \tfrac{1}{2} \varepsilon^{0}{}_{0\rho\sigma} (P^0 S^{\rho\sigma})_R = 0,$$

because of the antisymmetry of ε .

$$(W^1)_R = (\tfrac{1}{2} \varepsilon^{1}{}_{023} P^0 S^{23} + \tfrac{1}{2} \varepsilon^{1}{}_{032} P^0 S^{32})_R = m(S^{23})_R ,$$

similarly :

$$(W^2)_R = m(S^{31})_R \qquad \text{and} \qquad (W^3)_R = m(S^{12})_R \ .$$

Hence

$$(W^{\mu})_R = m(0,\ S^{23},\ S^{31},\ S^{12})_R = m(S^{\mu})_R \equiv m(0,\ \vec{s}_R) \ . \qquad \text{(I.148)}$$

We see that in the rest frame W^{μ} reduces to a three vector $m\,\vec{s}_R$, whose three components are equal to the three non-vanishing components of (m times) the intrinsic angular momentum tensor. In this particular system R the intrinsic angular momentum can therefore be described by a (pseudo- or axial) vector $\vec{s}_R$ (similarly as the magnetic field can be described by an axial vector). Now, if we go to an arbitrary frame, W^{μ} transforms as a four vector and $S^{\mu\nu}$ as a tensor. They are different things indeed, but in R "they touch each other" in the sense of Eq. (I.148).

If we allow, in our thoughts, our system of N spinless particles to shrink to almost one point, so that we no longer can distinguish its N constituents; then it becomes what we would call a particle with mass m, momentum P and intrinsic angular momentum or spin $S^{\mu\nu}$ [the square of the magnitude of the spin would be the invariant $s^2 = |\vec{s}_R|^2 = -S^\mu S_\mu$]. It should be stressed, however, that we are playing here with an analogy only : by this kind of argument we never would obtain half-integer values of s, even not by invoking quantum mechanics. But quantum mechanics provide us with an additional kind of angular momentum, namely, the spin which, as a kinematical quantity, behaves indeed as our $S^{\mu\nu}$, in particular it holds that $S^{\mu\nu}P_\nu = 0$. [See remark below Eq. (I.168)].

But it can no longer be considered as coming from the bodily rotation of N mass points clustered together. Therefore, from now on, we shall consider $S^{\mu\nu}$ as a really new thing which cannot be understood from analyzing the "structure and rotation" of the particle. If, e.g., one considers an electron as a rotating sphere of uniform mass distribution, then its moment of inertia is $T = \frac{2}{5} mr^2$, its angular momentum is $s = \frac{1}{2}\hbar = \omega \cdot T$ where $\omega = \frac{v}{r}$ is its angular velocity and v is the velocity at its equator. Then $v = \frac{r}{2}\frac{\hbar}{T} = \frac{5}{4}\frac{\hbar}{mr}$; putting $r = \frac{e^2}{mc^2}$ one finds $\frac{v}{c} = \frac{5}{4}\frac{\hbar c}{e^2} = \frac{5}{4} \cdot 137 \gg c$. This contradicts relativity !

Clearly W^μ and S^μ are identical with our previously defined quantities because they are the same in the rest frame. However, so far this is only true as long as we deal with one single spinning particle. We shall remove partly this restriction later on.

ii) we see that in Eq. (I.147) the intrinsic part $S^{\mu\nu}$ is projected out of $M^{\mu\nu}$. And now this equation can in fact be solved for $S^{\mu\nu}$ in the following way :

$$S^{\mu\nu} = \frac{1}{m^2}\varepsilon^{\mu\nu}{}_{\alpha\beta}\, P^\alpha W^\beta \qquad \text{(I.149)}$$

which is the inverse of (I.147). Since this is a covariant equation, it suffices to prove it in the rest system ($\vec{P} = 0$; $P^0 = m$). Indeed :

$$(S^{12})_R = \frac{1}{m^2}\,\varepsilon^{12}{}_{03}\,(P^oW^3)_R = \frac{1}{m^2}\,\varepsilon^{12}{}_{03}\,m\cdot m(S^{12})_R = (S^{12})_R$$

$$(S^{ok})_R = \frac{1}{m^2}\,\varepsilon^{ok}{}_{oi}\,(P^oW^i)_R = 0 \quad \text{because} \quad \varepsilon^{ok}{}_{oi} = 0.$$

And now we only need to introduce (I.147) :

$$W^{\beta} = \tfrac{1}{2}\varepsilon^{\beta}{}_{\lambda\rho\sigma}\,P^{\lambda}M^{\rho\sigma}$$

into (I.149) to obtain

$$S^{\mu\nu} = \left[\frac{1}{2m^2}\,\varepsilon^{\mu\nu}{}_{\alpha\beta}\,\varepsilon^{\beta}{}_{\lambda\rho\sigma}\,P^{\alpha}P^{\lambda}\right]\cdot M^{\rho\sigma} \equiv \sum{}^{\mu\nu}{}_{\rho\sigma}\cdot M^{\rho\sigma}$$

(I.150)

$$\sum{}^{\mu\nu}{}_{\rho\sigma} = \frac{1}{2m^2}\,\varepsilon^{\mu\nu}{}_{\alpha\beta}\,\varepsilon^{\beta}{}_{\lambda\rho\sigma}\,P^{\alpha}P^{\lambda} = \varepsilon^{\mu\nu}{}_{\alpha\beta}\left(\frac{P^{\beta}P_{\lambda}}{2m^2}\right)\varepsilon^{\lambda\alpha}{}_{\rho\sigma}$$

where $\sum^{\mu\nu}{}_{\rho\sigma}$ is the desired projection operator. Indeed :

$$S^{\mu\nu}P_{\nu} = 0$$

since $\varepsilon^{\mu\nu}{}_{\alpha\beta}\,\varepsilon^{\beta}{}_{\lambda\rho\sigma}\,P^{\alpha}P^{\lambda}P_{\nu} = 0$ because of the asymmetry of ε .

We have accomplished two things : we have split the angular momentum tensor $M^{\mu\nu}$ for one spinning particle (or else for a system of spinless particles) in a covariant way into the intrinsic part $S^{\mu\nu}$ and the orbital part $L^{\mu\nu} = M^{\mu\nu} - S^{\mu\nu}$. And we have shown that the intrinsic part is that part which in the CG-system survives and there (as well as in the rest system) has only three components S^{23}, S^{31}, S^{12} different from zero. They can be used to define a new (and different) quantity : the four vector $(S^{\mu})_R = (0, S^{23}, S^{31}, S^{12})_R$ and this four vector is identical to the polarization four vector (defined earlier) as well as to the S^{μ} defined by Eqs. (I.147). There is one restriction, however : this is only true for one single particle with spin.

We shall now loosen this restriction somewhat by going one step further and considering a system of s p i n n i n g particles with

$$\begin{cases} \text{momentum } p \\ \text{mass } m = \sqrt{p^{\mu} p_{\mu}} \\ \text{co-ordinate } x^{\mu} \\ \text{total angular momentum } M^{\mu\nu} = x^{\mu} p^{\nu} - x^{\nu} p^{\mu} + S^{\mu\nu}. \end{cases} \qquad (I.151)$$

We must, however, forget that the spinning particle was obtained from a system of spinless particles by a limiting process and consider the properties in (I.151) as those of a particle with true spin.

Suppose we have now a system of N such spinning particles (with an index $i = 1\ldots N$ labelling them) and let us consider the total angular momentum of that system :

$$M^{\mu\nu} = \sum_i M_i^{\mu\nu} = \sum_i (x_i^{\mu} p_i^{\nu} - x_i^{\nu} p_i^{\mu} + S_i^{\mu\nu}). \qquad (I.152)$$

This can be written in two parts, namely :

$$M^{\mu\nu} = \sum_i L_i^{\mu\nu} + \sum_i S_i^{\mu\nu}$$

but this does in general amount to nothing because we are not able to disentangle these two parts unless we know the state of motion of each particle separately. Why ? The answer comes from our previous considerations on a system of particles. Namely, we can go step by step from Eq. (I.137) to Eq. (I.150), but applying everything to the part

$$\sum_i L_i^{\mu\nu} = \sum_i (x^{\mu} p^{\nu} - x^{\nu} p^{\mu})_i$$

only. The result is that already $\sum_i L_i^{\mu\nu}$ splits into two terms :

$$\sum_i L_i^{\mu\nu} \equiv L^{\mu\nu} + S_{orbit}^{\mu\nu} = X^\mu P^\nu - X^\nu P^\mu + S_{orbit}^{\mu\nu}$$

where $S_{orbit}^{\mu\nu}$ is that part of $M^{\mu\nu}$, which comes from the orbital motion of the particles relative to the centre of gravity X, but which does not contain contributions from the spins of the particles. It is, however, also an intrinsic angular momentum of our system, although only a part of it. Therefore

$$M^{\mu\nu} = \sum_i L_i^{\mu\nu} + \sum_i S_i^{\mu\nu} = X^\mu P^\nu - X^\nu P^\mu + \left[S_{orbit}^{\mu\nu} + S_{spin}^{\mu\nu} \right] \quad (I.153)$$

where $\sum_i S_i^{\mu\nu} = S_{spin}^{\mu\nu}$.

We see : the intrinsic angular momentum is now

$$S^{\mu\nu} = \left[S_{orbit}^{\mu\nu} + S_{spin}^{\mu\nu} \right] = M^{\mu\nu} - (X^\mu P^\nu - X^\nu P^\mu) = M^{\mu\nu} - L^{\mu\nu}. \quad (I.154)$$

If we consider the system as a whole, then $M^{\mu\nu}$, X^μ, P^μ are known and therefore $\left[S_{orbit}^{\mu\nu} + S_{spin}^{\mu\nu} \right]$ can be defined in a covariant way, but not $S_{spin}^{\mu\nu}$ separately. For that it would be necessary to know $S_{orbit}^{\mu\nu}$, but this requires the knowledge of each x_i^μ and p_i^μ $(i = 1...N)$. Separating covariantly the true "spin part" $S_{spin}^{\mu\nu}$ off from the angular momentum $M^{\mu\nu}$ of a system means of course to accomplish this separation in using only the four vectors and tensors $M^{\mu\nu}$, P^μ, X^μ pertaining to the system as a whole. In this sense it is not generally possible to define the spin part of a system of particles covariantly.

Our projection operator $\Sigma^{\mu\nu}{}_{\rho\sigma}$ from Eq. (I.150) would fail here even to project out $S^{\mu\nu} = S_{orbit}^{\mu\nu} + S_{spin}^{\mu\nu}$. Namely :

$$\Sigma^{\mu\nu}{}_{\rho\sigma} M^{\rho\sigma} = \Sigma^{\mu\nu}{}_{\rho\sigma} L^{\rho\sigma} + \Sigma^{\mu\nu}{}_{\rho\sigma} S_{orbit}^{\rho\sigma} + \Sigma^{\mu\nu}{}_{\rho\sigma} S_{spin}^{\rho\sigma}. \quad (I.155)$$

The first term is zero and the second term equals $S_{orbit}^{\mu\nu}$, both because of the properties of $\Sigma^{\mu\nu}{}_{\rho\sigma}$ [see after Eq. (I.147), and the derivation of $\Sigma^{\mu\nu}{}_{\rho\sigma}$]. Therefore

$$\Sigma^{\mu\nu}{}_{\rho\sigma} M^{\rho\sigma} = S^{\mu\nu}_{orbit} + (\Sigma^{\mu\nu}{}_{\rho\sigma} S^{\rho\sigma}_{spin}) . \qquad (I.156)$$

The last term is, in general, not equal to $S^{\rho\sigma}_{spin}$. Namely, contracting it with P_ν gives

$$\left(\Sigma^{\mu\nu}{}_{\rho\sigma} S^{\rho\sigma}_{spin} \right) P_\nu = 0$$

since $\Sigma^{\mu\nu}{}_{\rho\sigma} P_\nu = 0$ [see after Eq. (I.150)].

On the other hand, since $S^{\mu\nu}_{spin} = \sum_i S^{\mu\nu}_i$, we have

$$S^{\mu\nu}_{spin} P_\nu = \left(\sum_i S^{\mu\nu}_i \right) P_\nu \neq \sum_i S^{\mu\nu}_i p_{i,\nu} = 0 . \qquad (I.157)$$

Therefore in general

$$\left. \begin{aligned} \Sigma^{\mu\nu}{}_{\rho\sigma} S^{\rho\sigma}_{spin} &\neq S^{\mu\nu}_{spin} \\ \Sigma^{\mu\nu}{}_{\rho\sigma} M^{\rho\sigma} &\neq S^{\mu\nu} \\ S^{\mu\nu} P_\nu &\neq 0 \end{aligned} \right\} \qquad (I.158)$$

The reason for all this is that it is essential that a Lorentz system exists, in which all time components of $M^{\mu\nu}$ vanish. [This is so because the operation

$$W^\mu = \tfrac{1}{2} \varepsilon^\mu{}_{\nu\rho\sigma} P^\nu S^{\rho\sigma}$$

has in the rest system the effect of annihilating the time components [see (I.148)]. This is, of course, permissible only if they are zero anyway. If not, then the following operation $\frac{1}{m^2} \varepsilon^{\mu\nu}{}_{\rho\sigma} P^\rho W^\sigma$, which cannot repair the damage made to $S^{\rho\sigma}$, will not restore the old quantity.] Such a system could be shown to exist in our consideration on a system of N spinless particles, it was the centre of gravity system. However, in a system of N spinning particles, no such system exists in general; writing namely :

$$M^{\mu\nu} = L^{\mu\nu} + S^{\mu\nu}_{orbit} + S^{\mu\nu}_{spin}$$

one sees that one can achieve $L^{\mu\nu} = 0$ and $S^{ok}_{orbit} = 0$ by going to the CG-system. But in that system $(S^{\mu\nu}_{spin})_{CG}$ will, in general, have time components because

$$(S^{\mu\nu}_{spin})_{CG} = \sum_{j=1}^{N} (S^{\mu\nu}_{j})_{CG}$$

and the individual $(S^{\mu\nu}_{j})_{CG}$ have time components. Namely, the time components of the $S^{\mu\nu}_{j}$ vanish only in the individual rest system R_j of the particle j and the value $(S^{\mu\nu}_{j})_{CG}$ is obtained from $(S^{\mu\nu}_{j})_{R_j}$ by a Lorentz transformation to the CG-system - by which time components are generated. As each particle has then, in general, its own Lorentz transformation, there is no hope that in

$$\sum_{j=1}^{N} (S^{\mu\nu}_{j})_{CG}$$

these time components should cancel ; this is expressed by Eq. (I.157). Therefore there is in general no Lorentz system in which the time components of $M^{\mu\nu}$ are zero and therefore our procedure does not work *). This is a consequence of considering the $S^{\mu\nu}_{i}$ as the t r u e s p i n, which cannot be reduced to an expression of the form

$$\sum \xi^{\mu}_{i} \pi^{\nu}_{i} - \xi^{\nu}_{i} \pi^{\mu}_{i}$$

over the internal co-ordinates ξ^{μ}_{i} and momenta π^{μ}_{i} of the particle. Otherwise, of course, we could consider this as a cluster of spinless particles again, sum in the old way over all particles of all clusters and have the old situation in which - as we know - the projection operator works.

---

*) This is, by the way, the reason why one cannot (in general) describe the magnetic field by a four vector : there does not necessarily exist a system where all components of the electric field (= the time component $F^{\mu\nu}$) vanish.

It seems therefore that everything we have achieved breaks down if we consider a system of spinning particles.

Luckily enough, there is an exception : it still works if we know the individual momenta of all N particles. This happens in two most important cases :

i) if only very few particles are present, e.g., in the decay $M \longrightarrow m+\mu$ or similar situations.

ii) if we have a beam of like particles with equal sharp momentum then we know in fact the individual momenta and they all are equal. Therefore in the rest system $(S^{\mu\nu}_{orbit})_R = 0$ because there all particles are at rest and no orbital motion around the origin remains. Finally, as all particles are at rest, also the individual $S^{\mu\nu}_j$ have no time components and thus their sum $(\sum_j S^{\mu\nu}_j)_R$ has no time components either. Hence in the rest system

$$(M^{\mu\nu})_R = (L^{\mu\nu})_R + (S^{\mu\nu}_{spin})_R = (L^{\mu\nu})_R + (\sum_j S^{\mu\nu}_j)_R ,$$

and, by a general Lorentz transformation,

$$M^{\mu\nu} = L^{\mu\nu} + S^{\mu\nu}_{spin} . \qquad (I.159)$$

[$S^{\mu\nu}_{orbit} = 0$ holds in every Lorentz system, since it is true in the rest system.] Indeed, our projection operators work here

α) $S^{\mu\nu}_{spin} P_\nu = N \cdot \sum S^{\mu\nu}_j p_\nu = 0$

because the particle's momenta are equal :

$p_\nu = \frac{P_\nu}{N}$ and $S^{\mu\nu}_j p_\nu = 0$

β) $\Sigma^{\mu\nu}{}_{\rho\sigma} M^{\rho\sigma} = S^{\mu\nu}_{spin}$,

this is true because the P^ν in the projection operator can be written $N \cdot p^\nu$.

(I.160)

2658

Then $\Sigma^{\mu\nu}{}_{\rho\sigma}$ as a whole annihilates the $L^{\mu\nu}$ part of $M^{\mu\nu}$ and, when it is written as [see (I.150)]

$$\Sigma^{\mu\nu}{}_{\rho\sigma} = \frac{1}{2N^2m^2}\, \varepsilon^{\mu\nu}{}_{\alpha\beta}\, \varepsilon^{\beta}{}_{\lambda\rho\sigma}\, p^{\alpha} p^{\lambda}$$

it clearly reproduces each individual $S_j^{\mu\nu}$ and thus also $S^{\mu\nu}_{spin} = \sum_j S_j^{\mu\nu}$.

This shows that in the important case of a well-defined beam of like particles (or else for a plane wave state describing a spinning particle) it is indeed possible to separate in a covariant way the true spin part $S^{\mu\nu}_{spin}$ of the total angular momentum from the $L^{\mu\nu}$-part. In this case $S^{\mu\nu}$ and $W^{\mu} = mS^{\mu}$ as defined by Eqs. (I.150) and (I.147) respectively, can likewise be used to describe the state of polarization of the beam. Then S^{μ} and W^{μ} are identical with the quantities defined in sub-sections 8)-(a) and 8)-(b) above [e.g., Eqs. (I.122) and (I.135)].

It should be stressed that neither $L^{\mu\nu}$ nor $S^{\mu\nu}_{spin}$ is conserved, but only $M^{\mu\nu}$.

Since a beam of like particles with sharp momentum is described in quantum mechanics as a plane wave state, we expect also a close relation between $S^{\mu\nu}$ and S^{μ} and quantum mechanical operators. We shall establish this relation for the case of Dirac particles [for details, see, e.g., Schweber, Relativistic Quantum Field Theory, 1961, p. 74-85. In the same book, p. 18-53, one also finds some further information about our polarization four vector W^{μ}, whose operator counterpart plays an essential role in the classification of the representations of the Lorentz group and thus in the classification of all possible types of relativistic free fields].

(d) The correspondence between the polarization four vector, angular momentum and γ -matrices in Dirac's theory

We choose that particular representation of the γ -relation

$$\gamma^{\mu}\gamma^{\nu} + \gamma^{\nu}\gamma^{\mu} = 2g^{\mu\nu} \qquad (I.161)$$

in which the components of the spinors split in a natural way into two large and two small ones in the non-relativistic limit :

$$\gamma^{0} = \gamma_{0} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; \quad \gamma^{k} = -\gamma_{k} = \begin{pmatrix} 0 & \sigma^{k} \\ -\sigma^{k} & 0 \end{pmatrix} ; \quad \gamma^{5} = -\gamma_{5} = i\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\sigma^{1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} ; \quad \sigma^{2} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} ; \quad \sigma^{3} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad (I.162)$$

where the γ's are of course 4x4 matrices, the σ's are 2x2. In this particular representation, one obtains for free particles the spinors

$$\psi(x) = e^{-ipx} \cdot U(p)$$

$$U(p) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} \qquad (I.163)$$

$$\begin{pmatrix} u_3 \\ u_4 \end{pmatrix} = \frac{\vec{\sigma}\vec{p}}{E+m} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

where $u_1(p)$ and $u_2(p)$ are arbitrary. Obviously $\begin{pmatrix} u_3 \\ u_4 \end{pmatrix}$ vanishes if $\vec{p} \to 0$.

We now consider two classes of operators

$$\Sigma^{\mu} \equiv i\gamma^5\gamma^{\mu}$$
$$\sigma^{\mu\nu} \equiv \frac{i}{2}(\gamma^{\mu}\gamma^{\nu} - \gamma^{\nu}\gamma^{\mu}). \tag{I.164}$$

One shows easily, by means of (I.162), that

$$\Sigma^{o} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}; \quad \Sigma^{k} = \begin{pmatrix} \sigma^{k} & 0 \\ 0 & -\sigma^{k} \end{pmatrix}$$
$$\sigma^{ok} = i\begin{pmatrix} 0 & \sigma^{k} \\ \sigma^{k} & 0 \end{pmatrix}; \quad \sigma^{jk} = \begin{pmatrix} \sigma^{\ell} & 0 \\ 0 & \sigma^{\ell} \end{pmatrix} \text{ (jk}\ell\text{ cycl.)} . \tag{I.165}$$

We define now

$$S^{\mu} \equiv \langle \Sigma^{\mu} \rangle \equiv \bar{\Psi} i\gamma^5\gamma^{\mu}\Psi \quad ; \quad \bar{\Psi} \equiv \Psi^{*}\gamma^{o}$$
$$S^{\mu\nu} \equiv \langle \sigma^{\mu\nu} \rangle \equiv \bar{\Psi}\frac{i}{2}(\gamma^{\mu}\gamma^{\nu} - \gamma^{\nu}\gamma^{\mu})\Psi \tag{I.166}$$

and show that they are identical with our old S^{μ} and $S^{\mu\nu}$ respectively. Since the above expressions are manifestly covariant, we only need to check the assertion in the rest system, where $u_3 = u_4 = 0$.

Since, according to (I.165), Σ^{o} and σ^{ok} both interchange $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ with $\begin{pmatrix} u_3 \\ u_4 \end{pmatrix}$, we have

$$\langle \Sigma^{o} \rangle_R = \langle \sigma^{ok} \rangle_R = 0 .$$

Furthermore Σ^k and σ^{jk} do not interchange $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ with $\begin{pmatrix} u_3 \\ u_4 \end{pmatrix}$, thus

$$\langle \Sigma^k \rangle_R = (u_1^* u_2^*)\, \sigma^k \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \langle \sigma^k \rangle_R$$

$$\langle \sigma^{jk} \rangle_R = (u_1^* u_2^*)\, \sigma^\ell \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \langle \sigma^\ell \rangle_R \qquad (jk\ell \quad \text{cycl.}) .$$

Hence

$$(S^\mu)_R = (0, \langle \vec{\sigma} \rangle)_R$$

$$(S^{\mu\nu})_R = \left\{ \begin{array}{l} \langle \sigma^{ok} \rangle = 0 \\ \\ \langle \sigma^{jk} \rangle = \langle \vec{\sigma} \rangle \end{array} \right\}_R \qquad \text{(I.167)}$$

That is : in the rest system S^μ and $S^{\mu\nu}$ reduce (in a loose sense) to our old quantities. Therefore this is true in all Lorentz systems and we have established the correspondence

$$S^{\mu\nu} \longleftrightarrow \sigma^{\mu\nu} = \tfrac{i}{2}(\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu)$$

$$S^\mu \longleftrightarrow \Sigma^\mu = i\gamma^5 \gamma^\mu \qquad \text{(I.168)}$$

Eq. (I.167) also shows that the assertion that for the true spin the relation $S^{\mu\nu} P_\nu = 0$ holds, is at least fulfilled for spin $\frac{1}{2}$ particles.

The operator Σ^μ is closely related to a covariant projection operator P_t with the property that for any spinor ψ, the projected spinor

$$\psi_t \equiv P_t \psi \qquad \text{(I.169)}$$

has the expectation value of the spin pointing in the direction $\vec{t}$, namel

$$\langle \sigma^{23}, \sigma^{31}, \sigma^{12} \rangle$$

parallel to $\vec{t}$. Here t is a four vector

$$t^{\mu} = (t^{o}, \vec{t}) \qquad \text{(I.}$$

whose zero component is chosen such that

$$p^{\mu} t_{\mu} = 0 . \qquad \text{(I.}$$

The projection operator then assumes the form

$$P_{t} = \tfrac{1}{2}(1 + \Sigma^{\mu} \quad t_{\mu} \cdot \frac{E}{|E|}) . \qquad \text{(I.}$$

For further information see : Hamilton, the Theory of Elementary Particle 1959, p. 124-129.