Whenever, one does complex calculation you should check your results by calculating **limits** where the answer is known.

Letting $t \to 0$ we get

$$\sqrt{1+\left(\frac{F}{m_0 c}\right)^2 t^2} \approx 1+\frac{1}{2}\left(\frac{F}{m_0 c}\right)^2 t^2$$

$$v \approx \frac{F}{m_0}t \quad , \quad a \approx \frac{F}{m_0} \quad , \quad \mathrm{x} \approx \frac{1}{2}\frac{F}{m_0}t^2 \quad \text{as expected}$$

Letting $t \to \infty$ we get

$$v \to c \quad \text{and} \quad x \to ct \quad \text{as expected}$$

Digression to 4-Vectors

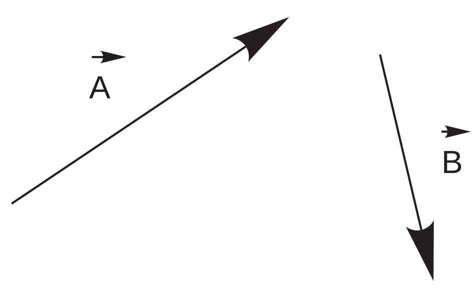
What is an ordinary vector in 3-dimensional space?

A vector has many levels of complexity and is a very abstract mathematical object. A vector is a mathematical(geometrical) object that is representable by two numbers in two dimensions, three numbers in three dimensions, and so on. One characterization is to specify its magnitude or length and orientation or direction - imagine that it is a directed line segment. As we shall see, quantum mechanics will be formulated in terms of vectors, but they will not be directed line segments.

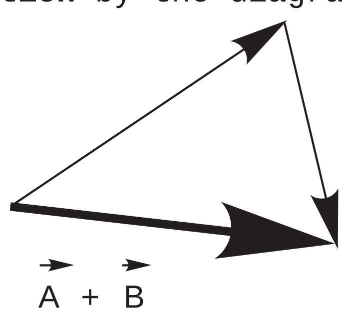
The Standard Language of Vectors

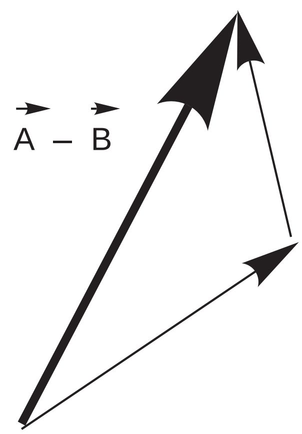
As we said, in ordinary space, we can represent a vector by a directed line segment(an arrow). A straightforward property of a vector is multiplication of the vector by a scalar (a real number) $\vec{C} = \alpha\vec{A}$. In this case the magnitude of the vector changes and the direction stays the same (it might reverse if $\alpha < 0$).

Now given two vectors as shown below



we define the sum and difference of the two vectors or the general property **vector addition** by the diagrams shown below:





Clearly vector addition as defined above, i.e.,

$$\vec{C} = \vec{A} + \vec{B}$$
$$\vec{D} = \vec{A} - \vec{B} = \vec{A} + (-\vec{B})$$

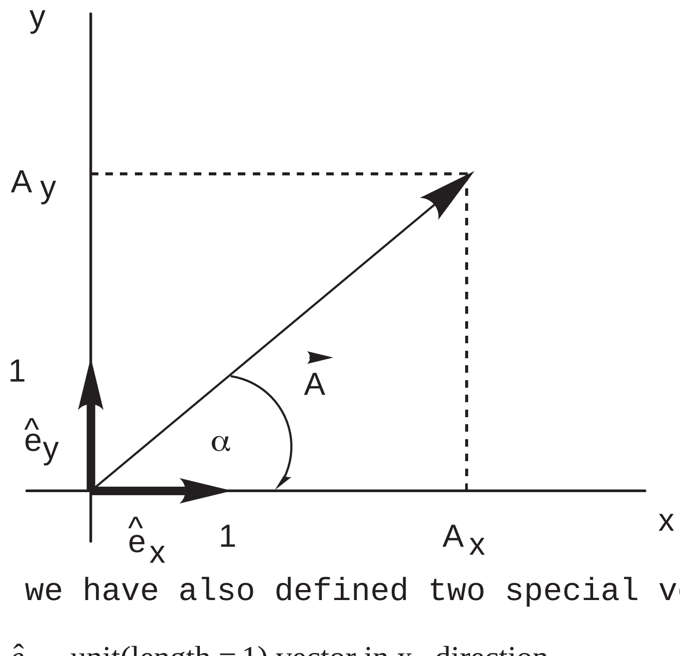
yields a new vector in each case. This new vector can have both a different direction and a different magnitude than either of the two vectors that are used to create it.

These two properties allow us to define a **linear combination** of vectors as $\vec{C} = \alpha\vec{A} + \beta\vec{B}$, which is also a well-defined vector.

Although this is a perfectly good way to proceed, it will not allow us to generalize the notion of a vector beyond ordinary space, which is an arena that will turn out to be much too confining in our effort to understand quantum mechanics later.

We need to formulate these same concepts in another way.

Consider the vector shown below:



In this figure, we have also defined two special vectors, namely,

$$\hat{e}_x = \text{unit(length} = 1) \text{ vector in x - direction}$$
$$\hat{e}_y = \text{unit(length} = 1) \text{ vector in y - direction}$$

In terms of these unit vectors we can write

$$\vec{A} = A_x\hat{e}_x + A_y\hat{e}_y$$

where

$$A_x\hat{e}_x = \text{vector of length } A_x \text{ in the x-direction}$$
$$A_y\hat{e}_y = \text{vector of length } A_y \text{ in the y-direction}$$

and the sum of these two vectors equals $\vec{A}$ because of the rule for adding vectors that we stated earlier.

We now define

$$A_x = \text{component of vector } \vec{A} \text{ in the x-direction}$$
$$A_y = \text{component of vector } \vec{A} \text{ in the y-direction}$$

From the diagram it is also clear that

$$A_x = A\cos\alpha \quad \text{and} \quad A_y = A\sin\alpha$$

where

A = length of the vector $\vec{A} = \sqrt{A_x^2 + A_y^2}$
(by Pythagorous theorem)

We can then redefine vector addition in terms of components and unit vectors as follows:

$$\vec{A} = A_x\hat{e}_x + A_y\hat{e}_y$$
$$\vec{B} = B_x\hat{e}_x + B_y\hat{e}_y$$
$$\vec{A} + \vec{B} = (A_x + B_x)\hat{e}_x + (A_y + B_y)\hat{e}_y$$
$$\vec{A} - \vec{B} = (A_x - B_x)\hat{e}_x + (A_y - B_y)\hat{e}_y$$

i.e., we can just **add and subtract components.**

We now define an important new mathematical object using unit vectors. It is the **scalar or inner product** and its symbol is a . (dot). We define this operation with a set of rules involving the unit vectors:

$$\hat{e}_x \cdot \hat{e}_x = 1 = \hat{e}_y \cdot \hat{e}_y$$
$$\hat{e}_x \cdot \hat{e}_y = 0 = \hat{e}_y \cdot \hat{e}_x$$

or

$$\hat{e}_i \cdot \hat{e}_j = \delta_{ij} = \begin{cases} 1 & \text{i = j} \\ 0 & \text{i} \neq \text{j} \end{cases} \quad = \text{ Kronecker delta}$$

The inner product satisfies the following relations;

$$(\alpha\hat{e}_i)\cdot(\beta\hat{e}_j) = \alpha\beta\hat{e}_i \cdot \hat{e}_j$$
$$(\alpha\hat{e}_i + \gamma\hat{e}_k)\cdot(\beta\hat{e}_j + \eta\hat{e}_m) = \alpha\beta\hat{e}_i \cdot \hat{e}_j + \alpha\eta\hat{e}_i \cdot \hat{e}_m + \gamma\beta\hat{e}_k \cdot \hat{e}_j + \gamma\eta\hat{e}_k \cdot \hat{e}_m$$

Using these defining relations we can determine the scalar product of any two vectors as follows

$$\vec{A} = A_x\hat{e}_x + A_y\hat{e}_y$$

$$\vec{B} = B_x\hat{e}_x + B_y\hat{e}_y$$

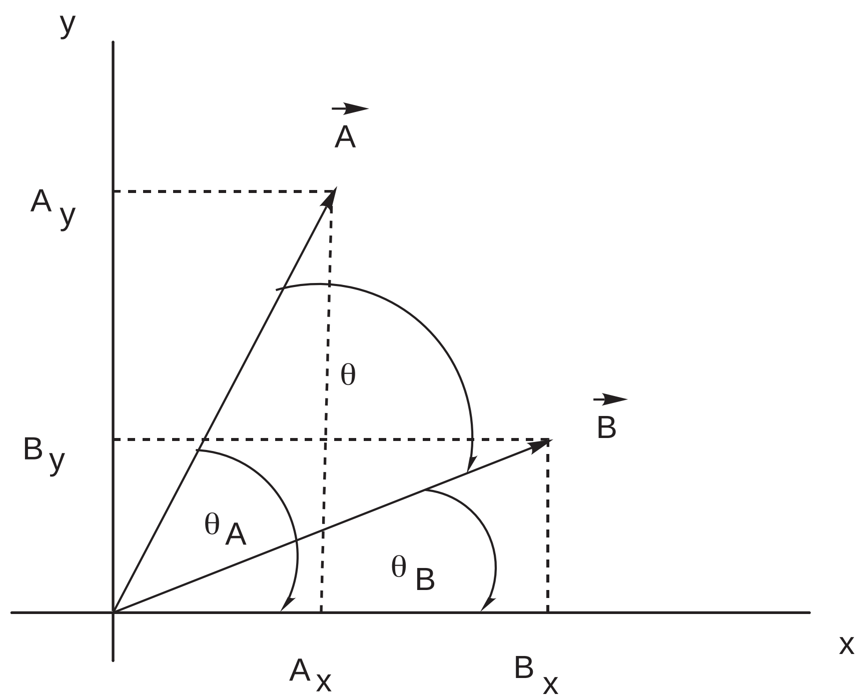
$$\begin{aligned}\vec{A}\cdot\vec{B} &= (A_x\hat{e}_x + A_y\hat{e}_y)\cdot(B_x\hat{e}_x + B_y\hat{e}_y)\\ &= A_xB_x\hat{e}_x\cdot\hat{e}_x + A_xB_y\hat{e}_x\cdot\hat{e}_y + A_yB_x\hat{e}_y\cdot\hat{e}_x + A_yB_y\hat{e}_y\cdot\hat{e}_y\\ &= A_xB_x(1) + A_xB_y(0) + A_yB_x(0) + A_yB_y(1)\\ &= A_xB_x + A_yB_y\end{aligned}$$

We note that

$$\vec{A}\cdot\vec{A} = A_xA_x + A_yA_y = A_x^2 + A_y^2 = A^2 = \textit{norm of } \vec{A}$$

$$A = \sqrt{\vec{A}\cdot\vec{A}} = \text{length of the vector } \vec{A}$$

Now looking at the diagram below we can derive another important result.



We have

$$\begin{aligned}\vec{A}\cdot\vec{B} &= A_xB_x + A_yB_y = AB\left(\cos(\theta_A)\cos(\theta_A) + \sin(\theta_A)\sin(\theta_A)\right)\\ &= AB\cos(\theta_A - \theta_B) = AB\cos\theta\end{aligned}$$

so that

$$\begin{aligned}\vec{A}\cdot\vec{B} &= AB\cos\theta\\ &= (\textit{length of } \vec{A})(\textit{length of } \vec{B})\cos(\textit{angle between } \vec{A} \textit{ and } \vec{B})\\ &= (\textit{length of } \vec{A})(\textit{length of } \vec{B} \textit{ in the direction of } \vec{A})\\ &= (\textit{length of } \vec{A})(\textit{projection of } \vec{B} \textit{ onto the direction of } \vec{A})\end{aligned}$$

Therefore, we have

$$\vec{B} = \vec{A} \rightarrow \theta = 0 \rightarrow \vec{A} \cdot \vec{A} = A^2 \quad \text{as before}$$

$$\vec{B} \text{ perpendicular(orthogonal) to } \vec{A} \rightarrow \theta = \frac{\pi}{2} = 90^\circ \rightarrow \vec{A} \cdot \vec{B} = 0$$

or vice versa

$$\textit{if } \vec{A} \cdot \vec{B} = 0 \text{, } \textit{then } \vec{A} \text{ is orthogonal to } \vec{B}$$

If two vectors satisfy $\vec{A} \cdot \vec{B} = 0$, then they are said to orthonormal = orthogonal. If a vector satisfies $\vec{A} \cdot \vec{A} = 1$, then it is said to normalized to one.

We also have for any vector

$$\vec{A} = A_x \hat{e}_x + A_y \hat{e}_y$$
$$\vec{A} \cdot \hat{e}_x = \left(A_x \hat{e}_x + A_y \hat{e}_y\right) \cdot \hat{e}_x = A_x = \text{x - component}$$
$$\vec{A} \cdot \hat{e}_y = \left(A_x \hat{e}_x + A_y \hat{e}_y\right) \cdot \hat{e}_y = A_y = \text{y - component}$$
$$\vec{A} = (\vec{A} \cdot \hat{e}_x)\hat{e}_x + (\vec{A} \cdot \hat{e}_y)\hat{e}_y$$

Generalizing to 3 dimensions we have

$$\vec{A} = A_x \hat{e}_x + A_y \hat{e}_y + A_z \hat{e}_z \; = \; \text{any vector in the vector space}$$

where the set of three orthonormal vectors $\left\{\hat{e}_x, \hat{e}_y, \hat{e}_z\right\}$ are called a **basis** for the vector space (any vector can be written as a linear combination of the basis vectors) and we have

$$\hat{e}_i \cdot \hat{e}_j = \begin{cases} 1 & \text{i = j} \\ 0 & \text{i} \neq \text{j} \end{cases} \qquad \text{i , j = x,y,z}$$

The number of required basis vectors is the number of numbers needed to characterize a general vector = the **dimension** of the space.

The entire collection of vectors we can generate from a basis set is called a **vector space.**

So in this room, I would need 3 numbers to characterize each vector. This room is a small part of a 3-dimensional vector space, which is called the **universe** at an instant of time.

Completely removing (x,y,z) from our notation(because it limits us to a maximum of 3 dimensions) we have

$$\vec{A} = \sum_{j=1}^{3} A_j \hat{e}_j$$

$$\hat{e}_k \cdot \vec{A} = \hat{e}_k \cdot \sum_{j=1}^{3} A_j \hat{e}_j = \sum_{j=1}^{3} A_j \hat{e}_k \cdot \hat{e}_j = \sum_{j=1}^{3} A_j \delta_{kj} = A_k = k^{th} \textit{ component}$$

so that

$$\vec{A} = \sum_{j=1}^{3} A_j \hat{e}_j = \sum_{j=1}^{3} \left(\hat{e}_j \cdot \vec{A}\right)\hat{e}_j$$

Returning to the Discussion of 4-Vectors

Consider a vector $\vec{A}$ representing some physical variable. Using cartesian unit vectors we can write

$$\vec{A} = \sum_i A_i \hat{e}_i$$

The components of the vector $A_i \quad i = 1,2,3$ are its representation in a given coordinate system. We must choose a coordinate system in order to define the unit vectors. The coordinate system is not an essential part of the physics however. We can just as well use any other coordinate system to define unit vector and the vector $\vec{A}$.

In particular, we consider another coordinate system with the same origin, but rotated from the first system. In another coordinate system we would write

$$\vec{A} = \sum_i A'_i \hat{e}'_i$$

Note that the vector $\vec{A}$ has not changed; only its representation (components) in the new system (new basis) has changed. We relate the two representations (components) as follows:

$$\sum_i A_i \hat{e}_i = \sum_i A'_i \hat{e}'_i$$

$$\hat{e}'_j \cdot \sum_i A_i \hat{e}_i = \hat{e}'_j \cdot \sum_i A'_i \hat{e}'_i = \sum_i A'_i \hat{e}'_i \cdot \hat{e}'_j \sum_i A'_i \delta_{ij} = A'_j$$

$$A'_j = \sum_i A_i (\hat{e}'_j \cdot \hat{e}_i)$$

The coefficients $(\hat{e}'_j \cdot \hat{e}_i)$ are numbers that are determined by the specific rotation. They are independent of the vector $\vec{A}$. We now redefine a vector:

> **A vector in 3 dimensions is a set of 3 numbers** $\{A_i\}$ **(components) which transform under a rotation of the coordinate system according to**
>
> $$A'_j = \sum_i A_i (\hat{e}'_j \cdot \hat{e}_i)$$

Any quantity which is unchanged by a coordinate transformation is called an **invariant** of the transformation. Since the principle of relativity requires that the results of physical theories (physical laws) be independent of the choice of coordinate system (must be inertial however), all physical laws must involve **only** invariants.

The dot product of two vectors is a scalar. Scalars are numbers that are independent of our choice of coordinate system. This gives us a method for **constructing** invariants. We can show that the dot product produces an invariant as follows:

$$\vec{A}'\cdot\vec{B}' = \left(\sum_i A_i\hat{e}_i\right)\cdot\left(\sum_j A_j\hat{e}_j\right) = \sum_{i,j} A_i B_j \hat{e}_i\cdot\hat{e}_j$$
$$= \sum_{i,j} A_i B_j \delta_{ij} = \sum_i A_i B_i = \vec{A}\cdot\vec{B}$$

In particular, the norm or length-squared of a vector, $A^2 = \vec{A}\cdot\vec{A}$, is a scalar invariant. We now define a rotation.

A rotation is any transformation which leaves $r^2 = \vec{r}\cdot\vec{r} = x^2 + y^2 + z^2$ **invariant**

In Minkowski 4-dimensional spacetime we define vectors in a different manner. Both the ordinary space 3-dimensional and the Minkowski 4-dimensional vector definitions are special cases of a more general definition. The ordinary 3-dimensional definition corresponds to Euclidean geometry.

In Minkowski 4-dimensional spacetime we write the spacetime 4-vector in this way

$$\overleftarrow{s} = (ct, x, y, z)$$

and the scalar product of the vector with itself (its norm) as

$$\overleftarrow{s}\cdot\overleftarrow{s} = c^2t^2 - x^2 - y^2 - z^2 \quad \text{(note the minus signs)}$$

This is a scalar invariant under Lorentz transformations(it is the spacetime interval). In fact, any set of 4 numbers $\overleftarrow{A} = (A_0, A_1, A_2, A_3)$ represents a Minkowski 4-vector if its norm defined by

$$\overleftarrow{A}\cdot\overleftarrow{A} = A_0^2 - A_1^2 - A_2^2 - A_3^2$$

is a scalar invariant. In addition, if a set of 4 numbers is a 4-vector then the components transform between frames via the Lorentz transformations as

$$A'_0 = \gamma(A_0 - \beta A_1)$$
$$A'_1 = \gamma(A_1 - \beta A_0)$$
$$A'_2 = A_2$$
$$A'_3 = A_3$$

for relative motion along the 1-axis.

It is in this sense that spatial and time variables are **not distinct entities** but are simply **different components** of the same vector and transform into each other under Lorentz transformations.

This corresponds to a non-Euclidean geometry.

Another 4-vector is $d\overleftarrow{s} = (cdt, dx, dy, dz)$ since it is the difference of two 4-vectors. Hence, its norm

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

is a Lorentz invariant

A related quantity of great importance is $d\tau^2 = \frac{ds^2}{c^2}$ (dividing an invariant by an invariant means that we still have an invariant). In particular,

$$d\tau^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

Consider a displacement $d\bar{s}$ between two events on the worldline of a moving particle. In the rest frame of the particle, $dx = dy = dz = 0$ and hence

$$d\tau = dt$$

in the particle rest frame (the events are separated only by time). $d\tau$ is the time interval between the two events measured in the rest frame and is thus the **proper time**. It is a **Lorentz invariant.**

Time Dilation (the easy way)

Consider and observer at rest in x',y',z',t' system. In this system the proper time between two events is $d\tau = dt'$. In the x,y,z,t system moving with velocity $\vec{v}$ relative to the first frame, the time interval between the same two events is given by

$$dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

But $d\tau$ is an invariant or its value is the same in all frames. We therefore have

$$dt'^2 = dt^2 - \frac{1}{c^2}(dx^2 + dy^2 + dz^2)$$

$$\left(\frac{dt'}{dt}\right)^2 = 1 - \frac{1}{c^2}\left(\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right)$$

$$= 1 - \frac{v^2}{c^2} = \frac{1}{\gamma^2}$$

Therefore, $dt = \gamma dt'$ which is the time dilation formula.

We did not need to introduce hypothetical experiments or discussions of simultaneity to obtain this result. That is an example of the power of using 4-vectors.

Other 4-Vectors

Using $d\bar{s} = (cdt, dx, dy, dz)$ and dividing by the Lorentz invariant $d\tau$ yields another 4-vector

$$\frac{d\bar{s}}{d\tau} = (c\frac{dt}{d\tau}, \frac{dx}{d\tau}, \frac{dy}{d\tau}, \frac{dz}{d\tau}) = \bar{u} = 4 - vector\ velocity$$

Its norm is an invariant so it can be calculated by in any frame. We pick the rest frame where

$$\bar{u} = (c,0,0,0) \rightarrow u^2 = c^2 = \text{invariant}$$

For a moving particle where the x,y,z,t system moves with velocity $-\vec{v}$ relative to the rest frame of the particle we have $dt = \gamma d\tau$ and thus

$$\bar{u} = (c\frac{dt}{d\tau},\frac{dx}{d\tau},\frac{dy}{d\tau},\frac{dz}{d\tau}) = \gamma\left(c,\vec{v}\right) \quad , \quad \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$

Since the rest mass m_0 is a Lorentz invariant, $m_0\bar{u}$ is a 4-vector with dimensions of momentum. We define the 4-momentum as

$$\bar{\rho} = m_0\bar{u} = m_0\gamma\left(c,\vec{v}\right) = \left(\frac{E}{c},\vec{p}\right)$$

We already saw that

$$\rho^2 = \left(\frac{E}{c}\right)^2 - \vec{p}^2 = m_0^2c^2 = \text{invariant}$$

Since the variables $E \, and \, \vec{p}$ are components of a 4-vector the must obey the Lorentz transformations

$$\frac{E'}{c} = \gamma\left(\frac{E}{c} - \beta p_x\right)$$

$$p'_x = \gamma\left(p_x - \beta\frac{E}{c}\right)$$

$$p'_y = p_y$$

$$p'_z = p_z$$

You will use these relations in Physics 8 to prove that a magnetic field is observed in frames moving relative to fixed charged particles whereas only electric fields are observed in the rest frame of the charged particles. Magnetic fields are a consequence of special relativity!!

Finally we confirm our identification of the energy. We define the 4-vector Minkowski force as

$$\bar{\phi} = \frac{d\bar{\rho}}{d\tau} = \left(\frac{d\gamma m_0 c}{d\tau},\frac{d\vec{p}}{d\tau}\right)$$

If dt is the time interval in the observer’s frame corresponding to the interval of proper time $d\tau$, then $dt = \gamma d\tau$ and we get

$$\bar{\phi} = \gamma\left(\frac{d\gamma m_0 c}{d\tau},\vec{F}\right) \quad , \quad \vec{F} = \frac{d\vec{p}}{dt}$$

With this construction, the 4-momentum is conserved(constant) when the 4-force is zero. This corresponds to energy and momentum conservation. If the 4-force is zero in one frame then it is zero in all frames and hence if energy and momentum are conserved in one frame they are conserved in all frames. In Newtonian physics

$$\vec{F}\cdot\vec{v} = \frac{dE}{dt}$$

where E = total energy. Let us look at the corresponding quantity in 4-dimensions

$$\overleftarrow{\phi}\cdot\bar{u} = \gamma\left(\frac{d\gamma m_0 c}{dt},\vec{F}\right)\cdot\gamma\left(c,\vec{v}\right) = \gamma^2\left[\frac{d\gamma m_0 c^2}{dt} - \vec{F}\cdot\vec{v}\right]$$

Now the scalar product is an invariant and thus we can evaluate it in the rest frame of the particle. In this frame $\vec{F}\cdot\vec{v} = 0$ since $\vec{v} = 0$. We also have

$$\frac{d\gamma m_0 c^2}{dt} = \gamma m_0 v\left(\frac{dv}{dt}\right) = 0$$

since $v = 0$. Therefore

$$\overleftarrow{\phi}\cdot\bar{u} = 0 = \gamma^2\left[\frac{d\gamma m_0 c^2}{dt} - \vec{F}\cdot\vec{v}\right]$$

$$\vec{F}\cdot\vec{v} = \frac{d\gamma m_0 c^2}{dt} \rightarrow E = \gamma m_0 c^2$$

as we indicated earlier. In this sense the momentum and energy variables are **not distinct entities** but are simply **different components** of the same vector and transform into each other under Lorentz transformations.

A Further Generalization

We can generalize the scalar product to any number of dimensions and any type of geometry. We have

$$\mathbf{A}\cdot\mathbf{B} = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}A_i B_j$$

where g_{ij} is the so-called metric object. We can represent it by a matrix. In ordinary 3-dimensional space we have

$$\left[g\right] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \rightarrow g_{ij} = \delta_{ij}$$

and hence

$$\mathbf{A}\cdot\mathbf{B} = \sum_{i=1}^{n} A_i B_i = \vec{A}\cdot\vec{B}$$

In Minkowski 4-space we have

$$\left[g\right] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

and hence

$$\mathbf{A}\cdot\mathbf{B} = \overleftarrow{A}\cdot\overleftarrow{B}$$

In the theory of gravitation (general relativity) we have

$$[g] = \begin{pmatrix} g_{00}(x,y,z,t) & g_{01}(x,y,z,t) & g_{02}(x,y,z,t) & g_{03}(x,y,z,t) \\ g_{10}(x,y,z,t) & g_{11}(x,y,z,t) & g_{12}(x,y,z,t) & g_{13}(x,y,z,t) \\ g_{20}(x,y,z,t) & g_{21}(x,y,z,t) & g_{22}(x,y,z,t) & g_{23}(x,y,z,t) \\ g_{30}(x,y,z,t) & g_{31}(x,y,z,t) & g_{32}(x,y,z,t) & g_{33}(x,y,z,t) \end{pmatrix}$$

and clearly the world is considerably more complicated (left for a senior seminar).

Now back to special relativity.

Let us return to the relations $\left(\frac{E}{c}\right)^2 - \vec{p}^2 = m_0^2c^2$ and $\frac{pc}{E} = \frac{v}{c} = \beta$. Notice that if $v = c$, then $E = pc$ and $m_0 = 0$. Therefore, particles with zero rest mass exist. They always move with the speed of light. Even though the have no mass they do have energy and momentum! An example of such a particle is the photon, the particle of light.

Radiation Pressure

When light(photons) which carries momentum and energy reflects off of a surface it transfers momentum and energy to the surface. Since a change in momentum corresponds to a force and a force on a surface area corresponds to pressure, light exerts radiation pressure on any reflecting surface. If we have normal incidence on the surface, then the total change in the photon momentum is

$$\Delta p = 2p = 2\frac{E}{c}$$

If there are n photons per unit area per second, then the total momentum change per second per unit area is $pressure = 2n\frac{E}{c} = 2\frac{I}{c}$ where $I = nE$ is the intensity of the light (the power per unit area). The average intensity of sunlight falling on the earth surface is $\approx 1000\ W/m^2 = 1000\ J/m^2\cdot\sec$. The radiation pressure on a mirror is then $pressure = 2\frac{I}{c} = 7\times10^{-4}\,N/m^2$. This is very small(atmospheric pressure is $10^6\,N/m^2$). On a cosmic scale, however, this radiation pressure is large, that is, it is able to help keep stars from collapsing under their own gravitational forces.

How big must the sail of a light-sail starship be to work effectively?

Suppose that the sail material has the property $mass\ per\ m^2 = \rho\ kg$ and that the ship has a mass of M kg. A crude calculation goes like this

$$\text{pressure at distance r from sun} = 7\times 10^{-4}\left(\frac{r_{earth}}{r}\right)^2 N/m^2$$

$$\text{force on sail} = \text{pressure}\times\text{area} = 7\times 10^{-4}\left(\frac{r_{earth}}{r}\right)^2 A$$

$$acceleration = a = \frac{force}{total\ mass} = \frac{7\times 10^{-4}\left(\frac{r_{earth}}{r}\right)^2 A}{M+\rho A}$$

Suppose we have a sail with an maximum area = USA = $10^{13}\ m^2$. For $r = n\times r_{earth}$, $\rho = 10^{-8}$, $M = 10^5$, α = fraction of area used we get

$$a = \frac{7\times 10^{-4}\left(\frac{1}{n}\right)^2 10^{13}\alpha}{10^5 + 10^5\alpha} = 7\times 10^4\left(\frac{\alpha}{1+\alpha}\right)\left(\frac{1}{n}\right)^2 \frac{m}{\sec^2} \quad , \quad 0\le\alpha\le 1 \quad , \quad n\ge 1$$

The acceleration will drop below 0.0001g for $\alpha = 1$ when $n = 20000$ or we are at a distance of 20000 earth radii or about 2×10^{12} miles from the sun. This is about 0.03 light-year. Depending on what we did earlier we could have a sizable speed by this point.

The Power of 4-Vectors

We now illustrate the power of 4-vectors. We consider a photon with energy $E = h\nu$ and momentum $p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}$ traveling in the $x-y$ plane at an angle ϕ with the x-axis. The 3-momentum is

$$\vec{p} = \frac{h\nu}{c}(\cos\phi, \sin\phi, 0)$$

and the energy-momentum 4-vector is

$$\overleftarrow{p} = \frac{h\nu}{c}(1, \cos\phi, \sin\phi, 0)$$

In another system moving relative to the first with velocity v along the common $x-x'$ axis the energy-momentum 4-vector is

$$\overleftarrow{p} = \frac{h\nu'}{c}(1, \cos\phi', \sin\phi', 0)$$

Now these two frames are related by the Lorentz transformation such that

$$\frac{E'}{c} = \frac{h\nu'}{c} = \gamma\left(\frac{E}{c} - \beta p_x\right) = \gamma\left(\frac{h\nu}{c} - \beta\frac{h\nu}{c}\cos\phi\right)$$

which says that

$$\nu' = \gamma\nu(1-\beta\cos\phi)$$

or

$$\nu = \frac{\nu'}{\gamma}\frac{1}{1-\beta\cos\phi} = \nu'\frac{\sqrt{1-\beta^2}}{1-\beta\cos\phi}$$

This is identical to our earlier result for $\phi = 0$ and $\nu_D = \nu$ and $\nu' = \nu_0$. Our derivation give an additional result however. In a direction perpendicular to the line of relative motion, $\phi = \frac{\pi}{2}$, we get $\nu_D = \nu_0\sqrt{1-\beta^2}$, which is called the **transverse** Doppler effect and is due to time dilation.

High Energy Particle Physics

A special reference frame is the center of mass or zero momentum system frame. It is very useful when discussing high energy particle reactions.

We consider a collision between two particles with rest masses m_1 and m_2. We assume that particle 1 is moving with velocity $\vec{u}$ in the laboratory system and that particle 2 is at rest in that system. We have the energy-momentum 4-vectors

$$\tilde{p}_1 = \left(\frac{E_1}{c}, p_1, 0, 0\right) \quad \text{and} \quad \tilde{p}_2 = \left(\frac{E_2}{c}, 0, 0, 0\right)$$

and the total energy-momentum

$$\tilde{P} = \tilde{p}_1 + \tilde{p}_2 = \left(\frac{E_1 + E_2}{c}, p_1, 0, 0\right)$$

In a new frame moving along the x-axis with speed V we have

$$P'_1 = \Gamma\left(p_1 - \frac{V}{c}\frac{E_1 + E_2}{c}\right) \quad , \quad P'_2 = 0 \quad , \quad P'_3 = 0$$

where $\Gamma = \left(1 - \frac{V^2}{c^2}\right)^{-1/2}$.

In the center of mass system, $V = V_{CM}$ and $\vec{P}' = 0$. This says that

$$V_{CM} = \frac{p_1 c^2}{E_1 + E_2}$$

The energy available for physical processes such as the production of new particles or inelastic events is the total energy in the center of mass system, E'. In the center of mass system the total energy-momentum 4-vector is

$$\left(\frac{E'}{c}, 0, 0, 0\right)$$

We can find E' by using the fact that the norm of the energy-momentum 4-vector is invariant

$$\left(\frac{E'}{c}\right)^2 = \left(\frac{E_1+E_2}{c}\right)^2 - p_1^2$$

or

$$E'^2 = E_1^2 + E_2^2 + 2E_1E_2 - p_1^2c^2 = E_1^2 + E_2^2 + 2E_1E_2 - \left(E_1^2 - m_1^2c^4\right)$$
$$= m_1^2c^4 + 2E_1E_2 + E_2^2$$

We have

$$E_1 = \gamma m_1 c^2 \quad \text{and} \quad E_2 = m_1 c^2 \qquad , \qquad \gamma = \left(1 - \frac{u^2}{c^2}\right)^{-1/2}$$

Therefore

$$E = \left(\gamma m_1 + m_2\right)c^2 = \text{total energy in laboratory system}$$

and

$$E' = \left(m_1^2 + m_2^2 + 2\gamma m_1 m_2\right)^{1/2} c^2$$

The fraction of energy available for physical processes is

$$\frac{E'}{E} = \frac{\left(m_1^2 + m_2^2 + 2\gamma m_1 m_2\right)^{1/2}}{\gamma m_1 + m_2}$$

For the special case $m_1 = m_2 = m$ we have

$$\frac{E'}{E} = \sqrt{\frac{2}{1+\gamma}}$$

At low velocity or low energy of the incident particle (the one that is moving), we have

$$\gamma \approx 1 \rightarrow \frac{E'}{E} = 1 \rightarrow \text{all energy available}$$

In this case, most of the energy is rest energy and kinetic energy is unimportant. In the high speed or high energy limit we have

$$\frac{E'}{E} = \sqrt{\frac{2}{1+\frac{E_1}{mc^2}}} \rightarrow \sqrt{\frac{2mc^2}{E_1}}$$

Thus, the useful fraction of energy decreases as $E_1^{-1/2}$. For example, in a 300 GeV accelerator (1 GeV = $10^9\ eV = 10^9 \times 1.6x10^{-12}\ J = 1.6x10^{-3}\ J$) an accelerated proton ($mc^2 \approx 1\,GeV$) colliding with a hydrogen target(protons) has

$$\frac{E'}{E} = \rightarrow \sqrt{\frac{2}{300}} = 0.082$$

or only 25 GeV is available for reactions!!! We will show how to fix this up shortly.

Let us look at **production reactions** in another way. Suppose that we have two particles that interact with each other(one is at rest -- the target) and produce N final particles. The high energy available from the incident particle is converted into mass of newly created particles. We ask the question: What is the minimum energy needed by the incident particle in order to produce the final state of N particles?

In the initial state we have

$$\left(\frac{E_{inc}}{c}, p_{inc}, 0, 0\right) + \left(m_{\text{target}}c, 0, 0, 0\right) = \left(\frac{E_{inc}}{c} + m_{\text{target}}c, p_1, 0, 0\right)$$

$$E_{inc}^2 = p_{inc}^2 c^2 + m_{inc}^2 c^4$$

In the final state we have

$$\left(\frac{\sum_{i=1}^{N} E_i}{c}, \sum_{i=1}^{N} \vec{p}_i\right) \quad \text{where} \quad E_i^2 = p_i^2 c^2 + m_i^2 c^4 \quad , \quad \text{i} = 1, 2, 3, 4, \ldots\ldots\ldots, \text{N}$$

Now, the norm of the energy-momentum 4-vector is invariant in time and across different frames. Therefore

norm in laboratory before = norm in center of mass after

This gives

$$\left(\frac{E_{inc}}{c} + m_{\text{target}}c\right)^2 - p_1^2 = \left(\frac{\sum_{i=1}^{N} E_{i,CM}}{c}\right)^2 - \left(\sum_{i=1}^{N} \vec{p}_{i,CM}\right)^2$$

By definition, however, $\sum_{i=1}^{N} \vec{p}_{i,CM} = 0$. After some algebra we have

$$E_{inc} = \frac{\left(\sum_{i=1}^{N} E_{i,CM}\right)^2 - \left(m_{\text{inc}}c^2\right)^2 - \left(m_{\text{target}}c^2\right)^2}{2m_{\text{target}}c^2}$$

This is a minimum when $\sum_{i=1}^{N} E_{i,CM}$ is a minimum or when

$$\sum_{i=1}^{N} E_{i,CM} = \sum_{i=1}^{N} m_i c^2$$

or all the particles are at rest in the center of mass system after the collision (what are they doing in the laboratory system).

Therefore the minimum energy needed by the incident particle (this is called the **threshold energy**) is

$$E_{inc,threshold} = \frac{\left(\sum_{i=1}^{N} m_{\text{i}}c^2\right)^2 - \left(m_{\text{inc}}c^2\right)^2 - \left(m_{\text{target}}c^2\right)^2}{2m_{\text{target}}c^2}$$

For example, consider the reaction $p+p \to p+p+\pi+\pi+\pi$ where a proton is incident on another proton producing two protons and three pi mesons. The threshold energy is

$$E_{p,threshold} = \frac{\left(2m_p + 3m_\pi\right)^2 - 2m_p^2}{2m_{\rm p}} c^2 = \left(m_{\rm p} + 6m_\pi + \frac{9}{2}\frac{m_\pi^2}{m_{\rm p}}\right) c^2$$

Clearly, this is a very non-intuitive answer!!!

Now let us consider the difference between a particle accelerator where one particle is accelerated and collides with a second particle at rest (as above=laboratory system) and two particle accelerators where each particle is accelerated in the same way (colliding beams=center of mass system). We have

Single Accelerator

$$\left(\frac{E_{\text{total lab}}}{c}, \vec{p}_{\text{total lab}}\right) = \left(\frac{E_1 + m_2c^2}{c}, \vec{p}_1\right) \quad , \quad E_1^2 = p_1^2c^2 + m_1^2c^4$$

Colliding Beams

$$\left(\frac{E_{\text{total cm}}}{c}, \vec{p}_{\text{total cm}}\right) = \left(\frac{2E}{c}, 0\right) \quad , \quad E = \text{energy of each particle}$$

In the first case the accelerator must produce energy E_1 and in the second case each accelerator must produce energy E.

The two accelerators are equivalent (same energy available for physical processes) if

$$\left(\frac{E_1 + m_2c^2}{c}, \vec{p}_1\right)^2 = \left(\frac{2E}{c}, 0\right)^2$$

Algebra gives the result

$$E = \frac{1}{2}\sqrt{m_1^2c^4 + m_2^2c^4 + 2m_2c^2E_1}$$

If we consider the case of very high energy accelerators where $E_1 >> m_ic^2$ we have

$$E = \frac{1}{2}\sqrt{2m_2c^2E_1}$$

Suppose we want to build a single 10 TeV accelerator ($1\,TeV = 10^3\,GeV$) so that $E_1 = 10^4\,GeV$. This is very difficult to design and requires the development of significant new equipment ($$$$$$$).

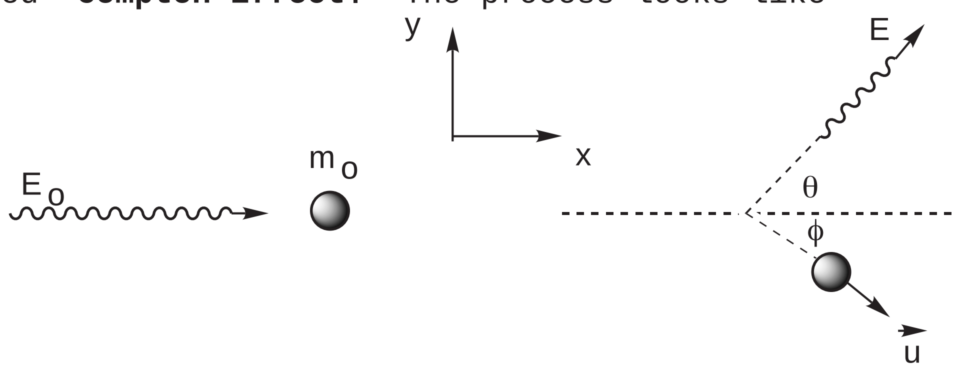
If instead we build two smaller accelerators and use them in the colliding beams configuration, then we get the same available energy with

$$E = \frac{1}{2}\sqrt{2E_1} = \sqrt{5000} = 71\,GeV$$

which we already know how to build. In fact, if we use an old single accelerator of this size that already exists, we then only have to build one small new accelerator ($$).

High Energy Collisions

Earlier we discussed low energy collisions between particles using conservation of energy and momentum. Let us look at the same processes at high energy. We consider a collision in which the incident particle has zero rest mass (photon) and the target particle is at rest. If the target particle is an electron, then this if the so-called **Compton Effect**. The process looks like



The photon momentum is $\frac{E_0}{c}$. After the collision the photon is scattered through an angle θ with energy E and the electron recoils at an angle ϕ with velocity $\vec{u}$. The final electron energy is

$$E_e = \gamma(u) m_0 c^2 = \frac{m_0 c^2}{\sqrt{1 - \frac{u^2}{c^2}}}$$

Conservation of energy gives $E_0 + m_0 c^2 = E + E_e$. Conservation of momentum gives (x and y directions)

$$\frac{E_0}{c} = \frac{E}{c}\cos\theta + p\cos\phi$$

$$0 = \frac{E}{c}\sin\theta - p\sin\phi$$

where

$$\vec{p} = \gamma m_0 \vec{u} \qquad \text{or} \qquad E_e^2 = p^2 c^2 + m_0^2 c^4$$

We want to eliminate reference to the electron and find the new photon energy(that is what is detected in the experiment).

$$\frac{E_0}{c} = \frac{E}{c}\cos\theta + p\cos\phi \rightarrow p\cos\phi = \frac{E_0}{c} - \frac{E}{c}\cos\theta \rightarrow p^2\cos^2\phi = \left(\frac{E_0}{c} - \frac{E}{c}\cos\theta\right)^2$$

$$0 = \frac{E}{c}\sin\theta - p\sin\phi \rightarrow p\sin\phi = \frac{E}{c}\sin\theta \rightarrow p^2\sin^2\phi = \frac{E^2}{c^2}\sin^2\theta$$

Adding these equations we get

$$p^2c^2 = E_e^2 - m_0^2c^4 = E_0^2 - 2E_0E\cos\theta + E^2$$

Using the energy conservation equation we have (after algebra)

$$E = \frac{E_0}{1 + \left(\frac{E_0}{m_0c^2}\right)(1 - \cos\theta)}$$

The first thing to note is that $E > 0$. This means that a free electron cannot absorb a photon completely; there will always be a scattered photon of some energy. If we convert to wavelengths using

$$E = h\nu = h\frac{c}{\lambda}$$

we get

$$\lambda - \lambda_0 = \frac{h}{m_0c}(1 - \cos\theta)$$

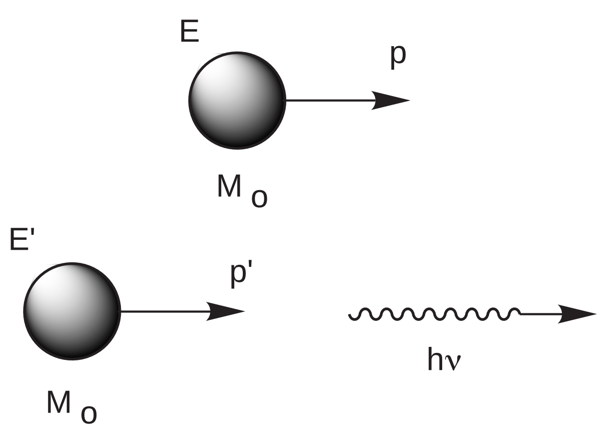
The shift in wavelength at a given angle is independent of the incident photon energy. You will do this experiment in Physics 14.

Doppler effect as a Collision with Photons

We consider an atom with a rest mass of M_0 . If held stationary and the atom emits a photon of energy $h\nu_0$, then its rest mass must change (it is losing energy)

$$M'_0c^2 = M_0c^2 - h\nu_0$$

Suppose that it is moving as shown in the top part of the figure below and then emits a photon as shown in the bottom part of the figure.



Before photon emission we have

$$E = \gamma M_0c^2 = \frac{M_0c^2}{\sqrt{1 - \frac{u^2}{c^2}}} \quad , \quad p = \gamma M_0u = \frac{M_0u}{\sqrt{1 - \frac{u^2}{c^2}}}$$

After emission, the atom has energy E' and momentum p' while the emitted photon has

$$E_\nu = h\nu = p_\nu c$$

Conservation of energy and momentum says that

$$E = E' + h\nu$$

$$p = p' + \frac{h\nu}{c}$$

Therefore

$$E'^2 - p'^2 c^2 = (E - h\nu)^2 - (pc - h\nu)^2 = M'^2_0 c^4 = \left(M_0 c^2 - h\nu_0\right)^2$$

Some algebra gives

$$\nu = \nu_0 \frac{2M_0 c^2 - h\nu_0}{2(E - pc)}$$

But

$$E - pc = \frac{M_0 c^2}{\sqrt{1 - \frac{u^2}{c^2}}}\left(1 - \frac{u}{c}\right) = M_0 c^2 \sqrt{\frac{c-u}{c+u}}$$

Thus,

$$\nu = \nu_0 \frac{2M_0 c^2 - h\nu_0}{2M_0 c^2 \sqrt{\frac{c-u}{c+u}}} = \nu_0 \left(1 - \frac{h\nu_0}{2M_0 c^2}\right)\sqrt{\frac{c+u}{c-u}}$$

The term $\frac{h\nu_0}{2M_0 c^2}$ represents a decrease in the photon energy due to the recoil energy of the atom. For massive atoms, this is negligible and thus

$$\nu = \nu_0 \sqrt{\frac{c+u}{c-u}}$$

which is the standard Doppler formula.

We note for the future that we have shown that two completely different pictures of light, wave and particle, lead exactly to the same prediction for the shift in the frequency of radiation from a moving source.

The Mass of a Photon

Pulsars are collapsed stars that emit regular bursts of energy at repetition frequencies from 30 to 0.1 Hz. The are collapsed stars with intense magnetic fields that are rotating rapidly. The pulsar in the Crab nebula has the a frequency of 30 Hz and pulses in the optical and x-ray regions, as well as at radio frequencies. The pulses are extremely sharp and their arrival times can be measured to an accuracy of microseconds or better. Experimentally, the radiation from the pulsar at all different wavelengths seems to arrive simultaneously (or all within the experimental resolving time).

Let us use this data to set a limit on the rest mass of the photon.

It takes light about 5000 years to arrive at the earth from the Crab nebula. Suppose that signals at two different frequencies travel with a small difference in speed, Δv, and thus arrive at slightly different times, T and $T+\Delta T$. Since $T=\frac{L}{v}$, where L = distance to the Crab nebula, we have

$$v=\frac{L}{T}\rightarrow \Delta v=-\frac{L}{T^2}\Delta T\rightarrow \frac{\Delta v}{v}=-\frac{\Delta T}{T}$$

No such velocity difference has been observed, but by estimating the sensitivity of the experiment we can set an upper limit on the quantity Δv.

ΔT can be measured to an accuracy of about $2\times 10^{-3}\,\mathrm{sec}$ and using $T=5\times 10^{3}\ years=1.5\times 10^{11}\,\mathrm{sec}$ we have

$$\left|\frac{\Delta v}{c}\right|=\left|\frac{\Delta T}{T}\right|<\frac{2\times 10^{-3}}{1.5\times 10^{11}}\approx 10^{-14}$$

where we have assumed that $v\approx c$.

Now we translate this limit on Δv into a limit on the possible rest mass of the photon.

If the photon had a nonzero rest mass, the velocity of light would be different from c. If we let m_p represent the rest mass of the photon, then we would have

$$E=\gamma m_p c^2$$

If we assume that the photon energy frequency relation $E=h\nu$ is still valid, then we have

$$(h\nu)^2=\left(m_p c^2\right)^2\frac{1}{1-\frac{v^2}{c^2}}\rightarrow \frac{v^2}{c^2}=1-\frac{\nu_0^2}{\nu^2}$$

where $h\nu_0=m_p c^2$. ν_0 plays the role of a characteristic frequency for the photon. $h\nu_0$ is the rest energy of the photon. If $\nu_0=0$, then we have $v=c$. Otherwise the velocity of light depends on frequency.

Now consider two frequencies ν_1 and ν_2. We then have

$$\frac{v_1^2}{c^2}-\frac{v_2^2}{c^2}=\nu_0^2\left(\frac{1}{\nu_2^2}-\frac{1}{\nu_1^2}\right)$$

$$\frac{1}{c^2}\left(v_1^2-v_2^2\right)=\frac{1}{c^2}\left(v_1-v_2\right)\left(v_1+v_2\right)=\nu_0^2\left(\frac{1}{\nu_2^2}-\frac{1}{\nu_1^2}\right)$$

$$\frac{1}{c^2}\Delta v(2c)=2\frac{\Delta v}{c}=\nu_0^2\left(\frac{1}{\nu_2^2}-\frac{1}{\nu_1^2}\right)$$

For observations made in the optical regions we can use

$$\nu_1 = 8\times 10^{14}\,Hz(blue) \quad \text{and} \quad \nu_2 = 5\times 10^{14}\,Hz(red)$$

Then we have

$$2\times 10^{-14} > 2\frac{\Delta\nu}{c} = \nu_0^2\left(\frac{1}{\nu_2^2} - \frac{1}{\nu_1^2}\right) = \frac{1}{10^{28}}\left(\frac{1}{5^2} - \frac{1}{8^2}\right)$$

$$\nu_0 < 10^7\,Hz$$

This gives and upper limit to the photon rest mass of

$$m_p = \frac{h\nu_0}{c^2} < 10^{-40}\,kg$$