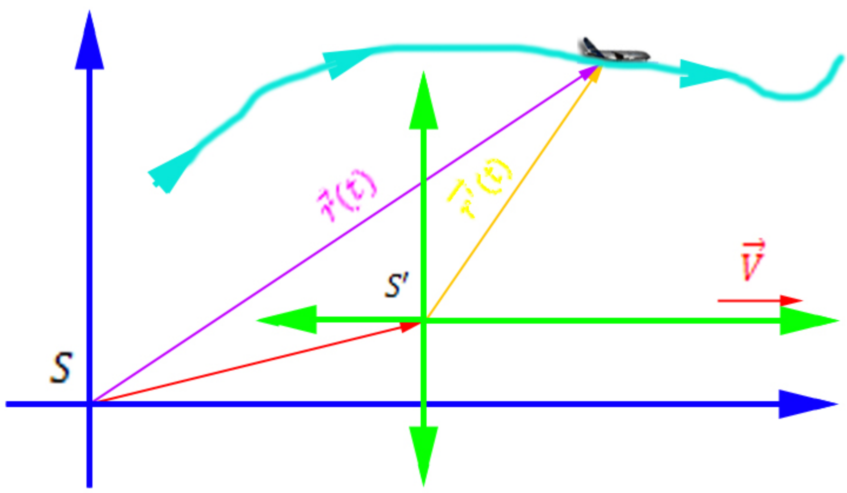
Relatividad especial

1. Concepto de la relatividad

Para empezar a entender qué significa todo el asunto de la relatividad, conviene tener una idea clara de qué es lo que estamos haciendo, usando relaciones y ecuaciones más o menos conocidas (o al menos más o menos intuitivas). Para ello comenzaremos con el concepto de relatividad más básico (al menos, el más básico de los modelos útiles de relatividad), que es la **relatividad de Galileo**, de la que se supone que el lector debería tener ya ciertos conocimientos. La intuición juega aquí un importante papel: supongamos que nos situamos en el arcén de una carretera, nos sentamos, y vemos los coches pasar: para nosotros esos coches llevarán una cierta velocidad, digamos, de 100 km/h. Ahora supongamos que nos ponemos a perseguir a esos coches con una moto que viaja a 70 km/h. Es evidente que los coches irán más deprisa que nosotros (se irán alejando), pero lo harán con una velocidad, respecto a nosotros en la moto, de 30 km/h, menor que la que tenían aparentemente cuando estábamos quietos. Supongamos un caso más: que vamos dentro de uno de esos coches. Para nosotros, en esa situación, el coche está parado, ya que no se mueve respecto a donde estamos. Sabríamos que nos movemos porque podemos ver a través de las ventanas, y porque vemos el cuentakilómetros, pero ¿qué pasaría si nos encerrasen en un coche con las ventanas tapadas y forma de habitación, a oscuras, que viaja a 100 km/h? En principio, no sabríamos que nos movemos. Es más, sabemos que la Tierra se mueve. Entonces ¿por qué cuando estamos tranquilamente tumbados en un sofá nos parece que estamos quietos? Pues precisamente, porque no hay nada alrededor que nos indique que nos movemos. Es decir, en pocas palabras, la velocidad de un objeto depende de por donde se mire. Así pues, la relatividad es algo más evidente de lo que puede parecer en un primer momento. La cuestión es, ¿qué más cosas dependen de desde dónde se miren además de la velocidad?

La idea, como se ve, es sencilla, al menos de intuir. Supondremos un sistema de referencia S en reposo, y un sistema de referencia S' que se mueve con una velocidad $\vec{V}$ con respecto a este último (el sistema S somos nosotros (el laboratorio) sentados en el arcén de la carretera, y el S' es el sistema de referencia del ocupante de uno de los coches, de modo que la velocidad relativa entre los sistemas, $\vec{V}$, será justamente la velocidad del coche $\left(\vec{V} = \vec{V}_{coche} - \vec{V}_{nosotros} = \vec{V}_{coche} - \vec{0} = \vec{V}_{coche}\right)$. Así pues, el sistema S verá un objeto moviéndose (por ejemplo, un avión que vuela por el cielo), con unas coordenadas $\vec{r}(t)$, mientras que el sistema S' verá unas coordenadas diferentes para el mismo objeto (al estar moviéndose), que llamaremos $\overrightarrow{r'}(t)$. Esto se ve bien en el esquema siguiente:



En este dibujo vemos lo siguiente: tenemos el sistema de referencia en reposo, S, cuyos ejes de coordenadas están representados por líneas azules. En el sistema de referencia en movimiento S', sus ejes están representados con líneas en verde, y se estaría moviendo con velocidad $\vec{V}$ con respecto de S (es decir, dentro de 5 o 10 segundos, posiblemente los ejes verdes habrían desaparecido del dibujo, ya que se mueven a la derecha, mientras que los ejes azules seguirían donde están eternamente). La línea roja que une ambos centros sería la coordenada relativa, que se estiraría a medida que pase el tiempo, puesto que el sistema S' se desplaza hacia la derecha a medida que pasa el tiempo. Por otro lado, suponemos que ambos están viendo el avión de la figura, cuya trayectoria (la del avión) viene dada por la línea azul celeste del dibujo. Las coordenadas de este vienen dadas por el vector $\vec{r}(t)$ para el sistema S, es decir, la línea morada, y las coordenadas del avión desde S' vienen dadas por el vector $\vec{r}'(t)$. Como es lógico, estas coordenadas son distintas en uno y otro sistema (y por tanto, las velocidades, que serán sus derivadas en el tiempo, intuimos que también serán distintas). Lo que nos queda saber es cómo, sabiendo cuales son unas coordenadas, podemos obtener las otras . Es decir, si el observador S' le pega un telefonazo a S le cuenta como ve el avión, (o lo que es lo mismo, le informa de cuales son las coordenadas $\vec{r}'(t)$ que él está viendo), que S sea capaz, sabiendo la velocidad relativa $\vec{V}$ entre ambos, determinar las coordenadas a las que para él se encuentra dicho avión, $\vec{r}(t)$.

La versión más simple de esto será cuando supongamos que el sistema S' lleva una velocidad relativa constante al sistema S, que llamaremos $\vec{V} = \overrightarrow{cte}$ (la segunda versión más simple es cuando esta velocidad esté situada paralela a uno de los ejes coordenados, como en el dibujo, en que la velocidad relativa es paralela al eje de las *"x"*, de modo que nos quitamos de encima los vectores, pero eso lo veremos más adelante). La pregunta es ¿qué podemos intuir o qué creemos que se mantendrá constante para ambos sistemas? Es evidente que la posición del objeto es distinta vista desde un sistema o desde el otro, y parece lógico que la velocidad del objeto que ven ambos sistemas tampoco es la misma. Pero ¿qué ocurre con cantidades como la masa, el color o la fuerza? Es lógico pensar que ambos verán al avión del mismo color. Evidentemente la masa que mida un observador (si fuese capaz de medirla), sería la misma que la que mide el otro (no por mirar un avión desde dos perspectivas pesará más o menos), pero ¿hay alguna cantidad cinemática que podamos decir que será invariante?

Aquí fue donde Galileo usó su intuición, y supuso que eran las <u>leyes de la dinámica</u> las que debían ser las mismas para dos observadores inerciales. Estas leyes de la dinámica, por entonces, eran básicamente la ley de Newton de las fuerzas, $\vec{F} = m\vec{a}$, y dado que parece obvio dar por hecho que ambos deberían medir la misma masa, lo que se debe cumplir en definitiva es que, se mida desde un sistema o desde el otro:

$$\boxed{\vec{F} = \vec{F}' \xrightarrow{con\, m=m'} \vec{a} = \vec{a}'}$$

Es decir, miden la misma aceleración. Lo siguiente que debemos hacer, es tratar de imponer unas transformaciones (para pasar de unas coordenadas a otras), que cumpliesen que, tal y como acabamos de escribir, la fuerza vista desde un sistema de referencia, o vista desde el otro, fuesen la misma en ambas, o lo que es lo mismo, que la aceleración de un objeto vista desde un sistema de referencia o desde otro sean iguales. Estas transformaciones (que al menos le sonarán al lector), son las siguientes:

$$\boxed{\begin{cases} \vec{r}'(t) = \vec{r}(t) - \vec{V}t \\ t' = t \end{cases}}$$

Por supuesto, Galileo no tuvo en cuenta la segunda expresión, $t' = t$, dado que para él, como parece lógico, el tiempo que transcurre observando un cierto objeto desde un cierto sistema de referencia es el mismo que transcurre viéndolo desde otro distinto (por eso la segunda expresión le parecería sumamente obvia). Es decir, el tiempo que pasaría en mi reloj si estuviese viendo los coches sentado en el arcén sería el mismo que marcaría el reloj del conductor del coche. No es extraño, claro, que nadie cuestionase esta especie de postulado, ni que siquiera se escribiese en las transformaciones.

Si expresamos estas ecuaciones en componentes vendrían dadas por:

$$\begin{cases} x'(t) = x(t) - V_x t \\ y'(t) = y(t) - V_y t \\ z'(t) = z(t) - V_z t \\ t' = t \end{cases}$$

Donde V_x, V_y y V_z son las componentes del vector velocidad relativa: $\vec{V} = \left(V_x, V_y, V_z\right)$, y el vector de posición sería $\vec{r}(t) = (x(t), y(t), z(t))$ o, si nos gusta más una notación más, digamos, numérica, podríamos escribir como: $\vec{r}(t) = (x_1(t), x_2(t), x_3(t))$, lo mismo nos da.

Ya hemos simplificado diciendo que la velocidad relativa sea constante. Hagamos una segunda simplificación para no rompernos la cabeza con interminables ecuaciones: supongamos que la velocidad relativa entre ambos sistemas sea paralela al eje *x*, es decir, la única coordenada que se transforma es la coordenada *x*, mientras que las transformaciones para *y* y *z* no varían, quedando las transformaciones siguientes:

$$\begin{cases} x'(t) = x(t) - Vt \\ y'(t) = y(t) \\ z'(t) = z(t) \\ t' = t \end{cases}$$

Donde hemos llamado $V_x \equiv V$ por simplicidad, dado que ahora sólo tenemos una componente del vector velocidad relativa: $\vec{V} = (V, 0,0)$. Como decíamos antes, Galileo buscó estas transformaciones tratando que la fuerza vista desde un sistema de referencia o desde otro fuese la misma. Así pues, comprobémoslo para no dejarle como un mentiroso: para ello comencemos obteniendo la velocidad de la partícula (no confundir la velocidad, v_x o simplemente v, con la velocidad relativa V: escribiremos siempre en mayúscula la velocidad relativa entre los sistemas y en minúsculas las componentes de la velocidad de la partícula vista desde cada sistema de referencia (la velocidad del avión en cada punto)). Además, omitiremos de aquí en adelante la dependencia temporal, dándola por hecha (de esta forma, escribiremos v_x en vez de $v_x(t)$), para que la notación no sea demasiado engorrosa:

$$\begin{cases} v'_x = \dfrac{dx'}{dt} = \dfrac{d}{dt}[x - Vt] = \dfrac{dx}{dt} - V = v_x - V \\ v'_y = \dfrac{dy'}{dt} = \dfrac{dy}{dt} = v_y \\ v'_z = \dfrac{dz'}{dt} = \dfrac{dz}{dt} = v_z \end{cases}$$

Puesto que $\frac{dx}{dt} \equiv v_x$. Esta es la famosa ley de adición de velocidades de Galileo. Vemos que las velocidades que miden los dos sistemas de referencia en los ejes *y* y *z* son lógicamente las mismas, mientras que la velocidad en el eje *x* depende de la velocidad relativa entre ellos. La expresión general que rige la ley de transformación de Galileo para las velocidades no es complicada (puede deducirse de lo anterior), y es:

$$\boxed{\vec{v}'(t) = \vec{v}(t) - \vec{V}}$$

Sin embargo, operando con rigor, deberíamos haber hecho lo siguiente (puesto que el sistema S' vería un tiempo *t'* mientras que el sistema S ve un tiempo *t*, suponiendo que ambos fuesen distintos):

$$v'_x(t) = \frac{dx'(t)}{dt'} = \frac{\partial x'(t)}{\partial t}\frac{\partial t}{\partial t'} = \frac{\partial}{\partial t}[x(t) - Vt] = \frac{dx(t)}{dt} - V = v_x - V$$

E igual para *y* y *z*. Para ello, no hemos hecho más que usar la regla de la cadena, y sabiendo que con estas transformaciones, *t=t'*, vemos que $\frac{\partial t}{\partial t'} = 1$, con lo que realmente queda igual que si no nos hubiésemos preocupado de este detalle, como hizo Galileo (aunque conviene verlo bien, dado que en un par de páginas habrá que tenerlo en cuenta, ya que los tiempos no serán iguales).

A continuación, calculemos la aceleración que ven cada uno de los dos sistemas. Para ello, aplicaremos de nuevo la regla de la cadena (más para hacerlo bien que por necesidad, dado que los tiempos son iguales y no sería necesario), de modo que:

$$a'_x(t) = \frac{d^2x'(t)}{dt'^2} = \frac{d}{dt'}v'_x(t) = \frac{d}{dt'}[v_x(t) - V] = \frac{\partial t}{\partial t'}\frac{\partial}{\partial t}[v_x - V] = \frac{dv_x}{dt} - \frac{\partial}{\partial t}V = a_x$$

Ya que $\frac{dv_x}{dt} \equiv a_x$, y al derivar la velocidad relativa, da cero como resultado, dado que habíamos supuesto por hipótesis que esta velocidad sería constante en el tiempo. El resultado sería el mismo para las componentes de la aceleración en los ejes *y* y *z*. Así pues, tenemos que:

$$a'_x(t) = a_x(t)$$

La expresión general, de haberlo hecho para una velocidad cualquiera, no necesariamente paralela al eje *x* (aunque sí constante), sería:

$$\boxed{\vec{a}'(t) = \vec{a}(t)}$$

O lo que es lo mismo, los dos observadores ven la misma aceleración del objeto. Si nos vamos ahora a la fuerza (particularizado cuando la velocidad es constante y paralela al eje *x*, por supuesto):

$$
\begin{array}{ccc}
F_x = ma_x & F_y = ma_y = 0 & F_z = ma_z = 0 \\
F'_x = ma'_x & F'_y = ma'_y = 0 & F'_z = ma'_z = 0
\end{array}
$$

Pero dado que las aceleraciones son iguales, llegamos finalmente a:

$$F_x = F'_x$$

Es decir, obtenemos como queríamos que las fuerzas que ven los dos observadores son las mismas. Por supuesto, de verlo en versión más general, y con lo ya deducido, vemos que:

$$\left.\begin{array}{r}\vec{F} = m\vec{a} \\ \vec{F}' = m\vec{a}'\end{array}\right\} \boxed{\boxed{\vec{F} = \vec{F}'}}$$

Ejercicio 1.1. *Calcular la transformación inversa de coordenadas (x en función de x')*

Ejercicio 1.2. *Partiendo de las transformaciones de Galileo para la posición:*

$$\begin{cases}\vec{r}'(t) = \vec{r}(t) - \vec{V}t \\ t' = t\end{cases}$$

Deducir la expresión general para las transformaciones de la velocidad, la aceleración, y la fuerza.

Solución: Ver las páginas anteriores

Ejercicio 1.3. *Sabiendo que la ecuación de ondas es:*

$$\frac{\partial^2 \Phi(x,t)}{\partial t^2} - \frac{1}{c^2}\frac{\partial^2 \Phi(x,t)}{\partial x^2} = 0$$

Demostrar que no es invariante bajo transformaciones de Galileo, es decir:

$$\frac{\partial^2 \Phi(x',t')}{\partial t'^2} - \frac{1}{c^2}\frac{\partial^2 \Phi(x',t')}{\partial x'^2} \neq \frac{\partial^2 \Phi(x,t)}{\partial t^2} - \frac{1}{c^2}\frac{\partial^2 \Phi(x,t)}{\partial x^2}$$

NOTA: Para ello tener en cuenta que al transformar el potencial deberíamos hacer lo siguiente:

$$\frac{\partial \Phi(x',t')}{\partial t'} = \frac{\partial \Phi(x',t)}{\partial t}\frac{\partial t}{\partial t'}$$

2. La relatividad según Einstein: Relatividad Especial. Transformaciones de Lorentz

Desde joven, Einstein se había hecho una interesante pregunta: *"¿cómo vería la luz un observador que viajase más o menos a la misma velocidad que la misma?"*. Es decir, qué se vería si se "persiguiese la luz" a suficiente velocidad como para que la situación fuese similar a perseguir a un coche con una moto. La idea, por supuesto, no tenía respuesta inmediata. Si desde el arcén de la carretera vemos un coche a 200 km/h, tan sólo vemos como un "destello", pero nada más. Pero si perseguimos al coche a una velocidad parecida a la que lleva el mismo, podemos ver a sus ocupantes, la matrícula, el color del coche… Y quién sabe si al viajar a velocidades cercanas a la luz, al perseguirla, podríamos ver de qué está hecha la luz, su composición, etc. Esta idea permaneció latente varios años, hasta que los descubrimientos de la época llevaron a plantearse ciertas conclusiones.

El experimento de Michelson Morley demostró que la velocidad de la luz toma el mismo valor se mire desde el sistema de referencia que se mire, que no depende del observador. Es decir, con esta afirmación, se deduce inmediatamente que la velocidad de la luz, *c*, es invariante e insuperable, dado que por mucho que la persigamos o nos alejemos de ella, no habría adición de velocidades como con Galileo, sino que seguiríamos viéndola igual. Desde luego, esto entraba en contradicción con la ley de adición de velocidades de Galileo, dado que uno esperaría que si viaja en sentido contrario al de la luz, la velocidad a la que la vería sería superior, y si uno "persigue" a la luz, la vería "ralentizada", con velocidad inferior a *c*.

Supongamos que ponemos un sistema de referencia en "la luz", y otro sistema de referencia moviéndose con velocidad relativa $-V$ con respecto al mismo, es decir, alejándose de la luz (en direcciones opuestas). Según las ecuaciones anteriores de Galileo, la velocidad que vería este último observador sería:

$$v'_x = c - (-V) = c + V > c$$

Que evidentemente sería mayor que la velocidad de la luz, lo cual no tendría sentido según las hipótesis anteriores, entrando en conflicto con las conclusiones de Michelson. Así pues, Einstein (o Lorentz) como tantos otros, vieron que algo no cuadraba, y había que elegir qué leyes o hipótesis no eran correctas: o bien la velocidad de la luz sí cambiaba según el observador que la viese, entrando en conflicto con los experimentos de Michelson, o bien las transformaciones de Galileo no eran correctas, pese a lo intuitivas que parecieron. El problema era que suponer esto último implicaba directamente que, si se buscaban unas transformaciones que hiciesen que cualquier sistema de referencia viese la luz con la misma velocidad, ya no verían la misma fuerza los dos observadores. Y esto no solamente implicaría esta simple observación, sino que echaría por tierra a lo que entonces se consideraba como una eminencia en el campo de la dinámica: Newton. Sus leyes, aparentemente perfectas, no habían sido discutidas durante demasiado tiempo, dando siempre excelentes resultados para aplicaciones que iban desde lo más simple a lo más complejo. Y aunque podría pensarse que meterse con Newton en física significa poner en juego la reputación del protestante, se optó por admitir que las leyes de Newton no eran las mismas según cual fuese el sistema de referencia elegido.

Pero entonces las nuevas transformaciones ¿qué deberían cumplir?. Principalmente, que la velocidad de la luz fuese la misma fuese cual fuese el sistema de referencia. Esto suponía que el intervalo espaciotemporal, intervalo entre dos sucesos, que veremos más adelante, fuese invariante. Pero sin ir tan lejos, y de forma más simple, que las ecuaciones que rigen la dinámica lumínica lo fuesen. Estas ecuaciones no son ni más ni menos que las ecuaciones de onda. Después de buscar unas transformaciones que hiciesen que esto sucediese (puede verse la deducción en, por ejemplo, el libro de Landau), se llegó a las siguientes transformaciones (donde hemos supuesto de nuevo que la velocidad relativa entre los dos sistemas es constante, y paralela al eje *x*, por sencillez) llamadas **transformaciones de Lorentz**:

$$
\begin{aligned}
x' &= \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}[x - Vt] \\
y' &= y \\
z' &= z \\
t' &= \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}\left[t - \frac{V}{c^2}x\right]
\end{aligned}
$$

Donde por comodidad se suele llamar $\boxed{\gamma \equiv \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}}$, incluso para más sencillez, también se suele usar el convenio de llamar $\beta \equiv v/c$, quedando $\gamma \equiv \frac{1}{\sqrt{1-\beta^2}}$, aunque aquí prescindiremos de esta última definición. De esta forma que estas ecuaciones quedan como:

$$
\boxed{\begin{cases} x' = \gamma[x - Vt] \\ y' = y \\ z' = z \\ t' = \gamma\left[t - \frac{V}{c^2}x\right] \end{cases}}
$$

Las conclusiones de estas nuevas transformaciones son evidentes: para empezar, el tiempo ya no es el mismo según desde donde se mire. El reloj del observador en reposo en el arcén de la carretera medirá un tiempo, mientras que el tiempo que medirá un observador dentro del coche será distinto al del primero. Afortunadamente, para los casos reales, el factor γ tiende a uno (cuando las velocidades son pequeñas comparadas con la de la luz), devolviendo de nuevo las transformaciones de Galileo. Así pues se deduce que las transformaciones de Galileo, y por tanto toda la dinámica de Newton, son aproximadamente válidas para velocidades pequeñas (comparadas con la luz, lo que incluye a casi cualquier velocidad), y sólo será necesario usar la relatividad especial cuando tengamos velocidades muy próximas a la de la luz. En resumen: en principio no deberíamos preocuparnos mucho por la teoría de la relatividad al ir, por ejemplo, a una cita, ya sea andando, en coche o en avión supersónico, dado que todas estas velocidades son extremadamente pequeñas comparadas con la luz, y los tiempos son, a efectos de cualquier reloj, incluido los más precisos, iguales. Pero ¿qué le ocurre a, por ejemplo, un electrón que viaja infinitamente más rápido que un vehículo? Una cosa es que en la vida cotidiana no sea necesario usar esta teoría, y otra que no valga para nada. De hecho, casi toda la física del último siglo está basada, en una medida u otra, en la teoría de la relatividad de Einstein.

Ejemplo 2.1. *Tomemos un ejemplo sencillo para ver esto con claridad: un vehículo viajando por una carretera desde Madrid a Burgos con velocidad V. Supongamos un observador en Madrid que, por algún extraño motivo, es capaz de ver a ese vehículo desde que sale de Madrid hasta que llega a Burgos. Tomemos como S el sistema de referencia del observador que se encuentra en el propio vehículo, y como S' al observador que se encuentra quieto en Madrid, mirando. Tomemos además, otro sistema S'' como el sistema de referencia asociado al propio vehículo (este último punto es simplemente para, al final, comprender bien lo que estamos haciendo). Podemos imaginar esto como unos ejes pinchados en el asiento del vehículo para el sistema S''(y un reloj) , otro ejes pinchados en la cabeza del conductor, con su reloj de pulsera, que marcarán las coordenadas en el sistema S y finalmente un trípode donde hemos colgado unos ejes, fijos Tierra en Madrid, con un reloj de cadena colgando, que compondrán las coordenadas en S' (todo conjunto de coordenadas debe comprender tanto las espaciales como la coordenada temporal para poder fijar exactamente un suceso dado).*

Supongamos que queremos medir las características del vehículo desde el sistema de referencia del conductor. Para el observador que se encuentra en el propio vehículo, este se hallará en reposo, es decir, la velocidad relativa entre el sistema S (del conductor), y el sistema S'' (el propio vehículo), será $V_{S-S''} = 0$. Así pues, las transformaciones entre un sistema y otro serán:

$$\begin{cases} x'' = \dfrac{1}{\sqrt{1-\dfrac{0^2}{c^2}}}[x - 0t] = x \\ t'' = \dfrac{1}{\sqrt{1-\dfrac{0^2}{c^2}}}\left[t - \dfrac{0}{c^2}x\right] = t \end{cases}$$

Es decir, ambos ven el mismo tiempo, y miden la misma posición. Esto parece obvio, ya que a fin de cuentas son el mismo objeto. Además es un alivio, ya que no es por estar tan cerca el uno del otro por lo que miden distancias y tiempos iguales, sino porque su velocidad relativa es cero. Podrían estar separados miles de kilómetros que, si su velocidad relativa es cero, se cumplen las mismas características que en el caso anterior (medirían obviamente distintas coordenadas, como ya se podía intuir de antemano: por ejemplo, Moscú estará situado de forma distinta visto desde Nueva York o Madrid, pero el tiempo que pase en Madrid o en Nueva York no varía, y los relojes que se fabriquen en un sitio u otro no tienen que llevar ninguna corrección relativista). Esto significa que, dado que la velocidad relativa entre Madrid y Nueva York es cero (no se mueven, en media, uno respecto al otro), el tiempo que midan uno u otro serán siempre el mismo.

A este tiempo que mide el conductor se le llamará el ***tiempo propio****, que se suele denominar con la letra* $\boldsymbol{\tau}$*, que es el tiempo que mide un observador cuyo sistema de referencia es solidario con el objeto de estudio (este sistema de referencia sería S'', que cumple que $V_{S-S''} = 0$). Hasta aquí no hay nada especial. Hecha la comprobación e introducido el tiempo propio, olvidémonos del sistema S'' (puesto que a efectos es el mismo que S, y los tomaremos como un único sistema de referencia S), y estudiemos el movimiento relativo entre el vehículo y el observador quieto en Madrid.*

Recordemos: hemos llamado S al sistema de referencia "enganchado" al vehículo, y S' al sistema de referencia del tipo quieto en Madrid, observando de algún modo al vehículo. Además, llamaremos V a la velocidad relativa entre un sistema y otro (que evidentemente es la velocidad del coche, puesto que es esta la velocidad a la que se mueve un sistema de referencia con respecto al otro, como vimos en el ejemplo de la introducción). Queremos ver como ve el sistema en reposo, es decir, el que está quieto en Madrid, al vehículo. Como el vehículo se está alejando del observador en reposo, la velocidad relativa entre ambos, vista desde el observador de Madrid, será $-V$, y las transformaciones serán:

$$\begin{cases} x' = \dfrac{1}{\sqrt{1-\dfrac{V^2}{c^2}}}[x+Vt] \\ t' = \dfrac{1}{\sqrt{1-\dfrac{V^2}{c^2}}}\left[t+\dfrac{V}{c^2}x\right] \end{cases}$$

Donde x es la posición del vehículo vista desde el punto de vista del propio vehículo. Como es lógico, para cualquier instante de tiempo, x=0, puesto que desde el punto de vista del vehículo, este no se mueve, al estar "observándose a sí mismo". Por otro lado, $t=\tau$ será el tiempo que mide el propio vehículo (tiempo propio), mientras que x' y t' son respectivamente la posición del vehículo y el tiempo que mide el observador que se encuentra quieto. Con esta simplificación, las transformaciones quedan:

$$\begin{cases} x' = \dfrac{V\tau}{\sqrt{1-\dfrac{V^2}{c^2}}} \\ t' = \dfrac{\tau}{\sqrt{1-\dfrac{V^2}{c^2}}} \end{cases}$$

Donde se ve bien que a medida que pasa el tiempo, la posición que mide el observador en reposo del coche, x', va aumentando (como parece lógico, pues el vehículo se aleja). Pues bien, así tomaría el observador quieto en Madrid los datos referentes al vehículo (la correspondencia entre la posición del vehículo y el tiempo entre un observador y otro). Evidentemente, para τ =0, el observador de Madrid mide que el objeto se encuentra en x'=0 (el eje de coordenadas), ya que t'=0 (en el instante inicial, los dos miden el mismo tiempo en su reloj, y miden la misma posición para el vehículo, lo cual parece obvio, dado que los dos sistemas de coordenadas están uno sobre el otro). En cambio, veamos qué ocurre, por ejemplo, a mitad de camino (suponiendo que la distancia Madrid-Burgos son 200 km, digamos que a mitad de camino hay unos 100km=10^5m):

a) Supongamos primero un caso realista: el vehículo viaja a 120km/h (33'33m/s). Para calcular el tiempo que ha pasado para el observador en el vehículo (atendiendo por ejemplo al cuentakilómetros) cuando ha recorrido la mitad del camino, y dado que lo hace a velocidad constante, no tenemos más que recurrir a:

$$\tau_{mitad} = \frac{x_{mitad}}{V} = \frac{10^5(m)}{33'3333(\frac{m}{s})} = 3000(s)$$

Según Galileo, el tiempo que ha pasado para el observador en reposo debería ser el mismo. Para Einstein, este será:

$$t' = \frac{\tau}{\sqrt{1-\frac{V^2}{c^2}}} = \frac{3000}{\sqrt{1-\frac{(33'3333)^2}{c^2}}} \approx 3000\ (s)$$

Dado que la diferencia entre un tiempo y otro no es siquiera apreciable por el mejor de los relojes (el resultado exacto, que ni siquiera la calculadora o el "Derive" es capaz de diferenciar, es 3000'00000000004s), con lo que ni siquiera nos deberíamos molestar en encontrar el primer decimal en esta operación. Así pues, vemos que en un caso relativamente normal, las transformaciones de Galileo son válidas (nótese que cuando en γ *ponemos una velocidad baja comparada con c, este factor tiende a uno y recuperamos las ecuaciones de Galileo).*

b) *En cambio, supongamos ahora un caso diferente: que el vehículo viaja a una velocidad mucho más alta: de* 3.10^7 *m/s (la velocidad de la luz sigue siendo diez veces mayor, pero veamos qué ocurre). En esta situación el tiempo que tardaría en recorrer el espacio indicado sería* $3'3333.10^{-3}$*, atendiendo a la misma expresión de antes:*

$$\tau_{mitad} = \frac{x_{mitad}}{V} = \frac{10^5(m)}{3.10^7\ (\frac{m}{s})} = 3'3333.10^{-3}(s)$$

Lógicamente, a esa velocidad no tarda apenas una milésima de segundo en recorrer la distancia propuesta. Así pues, teniendo ya el tiempo propio que mide el observador en el vehículo, calculemos el tiempo que mide el observador en Madrid:

$$t' = \frac{\tau}{\sqrt{1-\frac{V^2}{c^2}}} = \frac{3'3333.10^{-3}}{\sqrt{1-\frac{(3.10^7)^2}{c^2}}} = 3'3468.10^{-3}\ (s)$$

Es decir, a esta velocidad, el reloj del observador que va en el coche mide $3'3333.10^{-3}$ *segundos, mientras que el que está quieto en Madrid mide* $3'3468.10^{-3}$ *segundos. Es decir, si realmente lográsemos un vehículo que alcanzase tal velocidad, el conductor del vehículo ahorraría* $13'5.10^{-6}$ *segundos respecto al tiempo de aquellos que están en reposo frente a él, que en comparación es un 0'4% del tiempo que lograría acortar puesto que para un observador en reposo, como el de Madrid, ha pasado más tiempo entre la salida de Madrid y la llegada al punto medio de la trayectoria. Parece que no es mucho, pero es razonablemente alto comparando ambos tiempos, dado que la lógica nos dice que el tiempo no debería haber cambiado en absoluto, y menos aún en un intervalo de tiempo tan pequeño.*

Pero este no es el único ahorro, ya que además de esto, la posición que mediría el observador en reposo para el vehículo, en vez de ser de 100 km como sería de esperar, sería:

$$x' = \frac{V\tau}{\sqrt{1-\frac{V^2}{c^2}}} = 100502'77(m)$$

Es decir, que en vez de los 100000m=100km, el observador en reposo vería al otro objeto unos 500 metros más lejos de lo que realmente está, y medir 500 metros más en un experimento que dura $3'3333.10^{-3}s$ *es una cantidad bastante elevada.*

c) *Como último ejemplo, supongamos que el vehículo viaja a* $V = 0'99999c$ *, que es prácticamente la velocidad de la luz, pero sin llegar a serlo. En este caso tenemos:*

$$\tau_{mitad} = \frac{x_{mitad}}{V} = \frac{10^5(m)}{0'99999c(\frac{m}{s})} = 333'5.10^{-6}(s)$$

$$\gamma = \frac{1}{\sqrt{1-\frac{V^2}{c^2}}} = \frac{1}{\sqrt{1-0'99999^2}} = 223'6$$

$$x' = \frac{V\tau}{\sqrt{1-\frac{V^2}{c^2}}} = 223'6.10^5(m)$$

$$t' = \frac{\tau}{\sqrt{1-\frac{V^2}{c^2}}} = \frac{3'3333.10^{-3}}{\sqrt{1-\frac{(3.10^7)^2}{c^2}}} = 74'57.10^{-3}\ (s)$$

Es decir, ha recorrido 223 km (visto desde Madrid) cuando debería haber recorrido 100 (visto desde el vehículo), y el tiempo que pasa para el conductor es 224 veces menor que para el que se encontraba en reposo. El conductor en reposo, en este momento, ha ganado mucho tiempo, dado que para el observador en reposo han pasado $74'57.10^{-3}s$ *mientras que para él tan sólo* $333'5.10^{-6}s$. *Es decir, el tiempo que mide el observador del vehículo se ha acortado en un 22360%. En nuestro problema esto puede parecer poco relevante, dados los tiempos tan cortos que manejamos (recorrer 100km a casi la velocidad de la luz es completamente trivial si alcanzásemos dicha velocidad), pero supongamos que el vehículo sigue, digamos, hasta una distancia astronómica (se pega unas cuantas vueltas a la Tierra):*

Supongamos que el vehículo se pega 240 millones de vueltas a la Tierra, lo que equivale aproximadamente a unos $9'5.10^{15}m$ *recorridos. A la velocidad que lleva, los recorrería (según su propio reloj) en:*

$$\tau_{240millones\ de\ vueltas} = \frac{x_{240\ millones\ de\ vueltas}}{V} = \frac{9'5.10^{15}m}{0'99999c(\frac{m}{s})} = 31'67.10^6(s)$$

O lo que viene a ser aproximadamente un año. Es decir, el tipo del vehículo se pasaría dentro del mismo un año dando millones de vueltas a la Tierra. Al cabo de ese tiempo, el tiempo que habría pasado para el observador en reposo (supuesto que hubiese aguantado un año midiendo), sería:

$$t' = \frac{\tau}{\sqrt{1-\frac{V^2}{c^2}}} = \frac{31'67.10^6}{\sqrt{1-\frac{(3.10^7)^2}{c^2}}} = 7'08.10^9\ (s) \approx 223años$$

En otras palabras: en lo que para un observador ha pasado un año, al otro le ha dado tiempo a vivir casi tres vidas. Es decir, no podría haber seguido midiendo, y de haberlo logrado por algún milagro médico, se habría encontrado que al parar el vehículo, su amigo tendría apenas un año más que cuando se fue, 223 años atrás. Esto da una idea de la potencia de estas correcciones cuando manejamos velocidades suficientemente altas.

Es lógico que para medidas usuales la relatividad no tiene mucho sentido (ningún conductor debería preocuparse porque su reloj se atrase, ni ningún observador en reposo porque el conductor llegue tarde). De hecho, se puede demostrar por un simple desarrollo en serie que a pequeñas velocidades comparadas con *c* las transformaciones de Lorentz se reducen a las de Galileo (el lector con prisa que se salte esta demostración, o que la haga simplemente poniendo *V=0* como aproximación):

$$Transformaciones: \begin{cases} x' = \dfrac{1}{\sqrt{1-\dfrac{V^2}{c^2}}}[x - V\tau] \\ t' = \dfrac{1}{\sqrt{1-\dfrac{V^2}{c^2}}}\left[\tau - \dfrac{V}{c^2}x\right] \end{cases}$$

Llamemos $\beta \equiv \frac{V^2}{c^2}$ y desarrollemos en torno a β*=0*:

$$x' = \frac{1}{\sqrt{1-\beta^2}}[x - Vt] \approx \left.\frac{x - V\tau}{\sqrt{1-\beta^2}}\right|_{\beta_0=0} + \left.\frac{\beta(\beta-\beta_0)(x - V\tau)}{\left(1-\beta_0{}^2\right)^{7/2}}\right|_{\beta_0=0} + \cdots = x - V\tau$$

Ejercicio 2.1*. Demostrar que las transformaciones para el tiempo se reducen también a las de Galileo (es decir,* $t' = \tau$*), para velocidades bajas comparadas con las de la luz. (ejercicio teórico: sin solución)*

Para completar, hagamos un resumen de todo lo visto. En una columna escribiremos la transformación completa (más difícil de deducir y de manejar) para una velocidad relativa arbitraria, y en la columna derecha las transformaciones que manejaremos habitualmente, que serán aquellas en la que la velocidad relativa es paralela a uno de los ejes (tomamos por simplicidad el eje OX):

	Ecuación general (con V=cte)	***Velocidad paralela a OX (V=cte)***
Posición y Tiempo	$\vec{x}' = \vec{x} + (\gamma - 1)(\hat{V}\cdot\vec{x})\cdot\hat{V} - \gamma\vec{V}t$ $t' = \gamma\left(t - \dfrac{\vec{V}\cdot\vec{x}}{c^2}\right)$	$x' = \gamma[x - Vt]$ $y' = y$ $z' = z$ $t' = \gamma\left[t - \dfrac{V}{c^2}x\right]$
Velocidad	$\vec{v}' = \dfrac{\vec{v} + (\gamma - 1)(\hat{V}\cdot\vec{v})\cdot\hat{V} - \gamma\vec{V}}{\gamma\left(1 - \dfrac{\vec{V}\cdot\vec{v}}{c^2}\right)}$	$v_x{}' = \dfrac{v_x - V}{1 - \dfrac{Vv_x}{c^2}}$ $v_y{}' = \dfrac{v_y}{\gamma\left(1 - \dfrac{Vv_x}{c^2}\right)}$ $v_z{}' = \dfrac{v_z}{\gamma\left(1 - \dfrac{Vv_x}{c^2}\right)}$

Donde $\vec{x}$ significa el vector *x* (sus tres componentes), y $\hat{V}$ significa un vector unitario en dirección del vector $\vec{V}$, y "·" denota el producto escalar:

$$\vec{x} = (x, y, z) \quad \vec{V} = (V_x, V_y, V_z) \quad \hat{V} = \frac{1}{|\vec{V}|}(V_x, V_y, V_z) \quad \vec{V}\cdot\vec{x} = xV_x + yV_y + zV_z = |\vec{V}|.|\vec{x}|.cos\theta_{\widehat{\vec{V}\vec{x}}}$$

Ejercicio 2.1. *Obtener las transformaciones inversas de coordenadas, es decir, despejar x como función de x' y t', y t como función de x' y t'. (ejercicio teórico, sin solución)*

Ejercicio 2.2. *Obtener las expresiones de la derecha de la tabla a partir de las expresiones generales de la izquierda. (ejercicio teórico, sin solución)*

Ejercicio 2.3. *Demostrar que el tiempo que ve un observador en S' a pequeñas velocidades comparadas con la de la luz, es el mismo que el que observa un observador en reposo, S, es decir, que se cumplen las transformaciones de Galileo. (ejercicio teórico, sin solución, y ya hecho anteriormente)*

Ejercicio 2.4. *Obtener la ley de transformación de velocidades con las transformaciones de Lorentz. (ejercicio teórico cuya solución se muestra en la tabla de arriba)*

Ejercicio 2.5. *Demostrar que la fuerza F' que ve un observador respecto a un sistema de referencia S', no es necesariamente la misma fuerza F que la que ve un observador desde S. Para ello, habrá que calcular la ley de transformaciones de las aceleraciones, al igual que se calculó la de las velocidades (ejercicio teórico, sin solución: para cualquier duda, recomendamos acudir al libro de Landau)*

Ejercicio 2.6. *La ecuación de ondas en el vacío para una onda electromagnética (la luz) viene dada por:*

$$\frac{\partial^2 \Phi(x,t)}{\partial t^2} - \frac{1}{c^2}\frac{\partial^2 \Phi(x,t)}{\partial x^2} = 0$$

Demostrar que esta ecuación es invariante bajo las transformaciones de Lorentz (la idea es la misma que para demostrar lo contrario con Galileo). (ejercicio teórico: sin solución)

Ejercicio 2.7. *Supongamos que un sistema de referencia S' se mueve con respecto a otro en reposo S con velocidad c, alejándose del mismo (en vez de V, tenemos –V), y supongamos que el observador del sistema S ve un rayo de luz (de velocidad c). Según Galileo, el sistema de referencia S' vería la luz moviéndose con velocidad:*

$$v' = v - (-V) = c + c = 2c$$

Lo cual sería imposible. Calcular la velocidad a la que vería el sistema S' moverse a la luz si usásemos las transformaciones de Lorentz.

(Solución: velocidad de la luz en el segundo sistema de referencia: c, igual que en el primero)

3. Intervalos espacio-temporales

Ya hemos visto que con las transformaciones de Lorentz, se cumple que la velocidad de la luz es constante sea cual sea el observador inercial que la mire. En realidad el término "relatividad" no es muy acertado, puesto que más que buscar cosas relativas, de lo que se trata es de encontrar cantidades absolutas (como la velocidad de la luz, o la ecuación de ondas, o la masa del objeto, que no dependen del sistema de referencia que se elija para medirlas o expresarlas). La siguiente cantidad que podemos formar, también invariante, se llama **intervalo espacio-temporal**, y para enunciarla, primero debemos saber qué es un **suceso**. El concepto de suceso no es difícil de entender: significa unir las tres coordenadas espaciales y la coordenada temporal en un mismo objeto. Así, supongamos que tenemos dos laboratorios A y B, que están midiendo la explosión de una estrella. El primero de ellos medirá un cierto tiempo t_1, que (además de las correspondientes transformaciones que se puedan hacer), dependerá fundamentalmente de cuándo haya pulsado el cronómetro dicho observador. Igualmente, el segundo observador pulsará su cronómetro cuando le parezca, de forma que medirá un tiempo que probablemente nada tiene nada que ver con el de A. Por ejemplo, el observador en B pulsa su cronómetro, y pasados 300 segundos, la estrella estalla. Así pues, $t_1 = 300s$ para dicho observador. En cambio, y dado que B lo pulsó más tarde, el tiempo que tarda en medir el suceso este segundo observador puede ser, por ejemplo, $t_2 = 100s$. Por otro lado, dado que los laboratorios A y B no se encuentran en el mismo punto del planeta, es lógico que las coordenadas espaciales que miden cada uno desde su posición a la estrella serán diferentes, por lo que el vector de posición $\vec{x}_1$ y $\vec{x}_2$ probablemente no tendrán nada que ver (salvo que uno mida desde una habitación y el otro desde la habitación contigua, en cuyo caso seguirán sin ser iguales, pero serán muy parecidos). De este modo, definiremos los sucesos que ven cada uno como:

$$\begin{cases} s_1 = (ct_1, \vec{x}_1) \\ s_2 = (ct_2, \vec{x}_2) \end{cases}$$

Donde la c es simplemente para que todo esté medido en las mismas dimensiones. Así pues, con estos sucesos recién definidos, podemos sacarnos de la manga una cantidad que va a ser invariante bajo las transformaciones de Lorentz, que será el cuadrado del intervalo espacio-temporal:

$$\boxed{(\Delta s)^2 = \mathrm{c}^2(\Delta t)^2 - (\Delta \vec{x})^2}$$

Donde $\Delta s = s_2 - s_1$ es el intervalo espacio temporal entre dos sucesos 1 y 2 que acabamos de enunciar (vistos desde dos sistemas de referencia distintos), $\Delta \vec{x} = \vec{x}_2 - \vec{x}_1$ es el intervalo espacial entre los dos sucesos, y $\Delta t = t_2 - t_1$ el intervalo temporal entre esos dos sucesos. Podemos imaginar el suceso s_1 como, por ejemplo, una explosión, y el suceso s_2, como la recepción auditiva del mismo por parte de alguien en algún punto del espacio. Así pues, $\vec{x}_1$ y t_1 serán la posición y el tiempo en que se produce la explosión respecto al sistema de referencia S, y $\vec{x}_2$ y t_2 serán la posición y el tiempo que mide la persona que recibe el "ruido" de la explosión, también medido en el sistema de referencia S. Sin embargo, si nos situamos en otro sistema de referencia distinto, S', en el cual esos dos sucesos serían s'_1 y s'_2 (no tienen por qué estar en el mismo punto que los otros dos, ni por qué medir el mismo tiempo), pero se sigue cumpliendo que:

$$\boxed{\boxed{(\Delta s)^2 = (\Delta s')^2}}$$

La demostración de esta última afirmación es sencilla: no tenemos más que sustituir:

$$c^2(\Delta t)^2-(\Delta\vec{x})^2=c^2(\Delta t')^2-(\Delta\vec{x}')^2$$

Es decir:

$$c^2(\Delta t)^2-\Delta\vec{x}\cdot\Delta\vec{x}=c^2(\Delta t')^2-\Delta\vec{x}'\cdot\Delta\vec{x}'$$

$$c^2(t_2-t_1)^2-(\overrightarrow{x_2}-\overrightarrow{x_1})^2=c^2(t'_2-t'_1)^2-(\overrightarrow{x_2}'-\overrightarrow{x_1}')^2$$

$$\begin{aligned}c^2(t_2-t_1)^2-(x_2-x_1)^2-(y_2-y_1)^2-(z_2-z_1)^2\\=c^2(t'_2-t'_1)^2-(x'_2-x'_1)^2-(y'_2-y'_1)^2-(z'_2-z'_1)^2\end{aligned}$$

Y así sucesivamente. Dejamos el resto del ejercicio para el lector.

Ejercicio 3.1. *Usando las transformaciones de Lorentz, completar esta demostración, y ver que la cantidad $(\Delta s)^2=(\Delta s')^2$ son iguales sean cuales sean los sucesos que estamos midiendo. Para ello, utilizar las expresiones en que la velocidad relativa entre ambos sistemas es paralela al eje x, que son más sencillas de usar que las expresiones generales. (ejercicio teórico, sin solución)*

Al lector agudo, sin embargo, le chirriará esto que acabamos de hacer. El hecho de que *c* multiplique al intervalo temporal no es de extrañar, ya que con esto logramos que tanto $c^2(\Delta t)^2$ como $(\Delta\vec{x})^2$ estén en las mismas unidades, lo cual era esperable (no se podrían sumar si no fuese así). Sin embargo, el problema es que si hacemos esto con cuidado, deberíamos haber escrito lo siguiente:

$$\begin{aligned}(\Delta s)^2=\Delta s\cdot\Delta s=(s_2-s_1)\cdot(s_2-s_1)=(ct_2-ct_1,\vec{x}_2-\vec{x}_1)\cdot(ct_2-ct_1,\vec{x}_2-\vec{x}_1)\\=(c\Delta t,\Delta\vec{x})\cdot(c\Delta t,\Delta\vec{x})=c^2(\Delta t)^2+(\Delta\vec{x})^2\end{aligned}$$

Que no es lo que acabamos de decir. El problema es que no hemos usado el producto escalar habitual (que es lo que está hecho en rojo), sino otro producto escalar distinto. La cuestión es que abandonamos el producto escalar habitual (definido para el espacio euclídeo habitual, con su producto escalar euclídeo), y pasamos a usar el **espacio-tiempo de Minkovskii**, con su producto escalar, que es el definido según la ecuación anterior. La principal diferencia es que, mientras que en el espacio euclídeo el producto escalar siempre es una cantidad positiva (a fin de cuentas es la norma al cuadrado de un vector), en el espacio de Minkovskii esto ya no ocurre, pudiendo ser positivo, negativo o cero. Además, recordemos que en el producto escalar habitual, euclídeo, para que el producto escalar de un vector por sí mismo (es decir, la norma), fuese cero, este vector debía ser cero $\left(|\vec{x}|=0\Leftrightarrow\vec{x}=\vec{0}\right)$, mientras que en nuestro nuevo espacio, esto no tiene por qué ser necesariamente así. Para que el cuadrado del intervalo espaciotemporal sea cero debe cumplirse que:

$$(\Delta s)^2=c^2(\Delta t)^2-(\Delta\vec{x})^2=0\Leftrightarrow c^2(\Delta t)^2=(\Delta\vec{x})^2$$

O lo que es lo mismo, despejando y haciendo la raíz cuadrada:

$$\frac{(\Delta\vec{x})^2}{(\Delta t)^2}=c^2\Longrightarrow v=\frac{|\Delta\vec{x}|}{\Delta t}=c$$

Es decir, el intervalo espaciotemporal será igual a cero tan sólo cuando se trate de señales lumínicas, puesto que el espacio recorrido entre el tiempo que ha pasado, es decir, la velocidad, se ve que es exactamente c. Así pues, los dos sucesos estarían conectados por una señal lumínica.

Esto último es fácil de entender: Un suceso ocurre en A (por ejemplo, la explosión de una estrella), y lo denotamos con las coordenadas $s_1 = (ct_1, \vec{x}_1)$, y su efecto se recoge en otro punto del espacio, en otro tiempo distinto (aunque medidos en el mismo sistema de referencia), cuyas coordenadas vienen dadas por $s_2 = (ct_2, \vec{x}_2)$. Si algo ha sido causado y posteriormente ha sido recibido en otro lugar, es porque se ha transmitido esa información a lo largo del espacio por algún tipo de señal (por ejemplo, en nuestro caso, la señal lumínica que emite la estrella al explotar y que recoge el otro punto del espacio, o la señal sonora que emite un tipo gritando en un punto, y que recibe otro tipo algo más tarde en otro punto distinto). De esta forma, podemos definir la velocidad mínima de la señal necesaria para transmitir la información de un suceso a otro como el espacio recorrido por la señal entre el tiempo que ha necesitado para cubrir el trayecto (ambas medidas desde el sistema de referencia que estamos considerando), es decir:

$$\boxed{v_{señal} = \frac{|\Delta\vec{x}|}{\Delta t}}$$

Como acabamos de ver, el caso límite ocurre cuando el cuadrado del intervalo espaciotemporal es igual a cero, en cuyo caso: $v_{señal} = \frac{|\Delta\vec{x}|}{\Delta t} = c$

Es decir, ambos sucesos sólo podrían conectarse por una señal luminosa. Por ejemplo, si un tipo grita en un punto y se intenta recoger en otro, y el cuadrado del intervalo espaciotemporal para esos dos sucesos es cero, dado que la velocidad de la señal que transmite el grito es la velocidad del sonido, y acabamos de decir que al ser cero el cuadrado del intervalo espaciotemporal sólo pueden transmitirse si la velocidad de la señal es *c*, se dice que esos suceso **no están conectados causalmente**, es decir, que el tipo que intenta recibir el grito, simplemente, no lo oye (al menos no en el tiempo que se está buscando, dado que si se deja pasar el tiempo, posiblemente el cuadrado del intervalo espaciotemporal será mayor, y se podrá recibir la señal con la velocidad del sonido).

Ejercicio 3.2. *Un tipo emite un grito que se propaga por el espacio con la velocidad del sonido (340m/s), y otro tipo en otro lugar lo recibe en otro momento de tiempo (un poco más tarde). Medimos ambos sucesos desde un cierto sistema de referencia inercial S, recogiendo los siguientes datos:*

$$\begin{cases} s_1 = (900.10^8, 400.10^3, 0{,}0) \\ s_2 = (3000.10^8, 1000.10^3, 0{,}0) \end{cases}$$

Determinar si el cuadrado del intervalo espaciotemporal es positivo, negativo o cero, y a su vez determinar si el segundo tipo es capaz de escuchar el grito del primero. (Solución: $(\Delta s)^2$<0: no puede)

Ejercicio 3.3. *Supongamos los mismos datos que antes, pero el segundo observador deja pasar el tiempo, se toman medidas una vez, y dejando pasar el tiempo, se toma otra vez, recogiendo los siguientes datos para el segundo suceso s_2 (s_1 es el mismo que antes):*

a) $s_2 = (900'006.10^8, 1000.10^3, 0{,}0)$
b) $s_2 = (6'194117.10^{11}, 1000.10^3, 0{,}0)$

Determinar qué ocurre en cada uno de estos casos, determinando el signo del cuadrado del intervalo espaciotemporal y la posibilidad de que estén conectados por la onda que viaja a la velocidad del sonido, o se necesita otra señal (si es posible). Explicar el resultado.(Solución: en a), $(\Delta s)^2 = 0$: puede hacerlo mediante un rayo que viaje a la velocidad de la luz, pero no con el sonido; en b), $(\Delta s)^2 > 0$: y calculando, puede recibirlo con una velocidad $v \geq 340m/s$, así que sí podría ser)

Así pues, tal y como hemos definido el intervalo entre sucesos, Δs, acabamos de ver que puede ser tanto positivo como negativo (o cero), dependiendo de los valores que tomen Δt y $\Delta\vec{x}$ respectivamente. De este modo tenemos tres opciones para los posibles valores del cuadrado del mismo:

a) $(\Delta s)^2 = \mathrm{c}^2(\Delta t)^2 - (\Delta\vec{x})^2 > 0$

Los sucesos no pueden ser observados por ningún observador los dos a la vez (no pueden ser simultáneos), dado que de serlo, $\Delta t = t_2 - t_1 = 0$ y $-(\Delta\vec{x})^2 > 0$, lo cual es imposible. En cambio, sí pueden ocurrir en el mismo punto del espacio para algún observador, dado que en este caso $\Delta\vec{x} = 0$ y nos queda $c^2(\Delta t)^2 > 0$, que podría ocurrir de elegir correctamente el sistema de referencia. Llamaremos a este tipo de intervalos **intervalos de tipo tiempo**. Si despejamos la velocidad de la luz en esta ecuación nos queda:

$$\mathrm{c}^2(\Delta t)^2 - (\Delta\vec{x})^2 > 0$$
$$\mathrm{c}^2(\Delta t)^2 > (\Delta\vec{x})^2$$
$$\mathrm{c} > \frac{|\Delta\vec{x}|}{\Delta t} = v_{señal}$$

Así pues, diremos que los intervalos tipo tiempo son **causalmente posibles**, dado que existe una señal de velocidad menor que la de la luz capaz de unirlos (podemos emitir una señal con velocidad menor que *c* desde el suceso 1 que llegue al suceso 2 recorriendo el espacio $|\Delta\vec{x}|$ en un tiempo Δt)

b) $(\Delta s)^2 = \mathrm{c}^2(\Delta t)^2 - (\Delta\vec{x})^2 < 0$

Los sucesos no pueden verse en el mismo punto del espacio para ningún observador, dado que de ser así, $\Delta\vec{x} = \overrightarrow{x_2} - \overrightarrow{x_1} = 0$ y $\mathrm{c}^2(\Delta t)^2 < 0$, lo cual es imposible. En cambio, sí pueden verse simultáneamente por algún observador, puesto que si $\Delta t = 0$, nos queda $-(\Delta\vec{x})^2 < 0$, para lo cual no hay ningún problema. Llamaremos a este tipo de **intervalos de tipo espacio**. Haciendo las mismas cuentas que antes vemos que para que ocurra uno de estos sucesos, la velocidad de la señal que los une debería ser mayor que la de la luz, lo cual los hace **causalmente imposibles**, o lo que es lo mismo, diremos que están **desconectados causalmente**, es decir, que no se pueden dar como realidad física, puesto que para que el suceso 2 se entere de lo que ocurre en el suceso 1 la señal que viaja de uno a otro debería ir más rápido que la luz. Esto significa que la simultaneidad (temporal) no puede darse en la relatividad especial, es decir, *nunca podremos ver dos sucesos ocurriendo al mismo tiempo nos situemos en el sistema de referencia que nos situemos*.

c) $(\Delta s)^2 = \mathrm{c}^2(\Delta t)^2 - (\Delta\vec{x})^2 = 0$

Intervalo tipo luz: sólo una señal lumínica es capaz de unir estos dos intervalos si su $(\Delta s)^2 = 0$. Es decir, ningún objeto material podría interactuar entre s_1 y s_2 salvo que sea la propia luz (s_2 no vería la explosión producida en s_1 si esta no fuese transmitida por una señal a la velocidad de la luz). Despejando igual que antes, obtendríamos que $v_{señal} = c$, no admitiendo otra posibilidad. Estos sucesos, pues, están conectados causalmente, pero sólo por rayos de luz (el sonido por ejemplo es incapaz de unirlos: una explosión ocurrida en el suceso 1 no puede ser comunicada, por ejemplo, por teléfono, al suceso 2 si el intervalo es de tipo luz.

Esto nos da una conclusión importante: cualquier intervalo entre sucesos físico (causalmente posible) debe ser de tipo tiempo, lo que significa que si un suceso es físicamente posible, no podremos encontrar nunca ningún sistema de referencia en que dos sucesos sean simultáneos (ocurran al mismo tiempo). Es decir, **la simultaneidad no existe en la relatividad especial.**

Por último, suele ser útil tomar dos sucesos muy próximos entre sí, para los cuales pasaremos de $\Delta\vec{x}$ a $d\vec{x}$ y de Δt a dt (ocurren muy próximos entre sí), quedando el intervalo espaciotemporal como:

$$(ds)^2 = c^2(dt)^2 - (d\vec{x})^2$$

Aunque tanto en la definición del intervalo temporal finito como infinitesimal, es usual quitar tanto paréntesis, siempre teniendo en cuenta qué es lo que se está haciendo. Además, es conveniente no escribir *x* como vector, ya que primeramente se suele trabajar tan solo en una dimensión (se supone que el objeto de estudio mantiene siempre constantes las coordenadas *y* y *z*, como si estuviese situado en un plano), y posteriormente se introducirán los tensores, que acabarán finalmente con los vectores. De este modo la notación habitual para los intervalos es la siguiente:

$$\boxed{\begin{array}{l}\Delta s^2 = c^2\Delta t^2 - \Delta x^2\\ ds^2 = c^2dt^2 - dx^2\end{array}}$$

Por último, atendiendo a las transformaciones de Lorentz y a los intervalos entre dos sucesos, veamos la expresión para el tiempo propio de una cierta partícula, que ya introdujimos. Supongamos un objeto moviéndose por el espacio, que queremos medir. Supongamos que pinchamos un eje de coordenadas (un sistema de referencia) en el propio objeto. Pues bien, toda aquella información que mida este sistema de referencia del objeto lo llamaremos propio. Usemos los intervalos espaciotemporales en forma diferencial (dt en vez de Δt , etc.). Así, el tiempo transcurrido entre dos sucesos según este sistema de referencia, y el espacio transcurrido entre dos sucesos, en este caso, (que llamaremos *dx* y $dt = d\tau$) del objeto con respecto a otro sistema de referencia externo, que se mueve a velocidad *V*, serán *dx=0* (el objeto no se mueve respecto a sí mismo) y $dt = d\tau$. Usando el intervalo espaciotemporal:

$$ds^2 = c^2d\tau^2 - dx^2 = c^2d\tau^2$$

de donde: $\boxed{d\tau = \sqrt{\frac{ds^2}{c^2}}}$

Visto ahora desde otro observador externo, cambiando *ds* por su valor(es invariante):

$$d\tau = \sqrt{\frac{c^2dt'^2 - dx'^2}{c^2}} = \sqrt{\frac{c^2dt^2 - dx^2}{c^2}.\frac{dt^2}{dt^2}} = dt\sqrt{\frac{\frac{c^2dt^2}{dt^2} - \frac{dx^2}{dt^2}}{c^2}} = dt\sqrt{\frac{c^2 - V^2}{c^2}} = dt\sqrt{1 - \frac{V^2}{c^2}} = \frac{dt}{\gamma}$$

Podemos de esta forma relacionar el tiempo propio de un objeto (el que mide él mismo) con el tiempo que mediría un sistema de referencia que se mueve con velocidad *V* respecto a él como:

$$\boxed{d\tau = \frac{dt}{\gamma}}$$

Que también recibe el nombre de **contracción temporal**. Un cálculo análogo arrojaría la **contracción espacial**, cuya base es la misma, pero con el espacio propio, (válida para longitudes, áreas y volúmenes)

Ejemplo 3.1. Un cierto objeto explota y es recogido por un cierto sistema de referencia S, en el cual se miden las coordenadas $\overrightarrow{x_1} = (10^2, 0,0)(m)$ y $t_1 = 10^{-6}(s)$, y al cabo de un tiempo cesa la explosión, recogiendo los datos $\overrightarrow{x_2} = (10^3, 0,0)(m)$ y $t_2 = 2.10^{-6}(s)$. ¿Se trata de un intervalo tipo tiempo o de tipo espacio? ¿es posible encontrar algún sistema de referencia S' para el cual los dos sucesos ocurran al mismo tiempo (explota y cesa a la vez)? ¿Y en el cual los dos sucesos ocurren en el mismo lugar (aunque en distintos tiempos)?

El intervalo espaciotemporal de estos sucesos viene dado por:

$$\begin{aligned}\Delta s^2 = c^2\Delta t^2 - \Delta x^2 &= c^2(t_2 - t_1)^2 - (\overrightarrow{x_2} - \overrightarrow{x_1})^2 \\ &= (3.10^8)^2(2.10^{-6} - 10^{-6})^2 - (10^3 - 10^2)^2 - (0-0)^2 - (0-0)^2 \\ &= 9.10^{16}10^{-12} - 81.10^4 = 9.10^4 - 81.10^4 = -72.10^4\end{aligned}$$

Como el cuadrado del intervalo es negativo, será de tipo espacio. Esto significa que la señal que une los sucesos debería viajar a velocidad mayor que la de la luz para que el suceso 2 se enterase de lo que le ocurre al suceso 1 (para que la explosión llegue al punto 2 en el tiempo dado). En efecto, si la señal (la explosión) debe recorrer el espacio Δx en un tiempo Δt, la velocidad que debe llevar es:

$$v_{señal} = v_{explosión} = \frac{|\Delta\vec{x}|}{\Delta t} = \frac{10^3 - 10^2}{2.10^{-6} - 10^{-6}} = \frac{900}{10^{-6}} = 9.10^8 > c$$

Como la velocidad de la señal debe ser mayor que la de la luz, y esto no es posible, diremos que ambos sucesos están desconectados causalmente. Aún así, busquemos para qué sistema de referencia los dos ocurrirían en el mismo punto del espacio o en el mismo instante de tiempo (aunque esto implique que la señal que media entre ambos lleve una velocidad mayor que la de la luz). Veamos:

a) Si en otro sistema de referencia los sucesos ocurren en el mismo punto del espacio, es decir $(x'_2 - x'_1 = 0)$, y dado que el cuadrado del intervalo espaciotemporal es el mismo sea cual sea el observador inercial que lo mire(es invariante bajo transformaciones de Lorentz), debería cumplirse que:

$$\Delta s'^2 = \Delta s^2 = -72.10^4 = c^2\Delta t'^2 - \Delta x'^2 = c^2\Delta t'^2$$
$$c^2\Delta t'^2 = -72.10^4$$

Lo cual es imposible, luego es imposible encontrar un sistema de referencia S' en que los dos sucesos se vean en el mismo punto del espacio.

b) Si en otro sistema de referencia los sucesos ocurren simultáneamente $(t'_2 - t'_1 = 0)$, debería cumplirse que:

$$\Delta s'^2 = \Delta s^2 = -72.10^4 = -\Delta x'^2$$

Luego:

$$\Delta x' = \sqrt{72}.10^2$$

De aquí podemos sacar la velocidad a la que debería moverse este sistema de referencia S' con respecto al primero para que esto ocurra. Usemos las transformaciones de Lorentz:

$$\Delta x' = x'_2 - x'_1 = \gamma[x_2 - Vt_2] - \gamma[x_1 - Vt_1] = \sqrt{72}.10^2$$

De donde podemos despejar *V*, desarrollando y despejando la expresión anterior, que será:

$$\frac{1}{\sqrt{1-\frac{V^2}{c^2}}}[10^3 - V2.10^{-6}] - \frac{1}{\sqrt{1-\frac{V^2}{c^2}}}[10^2 - V10^{-6}] = \sqrt{72}.10^2$$

$$\frac{10^3 - V2.10^{-6} - 10^2 + V10^{-6}}{\sqrt{1-\frac{V^2}{c^2}}} = \sqrt{72}.10^2$$

$$\left(\frac{9.10^2 - V10^{-6}}{\sqrt{1-\frac{V^2}{c^2}}}\right)^2 = \left(\sqrt{72}.10^2\right)^2$$

$$\frac{81.10^4 + V^2 10^{-12} - 1'8.10^{-3}V}{1-\frac{V^2}{c^2}} = 72.10^4$$

$$81.10^4 + V^2 10^{-12} - 1'8.10^{-3}V = 72.10^4 - \frac{V^2}{c^2}72.10^4$$

$$81.10^4 c^2 + V^2 10^{-12} c^2 - 1'8.10^{-3}Vc^2 = 72.10^4 c^2 - V^2 72.10^4$$

$$7'29.10^{22} + V^2.9.10^4 - 1'62.10^{14}V = 6'48.10^{22} - V^2 72.10^4$$

$$81.10^4 V^2 - 2'62.10^{14}V + 8'1.10^{21} = 0$$

$$V^2 - 2.10^8 V + 10^{16} = 0$$

$$V = \frac{2.10^8 \pm \sqrt{4.10^{16} - 4.10^{16}}}{2} = 10^8 = \frac{c}{3}$$

Aunque nos podríamos haber ahorrado toda esta cuenta operando de otro modo. Sabemos que además de cumplirse esto, también se cumple que $t'_2 - t'_1 = 0$. Recurriendo de nuevo a las transformaciones de Lorentz:

$$\gamma\left[t_2 - \frac{V}{c^2}x_2\right] = \gamma\left[t_1 - \frac{V}{c^2}x_1\right]$$

$$2.10^{-6} - \frac{V}{c^2}10^3 = 10^{-6} - \frac{V}{c^2}10^2$$

$$18.10^{10} - V.10^3 = 9.10^{10} - V.10^2$$

$$9.10^{10} = 900V$$

$$V = 10^8 = \frac{c}{3}$$

Como era lógico, el resultado es el mismo, sólo que se obtiene de un modo más sencillo. Así pues, vemos que si un cierto sistema de referencia S' se mueve con respecto a S con velocidad *c/3*, para él los dos sucesos ocurrirían simultáneamente, aunque en distintas posiciones. Esto corrobora que este intervalo espaciotemporal es de tipo espacio. Aún así, estos sucesos no podrían ser físicamente reales, por ser tipo espacio.

Ejercicio 3.2. *Un pulso de luz es emitido en un cierto sistema de referencia* S *en las coordenadas* $\overrightarrow{x_1} = (3.10^{10}, 0{,}0)(m)$ *y* $t_1 = 1(s)$, *y al cabo de un tiempo es observado en otro punto distinto, de coordenadas* $\overrightarrow{x_2} = (10^{10}, 0{,}0)(m)$ *y* $t_2 = 5\ (s)$. *¿Se trata de un intervalo tipo tiempo o de tipo espacio? ¿Es posible encontrar algún sistema de referencia* S' *para el cual los dos sucesos ocurran al mismo tiempo (se emite y se recibe a la vez)? ¿Y en el cual los dos sucesos ocurren en el mismo lugar (aunque en distintos tiempos)?¿Es posible conectar causalmente estos dos intervalos?*

Solución: Intervalo tipo tiempo: podemos encontrar un sistema en que sucedan en el mismo tiempo, si la velocidad relativa entre uno y otro es $V = 0'06c$. *En cambio, no pueden suceder en el mismo punto del espacio. Dado que es un intervalo tipo tiempo, se pueden conectar ambos causalmente, por una señal que viaje a* $V = 0'06c$.

Ejercicio 3.3. *Una nave espacial emprende un viaje hacia un lejano planeta desde la Tierra. Al llegar, sus instrumentos de a bordo marcan que han pasado* $300.10^6 s$, *y que ha recorrido* $9'5125$ *años luz (recordar que un año luz es la distancia recorrida la velocidad por la luz en un año). Determinar cuánto vale la velocidad que ha llevado durante el viaje, el factor* γ, *su tiempo propio, su espacio propio, el tiempo que ha pasado para la Tierra, y el espacio que ha recorrido para la Tierra. Supongamos un suceso 1: la nave sale de la Tierra, y un suceso 2: la nave llega al planeta de destino, cada uno con unas coordenadas espaciotemporales medidas por una parte desde la Tierra y por otra parte desde la nave. Calcular el cuadrado del intervalo espacio-temporal con los datos de la Tierra y con los datos de la nave, suponiendo que en el momento en que la nave parte de la Tierra, tanto los relojes de la nave como los de la Tierra se ponen a cero, y comienzan a contar en cuanto el viaje tiene lugar, y determinar si el intervalo entre sucesos es temporal, espacial o lumínico (es decir, determinar si el suceso final y el inicial están conectados causalmente o no). Además, si es espacial o temporal, habrá un sistema de referencia en el que los dos sucesos podrán verse simultáneamente, u otro en que podrán verse en el mismo punto del espacio. Determinar si será una u otra cosa, y calcular la velocidad a la que tiene que moverse dicho sistema de referencia respecto a la nave para que esto ocurra.*

Solución: $V_{Tierra-Nave} = V_{nave} = 0'999954c$; $\gamma = 104'258406$; $\Delta\tau = 300.10^6 s$; $\Delta x_{propio} = 0$; $\Delta t_{Tierra} = 3'1277.10^{10} s$; $\Delta x_{Tierra} = 991'758$ *años luz;* $\Delta s^2 = \Delta s'^2 = 8'06.10^{33}$ *(salvo aproximaciones en los decimales, que pueden dar lugar a grandes errores, por lo que es conveniente trabajar o sin decimales o con los máximos decimales posibles). Como* $\Delta s^2 > 0$ *tenemos un intervalo tipo tiempo, que se puede conectar con una señal con v<c (en este caso están conectados con una señal llamada "nave", que viaja a velocidad menor que la luz); Además, existe un SDR en que ambos sucesos se ven en el mismo punto, que es precisamente la propia nave (para ella, a la salida y a la llegada, x=0), con la velocidad relativa ya calculada anteriormente (cero respecto a la nave,* $0'999954c$ *respecto a la Tierra)*

NOTA: Es muy recomendable completar correctamente y, sobre todo, entender perfectamente este último ejercicio antes de continuar con el resto de temas.

4. Cono de luz

Intentemos visualizar la posibilidad o imposibilidad de mostrar la causalidad entre los sucesos en el espacio-tiempo de Minkovskii (recordar: causalidad significa que la causa que produce un fenómeno ocurra antes que la consecuencia, es decir, que veamos la explosión más tarde que cuando tuvo lugar la propia explosión). Para verlo, tomemos primero el caso límite y dibujémoslo: veamos qué le ocurre a una señal que viaja a la velocidad de la luz, *c*, para la cual se cumple que, en dos sucesos:

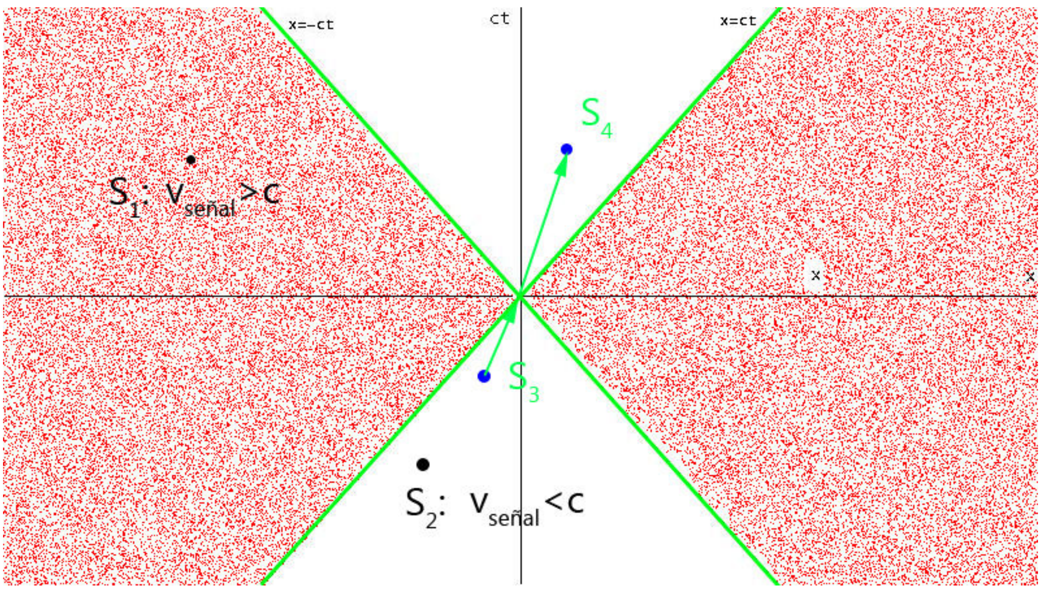
$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 = 0$$

Puesto que la velocidad de la luz es constante y será igual al espacio recorrido entre el tiempo transcurrido entre un suceso y otro, que a su vez es igual a la velocidad de la señal que los une:

$$v_{señal} = \frac{|\Delta x|}{\Delta t} = \pm c$$

Donde los signos más y menos vienen de hacer la raíz cuadrada del cuadrado de $\frac{|\Delta x|}{\Delta t}$, obteniendo dos soluciones para la velocidad posibles.

Así pues, vamos a dibujar un esquema en el que representaremos dos sucesos en el espacio tiempo, atendiendo a esta última cualidad. Lo haremos primero en dos dimensiones (*x* y c*t,* esto último por las dimensiones), y luego lo generalizaremos a tres (*x*, *y*, y *t*), donde pararemos, dado que su expresión en las cuatro dimensiones es más complicada e imposible de dibujar coherentemente.



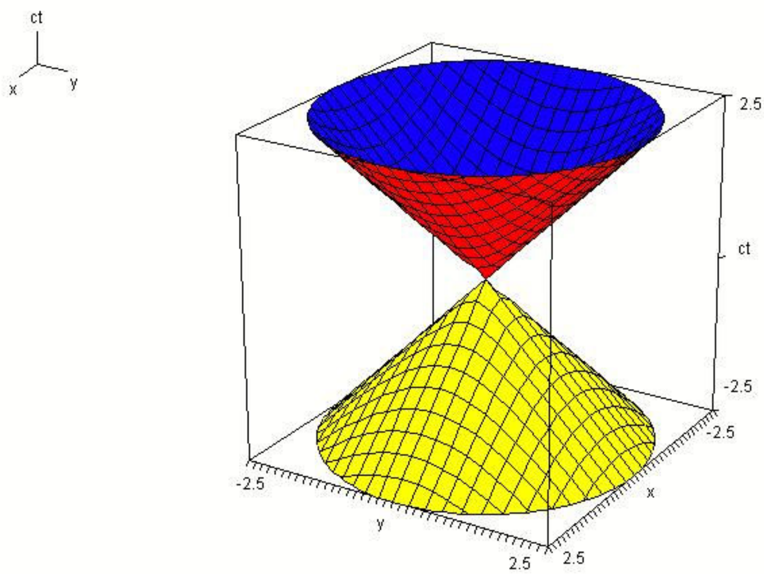
En este esquema tenemos las siguientes posibilidades

a) Si unimos dos sucesos dentro de la zona en blanco, como son s_2 con s_3, ó s_3 con s_4, o s_2 con s_4, las señales que unen dichos sucesos son posibles, dado que llevarían velocidad menor que la de la luz. Se sobreentiende que tan sólo son posibles señales que vayan desde abajo hasta arriba (no sería posible por ejemplo mandar una señal de s_4 a s_3), ya que lo que sucedería entonces sería que la señal retrocede en el tiempo, lo cual consideraremos como imposible. Así, se suele considerar, tomando como referencia una señal en el origen de coordenadas, como ***hoy*** justamente a esa señal, en el origen; como ***pasado absoluto*** a toda la región no sombreada por debajo de *hoy*, y como ***futuro absoluto*** a toda señal que esté por encima de *hoy*. Esto no significa más que si desde el origen (*hoy)* envío una señal de cualquier tipo a otro lugar, esta llegará en un futuro (unos pocos segundos después, o unos años después). Esto es lo que ocurre al ir, por ejemplo, desde el origen de coordenadas (suceso 0) al suceso s_4.

b) Señales lumínicas: serán todas aquellas señales que unan puntos que se encuentren justo en los límites de las líneas asintóticas (las líneas verdes). Todo par de sucesos que se encuentre sobre estas líneas sólo puede ser unido por señales con velocidad igual a la de la luz (es decir, con objetos no másicos, los fotones, que viajan a la velocidad de la luz), y en general cualquier par de sucesos que al unirlos tengan la misma pendiente que estas líneas. Puede demostrarse en este esquema que la pendiente de la recta que une los dos sucesos es precisamente la velocidad de la señal que viaja entre ellos.

c) Señales hiperlumínicas: serán aquellas que unen puntos fuera de la zona blanca, es decir, en la zona sombreada, como por ejemplo s_1. Para unirlas, sería necesaria una señal por encima de la velocidad de la luz, lo cual no es causalmente posible. Esta región, pues, se dice que no es causalmente posible, y no podemos unir sucesos con ella (al menos, no con señales reales).

El dibujo en más dimensiones es un tanto engorroso y no es fácil de interpretar (aunque sí de dibujar), y quedaría como sigue (en dos dimensiones más el tiempo):



A esto se le llama un **cono de luz**, y representa los sucesos espaciotemporales. Justo en la superficie del cono (en las líneas de la figura en 2D o en la superficie del cono), o los que tengan una pendiente igual que las de la superficie del cono, los sucesos pueden interaccionar entre sí si la señal que emite uno y recibe el otro va a la velocidad de la luz. Cualquier par de sucesos que se encuentren en el cono de luz son los sucesos de tipo tiempo (aquellos que pueden ocurrir en el mismo punto del espacio pero no pueden ser simultáneos), mientras que todo par de sucesos fuera del cono posee un intervalo espacial (no causal: pueden ocurrir simultáneamente en el tiempo dos sucesos para algún sistema de referencia).

Como prueba, puede intentar unirse cualquier punto con el origen de coordenadas, y se puede comprobar, como decíamos, que si el suceso está dentro del cono, su velocidad es menor que la de la luz, ya que la pendiente de la recta será mayor que *c*, mientras que si el suceso elegido está dentro del cono y lo unimos con una recta con el origen, su pendiente, es decir, su velocidad, será menor que *c*.

La señal que debemos mandar de un suceso a otro dentro del cono viajará, pues, a velocidad menor que la luz. En cambio, la señal que media los sucesos fuera del cono debería viajar a velocidad mayor que la de la luz. Estos sucesos no están **conectados causalmente**, es decir, de poder ser reales, ocurriría antes la causa que el efecto que la produce (primero oigo la llamada telefónica, y luego se produce dicha llamada). Las líneas que unen dos sucesos se llaman **líneas de universo**. Por ejemplo, la línea que une s_3 y s_4 en el dibujo en 2D es una línea de universo. Por ejemplo, dentro del cono, dos sucesos unidos s_1 y s_2 por una línea de universo son causalmente posibles (si ambos están dentro del cono). Imaginemos que el suceso s_1 significa que en un cierto punto del espacio y en un cierto tiempo alguien emite una llamada telefónica (de velocidad la velocidad del sonido, por ejemplo). El suceso es recogido en s_2, otro punto del espacio un poco más tarde. Este suceso es causalmente posible, como es lógico: primero se llama y luego se recibe la llamada. En cambio, dos sucesos fuera del cono significarían lo siguiente: una estrella detona y se ve en un cierto punto. En este caso, primero ocurriría la visualización de la detonación, y después la propia detonación. Esto, como es lógico, no es posible (desde ningún sistema de referencia), y por eso los sucesos tipo espacio no son físicamente posibles.

Recalcamos de nuevo los ejes dimensionalmente correctos: las dimensiones de *ct* son las siguientes: $[ct] = [c][t] = \frac{L}{T}T = L$, es decir, dimensiones de longitud. De este modo tenemos escritas las coordenadas espaciales y las coordenadas temporales en las mismas unidades. El siguiente paso es relativamente sencillo: puesto que tenemos un tiempo variable y relativo al observador, colocado en dimensiones de longitud, y unas coordenadas espaciales que a fin de cuentas actúan de la misma manera que el tiempo, ¿por qué no construir un vector que incluya tanto a las coordenadas espaciales como a las temporales? Este es el concepto del **espacio tiempo de Minkowskii**, los cuadrivectores, y el uso de tensores en la teoría especial de la relatividad.

Es muy recomendable un applet que se encuentra en la "wikipedia", nada más buscar transformaciones de Lorentz, en que se puede visualizar un eje coordenado, con el "hoy" como el centro, y las líneas de universo de un objeto que se mueve aceleradamente por el espacio-tiempo de Minkovskii. La dirección de la página es:

http://es.wikipedia.org/wiki/Transformación_de_Lorentz

5. ***Espacio tiempo de Minkovskii***

Aunque explicar qué es un espacio tiempo de Minkovskii podría llevar tiempo (habría que dar algunas nociones de geometría diferencial, hablar de topologías, cartas y demás), intentaremos introducirlo de manera natural (para más explicaciones, ver por ejemplo el libro de Landau o cualquier libro de geometría diferencial o similar).

Como decíamos en el capítulo anterior, tenemos las coordenadas espaciales y las temporales escritas en las mismas dimensiones y actuando unas sobre otras interrelacionadamente. Dada esta situación, podemos inventarnos una nueva terminología, los **cuadrivectores**, cuya función consiste básicamente en juntar todas estas coordenadas. El primero de todos estos cuadrivectores nos relaciona precisamente aquellas coordenadas, con las que construiremos el intervalo espaciotemporal Δs^2:

$$\boxed{x^\mu = (ct, \vec{x}) = (ct, x, y, z)}$$

Donde hemos incluido de nuevo la *c* en la coordenada temporal para que todas posean las mismas unidades. El superíndice μ no nos debería preocupar por ahora (veremos su significado en un par de líneas). La cuestión aquí es la siguiente: si hacemos ahora el producto escalar de este cuadrivector consigo mismo, nos interesaría que saliese precisamente el cuadrado del intervalo espaciotemporal. En cambio, vemos que el resultado, con el producto escalar cartesiano (euclídeo), sería:

$$x^\mu x^\mu = (ct, \vec{x})(ct, \vec{x}) = c^2t^2 + |\vec{x}|^2 = c^2t^2 + x^2 + y^2 + z^2$$

Que no es lo que buscamos, ya que no aparecerían los signos menos. Y en realidad, aunque construyamos distinto el cuadrivector, tampoco aparecerían, ya que el producto escalar en un espacio euclídeo común es siempre una cantidad positiva, y ya vimos anteriormente que el cuadrado del intervalo espaciotemporal no tenía por qué ser necesariamente así. Así pues, el siguiente paso es hablar de un producto escalar distinto al euclídeo: el de Minkovskii. Para ello, empezaremos por diferenciar dos tipos de formas de escribir los cuadrivectores:

a) ***Cuadrivector contravariante***: índices arriba: $\boxed{x^\mu = (ct, \vec{x})}$
b) ***Cuadrivector covariante***: índices abajo: $\boxed{x_\mu = (ct, -\vec{x})}$

Donde, por ahora, la única diferencia entre uno y otro es que en el covariante el índice está abajo, y las coordenadas espaciales llevan un signo menos, mientras que el contravariante lleva el índice arriba y las coordenadas espaciales no llevan signo menos. OJO: este convenio, al igual que el del cuadrado del intervalo espaciotemporal, puede ser distinto según el autor, así que prevenimos al lector sobre este hecho antes de incurrir en confusiones al mirar en distintos apuntes o libros.

De esta forma, ahora podremos hacer lo siguiente:

$$x^\mu x_\mu = c^2t^2 - |\vec{x}|^2 = c^2t^2 - x^2 - y^2 - z^2$$

Es decir, obtenemos precisamente el cuadrado del intervalo espaciotemporal buscado, Δs^2, como queríamos. Además, la cantidad $\Delta s^2 = x^\mu x_\mu$ se dice que es un **invariante Lorentz** (invariante bajo transformaciones de Lorentz), que ya demostramos en su momento (por supuesto, en la expresión anterior, cada coordenada habría que interpretarla como la diferencia de la final menos la inicial)

Pero, ¿qué significa el subíndice o superíndice μ? Este índice simplemente nos marca cada una de las coordenadas, de modo que:

i) para $\mu = 0 \Longrightarrow x^0 = ct$
ii) para $\mu = i \Longrightarrow x^i = x, y$ ó z para *i=1, 2, 3*

Creando, pues, el cuadrivector como sigue:

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \vec{x})$$

$$x_\mu = (x_0, x_1, x_2, x_3) = (ct, -x, -y, -z) = (ct, -\vec{x})$$

Es conveniente y usual hacer esta separación de coordenadas, de forma que se suelen usar letras griegas para denotar a los índices que incluyen a todo tipo de coordenadas, como μ, y a índices latinos, como "*i*", cuando se quiere hacer referencia a coordenadas espaciales. Visto así, es fácil ver que lo que estamos haciendo es lo siguiente:

$$\Delta s^2 = \sum_{\mu=0}^{3} x^\mu x_\mu = x^0 x_0 + \sum_{i=1}^{3} x^i x_i = = x^0x_0 + x^1x_1 + x^2x_2 + x^3x_3$$
$$= (ct)(ct) + (x)(-x) + (y)(-y) + (z)(-z) = c^2t^2 - x^2 - y^2 - z^2$$

Que es la expresión buscada. La cuestión es que (quizá por vaguería o por lucidez), a Einstein le debía aburrir colocar sumatorios por todas partes, así que se inventó un nuevo convenio, que se suele llamar **convenio de índices repetidos de Einstein**. Este convenio es muy simple: cada vez que se vea un índice covariante (abajo) con uno contravariante (arriba), iguales y juntos en dos objetos que se están multiplicando, se sobreentiende que detrás hay un sumatorio sobre todas sus componentes. Así, lo que nos quedaría sería, por ejemplo, la siguiente notación:

$$\sum_{\mu=0}^{3} x^\mu x_\mu \equiv x^\mu x_\mu$$

Pero en cambio, si queremos ver qué significa lo siguiente: $x^{\mu}x_{\beta}$ no significa nada en especial, ya que al no ser iguales no están sumando, simplemente son un montón de coordenadas juntas. Supongamos una serie de diferentes situaciones para intentar ilustrar esto un poco:

Ejemplo 5.1. Ilustrar las siguientes cantidades:

$$a)\ W \equiv w^{\mu}w_{\mu} = \sum_{\mu=0}^{n} w^{\mu}w_{\mu} = w^1w_1 + w^2w_2 + w^3w_3 + \cdots + w^nw_n$$

Donde $\mu = 1 \ldots n$ (aunque w^{μ} no tiene por qué ser necesariamente un cuadrivector), pero vale para ilustrar el convenio de índices repetidos. En cambio:

$$b)\ W^{\mu}_{\ \beta} \equiv w^{\mu}w_{\beta} = \begin{cases} si\ \mu = 0 & \beta = 0 & \Longrightarrow & W^0_{\ 0} = w^0w_0 \\ si\ \mu = 0 & \beta = 1 & \Longrightarrow & W^0_{\ 1} = w^0w_1 \\ si\ \mu = 1 & \beta = 0 & \Longrightarrow & W^1_{\ 0} = w^1w_0 \\ \quad \ldots & \ldots & & \ldots \\ si\ \mu = n & \beta = p & \Longrightarrow & W^n_{\ p} = w^nw_p \end{cases}$$

Es decir, esa cosa $W^{\mu}_{\ \beta}$ toma distintos valores según valgan los índices, pero no sabemos nada más sobre él en principio. Como mucho, esto puede recordarnos a una matriz. Supongamos que $\mu = 0,1,2$ (n=2) y que $\beta = 0,1,2$ (p=2):

$$W^{\mu}_{\ \beta} = \begin{pmatrix} w^0w_0 & w^1w_0 & w^2w_0 \\ w^0w_1 & w^1w_1 & w^2w_1 \\ w^0w_2 & w^1w_2 & w^2w_2 \end{pmatrix}$$

Donde los índices significarían fila y columna respectivamente. Pero no podemos ir mucho más lejos por ahora.

$$c)\ T \equiv \rho^{\beta}_{\ \mu}x^{\mu}x_{\beta} = \sum_{\mu=0}^{n}\sum_{\beta=0}^{m} \rho^{\beta}_{\ \mu}x^{\mu}x_{\beta} = \rho^0_{\ 0}x^0x_0 + \rho^1_{\ 0}x^0x_1 + \cdots + \rho^m_{\ 0}x^mx_0 + \cdots + \rho^m_{\ n}x^nx_m$$

Donde hemos supuesto que los índices de μ y β van de 0 a n y de 0 a m respectivamente: al final obtenemos un número (escalar), una vez hechos los productos.

$d)$ Multiplicación de matrices: caso sencillo (matrices de 2x2):

$$C^{\mu}_{\ \rho} \equiv \begin{pmatrix} c^1_{\ 1} & c^1_{\ 2} \\ c^2_{\ 1} & c^2_{\ 2} \end{pmatrix} = \begin{pmatrix} a^{11} & a^{12} \\ a^{21} & a^{22} \end{pmatrix} . \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a^{11}.b_{11} + a^{12}.b_{21} & a^{11}.b_{12} + a^{12}.b_{22} \\ a^{21}.b_{11} + a^{22}.b_{21} & a^{21}.b_{12} + a^{22}.b_{22} \end{pmatrix}$$

También podríamos ponerlo como: $c^{\mu}_{\ \rho} = a^{\mu\beta}.b_{\beta\rho}$
Donde por ejemplo, si queremos hallar el elemento $c^1_{\ 2}$ nos quedaría

$$c^1_{\ 2} = a^{1\beta}.b_{\beta 2} = \sum_{\beta=1}^{2} a^{1\beta}.b_{\beta 2} = a^{11}.b_{12} + a^{12}.b_{22}$$

Y así con cada uno de los elementos, podríamos escribir cada uno de los términos de nuestra nueva matriz. Aquí se empieza a ver un poco la utilidad de los tensores a la hora de escribir largas cadenas de matrices por componentes, de forma más compacta y más manejable (para quien sabe manejarlo).

6. Transformación de Lorentz

Tal y como acabamos de ver, podemos generalizar todo el formalismo visto antes con unos elementos nuevos, a los que llamaremos **tensores**, aunque su definición esperará aún un rato para hacerla suficientemente rigurosa. Supongamos que queremos, para empezar, aprender a cambiar coordenadas covariantes por contravariantes, es decir, buscar la forma de pasar de $x^\mu = (ct, \vec{x})$ a $x_\mu = (ct, -\vec{x})$. Para ello, hay una forma sencilla de hacerlo: necesitamos multiplicar al vector por algo que me de el otro vector (algo así como una aplicación lineal), pero que sólo afecte a las coordenadas espaciales, y no a las temporales. Este elemento que vamos a buscar se llama el **tensor métrico** o simplemente la **métrica**, y viene dado por:

$$\boxed{\eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \equiv diag[1, -1, -1, -1]}$$

NOTA: En otras notaciones, el tensor sería justamente al revés: $\eta^{\mu\nu} = diag[-1, +1, +1, +1]$, todo depende del convenio elegido para definir las coordenadas, aunque teniendo cuidado no es relevante, aunque conviene tener cuidado si se quieren manejar dos tipos de apuntes o libros a la vez.

Donde los subíndices y los superíndices denotan la fila y columna a la que nos referimos, es decir:

a) $\eta^{\mu\nu} = 1$ si $\mu = \nu = 0$
b) $\eta^{\mu\nu} = -1$ si $\mu = \nu = i$, con *i=1,2,3*, es decir, $\eta^{ii} = -1$
c) $\eta^{\mu\nu} = 0$ si $\mu \neq \nu$

El tensor inverso de este (con los dos índices abajo), queda exactamente igual que el que acabamos de escribir, aunque por ahora no demostraremos ni cómo ni por qué es así:

$$\boxed{\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \equiv diag[1, -1, -1, -1]}$$

Con este tensor métrico, podemos transformar coordenadas covariantes en contravariantes, es decir, subir y bajar índices. De nuevos usaremos el convenio de suma de índices repetidos de Einstein:

$$x^\mu = \eta^{\mu\nu} x_\nu$$

Que en el convenio habitual quiere decir:

$$x^\mu = \sum_{\nu=0}^{3} \eta^{\mu\nu} x_\nu = \eta^{\mu 0} x_0 + \eta^{\mu 1} x_1 + \eta^{\mu 2} x_2 + \eta^{\mu 3} x_3 = \eta^{\mu 0} x_0 + \eta^{\mu 1} x_1 + \eta^{\mu 2} x_2 + \eta^{\mu 3} x_3$$

Es decir, si nos fijamos en las cuatro componentes con la expresión anterior:

μ	x^μ	x^μ	**Ya que:**
0	$x^0 = \eta^{00}x_0 + \eta^{01}x_1 + \eta^{02}x_2 + \eta^{03}x_3 = x_0 = ct$	$x^0 = ct$	$\eta^{00} = 1$, resto cero
1	$x^1 = \eta^{10}x_0 + \eta^{11}x_1 + \eta^{12}x_2 + \eta^{13}x_3 = -x_1$	$x^1 = -x_1$	$\eta^{11} = -1$, resto cero
2	$x^2 = \eta^{20}x_0 + \eta^{21}x_1 + \eta^{22}x_2 + \eta^{23}x_3 = -x_2$	$x^2 = -x_2$	$\eta^{22} = -1$, resto cero
3	$x^3 = \eta^{30}x_0 + \eta^{31}x_1 + \eta^{32}x_2 + \eta^{33}x_3 = -x_3$	$x^3 = -x_3$	$\eta^{33} = -1$, resto cero

Conviene bien familiarizarse con esta tabla, y entender exactamente por qué hacemos lo que hacemos, y cómo se obtiene la tercera columna, puesto que en adelante este tipo de cálculos serán cada vez más frecuentes (y por desgracia menos intuitivos). Pues bien, con este formalismo nuevo, podemos escribir la matriz de transformación de coordenadas de Lorentz de forma relativamente simple. Una matriz de transformación es, en pocas palabras, lo que acabamos de escribir, el elemento $\eta^{\mu\nu}$, ya que me permite pasar de unas coordenadas a otras. Pero nótese que este elemento en realidad lo que hace es simplemente pasar de una forma de las coordenadas (covariantes) a otra (contravariantes), es decir, lo único que hace es subir o bajar índices, y poner o quitar signos en las coordenadas espaciales, si se lo enchufamos al cuadrivector x^μ. Así que, en realidad, $\eta^{\mu\nu}$ no es una verdadera transformación de coordenadas. Lo que queremos es saber como pasar de unas coordenadas x^μ a otras en otro sistema de referencia x'^μ, es decir, lo que hicimos al principio del tema (transformaciones de Galileo o de Lorentz), pero desde el punto de vista de nuestro nuevo formalismo.

Observemos un ejemplo de transformación conocido antes de emprender las transformaciones de Lorentz, y obtengamos una matriz conocida de cambio de base de unas coordenadas a otras.

***<u>Ejemplo 6.1</u>**. En una rotación alrededor del eje z de un cierto ángulo θ (una rueda de un coche), hallar las transformaciones de coordenadas de un observador que es solidario con la rueda (un tornillo clavado en la rueda) frente a uno situado con sus ejes fijos pinchados en el centro de la misma:*

Como sabemos, la matriz que hace cambiar de un eje de referencia al otro no es más que una matriz de rotación que, por ser en el eje z, viene dada por:

$$R_z(\theta) = \begin{pmatrix} cos\theta & -sen\theta & 0 \\ sen\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Y las coordenadas en los dos sistemas de referencia serán:

$$S \Longrightarrow \vec{x} = (x, y, z)$$
$$S' \Longrightarrow \vec{x}' = (x', y', z')$$

El paso de una a otra vendrá dado por:

$$\vec{x}' = R_z(\theta)\vec{x}$$

Que matricialmente:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} cos\theta & -sen\theta & 0 \\ sen\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} xcos\theta - ysen\theta \\ xsen\theta + ycos\theta \\ z \end{pmatrix}$$

En componentes:

$$\begin{cases} x' = x\cos\theta - y\,\mathrm{sen}\theta \\ y' = x\,\mathrm{sen}\theta + y\cos\theta \\ z' = z \end{cases}$$

Pues bien, en nuestra nueva notación:

$$\begin{cases} R_z(\theta) \equiv \Lambda^{\mu}{}_{\nu} \\ \vec{x} \equiv x^{\mu} \\ \vec{x}' \equiv x'^{\mu} \end{cases} \qquad \text{Con } \Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\mathrm{sen}\theta & 0 \\ \mathrm{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Y nuestra transformación vendrá dada por:

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}$$

Con el convenio de suma de índices repetidos de Einstein. En notación más habitual (con sumatorios):

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} = \sum_{\nu=1}^{3} \Lambda^{\mu}{}_{\nu} x^{\nu} = \Lambda^{\mu}{}_{1} x^1 + \Lambda^{\mu}{}_{2} x^2 + \Lambda^{\mu}{}_{3} x^3$$

Para cada valor de μ (en este caso, al tratarse de rotaciones espaciales, $\mu = 1,2,3$). Por componentes:

μ	x'^{μ}	x'^{μ}
1	$x'^1 = \Lambda^1{}_1 x^1 + \Lambda^1{}_2 x^2 + \Lambda^1{}_3 x^3 = x^1\cos\theta - x^2\mathrm{sen}\theta$	$x'^1 = x' = x\cos\theta - y\,\mathrm{sen}\theta$
2	$x'^2 = \Lambda^2{}_1 x^1 + \Lambda^2{}_2 x^2 + \Lambda^2{}_3 x^3 = x^1\mathrm{sen}\theta + x^2\cos\theta$	$x'^2 = y' = x\cos\theta - y\,\mathrm{sen}\theta$
3	$x'^3 = \Lambda^3{}_1 x^1 + \Lambda^3{}_2 x^2 + \Lambda^3{}_3 x^3 = x^3$	$x'^3 = z' = x\cos\theta - y\,\mathrm{sen}\theta$

Es decir, lo mismo. Conviene saber hacer esta tabla y este ejemplo, ya que es básico para lo que sigue.

Viendo lo visto, pasemos ahora a la representación de las transformaciones de Lorentz, o lo que es lo mismo, del llamado **grupo de Lorentz** (llamado grupo porque cumple las propiedades matemáticas de un grupo, que no enumeraremos aquí por ahora)

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3 \\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3 \\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3 \\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3 \end{pmatrix}$$

$$\Lambda^{\mu}{}_{\nu} = \begin{pmatrix} \gamma & -\gamma\frac{V_1}{c} & -\gamma\frac{V_2}{c} & -\gamma\frac{V_3}{c} \\ -\gamma\frac{V_1}{c} & 1-(1-\gamma)\frac{(V_1)^2}{V^2} & (\gamma-1)\frac{V_1V_2}{V^2} & (\gamma-1)\frac{V_1V_3}{V^2} \\ -\gamma\frac{V_2}{c} & (\gamma-1)\frac{V_2V_1}{V^2} & 1-(1-\gamma)\frac{(V_2)^2}{V^2} & (\gamma-1)\frac{V_2V_3}{V^2} \\ -\gamma\frac{V_3}{c} & (\gamma-1)\frac{V_3V_1}{V^2} & (\gamma-1)\frac{V_3V_2}{V^2} & 1-(1-\gamma)\frac{(V_3)^2}{V^2} \end{pmatrix}$$

O de forma más compacta, en notación tensorial:

$$\boxed{\Lambda^i{}_j = \begin{cases} \Lambda^0{}_0 = \gamma \\ \Lambda^0{}_i = -\gamma \dfrac{V_i}{c} \\ \Lambda^i{}_0 = -\gamma \dfrac{V^i}{c} \\ \Lambda^i{}_j = \delta^i{}_j + (\gamma - 1)\dfrac{V^i V_j}{V^2} \end{cases}}$$

Donde $\delta^i{}_j$ es la delta de Dirac (en 4 dimensiones), que vale uno si ambos índices son iguales, y cero si sin distintos.

***<u>Ejercicio 6.1</u>**. Calcular en base a lo anterior la matriz de transformación de Lorentz para cuando la velocidad es paralela al eje OX, es decir, con $\vec{V} = (V, 0,0)$ (la que está en todos los libros)*

Con esto, ya podemos dar una definición más exacta de lo que es un tensor, el orden del mismo, etc., dado que tenemos una matriz de cambio de base o de transformación de coordenadas con la que podemos operar, y además podremos saber cómo saber qué objetos son tensores y cuales no:

Un tensor n-contravariante y m-covariante es un elemento que se transforma n veces como lo hacen las coordenadas contravariantes y m veces como lo hacen las coordenadas covariantes. Empecemos por los casos del ejemplo anterior (el de la rotación): los tensores de orden uno (tienen un solo índice, como los vectores): cuadrivectores covariantes y contravariantes:

$\mathbf{a)}$ $\mathbf{Tensor\ 1 - contravariante}$: $\boxed{\boldsymbol{x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu}}}$

$\mathbf{b)}$ $\mathbf{Tensor\ 1 - covariante}$: $\boxed{\boldsymbol{x'_{\mu} = \Lambda^{\nu}{}_{\mu} x_{\nu}}}$

Este tipo de tensores no son ni más ni menos que los vectores de toda la vida. Si el índice μ recorre los índices del 1 al 3, tendremos un vector con *(x,y,z)* como los usuales. Si en cambio el índice μ recorre desde el cero hasta el 3, tendremos un cuadrivector como los que ya hemos visto, de la forma *(ct,x,y,z)*. El índice nos dice qué coordenada de las que posee el vector estamos mirando. Es evidente que en un vector sólo necesitamos un índice (en una matriz, por ejemplo, necesitaríamos dos: fila y columna).

Ya vimos que la transformación entre vectores se realizaba mediante una matriz de cambio de base, a la que hemos llamado $\Lambda^{\mu}{}_{\nu}$ (los índices recorren filas y columnas, respectivamente, para darnos todos los posibles elementos de la matriz), como en el ejemplo anterior, y por tanto su transformación será de la forma anterior. Pero esto ya lo sabíamos. Veamos ahora otras tres posibilidades más: tensores de orden 2 (dos índices, con una matriz de transformación de coordenadas, al igual que antes, $\Lambda^{\mu}{}_{\nu}$):

c) $\textbf{Tensor 2} - \textbf{contravariante}$: $\boxed{T'^{\mu\nu} = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}T^{\rho\sigma}}$

d) $\textbf{Tensor 2} - \textbf{covariante}$: $\boxed{T'_{\mu\nu} = \Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu}T_{\rho\sigma}}$

e) $\textbf{Tensor 1} - \textbf{covariante 1} - \textbf{contravariante (mixto)}$: $\boxed{T'^{\mu}{}_{\nu} = \Lambda^{\mu}{}_{\rho}\Lambda^{\sigma}{}_{\nu}T^{\rho}{}_{\sigma}}$

Y por último antes de empezar a explicar de qué va esto, veamos el caso más simple de todos: el tensor 0-covariante y 0-contravariante, que se transforma como cero veces un vector covariante y como cero veces un vector contravariante, es decir, no se transforma. A estos objetos los llamaremos **escalares**, y cumplen la propiedad de ser invariantes bajo transformaciones. Por ejemplo bajo la transformaciones de Lorentz, el cuadrado del intervalo espaciotemporal es un escalar (no varía al cambiar de sistema de referencia). Pero hay muchos más: la masa, por ejemplo, y otros muchos que veremos en un par de páginas.

Vamos a intentar dar unas pistas sobre lo que estamos haciendo aquí, ya que esto puede empezar a dar un poco de miedo. Para empezar, recordemos que en estas últimas definiciones y en lo que sigue, usaremos siempre el convenio de índices repetidos de Einstein, que ya daremos por hecho. Sin embargo, a veces conviene volver de tanto en cuando al convenio habitual, aunque sea para sentirnos en casa por un momento. Por ejemplo, veamos qué significa eso de la transformación de un tensor del tipo 1-covariante (o contravariante, la idea es la misma):

$$x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} \xrightarrow{Conv.Einstein} x'^{\mu} = \sum_{\nu=0}^{n} \Lambda^{\mu}{}_{\nu}x^{\nu}$$

Y dado que el índice μ no está repetido, esto significa que no está sumando, por lo que tomará el valor que nosotros le demos (uno por cada componente del vector x'^{μ}, como ya hemos hecho anteriormente). Es decir, la expresión que tenemos arriba lo único que nos dice es, dada una matriz de transformación $\Lambda^{\mu}{}_{\nu}$ y un cierto vector (cuadrivector o, en general, un vector de las componentes que queramos), llamado x^{ν}, cómo conseguir pasar de un sistema de referencia *S* que ve x^{ν} a otro sistema de referencia *S'* que ve x'^{μ}. Esto es lo que hemos hecho en el ejemplo anterior, con la matriz de rotación.

Tomemos ahora la transformación de un tensor de orden dos, por ejemplo el 2-contravariante. Para empezar, podemos intentar imaginar qué es un objeto de orden 2, es decir, con dos índices. ¿A qué suena un objeto que para saber cuánto vale debemos darle un índice, *i,* y además otro índice, digamos *j*? Pues sí, efectivamente, podríamos representarnos mentalmente un objeto con dos índices (da igual que sea contravariante, covariante o mixto), como una matriz, donde los índices indicarán la fila y la columna correspondiente donde buscar al elemento. Así pues, supongamos que un observador en un sistema de referencia *S* ve una matriz $T^{\rho\sigma}$, donde ρ y σ indicarían la fila y la columna respectivamente. Imaginemos que tenemos una transformación de coordenadas $\Lambda^{\mu}{}_{\rho}$ (que puede ser la de Lorentz, o la de Galileo, o una simple matriz de rotación como en el ejemplo anterior). Pues bien, lo que nos dicen las expresiones anteriores es que si quiero ver ahora esa matriz desde otro sistema de coordenadas *S'*, y sabiendo la matriz de cambio de base $\Lambda^{\mu}{}_{\rho}$ entre ambas, puedo hacer:

$$T'^{\mu\nu} = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}T^{\rho\sigma} \xrightarrow{Conv.Einstein} T'^{\mu\nu} = \sum_{\rho=0}^{n}\sum_{\sigma=0}^{m} \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}T^{\rho\sigma}$$

Nótese que los índices ρ y σ en esta expresión no tienen una realidad como tal, puesto que están siendo sumados a todas sus componentes (de 0 a n y de 0 a m respectivamente). A estos índices se les llama **índices mudos** o **índices contraídos**, ya que su única función en esa expresión es servir para recorrerlos de 0 a *n* o de 0 a *m* respectivamente. Sin embargo los índices μ ν son los que me dan cada elemento de la transformación. Por ejemplo, si me pregunto cuánto valdrá el elemento T'^{13} (el que ocupa la "fila uno y columna tres de mi nueva matriz de elementos) en el sistema de referencia *S'*, este será:

$$T'^{13} = \sum_{\rho=0}^{n}\sum_{\sigma=0}^{m} \Lambda^{1}{}_{\rho}\Lambda^{3}{}_{\sigma}T^{\rho\sigma}$$

Quedando ya una suma de términos y productos que, sabiendo qué es el tensor en el sistema *S* y sabiendo la matriz de cambio de base, se puede realizar numéricamente (con números).

Por supuesto, para crear las leyes de transformación de otros tensores no hay más que aplicar convenientemente la matriz de cambio de base tantas veces como sea necesaria. Por ejemplo veamos el tensor 1 covariante 3 contravariante. Su ley de transformación (para que sea tensor) será:

$$\boxed{\boldsymbol{T'_{\tau}{}^{\mu\nu\varepsilon} = \Lambda^{\mu}{}_{\gamma}\Lambda^{\nu}{}_{\beta}\Lambda^{\varepsilon}{}_{\alpha}\Lambda^{\rho}{}_{\tau}T_{\rho}{}^{\alpha\beta\gamma}}}$$

Y así sucesivamente. Puede pensarse, según esto, que cualquier elemento de *n* índices arriba y de *m* índices abajo ya es un tensor de orden (*n,m*). Sin embargo, esto no es así. Notemos que no hemos dicho nada de lo que hemos llamado tensor $T_{\rho}{}^{\alpha\beta\gamma}$, o cualquiera de los anteriores. Pero estos también pueden tener sus propias expresiones. Veamos un par de ejemplos: un tensor de orden cero (un escalar), y un objeto de dos índices que no es un tensor de dos índices:

Ejemplo 6.1*. El cuadrado del intervalo espaciotemporal es un escalar bajo transformaciones de Lorentz (es invariante bajo una transformación de Lorentz):*

$$\Delta s^2 = x^{\mu}x_{\mu}$$

Vemos que a Δs^2 no le hemos puesto ningún índice, dado que su única función es dar un valor numérico. Ya vimos en los ejemplos que Δs^2 no es más que un número (un escalar), y por tanto no necesita índices (no es un vector como lo puede ser x^{μ}, que necesita el índice μ puesto que con él seleccionamos las coordenadas x, y ó z). Veamos como se transformaría:

$$\Delta s'^2 = x'^{\mu}x'_{\mu} = \Lambda^{\rho\tau}x_{\tau}\,\Lambda_{\ \rho\tau}x^{\tau} \ = \Lambda_{\ \rho\tau}\Lambda^{\rho\tau}x_{\tau}\,x^{\tau}$$

Donde hemos tenido que poner iguales los primeros índices de la transformación de coordenadas, dado que si no nos quedarían un par de índices sueltos que no aparecen antes del igual. Pero

$$x_\tau\, x^\tau \xrightarrow{Conv.Einstein} x_\tau\, x^\tau\ = \sum_{\tau=0}^{3} x_\tau\, x^\tau\ = \Delta s^2$$

Sigue siendo el cuadrado del intervalo espaciotemporal (me da igual tener los índices τ que μ, ya que son índices mudos y se están sumando). Esta igual dad que acabamos de escribir, sabiendo que el cuadrado del intervalo espaciotemporal es invariante bajo transformaciones de Lorentz, nos lleva a una interesante propiedad:

$$\Delta s^2 = \Delta s'^2 = x'^\mu x'_\mu = \Lambda_{\ \rho\tau}\Lambda^{\rho\tau} x_\tau\, x^\tau\ = \Lambda_{\ \rho\tau}\Lambda^{\rho\tau}\Delta s^2$$
$$\Delta s^2 = \Lambda_{\ \rho\tau}\Lambda^{\rho\tau}\Delta s^2$$
$$\Lambda_{\ \rho\tau}\Lambda^{\rho\tau} = 1$$

Pero vayamos un poco más lejos. Ya vimos que el tensor métrico $\eta_{\mu\nu}$ nos servía para subir o bajar índices en los cuadrivectores. Pues bien, aquí también podemos utilizarlo para subir o bajar índices, haciendo cosas del tipo:

$$\boxed{\boldsymbol{\eta_{\mu\rho}\Lambda^{\rho\tau} = \Lambda_\mu^{\ \ \tau}}}$$

Pues bien, en la ecuación anterior, $\Lambda_{\ \rho\tau}\Lambda^{\rho\tau} = 1$, apliquémosle la métrica dos veces para bajar uno de los índices de las transformaciones y subir el otro, por supuesto multiplicando a ambos lados de la ecuación por lo mismo, como es habitual:

$$\eta_{\rho\alpha}\eta^{\rho\alpha}\Lambda_{\ \rho\tau}\Lambda^{\rho\tau} = \eta_{\rho\alpha}\eta^{\rho\alpha}$$
$$\eta_{\rho\alpha}\Lambda^{\alpha}_{\ \ \tau}\Lambda^{\rho\tau} = 1$$
$$\Lambda^{\alpha}_{\ \ \tau}\Lambda_\alpha^{\ \ \tau} = 1$$

De donde podemos sacar como conclusiones que $\boxed{\boldsymbol{\det \Lambda^\mu_{\ \ \nu} = \pm 1}}$ y que $\boxed{\boldsymbol{(\Lambda^{\alpha}_{\ \ \tau})^{-1} = \Lambda_\tau^{\ \ \alpha}}}$ (el orden de los índices tiene su importancia, aunque si los vemos como una matriz esto no pueda visualizarse bien, dado que sería lo mismo $\Lambda^\mu_{\ \ \nu}$, $\Lambda_\nu^{\ \ \mu}$, $\Lambda^{\mu\nu}$ ó $\Lambda_{\mu\nu}$)

NOTA: Los índices mudos (repetidos) no nos importan, por lo que estamos poniendo las letras que nos viene en gana, ya que tan sólo valen para hacer sumatorios

Ejemplo 6.2. *Ahora supongamos que queremos construir un objeto con dos índices que no sea un tensor. Imaginemos que construimos el tensor siguiente (el que nos ha dado la gana, no tiene significado):*

$$T^{\mu\nu} \equiv x^\mu x^\nu - x^\nu \frac{\partial F^\mu}{\partial x^0}$$

Donde F^μ es un cuadrivector (sea cual sea), que sabemos que se transforma como un tensor de orden uno. La derivada que aparece es simplemente una derivada temporal (la hemos introducido aquí para hacer un tensor con una notación elegante, e ir viendo ya lo que vendrá más adelante, pero podríamos haber hecho cualquier otro), de la forma:

$$\frac{\partial F^\mu}{\partial x^0} = \frac{\partial F^\mu}{\partial (ct)} = \frac{1}{c}\frac{\partial F^\mu}{\partial t}$$

Trataremos la derivada parcial en el tiempo como una derivada en el sentido usual, aunque más adelante cambiaremos un poco los conceptos. Pues bien, si queremos hacer la transformación de coordenadas al supuesto tensor $T^{\mu\nu}$ definido de la forma que acabamos de poner, nos queda, aplicando transformaciones primero al lado izquierdo de la definición y luego al derecho, lo siguiente:

$$\begin{cases} T'^{\mu\nu} = \Lambda_{\mu\alpha}\Lambda_{\mu\beta}\, T^{\alpha\beta} \\ x'^{\mu}x'^{\nu} - x'^{\nu}\dfrac{1}{c}\dfrac{\partial F'^{\mu}}{\partial t} = \left(\Lambda^{\mu}{}_{\alpha}x^{\alpha}\right)\left(\Lambda^{\nu}{}_{\beta}x^{\beta}\right) - \left(\Lambda^{\nu}{}_{\beta}x^{\beta}\right)\dfrac{1}{c}\dfrac{\partial\left(\Lambda^{\mu}{}_{\alpha}F^{\alpha}\right)}{\partial t} \end{cases}$$

Es decir, le enchufamos una matriz de cambio de base a cada cosa que deba ser transformada. Y ambas deberían ser iguales si realmente es un tensor. Ordenando un poco la segunda expresión (los elementos que aparecen aquí son conmutativos respecto a la multiplicación, de forma que podemos hacer cosas tipo $x^{\alpha}\Lambda^{\nu}{}_{\beta} = \Lambda^{\nu}{}_{\beta}x^{\alpha}$), y haciendo la derivada:

$$\Lambda^{\mu}{}_{\alpha}\,\Lambda^{\nu}{}_{\beta}x^{\alpha}x^{\beta} - \Lambda^{\nu}{}_{\beta}x^{\beta}\frac{1}{c}\left[F^{\alpha}\frac{\partial\left(\Lambda^{\mu}{}_{\alpha}\right)}{\partial t} + \Lambda^{\mu}{}_{\alpha}\frac{\partial\left(F^{\alpha}\right)}{\partial t}\right]$$

Que no es más que la derivada del producto, suponiendo que en una transformación general la derivada de cada elemento de la misma o de alguno sea distinta de cero (la matriz de transformación dependa del tiempo). Volvemos a organizar todo eso:

$$\Lambda^{\mu}{}_{\alpha}\,\Lambda^{\nu}{}_{\beta}x^{\alpha}x^{\beta} - \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta}x^{\beta}\frac{1}{c}\frac{\partial(F^{\alpha})}{\partial t} - \Lambda^{\nu}{}_{\beta}x^{\beta}\frac{1}{c}F^{\alpha}\frac{\partial\left(\Lambda^{\mu}{}_{\alpha}\right)}{\partial t}$$

Que podemos reescribir, sacando factor común, como:

$$\Lambda^{\mu}{}_{\alpha}\,\Lambda^{\nu}{}_{\beta}\left(x^{\alpha}x^{\beta} - x^{\beta}\frac{1}{c}\frac{\partial(F^{\alpha})}{\partial t}\right) - \Lambda^{\nu}{}_{\beta}x^{\beta}\frac{1}{c}F^{\alpha}\frac{\partial\left(\Lambda^{\mu}{}_{\alpha}\right)}{\partial t}$$

Luego nuestras expresiones anteriores quedan como:

$$\begin{cases} T^{\mu\nu} = \Lambda_{\mu\alpha}\Lambda_{\mu\beta}\, T^{\alpha\beta} \\ x'^{\mu}x'^{\nu} - x'^{\nu}\dfrac{1}{c}\dfrac{\partial F'^{\mu}}{\partial t} = \Lambda^{\mu}{}_{\alpha}\,\Lambda^{\nu}{}_{\beta}\left(x^{\alpha}x^{\beta} - x^{\beta}\dfrac{1}{c}\dfrac{\partial(F^{\alpha})}{\partial t}\right) - \Lambda^{\nu}{}_{\beta}x^{\beta}\dfrac{1}{c}F^{\alpha}\dfrac{\partial\left(\Lambda^{\mu}{}_{\alpha}\right)}{\partial t} \end{cases}$$

Y notemos que en la segunda expresión, sí está realmente la transformación de un tensor, pero le acompaña un sumando que sobra, no tensorial:

$$\begin{cases} T^{\mu\nu} \equiv x^{\mu}x^{\nu} - x^{\nu}\dfrac{\partial F^{\mu}}{\partial x^{0}} \\ T'^{\mu\nu} = \Lambda_{\mu\alpha}\Lambda_{\mu\beta}\, T^{\alpha\beta} \\ x'^{\mu}x'^{\nu} - x'^{\nu}\dfrac{1}{c}\dfrac{\partial F'^{\mu}}{\partial t} = \Lambda^{\mu}{}_{\alpha}\,\Lambda^{\nu}{}_{\beta}\,T^{\alpha\beta} - \Lambda^{\nu}{}_{\beta}x^{\beta}\dfrac{1}{c}F^{\alpha}\dfrac{\partial\left(\Lambda^{\mu}{}_{\alpha}\right)}{\partial t} \\ \boldsymbol{T'^{\mu\nu} = T^{\mu\nu} - \Lambda^{\nu}{}_{\beta}x^{\beta}\dfrac{1}{c}F^{\alpha}\dfrac{\partial(\Lambda^{\mu}{}_{\alpha})}{\partial t}} \end{cases}$$

Este objeto, así definido, no es un tensor (salvo que la matriz de transformación de coordenadas no dependa del tiempo, en cuyo caso sí lo sería bajo dicha transformación)

Conviene entender bien los pasos seguidos en estos dos ejemplos, así como tener una idea clara de lo que se ha hecho en todas las páginas anteriores, antes de seguir. Podemos decir que este punto es una especie de *"checkpoint"*, que debe ser superado para continuar correctamente.

7. Otros cuadrivectores y tensores

Hasta ahora el único cuadrivector que hemos visto (en sus versiones covariante y contravariante) es el cuadrivector posición $x^\mu = (ct, \vec{x}) = (x^0, x^1, x^2, x^3) = (ct, x^1, x^2, x^3) = (ct, x, y, z)$. Sin embargo, podemos construir muchos otros cuadrivectores (que como tal, bajo una transformación se transformaran como el cuadrivector posición). Los iremos enumerando uno por uno, al menos los más importantes:

a) Cuadrivelocidad

En nuestra nueva notación, el concepto de cuadrivelocidad no puede ser sencillamente la derivada del cuadrivector posición en el tiempo, dado que el tiempo lo estamos tratando como una coordenada más. Así pues, debemos dar una nueva noción de cuadrivelocidad (ojo: la velocidad de la partícula, o la velocidad relativa entre sistemas seguirán siendo las mismas que toda la vida, eso no cambia, simplemente queremos construir otro cuadrivector). Definiremos pues la cuadrivelocidad como la derivada del cuadrivector posición con respecto al *tiempo propio*, es decir:

$$u^\mu \equiv \frac{dx^\mu}{d\tau} = \gamma \frac{dx^\mu}{dt}$$

Podemos escribir esta cuadrivelocidad de muchas maneras. Por ejemplo, desarrollando la expresión anterior por componentes como sigue, que es la forma más usual:

$$u_\mu = \gamma\left(\frac{dx^0}{dt}, \frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right) = \gamma\left(\frac{d(ct)}{dt}, \frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right) = \gamma\left(c, \frac{dx^1}{dt}, \frac{dx^2}{dt}, \frac{dx^3}{dt}\right)$$

O lo que es lo mismo, sabiendo que $\frac{dx^i}{dt}$=v^i y que $v^i \equiv \vec{v} = (v^1, v^2, v^3)$:

$$\boxed{u^\mu = \gamma(c, \vec{v})}$$

Que es la expresión más habitual para la cuadrivelocidad, aunque las anteriores expresiones también son útiles.

Con esta cuadrivelocidad así definida, podemos obtener una nueva cantidad conservada (invariante Lorentz): el cuadrado de la cuadrivelocidad:

$$u^\mu u_\mu \equiv \sum_{\mu=0}^{3} u^\mu u_\mu = \gamma^2 \frac{dx^\mu dx_\mu}{dt.dt} = \gamma^2 \frac{ds^2}{dt^2} = \gamma^2 \frac{c^2dt^2 - dx^2}{dt^2} = \frac{(c^2 - V^2)}{\left(\sqrt{1-\frac{V^2}{c^2}}\right)^2} = c^2 \frac{1-\frac{V^2}{c^2}}{1-\frac{V^2}{c^2}} = c^2$$

Es decir, el cuadrado de la cuadrivelocidad vale lo mismo sea cual sea el sistema de referencia, y es igual a c^2:

$$\boxed{u'^\mu u'_\mu = u^\mu u_\mu = c^2}$$

b) Cuadrimomento

A continuación vamos a generalizar el momento lineal a nuestro nuevo caso, construyendo el cuadrimomento. Veremos más adelante como el momento en la relatividad especial no es simplemente el producto de la masa por la velocidad, pero dado que ahora simplemente estamos definiendo nuevos cuadrivectores, esto no nos importará mucho por ahora. Lo deduciremos en su momento:

$$\boxed{P^{\mu} = \gamma(E/c, \vec{p})}$$

Operando de forma análoga a lo que teníamos antes, llegamos a la siguiente conclusión:

$$\boxed{P'^{\mu}P'_{\mu} = P^{\mu}P_{\mu} = m^2}$$

Es decir, el cuadrimomento también se conserva para cualquier sistema de referencia. Esto es el análogo de lo que ya sabíamos para la mecánica clásica no relativista, pero generalizado. Por ahora nos conformaremos con esta información, aunque más adelante le daremos mucho más uso al cuadrimomento. Lo importante de esta última ecuación ya no es sólo que se conserve el cuadrado del cuadrimomento, sino que lo hace **para cualquier sistema de referencia**. Es decir, si analizamos un objeto moviéndose desde un cierto sistema de referencia S, y calculamos el cuadrado del cuadrimomento (la energía dependerá del sistema, y la velocidad también), obtenemos m^2. Si ahora dejamos pasar un rato (dejamos que el objeto siga su movimiento), y cogemos otro sistema de referencia, S', y calculamos el cuadrimomento (la energía será distinta al anterior, primero por cambiar el sistema de referencia, y segundo por haber dejado pasar el tiempo, e igual con la velocidad), entonces comprobamos que el cuadrado del cuadrimomento vuelve a devolvernos la cantidad m^2. Esta es la clave para las reacciones entre partículas, que veremos más adelante.

Aun a riesgo de adelantarnos, y dado que se supone que ya es una relación conocida, recordemos que la energía de una partícula en movimiento (energía absoluta), viene dada por la expresión:

$$\boxed{E = \sqrt{m^2c^4 + m^2\vec{p}^2}}$$

Además, para los casos en que tengamos una partícula libre (que no interacciona con nada, es decir, que no hay un potencial que le aporte o reste energía), la energía será igual a la energía cinética, que en función del momento viene dada por:

$$\boxed{E = T = \frac{1}{2}m\vec{v}^2 = \frac{\vec{p}^2}{2m}}$$

En cambio, para una partícula no másica (el fotón), recordemos que su energía viene dada por:

$$E = h\nu = h\frac{1}{T} = h\frac{\omega}{2\pi} = \hbar\omega$$

$$\boxed{E = \hbar\omega}$$

c) ***Cuadrigradiente y cuadridivergencia:***

En nuestro nuevo formalismo, podemos introducir nuevos operadores que sustituyen al gradiente y la divergencia de un vector. Recordemos brevemente un par de cosas sobre estos operadores en el espacio habitual, para asegurar su buen uso en el de Minkovskii:

Al operar con el gradiente sobre un objeto, **subimos su orden tensorial en uno**, mientras que al operar con la divergencia, hacemos lo contrario: **bajamos su orden tensorial en uno**. Por muy bien que quede esta frase, tiene una explicación sencilla. Recordemos que al hacer el gradiente de un escalar (una función, por ejemplo), el resultado era un vector (hemos pasado de tensor de orden cero (cero índices): el escalar, a un tensor de orden uno (un índice, del vector). En cambio, al aplicar el gradiente a un vector (tensor de orden uno), obteníamos una matriz (tensor de orden dos). Veámoslo en un sencillo ejemplo:

Ejemplo 7.1. *Calcular el gradiente de la función $f(x,y,z) = xy^2z + x$*

i) El gradiente de una función es el vector $\vec{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \equiv (\partial_x, \partial_y, \partial_z)$ donde hemos redefinido así las derivadas parciales por sencillez, y será el convenio que usaremos en el resto del texto, de modo que $\partial_x \equiv \frac{\partial}{\partial x}$. Al aplicarle el gradiente a la función obtenemos:

$$\begin{aligned}\vec{\nabla} f &= (\partial_x, \partial_y, \partial_z)(xy^2z + x) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) \\ &= \left(\frac{\partial(xy^2z + x)}{\partial x}, \frac{\partial(xy^2z + x)}{\partial y}, \frac{\partial(xy^2z + x)}{\partial z}\right) \\ &= (y^2z + 1, 2xyz, xy^2)\end{aligned}$$

Es decir, al hacer el gradiente de un escalar (orden cero) obtenemos un vector (orden uno). Hagamos ahora el gradiente de este vector que acabamos de obtener, al que llamaremos el vector $\vec{A}(x,y,z) = (y^2z + 1, 2xyz, xy^2)$. Una forma sencilla de entender esto, para hacer las operaciones correctamente, es colocar cada vector en horizontal o vertical según sea el caso. Por ejemplo, para el vector gradiente, deberíamos colocarlo en vertical y el vector en horizontal (más tarde, identificaremos el vector en horizontal con el covariante y el mismo vector puesto en vertical como contravariante, que es una forma sencilla de visualizar todo esto):

$$\vec{\nabla}\vec{A} = \begin{pmatrix} \partial_x \\ \partial_y \\ \partial_z \end{pmatrix} (A_x, A_y, A_z) = \begin{pmatrix} \partial_x A_x & \partial_x A_y & \partial_x A_z \\ \partial_y A_x & \partial_y A_y & \partial_y A_z \\ \partial_z A_x & \partial_z A_y & \partial_z A_z \end{pmatrix}$$

Donde, colocados así los vectores, no hemos hecho más que recurrir a la multiplicación de matrices del bachillerato (matriz de 3x1 por matriz de 1x3 da como resultado una matriz de 3x3)

En nuestro ejemplo, para completarlo, hacemos las cuentas:

$$\vec{\nabla}\vec{A} = \begin{pmatrix} 0 & 2yz & y^2 \\ 2yz & 2xz & 2xy \\ y^2 & 2xy & 0 \end{pmatrix}$$

Donde se ve que el gradiente lo que ha hecho es pasar de un tensor de orden uno (un vector: un índice) a un tensor de orden dos (una matriz: dos índices: fila y columna).

Podríamos ahora hacer el gradiente de esta matriz, obteniendo un tensor de orden 3, pero tendría un dibujo más complicado (por ejemplo, podríamos representarlo como un cubo, donde los tres índices marcan la anchura, altura (hasta aquí igual que la matriz), y el tercer índice la profundidad). Es decir, podríamos ponerlo como una matriz detrás de otra, pero esto se ve peor y es preferible no extender demasiado esta analogía, ya que el de orden 4 ya sí que no tiene ninguna analogía con matrices, y menos los de orden superior.

ii) *Vamos con la divergencia. Supongamos que empezamos con una matriz (tensor de orden dos, dos índices), y le aplicamos la divergencia. Esto significa, según lo que hemos dicho antes, que el resultado tendrá un orden tensorial menos (daría un vector). Si de nuevo colocamos los vectores y matrices correctamente (si con el gradiente colocamos en vertical al operador $\vec{\nabla}$, ahora lo colocaremos en horizontal), tenemos (para la matriz de antes, por ejemplo, a la que llamaremos $\mathcal{M} \equiv \mathcal{M}^i{}_j$). De nuevo, para las operaciones, recurriremos tal y como lo hemos puesto a segundo de bachiller:*

$$
\begin{aligned}
div\mathcal{M} = \vec{\nabla} \cdot \mathcal{M} = (\partial_x, \partial_y, \partial_z)\begin{pmatrix} 0 & 2yz & y^2 \\ 2yz & 2xz & 2xy \\ y^2 & 2xy & 0 \end{pmatrix} \\
= \Big(\partial_x(0) + \partial_y(2yz) + \partial_z(y^2), \partial_x(2xz) + \partial_y(2xz) + \partial_z(2xy), \partial_x(y^2) \\
+ \partial_y(2xy) + \partial_z(0)\Big) = (4z, 2z + 2x, 2x) = 2(2z\,, z + x\,, x)
\end{aligned}
$$

Obtenemos, como dijimos, un vector (tensor de orden uno). SI ahora queremos volver a realizar la divergencia, llamando de nuevo $\vec{A}$ al vector que acabamos de obtener:

$$
\vec{\nabla} \cdot \vec{A} = (\partial_x, \partial_y, \partial_z) \cdot \begin{pmatrix} 4z \\ 2z + 2x \\ 2x \end{pmatrix} = \partial_x(4z) + \partial_y(2z + 2x) + \partial_z(2x) = 0
$$

Obtenemos un escalar (tensor de orden uno (ha dado cero como podía dar cualquier cosa (¿o no?)). Esta notación de escribir los vectores en horizontal o vertical según el caso, así como el ver los tensores de orden cero, uno y dos como escalares, vectores y matrices, es sólo una representación intuitiva para trabajar más cómodamente con todo esto, pero recordemos que, por ejemplo, un tensor es mucho más que una matriz (debe cumplir unas determinadas leyes de transformación, etc.)

En nuestro nuevo formalismo tensorial, lo que acabamos de hacer es lo siguiente: llamaremos al **cuadrigradiente y cuadridivergencia $\boldsymbol{\partial_\mu}$ y $\boldsymbol{\partial^\mu}$** respectivamente. Por supuesto, al operar en relatividad especial, ahora el índice μ recorre desde el cero hasta el 3, no como antes, que sólo iba de 1 a 3. Así pues, aplicado a un vector covariante o contravariante, nos queda en ambos casos divergencias o gradientes según sea el caso. Y recordemos que sólo suman índices repetidos arriba y abajo.

Además, no es difícil de ver que ahora en esta notación, tenemos que:

$$\boxed{\partial_\mu \equiv \frac{\partial}{\partial x^\mu} \quad \partial^\mu \equiv \frac{\partial}{\partial x_\mu}}$$

Veamos unos ejemplos sencillos:

Ejemplo 7.2. *Desarrollar los siguientes productos:*

a) Veamos la acción de ∂_μ sobre x^μ, con índices repetidos (podemos hacerlo, ya que está uno arriba y otro abajo)

$$\partial_\mu x^\mu = \sum_{\mu=0}^{3} \partial_\mu x^\mu = \frac{\partial x^0}{\partial x^0} + \frac{\partial x^1}{\partial x^1} + \frac{\partial x^2}{\partial x^2} + \frac{\partial x^3}{\partial x^3} = 4$$

b) Si en vez de sobre x^μ actuamos sobre otro cuadrivector, como por ejemplo p^μ:

$$\partial_\mu p^\mu = \sum_{\mu=0}^{3} \partial_\mu p^\mu = \frac{\partial p^0}{\partial x^0} + \frac{\partial p^1}{\partial x^1} + \frac{\partial p^2}{\partial x^2} + \frac{\partial p^3}{\partial x^3} = \frac{\partial\left(\frac{E}{c}\right)}{\partial(ct)} + \vec{\nabla}\cdot\vec{p} = \frac{1}{c^2}\frac{\partial E}{\partial t} + \vec{\nabla}\cdot\vec{p}$$

Donde vemos que vamos a usar $\vec{\nabla}$ (operador nabla espacial), para las coordenadas espaciales, y $\vec{p}$ como las componentes espaciales del cuadrivector cuadrimomento.

c) Actuemos de nuevo con ∂_μ, pero esta vez sobre p_μ. Podemos hacer esta operación sobre índices iguales (aunque habría que usar entonces la métrica para subir índices), o con índices distintos. De esta forma quedaría:

$$\partial_\mu p_\nu = \begin{pmatrix} \frac{\partial p_0}{\partial x^0} & \frac{\partial p_1}{\partial x^0} & \frac{\partial p_2}{\partial x^0} & \frac{\partial p_3}{\partial x^0} \\ \frac{\partial p_0}{\partial x^1} & \frac{\partial p_1}{\partial x^1} & \frac{\partial p_2}{\partial x^1} & \frac{\partial p_3}{\partial x^1} \\ \frac{\partial p_0}{\partial x^2} & \frac{\partial p_1}{\partial x^2} & \frac{\partial p_2}{\partial x^2} & \frac{\partial p_3}{\partial x^2} \\ \frac{\partial p_0}{\partial x^3} & \frac{\partial p_1}{\partial x^3} & \frac{\partial p_2}{\partial x^3} & \frac{\partial p_3}{\partial x^3} \end{pmatrix} = \mathcal{M}_{\mu\nu}$$

Es decir, el resultado es un tensor de orden 2, dado que aparece un elemento para cada índice μ, y para cada uno de ellos, otro para el índice ν, en total 16 elementos posibles.

d) *Cuadricorriente.*

Escribiremos a continuación el vector cuadricorriente, j^μ, que es la generalización de la corriente $\vec{j}$ del electromagnetismo. Esta cuadricorriente se expresa como

$$\boxed{j^\mu = (c\rho, \vec{j})}$$

Donde ρ es la densidad de la distribución de carga, y $\vec{j}$ el vector densidad de corriente ya introducido en los cursos de electromagnetismo. Por recordar, la densidad de corriente $\vec{j}$ es la intensidad de la corriente por unidad de área (a la que llamaremos en forma diferencial en función de la longitud *x* y la anchura *y*: $dA = dxdy$). Y esta intensidad de corriente será la carga por unidad de tiempo. Por su parte, la densidad de carga de la distribución, ρ, es la carga (que en forma diferencial será *dq*), por unidad de volumen ($dV = dxdydz$). Así pues, podemos hacer lo siguiente:

$$\begin{cases} \vec{j} = \dfrac{dI}{dA} = \dfrac{\frac{dq}{dt}}{d\vec{x}d\vec{y}} = \dfrac{dq}{d\vec{x}d\vec{y}dt} \\ \\ \rho = \dfrac{dq}{dV} = \dfrac{dq}{d\vec{x}d\vec{y}d\vec{z}} \end{cases}$$

Podemos introducir el elemento de carga *dq*, despejado de la expresión de la densidad de carga ρ, en la expresión de la corriente, de modo que obtenemos:

$$\left.\begin{array}{r} \vec{j} = \dfrac{dq}{d\vec{x}dydt} \\ dq = \rho d\vec{x}d\vec{y}d\vec{z} \end{array}\right\} \vec{j} = \frac{\rho d\vec{x}d\vec{y}d\vec{z}}{d\vec{x}d\vec{y}dt} = \rho\frac{d\vec{z}}{dt} = \rho\vec{V}$$

Esta es una de tantas deducciones para obtener lo que acabamos de obtener (todas son igual de válidas). La cuestión es que, con esto, podemos escribir lo siguiente, mucho más compacto:

$$\boxed{j^\mu = \rho\left(c, \vec{V}\right)}$$

e) *Cuadripotencial*

Por último, aunque podemos introducir muchos más cuadrivectores, introduciremos una generalización del potencial escalar y potencial vector, juntándolos en uno solo, de modo que lo que nos queda es (completamos con “c” para dimensionalizar):

$$\boxed{A^\mu = \left(c\phi, \vec{A}\right)}$$

Donde ϕ es el potencial escalar (proveniente del campo eléctrico), y $\vec{A}$ el potencial vector, proveniente del campo magnético. Recordemos que estos potenciales se definen de forma que:

$$\phi = -div\vec{E}$$
$$\vec{B} = rot\vec{A}$$

Por supuesto, hemos escrito todos los cuadrivectores contravariantes. Para obtener los covariantes, no tenemos más que multiplicar por el tensor métrico, que nos bajará el índice abajo, e introducirá un menos en las componentes espaciales.

8. Ecuaciones covariantes

Una gran parte de la ventaja de todo este formalismo que acabamos de introducir es que ahora tenemos una base matemática para expresar ciertas ecuaciones de una forma mucho más compacta. Pero la mayor ventaja es que toda ecuación que quede expresada como una igualdad de escalares, es **invariante Lorentz** (por ejemplo, la ecuación de ondas ya vimos que debía ser invariante Lorentz: podremos ponerla de esa forma). Veamos algunas de estas ecuaciones:

a) Ecuación de continuidad:

$$\boxed{\partial_\mu j^\mu = 0} \Leftrightarrow \frac{\partial \rho}{\partial t} = \vec{\nabla}\,\vec{\mathrm{A}}$$

b) D'Alambertiano:

$$\boxed{\Box = \partial_\mu \partial^\mu} \Leftrightarrow \Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \vec{\nabla}^2$$

c) Tensor de Maxwell:

$$\boxed{F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu}$$

d) Con el tensor de Maxwell, podemos escribir la ***ecuación de movimiento electromagnética:***

$$\boxed{m\frac{du^\mu}{d\tau} = eF^{\mu\nu}u_\nu} \Leftrightarrow \vec{F} = m\vec{a} = q\vec{E} + q\left(\vec{V} \times \vec{B}\right)$$

e) Tensor de energía momento:

$$\boxed{T^{\mu\nu} = -\frac{1}{\mu_0}\frac{\partial A^\lambda}{\partial x_\mu}F^\nu{}_\lambda + \frac{1}{4\mu_0}\eta^{\mu\nu}F_{\varepsilon\rho}F^{\varepsilon\rho}}$$

ó

$$\boxed{T^{\mu\nu} = \frac{1}{\mu_0}\left(F^{\mu\gamma}F_\gamma{}^\nu + \frac{1}{4}\eta^{\mu\nu}F_{\varepsilon\rho}F^{\varepsilon\rho}\right)}$$

Donde μ_0 es la constante de permitividad magnética

Todas estas ecuaciones, llamadas ecuaciones covariantes (o contravariantes), cumplen varias funciones: por una parte, permiten escribir las ecuaciones de forma más compacta y elegante. Pero por otra, al estar escritas así, cumplen que son invariantes bajo transformaciones de Lorentz, es decir, si un observador posee una de estas ecuaciones en un sistema de referencia, otro observador ajeno medirá la misma ecuación, sin necesidad de corrección alguna. Por supuesto, están aquí escritas simplemente para observarlas, no es necesario fijarse mucho en ellas (como mucho, uno puede quedarse en la cabeza con las dos primeras, que quizá tengan algo de utilidad en estos momentos). Además de estas, pueden escribirse otras muchas ecuaciones, pero con las anteriormente expuestas es suficiente por ahora para ilustrar la potencia de nuestra nueva notación.

9. Dinámica relativista

Ya tenemos todos los cimientos para crear la teoría, pero aún no hemos visto como se mueven las partículas con ella, cuáles son las ecuaciones de movimiento, etc. Vamos a intentar comprender qué queremos hacer en este apartado con un par de analogías sencillas:

a) ***Dinámica de Newton***: para llevarla a cabo, era necesario introducir primeramente los vectores, para luego con ellos crear los de aceleración, fuerza, momento, etc., y finalmente relacionarlos todos mediante la ecuación de Newton, $\boxed{\sum \vec{F} = m\vec{a}}$, de la que se podían derivar muchas otras, como las de conservación del momento lineal, la energía, etc., y sobre todo, podíamos sacar las ecuaciones de movimiento.

b) ***Dinámica Lagrangiana***: para poder usarla, en su momento necesitamos introducir el concepto de "*Lagrangiano*", y el de *"coordenadas generalizadas"*, como se buscaban, y como se construían. Por el camino, tuvimos que servirnos de la ley de Hamilton de mínima acción (S), que simplemente decía que $\boxed{\delta S = 0}$, donde la acción era el producto del lagrangiano por el tiempo, y como resultado obteníamos las ecuaciones de Euler Lagrange, que eran: $\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) = \frac{\partial \mathcal{L}}{\partial q}$. O, si tenemos varios grados de libertad, por ejemplo n grados de libertad (desde *i=1…n)*:

$$\boxed{\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial q_i}}$$

Donde q_i es cada uno de ellos (para *i=1…n*) y $\dot{q}$ denota derivada total en el tiempo.

Además, recordemos que el lagrangiano debía estar expresado en términos de las coordenadas generalizadas y de sus derivadas temporales, y en todo caso también por el tiempo, de la forma:

$$\boxed{\mathcal{L} \equiv \mathcal{L}(q, \dot{q}, t) = T(q, \dot{q}, t) - V(q, \dot{q}, t)}$$

Donde $\mathcal{L}$ era el lagrangiano del sistema, y q cada una de las coordenadas generalizadas, T la energía cinética y V la energía potencial. De estas últimas expresiónes, podíamos obtener ecuaciones de movimiento, cantidades conservadas, etc.

c) ***Mecánica Hamiltoniana***: para ella era necesario la introducción del *"Hamiltoniano"*, que debía expresarse en función de las coordenadas generalizadas y sus momentos conjugados, y era:

$$\boxed{\mathcal{H}(q_i, p_i, t) = \sum_{i=1}^{n} \dot{q}_i p_i - \mathcal{L}(q_i, p_i, t)}$$

Donde el momento quedaba definido como:

$$\boxed{p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}}$$

Y las ecuaciones de Hamilton, que eran:

$$\boxed{\frac{\partial \mathcal{H}}{\partial q_i} = -\dot{p}_i \qquad \frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i \qquad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}}$$

De esta forma podíamos estudiar cualquier sistema desde cualquiera de los tres formalismos, obteniendo finalmente los mismos resultados (las mismas ecuaciones de movimiento). Hagamos un ejemplo para clarificar las cosas, es decir, para repasar prácticamente la mitad de la mecánica clásica: una partícula que cae bajo la acción de la gravedad:

Ejemplo 9.1. *Estudiar la caída vertical de un objeto sometido a la acción de la gravedad en los tres formalismos anteriores:*

a) Por mecánica Newtoniana

Tomemos la ecuación de Newton del movimiento. En este caso, tan sólo tenemos una partícula, así que tan sólo hay que utilizarla una vez. Además la única fuerza del sistema es la fuerza de la gravedad, que es:

$$\vec{F} = \vec{F}_g = mg(-\vec{j}) = -mg(\vec{j})$$

en un sistema de coordenadas con el eje de las "y" positivas apuntando hacia arriba (en contra de la gravedad). Por otro lado, la aceleración será la segunda derivada del vector posición con respecto al tiempo, y dado que este vector posición será el vector $\vec{r}(t) = r(t)\vec{j}$, la aceleración será

$$\vec{a}(t) = \frac{d^2\vec{r}(t)}{dt^2} = \frac{d^2r(t)}{dt^2}\vec{j}$$

De esta forma, la ecuación de Newton queda:

$$\sum \vec{F} = m\vec{a} \quad \Leftrightarrow \quad \boxed{-mg = \frac{d^2r(t)}{dt^2}}$$

Una vez sustituidos los datos y eliminados los vectores (al estar en la misma dirección, estos podemos quitarlos). La resolución de la ecuación diferencial anterior (por el método que se quiera), arroja como resultado la ecuación de movimiento, que será:

$$r(t) = a + bt - \frac{1}{2}gt^2$$

Que podemos completar con las condiciones iniciales, de forma que obtenemos la clásica ecuación del movimiento en caída libre, dada por:

$$\boxed{y(t) = y_0 + v_0t - \frac{1}{2}gt^2}$$

Donde hemos llamado $\vec{r}(t) = (0, r(t), 0) = (0, y(t), 0)$ e igual para la velocidad y coordenadas iniciales.

b) Mecánica lagrangiana

Este es un sistema con un único grado de libertad (como puede deducirse fácilmente). Tomaremos este grado de libertad como "y", la altura, de forma que para construir el lagrangiano necesitamos T y V:

$$\left.\begin{array}{r} T = \frac{1}{2}m\dot{y}^2 \\ V = mgy \end{array}\right\} \Longrightarrow \mathcal{L}(q,\dot{q},t) = \mathcal{L}(y,\dot{y},t) = \frac{1}{2}m\dot{y}^2 - mgy$$

Y las ecuaciones e Euler Lagrange arrojarían lo siguiente:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{y}}\right) = \frac{\partial \mathcal{L}}{\partial y} \Longrightarrow m\ddot{y} = -mg \Longrightarrow \boxed{\ddot{y} = -g} \Longrightarrow \boxed{y(t) = y_0 + v_0 t - \frac{1}{2}gt^2}$$

que es la misma ecuación anterior, que resolveríamos de la misma manera con las mismas condiciones iniciales (obtenemos el mismo resultado: la partícula no se moverá diferente por usar un formalismo u otro).

c) Mecánica Hamiltoniana

Para llevar a cabo este formalismo, debemos comenzar por construir el Hamiltoniano. Para ello, haremos lo siguiente:

$$\mathcal{H}(q_i,p_i,t) = \sum_{i=1}^{n} \dot{q}_i\, p_i - \mathcal{L}(q_i,p_i,t)$$

Con $\boldsymbol{\mathcal{L}(q_i,p_i,t) = \frac{p^2}{2m} - mgy}$ *y* $\boldsymbol{p_i \equiv p = \frac{\partial \mathcal{L}}{\partial \dot{y}} = m\dot{y}}$ *despejando con la relación anterior del lagrangiano obtenido en el apartado b. Escribiendo* $\boldsymbol{\dot{q}_i p_i = \dot{y}p = \frac{p^2}{m}}$, *usando de nuevo la definición de p, llegamos a que el hamiltoniano será:*

$$\mathcal{H}(y,p_i,t) = p - \mathcal{L}(q_i,p_i,t) = \frac{p^2}{m} - \left(\frac{p^2}{2m} - mgy\right) = \frac{p^2}{2m} + mgy$$

Ya tenemos nuestro Hamiltoniano. Vayamos a las ecuaciones de Hamilton. Vemos que lo que nos ha quedado no es ni más ni menos que la energía de la partícula en caída, así que podríamos identificar $\mathcal{H} = E$. *Las ecuaciones de Hamilton quedan como:*

$$\frac{\partial \mathcal{H}}{\partial y} = -\dot{p} \qquad \frac{\partial \mathcal{H}}{\partial p} = \dot{y} \qquad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial \mathcal{L}}{\partial t}$$

La última de ellas se cumple automáticamente, dado que ninguno de los dos depende del tiempo. En general, $\mathcal{H} = E$ *precisamente cuando ocurra que* $\boldsymbol{\frac{\partial \mathcal{L}}{\partial t} = 0}$, *es decir,* $\boldsymbol{\mathcal{L} = (cte)_t}$. *Las otras:*

$$\begin{cases} \dfrac{\partial \mathcal{H}}{\partial y} = -\dot{p} & \Longrightarrow \quad mg = -\dot{p} \\ \dfrac{\partial \mathcal{H}}{\partial p} = \dot{y} & \Longrightarrow \quad \dfrac{p}{m} = \dot{y} \end{cases}$$

De donde no es muy difícil salir. Estas son dos ecuaciones acopladas de sencilla interpretación: la segunda nos dice precisamente lo que nos acaba de quedar para el momento conjugado p, mientras que de la primera podemos sacar las ecuaciones de movimiento, sin más que sustituir p por su correspondencia con $\dot{y}$*, obteniendo:*

$$mg = -\dot{p} \;\Longrightarrow\; mg = -m\ddot{y} \;\Longrightarrow\; \boxed{\ddot{y} = -g} \Longrightarrow \boxed{y(t) = y_0 + v_0 t - \frac{1}{2}gt^2}$$

Dado que si $p = m\dot{y}$*, entonces* $\dot{p} = m\ddot{y}$*. Hemos obtenido la misma ecuación que en los casos anteriores, así que su solución será exactamente la misma (imponiendo finalmente las condiciones iniciales). Aunque en este caso parece absurdo usar Lagrange o Hamilton para estudiar esto, en otro tipo de problemas (como problemas con muelles o electromagnéticos), las formulaciones de Lagrange y Hamilton son mucho más potentes que la de Newton. En particular, trataremos el caso de la relatividad desde estos dos últimos puntos de vista.*

Ejercicio 9.1. *Estudiar con los tres formalismos el movimiento de una partícula libre (sin fuerzas ni potenciales), y el de una partícula enganchada a un muelle* $\left(F = -k\Delta x; E_p = \frac{1}{2}k\Delta x^2\right)$.

Solución: Partícula libre: $\boxed{\ddot{r} = 0} \Longrightarrow \boxed{r(t) = at + b = V_0 t + r_0}$ *(movimiento rectilíneo); Partícula en un muelle:* $\boxed{\ddot{x} = -\omega^2 x} \Longrightarrow \boxed{x(t) = A\cos(\omega t + \varphi)}$*; con* $\omega \equiv \sqrt{\frac{k}{m}}$

Pues bien, una vez expuestos los conceptos de una cosa y la otra, intentemos ahora desarrollar una teoría para la dinámica en relatividad, basándonos fundamentalmente en la mecánica hamiltoniana. Para ello, comenzaremos por recurrir al principio de mínima acción de Hamilton, que nos dice que la variación de la acción es nula:

$$\delta S = 0 \qquad S = \int \mathcal{L} dt$$

Donde S (acción) es simplemente el producto del lagrangiano por el tiempo. Pero no tenemos el lagrangiano de una partícula relativista (lo que sí podríamos es demostrar este principio para el ejemplo anterior, cuyo lagrangiano conocemos). Así pues, vamos a intuir: sabemos que al hacer la variación de la acción, esto debe darnos cero, con lo cual empecemos por el caso más sencillo: el de una partícula libre (no sujeta a ninguna interacción tipo gravedad), dado que sabemos qué tiene que darnos el lagrangiano cuando la velocidad de la partícula es pequeña comparada con la de la luz (lagrangiano de una partícula libre clásica). Así pues, presuponemos que el lagrangiano debe ser una cierta constante, que llamaremos α, y desarrollaremos a partir de aquí las hipótesis anteriores:

$$S = \int \mathcal{L} dt = \int \alpha dt = \alpha t$$

Si tomamos como τ el tiempo propio de una partícula, y t el tiempo visto desde otro sistema de referencia que se mueve a velocidad *v* con respecto a este, la ecuación anterior, es decir, la acción vista desde el punto de vista de la partícula será:

$$S = \int \alpha d\tau = \int \alpha \frac{dt}{\gamma} = \int \alpha \sqrt{1 - \frac{V^2}{c^2}} dt$$

Recordando la definición de tiempo propio. Si ahora identificamos el lagrangiano en esta última expresión vemos que, sin mucha dificultad, el lagrangiano de una partícula libre relativista sería:

$$\boxed{\mathcal{L} = \alpha \sqrt{1 - \frac{V^2}{c^2}}}$$

Pero sabemos qué tiene que dar eso en mecánica clásica no relativista (a velocidades bajas): dicho lagrangiano clásico vendría dado por:

$$\mathcal{L}_{cl} = T - V = T = \frac{p^2}{2m}$$

Y ambas expresiones deben coincidir en el límite no relativista. Así pues, desarrollemos el lagrangiano relativista y comparémoslo con el no relativista. Llamemos $\beta \equiv \frac{V}{c}$ por sencillez, y desarrollemos en torno al punto $\beta = 0$:

$$\mathcal{L}_{rel}|_{\beta \to 0} = \mathcal{L}_{rel}(0) + \left.\frac{\partial \mathcal{L}_{rel}(\beta)}{\partial \beta}\right|_{\beta \to 0} . \beta + \left.\frac{\partial^2 \mathcal{L}_{rel}(\beta)}{\partial \beta^2}\right|_{\beta \to 0} . \beta^2 + \left.\frac{\partial^3 \mathcal{L}_{rel}(\beta)}{\partial \beta^3}\right|_{\beta \to 0} . \beta^3 + \cdots + \mathcal{O}(\beta^4)$$

Calculando estas expresiones, y despreciando los términos superiores al cuarto orden, nos queda:

$$\mathcal{L}_{rel}|_{\beta \to 0} = \alpha - \frac{1}{2} \frac{\alpha}{c^2} V^2$$

Si comparamos ahora con el lagrangiano no relativista, vemos que el segundo término de la ecuación está claro que es la energía cinética de la partícula, de donde se puede sacar que:

$$-\frac{1}{2} \frac{\alpha}{c^2} V^2 = \frac{1}{2} m V^2 \quad \Longrightarrow \alpha = -mc^2$$

Y queda definitivamente:

$$\boxed{\mathcal{L}_{rel} = -mc^2 \sqrt{1 - \frac{V^2}{c^2}}}$$

<u>Ejercicio 9.2.</u> *Desarrollar el lagrangiano relativista en torno al punto $\beta = 0$, y obtener las expresiones anteriores (ejercicio consistente en hacer un simple desarrollo de Taylor).*

El problema es que el lagrangiano no tiene una realidad física concreta. Sin embargo, el Hamiltoniano sí lo tiene, ya que está estrechamente relacionado con la energía. Así pues, pasemos a mecánica Hamiltoniana a ver qué podemos sacar de todo esto:

$$\mathcal{H}(q_i, p_i, t) = \sum_{i=1}^{n} \dot{q}_i p_i - \mathcal{L}(q_i, p_i, t) = pV - \mathcal{L}(q_i, p_i, t) = pV + mc^2 \sqrt{1 - \frac{V^2}{c^2}} = \sqrt{m^2c^4 + c^2p^2}$$

Donde $p = \frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial \mathcal{L}}{\partial V} = \gamma mV$

$$\boxed{\mathcal{H} = \sqrt{m^2c^4 + c^2p^2} = \gamma mc^2}$$

__Ejercicio 9.3__. *Demostrar que $p = \gamma mV$, derivando el lagrangiano relativista, y demostrar que el Hamiltoniano es $\mathcal{H} = \sqrt{m^2c^4 + c^2p^2}$, despejando la velocidad en función del momento. Demostrar que, sustituyendo p, $\mathcal{H} = \gamma mc^2$*

Si ahora vemos qué le ocurre al Hamiltoniano relativista en comparación con el Hamiltoniano no relativista (clásico), en principio deberíamos obtener lo mismo: la energía. Veamos:

$$\mathcal{H}_{clas} = T = \frac{1}{2}mV^2 \iff \mathcal{H}_{rel}|_{\beta \to 0} = mc^2 + \frac{1}{2}mV^2 + \cdots$$

__Ejercicio 9.4__. *Obtener esta última expresión, sustituyendo $p = \gamma mV$ en el Hamiltoniano, o usando $\mathcal{H} = \gamma mc^2$, y desarrollando en torno a $\beta = 0$*

Aquí parece haber una inconsistencia, o al menos así se pensó en el momento. Parece que hay alguna energía nueva, cuyo valor es mc^2, que no es atribuible a nada, puesto que estamos suponiendo que no hay ninguna interacción sobre la partícula. Finalmente, la explicación a este enigma fue que la partícula, por el hecho de tener masa, tiene energía, o lo que es lo mismo, **la masa y la energía están relacionadas y se pueden considerar la misma cosa**. Por ejemplo, un objeto parado (por supuesto, con respecto a otro), con velocidad cero, tiene energía, y esta energía vale:

$$\boxed{E_0 = mc^2}$$

La llamada **energía en reposo**. El motivo por el cual esto no se podía ver antes era porque estábamos siempre haciendo un intervalo de energías, como cuando medimos la energía gravitatoria (energía entre un punto y otro). Lo que realmente hacemos es:

$$\Delta E_g = E_{gB} - E_{gA} = mc^2 + mgh - mc^2 - mgh = mg\Delta h$$

Lo que no implica que la energía en reposo exista, y sea una realidad física comprobada experimentalmente. Por ejemplo, todos los experimentos actuales de producción de nuevas partículas, así como muchas reacciones en principio inexplicables, o los reactores de fusión (y por supuesto, las bombas atómicas), llevan este principio detrás. Es decir, esta sencilla ecuación posee unas grandes implicaciones físicas y sociales.