1. El espacio-tiempo de Minkowski

Figura 1: Hermann Minkowski (1864-1909).

Una consecuencia inmediata de la transformación de Lorentz es que, mientras que distintos observadores inerciales adjudican distintas posiciones y tiempos a un mismo evento, el valor del *intervalo*

$$s^2 = -c^2t^2 + x^2 + y^2 + z^2 \tag{1}$$

es el mismo para todos. La gran innovación de Minkowski fue pensar a las coordenadas espaciales y a ct (la c se agrega para que todas las coordenadas tengan las mismas dimensiones) como las coordenadas x^{μ}[1] de un nuevo espacio (el *espacio-tiempo*) donde s^2 define una (pseudo) distancia[2]. Visto desde este punto de vista la ecuación [?] no es otra cosa que la manera de escribir el Teorema de Pitágoras en este espacio.

El intervalo se puede escribir como una forma cuadrática

$$s^2 = \eta_{\mu\nu}x^{\mu}x^{\nu} \tag{2}$$

donde $x^0 = ct$, $x^i = (x, y, z)$ para $1 \leq i \leq 3$. Estamos usando la convención de Einstein y la posición de los índices importa. La matriz

$$\eta = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3}$$

define la (pseudo) *métrica de Minkowski*. La invariancia del intervalo implica la invariancia de la métrica, ya que si ante una transformación de Lorentz tenemos que

$$x^{\mu} \to x'^{\mu'} = \frac{\partial x'^{\mu'}}{\partial x^{\nu}}x^{\nu} \tag{4}$$

entonces

$$s'2 = \eta_{\mu'\nu'}x'^{\mu'}x'^{\nu'} = \eta_{\mu'\nu'}\frac{\partial x'^{\mu'}}{\partial x^{\rho}}\frac{\partial x'^{\nu'}}{\partial x^{\sigma}}x^{\rho}x^{\sigma} \tag{5}$$

pero también

$$s'^2 = s^2 = \eta_{\rho\sigma}x^{\rho}x^{\sigma} \tag{6}$$

[1] En este curso, índices griegos $\mu, \nu, \rho, \ldots$ van de 0 a 3 y $x^0 = ct$; índices latinos $i, j, k, \ldots$ van de 1 a 3 y denotan las coordenadas espaciales $x^i = (x, y, z)$.

[2] El *pseudo* es porque para un matemático es anatema que una distancia pueda ser negativa u (¡horror!) imaginaria. Nosotros no somos matemáticos...

y como esto vale para cualquier conjunto de coordenadas, debe ser

$$\eta_{\rho\sigma} = \eta_{\mu'\nu'} \frac{\partial x'^{\mu'}}{\partial x^{\rho}} \frac{\partial x'^{\nu'}}{\partial x^{\sigma}} \tag{7}$$

La invariancia del intervalo conduce a la cuestión de si hay otras magnitudes que tengan leyes de transformación definidas frente a transformaciones de Lorentz. Por ejemplo, si un termómetro mide la temperatura en un punto dado en un momento dado, todos los observadores inerciales van a estar de acuerdo en cuál es la lectura del termómetro, pero no van a estar de acuerdo en dónde estaba el termómetro ni cuándo se realizó la medición. Es decir, si para un observador S la lectura del termómetro en el punto x^i en el instante ct está dada por la función $T\left(x^{\mu}\right)$, otro observador S' dirá en cambio que las lecturas del termométro están dadas por la función

$$T'\left(x'^{\mu'}\right) = T\left(x^{\mu}\right) \tag{8}$$

donde x^{μ} son las coordenadas en que se convierte $x'^{\mu'}$ bajo la transformación de Lorentz que vincula a S' con S. Es importante no perder de vista que la función $T'\left(x'^{\mu'}\right)$ podría ser completamente distinta de la función $T\left(x'^{\mu'}\right)$ que resultaría de simplemente reemplazar unas coordenadas por otras en la función T.

Cuando una magnitud se transforma como en la ecuación 8 decimos que es una *magnitud escalar*. El intervalo es un caso muy particular de escalar porque no sólo es invariante (todos los observadores concuerdan en cuál es el intervalo entre dos eventos dados) sino que también es invariante *de forma*, es decir, todos los observadores escriben a la métrica de Minkowski como la misma matriz 3

El concepto que sigue en nivel de complejidad es el de *vector*. Definimos un vector contravariante como un conjunto de cuatro números V^{μ} que se transforman frente a una transformación de Lorentz como se transforman los X^{μ}. Como la transformación de Lorentz es lineal, vale que

$$x^{\mu} = \frac{\partial x^{\mu}}{\partial x'^{\mu'}} x'^{\mu'} \tag{9}$$

Entonces, un vector V^{μ} se transforma como

$$V^{\mu} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} V'^{\mu'} \tag{10}$$

Las componentes del gradiente de un escalar *NO* son un vector contravariante, ya que

$$\frac{\partial \phi}{\partial x^{\mu}} = \frac{\partial x'^{\mu'}}{\partial x^{\mu}} \frac{\partial \phi'}{\partial x'^{\mu'}} \tag{11}$$

Definimos un vector *covariante* como un conjunto de cuatro números V_{μ} que se transforman como el gradiente de un escalar.

La *contracción* de un vector covariante V_{μ} y uno contravariante W^{μ},

$$\phi = V_{\mu} W^{\mu} = V_0 W^0 + V_i W^i \tag{12}$$

define un escalar. Nótese el uso de la *convención de Einstein*: índices repetidos, uno arriba y otro abajo, están sumados sobre su rango, de 0 a 3 si son griegos o de 1 a 3 si son latinos. Además, en atención a lo que viene, hacemos notar que

Los signos en la ecuación 12 son correctos.

Más generalmente, se define un *tensor* m veces contravariante y n veces covariante como un objeto que se transforma como el producto de m vectores contravariantes y n vectores covariantes. La ecuación 7 implica que la métrica de Minkowski $\eta_{\mu\nu}$ es un tensor simétrico dos veces covariante. La contracción de la métrica de Minkowski con un vector contravariante da un vector covariante

$$V_{\mu} = \eta_{\mu\nu} V^{\nu} \tag{13}$$

Por lo tanto $V_0 = -V^0$, $V_i = V^i$. Decimos que la métrica "baja" el índice del vector. Viceversa, todo vector contravariante se puede pensar como proveniente de uno covariante

$$V^{\mu} = \left[\eta^{-1}\right]^{\mu\nu} V_{\nu} \tag{14}$$

O sea $V^0 = -V_0$, $V^i = V_i$. Como finalmente

$$\eta^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \eta \tag{15}$$

se omite el "-1", y se considera a $\eta^{\mu\nu}$ como la forma dos veces contravariante del tensor métrico. Si "subimos" un único índice obtenemos la forma mixta, que es el tensor de Kronecker

$$\eta^{\mu}_{\nu} = \eta^{\mu\rho}\eta_{\rho\nu} = \delta^{\mu}_{\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{16}$$

Nótese que la posición de los índices importa, ya que la componente temporal de un vector cambia de signo cuando el índice sube o baja. Entonces la contracción de dos vectores resulta en

$$V_\mu U^\mu = V_0 U^0 + V_i U^i = -V^0 U^0 + V^i U^i \tag{17}$$

La propiedad más importante de las leyes de transformación que hemos introducido es la *homogeneidad*: si un escalar, un vector o un tensor se anulan en un referencial, entonces se anulan en todos. Por el mismo motivo, si dos vectores son iguales en un referencial, entonces son iguales en todos, ya que la diferencia entre ellos también es un vector y se anula en el referencial original. Las leyes de la física, si están bien conformadas, no sólo deben ser dimensionalmente correctas, sino también deben asociar cantidades del mismo tipo tensorial. Este principio del buen arte se conoce como *Principio de Curie*. A las leyes fundamentales se les pide algo totavía más exigente, a saber, que no sólo sean invariantes, sino también invariantes *de forma*. Por ejemplo, las ecuaciones de Maxwell

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \mathbf{\Delta}\right] A^\mu = \frac{4\pi}{c} J^\mu \tag{18}$$

no sólo son invariantes porque implican la identidad entre dos magnitudes vectoriales, son invariantes de forma porque todos los observadores las escriben con el mismo valor de c.

2. Cinemática relativista

Como en la física clásica, la cinemática relativista se ocupa de la *descripción* del movimiento, dejando de lado sus causas. En la cinemática clásica, el concepto fundamental es el de *velocidad*, que es la tasa de cambio de la posición respecto del tiempo. Al pasar a la relatividad, es necesario preguntarse ¿el tiempo medido por quién? En general, se privilegia el tiempo medido por un reloj que estuviera acompañando el movimeinto del cuerpo en cuestión. Pero debemos tener cuidado, ya que un sistema de coordenadas adherido al cuerpo podría no ser un sistema inercial. Para salvar esta dificultad se introduce el concepto de *sistema en reposo instantáneo*, que es el sistema inercial en el que ocurre que el cuerpo, en ese momento preciso, está en reposo. No lo estará ni un instante antes ni un instante después, pero justo en ese momento, sistema y cuerpo se mueven con la misma velocidad (ver fig. 2).

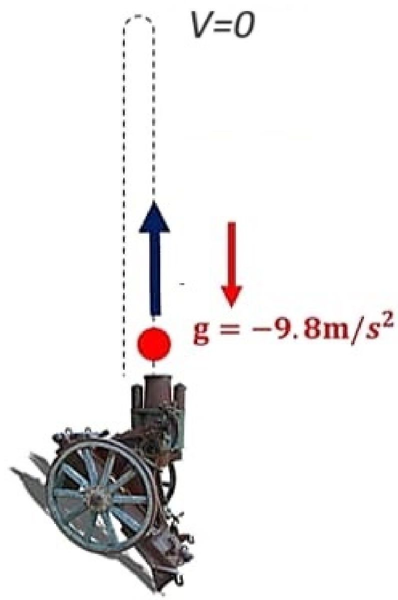

Figura 2: Una bala en tiro vertical está en reposo instantáneo en el punto más alto de su trayectoria.

Definimos la velocidad (o *tetra*-velocidad) como

$$u^\mu = \frac{dx^\mu}{cd\tau} \tag{19}$$

donde τ es el tiempo medido en el sistema en reposo instantáneo. Nótese que con esta definición la velocidad es definida adimensional. La invariancia del intervalo nos ahorra el problema de construir explícitamente la transformación de Lorentz al sistema en reposo instantáneo, ya que en este sistema, por definición, el cuerpo no se desplaza. Entonces

$$ds^2 = -c^2 dt^2 + \vec{dx}^2 = -c^2 d\tau^2 \tag{20}$$

o en otros términos

$$cd\tau = \sqrt{-ds^2} \tag{21}$$

Por supuesto la trayectoria de un cuerpo real está contenida en el cono de la luz, de manera que $-ds^2 \geq 0$. Con esta definición

$$u^2 = u_\mu u^\mu = \frac{1}{(-ds^2)} \eta_{\mu\nu} dx^\mu dx^\nu = -1 \tag{22}$$

de modo que la velocidad siempre es constante en módulo, pero puede cambiar su dirección. Por supuesto, en el referencial en reposo instantáneo la velocidad siempre es $u^0 = 1$, $u^i = 0$.

La velocidad ordinaria de la partícula es

$$v^i = \frac{dx^i}{dt} \tag{23}$$

usando que

$$cd\tau = cdt\sqrt{1 - \frac{v^2}{c^2}} \tag{24}$$

es fácil ver que

$$u^0 = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad u^i = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \frac{v^i}{c} \tag{25}$$

El concepto que sucede al de velocidad es la aceleración, es decir, la tasa de cambio de la velocidad. Como antes, referimos esta tasa al tiempo medido en el referencial en reposo instantáneo, de manera que

$$a^{\mu}=\frac{du^{\mu}}{cd\tau} \tag{26}$$

Derivando la relación 22 obtenemos

$$u_{\mu}a^{\mu}=\frac{1}{2}\frac{d}{cd\tau}u_{\mu}u^{\mu}=0 \tag{27}$$

De manera que la aceleración siempre es "perpendicular" a la velocidad, lo cual recuerda al movimiento circular uniforme de la cinemática clásica. Hay que tener cuidado porque el concepto de perpendicular en un espacio de Minkowski puede ser poco intuitivo, por ejemplo, un vector nulo es siempre perpendicular a sí mismo.

La relación 27 implica que la aceleración es siempre un vector espacial (porque la velocidad es siempre temporal, salvo que el cuerpo se esté moviendo con la velocidad de la luz) y que no puede haber un movimiento uniformemente acelerado, salvo el caso trivial en que $a^{\mu}=0$. Efectivamente, derivando 27 encontramos

$$0=\frac{d}{cd\tau}u_{\mu}a^{\mu}=u_{\mu}\frac{d}{cd\tau}a^{\mu}+a^{2} \tag{28}$$

Por lo tanto si $da^{\mu}/cd\tau=0$ debe ser $a^{2}=0$, que implica $a^{\mu}=0$ porque a es espacial.

Se define la *aceleración propia* como el valor instantáneo de la aceleración en el sistema en reposo. La aceleración propia sí puede ser constante; esto no contradice lo que acabamos de decir porque al comparar las aceleraciones propias en tiempos distintos estamos comparando aceleraciones medidas en sistemas distintos.

3. Mecánica relativista

En mecánica, las ecuaciones de movimiento de un punto material se deducen de la acción

$$S=\int dt\ L\left(q,\dot{q},t\right) \tag{29}$$

donde q y $\dot{q}$ son las posiciones y velocidades de la partícula a lo largo de una trayectoria parametrizada por el tiempo t. La trayectoria que efectivamente sigue la partícula es la que es un extremo de la acción, y satisface las ecuaciones de Euler-Lagrange

$$\frac{dp}{dt}=\frac{\partial L}{\partial q} \tag{30}$$

Donde p es el *impulso*

$$p=\frac{\partial L}{\partial \dot{q}} \tag{31}$$

En mecánica relativista asociamos a un punto material la acción

$$S=-mc\int d\lambda\ \sqrt{-\eta_{\mu\nu}\frac{dx^{\mu}}{d\lambda}\frac{dx^{\nu}}{d\lambda}} \tag{32}$$

donde λ es un parámetro (con unidades de longitud) a lo largo de la trayectoria. El factor mc se agrega para que la acción tenga unidades de energía por tiempo, como en mecánica ordinaria. La elección del parámetro es arbitraria: si reparametrizamos la trayectoria haciendo que $\lambda=\lambda\left(\lambda'\right)$, entonces

$$d\lambda'=\frac{d\lambda'}{d\lambda}d\lambda \tag{33}$$

pero también

$$\frac{dx^{\mu}}{d\lambda'}=\frac{d\lambda}{d\lambda'}\frac{dx^{\mu}}{d\lambda} \tag{34}$$

de manera que recuperamos el valor de la acción en la parametrización original.

En particular, podemos usar como parámetro el propio tiempo coordenado, es decir $\lambda=cdt$, en cuyo caso

$$S = -mc^2 \int dt \sqrt{1 - \frac{v^2}{c^2}} \tag{35}$$

donde v es la velocidad ordinaria, $v^i = dx^i/dt$. En el límite no relativista $v^2/c^2 \ll 1$

$$S \approx -mc^2 \int dt \left[1 - \frac{1}{2}\frac{v^2}{c^2} + \dots\right] \tag{36}$$

La constante no contribuye, y vemos que se recupera naturalmente la acción para una partícula libre no relativista (de paso, aprendemos que el parámetro m representa la masa de la partícula).

Volviendo a un parámetro general, definimos los impulsos

$$p_\mu = \frac{\partial L}{\partial x^\mu_{,\lambda}} \tag{37}$$

Haciendo la cuenta, encontramos

$$p_\mu = (-mc) \frac{1}{2\sqrt{-\eta_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda}}} (-2)\, \eta_{\mu\nu} \frac{dx^\nu}{d\lambda} \tag{38}$$

que se simplifica a

$$p_\mu = mcu_\mu \tag{39}$$

Como en este caso el lagrangiano no depende de las coordenadas, las ecuaciones de movimiento se reducen a la conservación del impulso

$$\frac{dp_\mu}{d\lambda} = 0 \tag{40}$$

Supongamos que en vez de un único móvil tuviéramos un gas de partículas llenando una región del espacio-tiempo, de manera que una partícula tiene impulso p^μ cuando atraviesa el evento x^μ. Entonces podemos escribir la ecuación de movimiento como

$$mc\frac{cd\tau}{d\lambda} u^\nu u^\mu_{,\nu} = 0 \tag{41}$$

donde $u^\mu_{,\nu} = \partial u^\mu / \partial x^\nu$. Como el primer factor no es nulo, debe ser

$$u^\nu u^\mu_{,\nu} = 0 \tag{42}$$

que es la *ecuación geodésica* y nos va a tener bastante ocupados el resto del curso.

Usando la parametrización 25 para la tetravelocidad encontramos que los impulsos

$$p^i = \frac{mv^i}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{43}$$

de modo que pueden considerarse como la generalización relativista del impulso "masa por velocidad" de la mecánica clásica. La componente temporal

$$p^0 = \frac{mc}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{44}$$

En el límite no relativista

$$p^0 = \frac{1}{c}\left[mc^2 + \frac{1}{2}mv^2 + \dots\right] \tag{45}$$

Como el segundo sumando es la energía cinética ordinaria, esto sugiere asociar $p^0 = E/c$, donde E es la energía de la partícula. Vemos que la energía no tiende a cero cuando la velocidad tiende a cero, sino que en ese límite obtenemos la *energía en reposo*

$$E = mc^2 \tag{46}$$

que es probablemente la fórmula más famosa de toda la física

Figura 3: La tripulación del portaaviones Enterprise, el primer portaaviones nuclear del mundo, escribiendo la fórmula de Einstein 43 sobre la cubierta.

4. Acerca de la literatura

El tema de esta clase no presenta muchas variaciones entre los libros de texto usuales, el Schutz entre ellos. El tema de la mecánica relativista es una buena excusa para volver al Landau, [1].

Referencias

[1] L. Landau y E. Lifshitz, *Teoría clásica de los campos*, Reverté (2008).