Relatividad General

Andrés Reyes
Semestre 2022-I

§1. Electrodinámica & Relatividad Especial.

Las ecuaciones de Maxwell

Comenzaremos nuestro curso con un breve repaso de las ecuaciones de Maxwell

$$\vec{\nabla}\cdot\vec{E} = \rho/\varepsilon_0 \quad \text{(ley de Gauss)}, \qquad \vec{\nabla}\cdot\vec{B} = 0 \quad \text{("no existencia de monopolos")}$$

$$\vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t} \quad \text{(ley de Faraday)}, \qquad \vec{\nabla}\times\vec{B} = \mu_0\varepsilon_0\frac{\partial\vec{E}}{\partial t} + \mu_0\vec{J}$$
(ley de Ampére - Maxwell).

Ecuación de onda $\longrightarrow$ A partir de las ecs. de Maxwell en el vacío $(\rho=0, \vec{J}=0)$

Tomemos la derivada con respecto a t en la ley de Ampére - Maxwell:

$$\vec{\nabla}\times\frac{\partial\vec{B}}{\partial t} = \mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2}$$

Por otro lado, usando la ley de Faraday, tenemos:

$$\vec{\nabla}\times\frac{\partial\vec{B}}{\partial t} = -\vec{\nabla}\times(\vec{\nabla}\times\vec{E}) \quad \Rightarrow \quad \vec{\nabla}\times(\vec{\nabla}\times\vec{E}) + \mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2} = 0.$$

Ahora, podemos hacer uso de la siguiente identidad,

$$\vec{\nabla}\times(\vec{\nabla}\times\vec{A}) = \vec{\nabla}(\vec{\nabla}\cdot\vec{A}) - \nabla^2\vec{A},$$

para llegar a

$$\vec{\nabla}(\underbrace{\vec{\nabla}\cdot\vec{E}}_{=0\ \text{(ley de Gauss)}}) - \nabla^2\vec{E} + \mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2} = 0 \rightsquigarrow \mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2} - \nabla^2\vec{E} = 0.$$

Si ahora repetimos el procedimiento, pero partiendo de la ley de Faraday, tenemos:

Tomando la derivada respecto a t en la ley de Faraday,

$$\vec{\nabla}\times\frac{\partial\vec{E}}{\partial t} = -\frac{\partial^2\vec{B}}{\partial t^2}$$

Usando la ley de Ampère-Maxwell $\longrightarrow$ $\vec{\nabla}\times\left(\frac{1}{\mu_0\varepsilon_0}\vec{\nabla}\times\vec{B}\right) = -\frac{\partial^2\vec{B}}{\partial t^2}$ $\longrightarrow$

$$\vec{\nabla}\times(\vec{\nabla}\times\vec{B}) = \vec{\nabla}(\underbrace{\vec{\nabla}\cdot\vec{B}}_{\substack{=0 \\ \text{(no monopolos)}}}) - \nabla^2\vec{B} = -\mu_0\varepsilon_0\frac{\partial^2\vec{B}}{\partial t^2}$$

En resumen, haciendo uso de las 4 ecs. de Maxwell (en el vacío), hemos obtenido las siguientes ecuaciones:

$$\mu_0\varepsilon_0\frac{\partial^2\vec{E}}{\partial t^2} - \nabla^2\vec{E} = 0$$

$$\mu_0\varepsilon_0\frac{\partial^2\vec{B}}{\partial t^2} - \nabla^2\vec{B} = 0$$

Estas ecs. nos dicen que, en el vacío, los campos $\vec{E}$ y $\vec{B}$ se propagan en forma de ondas que viajan a velocidad $v = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$.

Un hecho sorprendente es que esta velocidad resulta ser justo la velocidad de la luz. En efecto, el valor de las constantes ε_0 y μ_0 (SI) está dado por

$$\varepsilon_0 \approx 8.854187817\times10^{-12}\ C^2/N\cdot m^2$$

$$\mu_0 \approx 4\pi\times10^{-7}\ Wb/A\cdot m.$$

$\searrow$ $\frac{1}{\sqrt{\mu_0\varepsilon_0}} \approx 2.99792458\times10^{8}\ m/s.$, i.e. $\boxed{\frac{1}{\sqrt{\mu_0\varepsilon_0}} = c}$

Uno de los grandes problemas de la física a finales del siglo XIX consistió en entender las implicaciones de dicha relación.
Si la "permitividad" del vacío (ε_0) y la "permeabilidad" del vacío (μ_0) son constantes universales, cuyos valores reflejan propiedades de las interacciones electromagnéticas (interacciones entre cargas eléctricas, corrientes, imanes, etc.), ¿cómo explicar que den lugar a una <u>velocidad</u> que, como es obvio desde el punto de vista de la física newtoniana, es una cantidad que depende del marco de referencia en el qué esté siendo considerada? Como sabemos, esto da lugar a un conflicto irreconciliable entre la teoría electromagnética de Maxwell y la física newtoniana. Varias propuestas fueron formuladas (notablemente por Lorentz) para resolver las discrepancias, pero solo la de Einstein resultaría siendo correcta.
Para evitar entrar en un laberinto histórico relativo al desarrollo de estas ideas, tomaremos un "atajo" mencionando los siguientes puntos clave:

- La existencia del "éter", que era (supuestamente) aquel medio en el que se propagaba la luz, y con respecto al cual se llegó a pensar que la velocidad de propagación de esta era "c", terminó siendo refutada. Aquí, por supuesto es el experimento de Michelson-Morley el que juega un papel fundamental (¡a pesar de que no lo jugó tanto para Einstein!).
- Las ecuaciones de Maxwell <u>no</u> son invariantes bajo transformaciones de Galileo. Además, traen consigo información sobre la vel. de la luz, pero sin dar ningún tipo de información acerca del marco de referencia.

En realidad, una vez refutado el éter, estamos hablando acerca de la propagación de la luz en el vacío. Como el vacío mismo conlleva a la ausencia de puntos de referencia, pareciera que la única forma de dar sentido a una velocidad que se refiere al vacío es haciendo que su valor sea independiente del estado de movimiento del observador. Esto nos lleva a considerar a "c" como una velocidad universal. Así, si las ecs. de Maxwell no hacen referencia a un marco específico, podemos llegar a la conclusión de que todos los marcos inerciales deberían ser, de cierta forma, equivalentes.

- Desde un punto de vista moderno, tenemos forma de verificar experimentalmente relaciones de la forma $E = pc$ (efecto fotoeléctrico, efecto Compton) o, con más generalidad, $E^2 = (mc^2)^2 + (pc)^2$ (física de partículas).

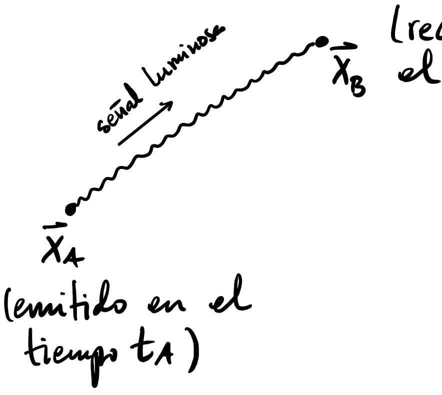
Veamos qué tan lejos podemos llegar adoptando el siguiente principio:

"La velocidad de propagación en todo marco inercial y en el vacío es siempre constante e igual a c".

Consideremos 2 eventos, A y B, que corresponden a la emisión de un rayo de luz en un instante dado (A) y su recepción posterior en otro lugar del espacio (B).
Consideraremos la situación desde el punto de vista de dos sistemas inerciales distintos:

Sistema 1 ("S"): Coordenadas de los eventos $\rightarrow (\vec{x}_A, t_A), (\vec{x}_B, t_B)$

Sistema 2 ("S'"): Coordenadas de los eventos $\rightarrow (\vec{x}_A', t_A'), (\vec{x}_B', t_B')$



Como la velocidad de la luz, para ambos sistemas de referencia, debe ser igual a "c", debemos tener

$$\frac{\|\vec{x}_B - \vec{x}_A\|}{(t_B - t_A)} = c = \frac{\|\vec{x}_B' - \vec{x}_A'\|}{(t_B' - t_A')},$$

lo cual podemos reescribir de la siguiente forma:

$$c^2(t_B - t_A)^2 - \|\vec{x}_B - \vec{x}_A\|^2 = 0 = c^2(t_B' - t_A')^2 - \|\vec{x}_B' - \vec{x}_A'\|^2 \qquad (*)$$

Antes de explorar las consecuencias de $(*)$, debemos establecer el tipo de transformaciones de coordenadas que vamos a considerar:

$$t' = t'(t, \vec{x}) \quad ; \quad x' = x'(t, \vec{x}).$$

En <u>relatividad especial</u> (o, dicho de otra forma, en ausencia de campos gravitacionales) consideraremos exclusivamente <u>transformaciones lineales</u>.

En su artículo de 1905, Einstein dice que las ecs. que rigen este tipo de transformaciones (entre marcos inerciales) deben "ser lineales, en razón de las propiedades de homogeneidad que le atribuimos al espacio y al tiempo".

Para poder avanzar rápidamente, asumiremos la linealidad de estas transformaciones. Sin embargo, es importante mencionar que la justificación de dicho "Ansatz" requiere un análisis detallado, que involucra, además de la hipótesis de homogeneidad del espacio-tiempo, sutilezas de las definiciones de marco inercial y de lo que se entiende por definición de tiempo en un tal sistema, el principio de inercia, etc.

Para un discusión detallada de este punto, recomendamos Torretti, §3.6.

Volviendo a la ec. (*), adoptemos las siguientes convenciones:

$$x_A = (c\,t_A, \vec{x}_A) \quad , \quad x_B = (c\,t_B, \vec{x}_B)$$

$$\rightsquigarrow \quad z := x_A - x_B \,.$$

z es un vector de 4 componentes: $z = (z^0, \vec{z}\,)$; $z^0 = c(t_A - t_B)$, $\vec{z} = \vec{x}_A - \vec{x}_B$

Usaremos una notación similar para el sistema S'.

Ahora, decir que la señal (en S) sale de $\vec{x}_A$ (en el tiempo t_A) y llega a $\vec{x}_B$ (en el tiempo t_B) viajando a vel $= c$, es lo mismo que escribir

$$c^2(t_B - t_A)^2 - \|\vec{x}_B - \vec{x}_A\| = 0$$

$$\Leftrightarrow \quad (z^0)^2 - \|\vec{z}\,\|^2 = 0$$

$$\Leftrightarrow \quad (z^0)^2 - (z^1)^2 - (z^2)^2 - (z^3)^2 = 0 \qquad \left[\ \vec{z} = (z^1, z^2, z^3)\ \right]$$

Reescribiendo esta última relación en notación matricial, tenemos:

$$(z^0, z^1, z^2, z^3)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} z^0 \\ z^1 \\ z^2 \\ z^3 \end{pmatrix} = 0$$

La matriz $g := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ da lugar a una <u>forma bilineal</u> en el espacio vectorial ($\mathbb{R}^4$) en el que está definido el vector z.

Recordemos la def. de forma bilineal (en el caso real):

<u>Def.</u> Sean V un espacio vectorial real y B un mapa $B: V\times V \longrightarrow \mathbb{R}$ que es lineal en cada entrada, i.e., tq

$$B(\lambda x+\mu y, z) = \lambda B(x,z) + \mu B(y,z), \qquad \forall\ x,y,z \in V$$
$$B(x, \lambda y+\mu z) = \lambda B(x,y) + \mu B(x,z), \qquad \lambda,\mu \in \mathbb{R}.$$

Entonces decimos que B es una <u>forma bilineal</u> sobre V.

▷ Si $\dim V = n < \infty$ y escogemos una <u>base</u> $\{e_1, e_2, \dots, e_n\}$ para V, es posible mostrar que existe una matriz $A \in M_n(\mathbb{R})$ tal que

$$B(x,y) = \sum_{i,j=1}^{n} x^i A_{ij}\, y^j ,$$

donde $x, y \in V$ son vectores arbitrarios que, expresados en la base $\{e_i\}_i$, tienen coordenadas x^i e y^i, respectivamente:

$$x = \sum_{i=1}^{n} x^i e_i \quad , \quad y = \sum_{i=1}^{n} y^i e_i .$$

En este caso ($\dim V = n$), tenemos un isomorfismo $V \cong \mathbb{R}^n$ que nos permite "pensar" en x como el vector (de $\mathbb{R}^n$) $\vec{x} \equiv (x^1, \dots, x^n)$.
De esta forma, podemos "representar" a B (un mapa bilineal) en términos de la matriz A.

<u>Def</u> Sean V un espacio vectorial real ($\dim V < \infty$) y $B: V\times V \to \mathbb{R}$ una forma bilineal en V. Decimos que B es

(i) simétrica, si $B(x,y) = B(y,x)\ \ \forall\ x,y \in V$

(ii) anti-simétrica, si $B(x,y) = -B(y,x)\ \ \forall\ x,y \in V$

(iii) no-degenerada, si $B(x,y) = 0\ \ \forall y \in V\setminus\{0\} \Rightarrow x = 0$

(iv) definida positiva, si $B(x,x) > 0\ \ \forall\ x \neq 0.$

<u>Def.</u> Una forma bilineal, simétrica, definida-positiva $B: V\times V \to \mathbb{R}$, para V espacio vectorial real, se llama <u>producto</u> <u>interior</u>.

→ Es muy importante notar que, dado un espacio vectorial V, existe una infinidad de formas bilineales definidas sobre este. Lo mismo aplica para los productos interiores.

→ Por ahora, denotaremos el producto interior usual de $\mathbb{R}^n$ (el "producto punto") de la siguiente forma:

$$\langle \vec{x}, \vec{y} \rangle \equiv x^1 y^1 + x^2 y^2 + \cdots + x^n y^n.$$

<u>Ejercicio</u>. V = esp. vec. real, $\dim V = n$. Sean $B: V\times V \to \mathbb{R}$ una forma bilineal. Sea $\{e_j\}_j$ una base de V. Mostrar que existe una matriz $A \in M_n(\mathbb{R})$ tal que

$$B(x,y) = \langle \vec{x}, A\vec{y} \rangle \quad \forall\ x,y \in V,$$

donde $x = \sum_{i=1}^n x^i e_i$, $y = \sum_{i=1}^n y^i e_i$, con "$\langle\ ,\ \rangle$" denotando el producto interno estándar (euclideo) en $\mathbb{R}^n$.

¿Cuáles son las componentes de la matriz A?

——————— // ———————

Volviendo a nuestra discusión, notemos que la matriz g determina una forma bilineal que, expresada en términos del prod. interior $\langle\ ,\ \rangle$ (el estándar / euclídeo) de $\mathbb{R}^4$, está dada por:

$$B(x,y) := \langle x, gy \rangle, \quad x,y \in \mathbb{R}^4, \quad g = \begin{pmatrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}.$$

Así, la condición $(z^0)^2 - (z^1)^2 - (z^2)^2 - (z^3)^2 = 0$ se puede expresar como

$$\langle z, gz \rangle = 0.$$

Sea $\Lambda : \mathbb{R}^4 \longrightarrow \mathbb{R}^4$ la transformación lineal que nos lleva del sistema S al sistema S'. Debemos tener $z' = \Lambda z$.

Así, la transformación Λ será admisible (desde el punto de vista físico) si se cumple:

$$\langle z, g z \rangle = 0 \quad \Rightarrow \quad \langle \Lambda z, g \Lambda z \rangle = 0. \qquad (**)$$

<u>Ejercicio:</u> Justificar la afirmación anterior.

Exploremos las consecuencias de $(**)$ para el caso de una dimensión espacial y una temporal (i.e., $z = (z^0, z^1)$).

En este caso la matriz Λ es una matriz 2×2, digamos que

$\Lambda = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, con A, B, C, D constantes reales, para las cuales debemos encontrar una condición que caracterice a Λ cuando represente una "transformación admisible".

¿Cuáles son, entonces, las condiciones sobre A, B, C y D?

Si $\langle z, gz \rangle = 0$, tenemos

$$\langle z, g z \rangle \equiv (z^0, z^1)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} z^0 \\ z^1 \end{pmatrix} = (z^0, z^1)\begin{pmatrix} z^0 \\ -z^1 \end{pmatrix} = (z^0)^2 - (z^1)^2 \overset{!}{=} 0$$

$$\Rightarrow \quad z^0 = \pm z^1.$$

Queremos $\langle \Lambda z, g \Lambda z \rangle = 0$, es decir, para $z' = \Lambda z = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} z^0 \\ z^1 \end{pmatrix} = \begin{pmatrix} Az^0 + Bz^1 \\ Cz^0 + Dz^1 \end{pmatrix}$,

requerimos: $(z^{0'})^2 - (z^{1'})^2 = 0$, donde
$$\left.\begin{aligned} z^{0'} &= A z^0 + B z^1 \\ z^{1'} &= C z^0 + D z^1 \end{aligned}\right. \qquad (\ddagger)$$

De $\langle z, g z\rangle = 0$ tenemos $z^0 = \pm z^1$. Reemplazando $(\ddagger)$ en $(z^{0'})^2-(z^{1'})^2=0$, obtenemos:

$$(Az^0 + Bz^1)^2 - (Cz^0 + Dz^1)^2 \overset{!}{=} 0 \quad \Longleftrightarrow$$

$$(Az^0)^2 + (Bz^1)^2 + 2AB z^0 z^1 - (Cz^0)^2 - (Dz^1)^2 - 2CD z^0 z^1 = 0 \quad \Longleftrightarrow$$

$$(z^0)^2 (\underbrace{A^2 + B^2 - C^2 - D^2}_{=:\alpha}) + 2 z^0 z^1 (\underbrace{AB - CD}_{=:\beta}) = 0 \quad \Longleftrightarrow$$

$$(z^0)^2 \alpha + 2 z^0 z^1 \beta = 0$$

Usando $z^0 = \pm z^1$, tenemos $(z^1)^2(\alpha \pm 2\beta) = 0 \rightsquigarrow \alpha \pm 2\beta = 0$

$\Rightarrow \alpha = 0, \beta = 0.$

Lo anterior implica:

i) $\lambda := A^2 - C^2 \overset{!}{=} D^2 - B^2$,

ii) $AB - CD = 0$

Por otro lado, tenemos:

$$\Lambda^T g \Lambda = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & C \\ B & D \end{pmatrix} \begin{pmatrix} A & B \\ -C & -D \end{pmatrix}$$

$$= \begin{pmatrix} A^2 - C^2 & AB - CD \\ AB - CD & B^2 - D^2 \end{pmatrix} \overset{i), ii)}{=} \begin{pmatrix} \lambda & 0 \\ 0 & -\lambda \end{pmatrix} = \lambda g$$

Vemos entonces que, en general, la condición que obtendremos será de la forma

$$\Lambda^T g \Lambda = \lambda g \quad , \quad \text{con } \lambda \in \mathbb{R}.$$

(este mismo tipo de relación se obtiene en el caso de $\mathbb{R}^4$).

Notemos que un tipo de transformación entre marcos inerciales es una rotación 3-dimensional (rotación espacial). Si $R \in SO(3)$, entonces

$$\mathcal{R} = \left(\begin{array}{c|c} 1 & 0\;0\;0 \\ \hline \begin{matrix}0\\0\\0\end{matrix} & R \end{array}\right)$$

es una transformación que satisface $\mathcal{R}^T g \mathcal{R} = \lambda g$, con $\lambda = 1$.

¿Qué pasaría si hubiese otra forma de implementar una rotación espacial a través de una transformación Λ en el espacio-tiempo $\mathbb{R}^4$, pero con $\Lambda^T g \Lambda = \lambda g$, y $\lambda \neq 1$?

→ Esto implicaría que, al rotar el sistema de referencia (espacialmente) se generaría un cambio de escala, lo cual no tiene mucho sentido desde el punto de vista físico. Argumentando de esta manera, podemos llegar a convencernos de que imponer la condición $\lambda = 1$ es algo razonable. Lo que haremos aquí será imponer dicha condición en la forma de un <u>postulado</u>, cuya "justificación" se dará a posteriori, en la medida que nos permita desarrollar una teoría coherente, cuyas predicciones estén de acuerdo con los fenómenos observados.

Antes de formular los postulados básicos de la relatividad especial, notemos que aparte de la matriz Λ, a una transformación de coordenadas le podemos añadir una traslación:

$$x \longmapsto x' = \Lambda x + a \qquad (a \in \mathbb{R}^4),$$

ya que, al comparar eventos, este tipo de términos aditivos desaparecen.

<u>Definición</u> Aquellas transformaciones $\Lambda \in M_4(\mathbb{R})$ que satisfacen $\Lambda^T g \Lambda = g$, con $\Lambda^0{}_0 \geq 1$ y $\det \Lambda = 1$, se denominan transformaciones de Lorentz <u>propias</u>, <u>ortocronas</u>. El conjunto de todas estas transformaciones será denotado $L_+^{\uparrow}$.

→ La razón para restringir nuestra atención a las transformaciones propias y ortocronas es que "simetrías" tales como las transformaciones de paridad no son preservadas en general (por ejemplo, la interacción débil viola la paridad de forma maximal).

Examinemos ahora el efecto de aplicar 2 transformaciones (Λ_1, a_1), (Λ_2, a_2), una después de la otra:

$$x \xmapsto[(\Lambda_1, a_1)]{} x' = \Lambda_1 x + a_1 \xmapsto[(\Lambda_2, a_2)]{} x'' = \Lambda_2 x' + a_2$$

El efecto neto de aplicar 2 transformaciones es el de dar lugar a una transformación del mismo tipo. En efecto, tenemos:

$$
\begin{aligned}
x'' &= \Lambda_2 x' + a_2 \\
&= \Lambda_2 (\Lambda_1 x + a_1) + a_2 \\
&= \Lambda_2 \Lambda_1 x + \Lambda_2 a_1 + a_2 .
\end{aligned}
$$

Esto nos lleva a la siguiente "regla de composición":

$$(\Lambda_2, a_2)\cdot(\Lambda_1, a_1) = (\Lambda_2\Lambda_1 \,,\, \Lambda_2 a_1 + a_2) \qquad (*)$$

Definición

El conjunto $\mathcal{P} = \{ (\Lambda, a) \in M_4(\mathbb{R}) \times \mathbb{R}^4 \mid \Lambda \in L_+^\uparrow \}$, provisto de la operación definida en (*) se conoce como el grupo de Poincaré.

Ejercicio. Mostrar que, en efecto, la operación (*) da lugar a una estructura de grupo sobre $\mathcal{P}$.

Postulado En ausencia de gravedad, las leyes de la naturaleza deben ser invariantes bajo transformaciones de Poincaré $(\Lambda, a) \in \mathcal{P}$.

Observación. Si bien es cierto que podríamos formular la teoría partiendo de 2 postulados básicos ((i) vel luz $= c$ en todo marco inercial y (ii) intervalo espacio-temporal entre eventos, $(z^0)^2 - \vec{z}^{\,2} \equiv$ invariante), el postulado propuesto aquí es mucho más general, en el sentido de que aplica no solo para la cinemática de partículas "clásicas", sino para cualquier interacción fundamental (salvo la gravitación).
Para un análisis de las relaciones entre diferentes postulados, consultar R. Torretti (Relativity and Geometry).