Sobre la definición de tensor

Resumen

Del repaso hecho a los textos, citados en las referencias, y que tratan sobre relatividad general se ve que todos hacen una introducción a los tensores con mayor o menor pedagogía. La visión matemática de Pope, la relación de los tensores con los sistemas de referencia no privilegiados que hace Plebansky y Krasinski, la clara definición que hace Schutz, tocadas parcialmente por Narlikar, Logunov, Janssen y Hooft, permiten aclarar el concepto sobre este objeto geométrico que es el tensor.

Narlikar

En el texto de **Narlikar** [3, p. 32] encontramos la ecuación del cuadrado de la distancia entre dos sucesos en el espaciotiempo, que incluye la notación de las componentes del tensor métrico y de los vectores con subíndices y superíndices

$$\mathrm{d}s^2 = g_{ik}\ \mathrm{d}x^i\ \mathrm{d}x^k, \tag{1}$$

explica que la notación para las coordenadas han cambiado indicando que la componente x^0 corresponde al tiempo y las componentes x^1, x^2 y x^3 a las espaciales. Respecto del tensor métrico no indica cual es la naturaleza de dicho objeto, pero habla de sus propiedades como que sus componentes son funciones de las coordenadas y que la matriz $||g_{ik}||$ tiene de signatura -2.

Se sigue en [3, p. 36] la definición de vectores contravariantes aquellos cuyas componentes, ante un cambio de coordenadas, transforman de acuerdo a la ley lineal

$$A'^k = \frac{\partial x'^k}{\partial x^i}\ \frac{\mathrm{d}x^i}{\mathrm{d}\lambda} = \frac{\partial x'^k}{\partial x^i}\ A^i, \tag{2}$$

para ello supone que una curva viene dada por una sucesión de puntos funciones de un parámetro $x^i \equiv x^i(\lambda)$, siendo la dirección de la tangente a la curva el vector $A^i \equiv \mathrm{d}x^i/\mathrm{d}\lambda$, y como la dirección de la tangente es invariante, ante un cambio de coordenadas no altera el concepto pero si tendrá unas nuevas componentes $A'^i \equiv \mathrm{d}x'^i/\mathrm{d}\lambda$.

En [3, p. 37] define los vectores covariantes como cantidades que transforman de acuerdo a la regla

$$B'_k = \frac{\partial x^i}{\partial x'^k}\ B_i, \tag{3}$$

y para ello considera una función escalar $\phi(x^k) =$ constante que describe una hipersuperficie y cuya normal tiene la dirección dada por cuatro componentes $B_i = \partial\phi/\partial x^i$, que no cambia ante una cambio de coordenadas pero si que lo hacen sus nuevas componentes $B'_i = \partial\phi/\partial x'^i$.

Narlikar dice que el concepto de vector puede ser generalizado al de tensor y de este modo un tensor contravariante de rango 2 se caracteriza por la siguiente transformación

$$T'^{ik} = \frac{\partial x'^i}{\partial x^m}\ \frac{\partial x'^k}{\partial x^n}\ T^{mn}, \tag{4}$$

un tensor covariante de rango 2 por

$$T'_{ik} = \frac{\partial x^m}{\partial x'^i}\ \frac{\partial x^n}{\partial x'^k}\ T_{mn} \tag{5}$$

y un tensor mixto de rango 2

$$T'^i{}_k = \frac{\partial x'^i}{\partial x^m}\ \frac{\partial x^n}{\partial x'^k}\ T^m{}_n \tag{6}$$

Logunov

Logunov [2, p 17]introduce sin ningún tipo de explicación el tensor métrico γ^{ik}, cuando transforma las ecuaciones de un punto material expresada en coordenadas cartesianas a curvilíneas, así como los superíndices de las coordenadas.

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left[\left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2\right] = 0, \tag{7}$$

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\gamma^{ik}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^k} = 0. \tag{8}$$

En su capítulo 25 sobre análisis tensorial [2, p 179-184] define vector contravariante y vector covariante a partir de una transformación de coordenadas, de forma que un vector contravariante, $a^i(x)$ (indica las componentes del vector) se define mediante un conjunto de n funciones $[a^1(x)...a^n(x)]$ que al pasar de un sistema a otro se transforma según la ley

$$a'^i(x') = \frac{\partial x'^i}{\partial x^k}a^k(x(x')). \tag{9}$$

En el caso del vector covariante se refiere a un conjunto de n funciones reales $a_i(x) = [a_1(x)...a_n(x)]$ que al aplicar una transformación de coordenadas se transforman como lo hace el gradiente de un escalar $\partial\Psi/\partial x^i$

$$a'_i(x') = \frac{\partial x^k}{\partial x'^i}a_k(x(x')). \tag{10}$$

implícitamente supone que el gradiente de un escalar es un vector covariante [1].
De estas definiciones pasa a la definición de un ente mas general, en comparación con el vector, denominando un tensor k veces covariante y l veces contravariante $T^{j_1...j_l}_{i_1...i_k}(x)$ a un ente geométrico definido en cada sistema local de coordenadas mediante un conjunto de n^{l+k} funciones tomadas en un orden determinado y que al pasar de un sistema de coordenadas se transforma según la ley

$$T^{j_1...j_l}_{i_1...i_k}(x') = \frac{\partial x^{a_1}}{\partial x'^{i_1}}...\frac{\partial x^{a_k}}{\partial x'^{i_k}}\,\frac{\partial x'^{j_1}}{\partial x^{b_1}}...\frac{\partial x'^{j_l}}{\partial x^{b_l}}\;T^{b_1...b_k}_{a_1...a_l}(x(x')) \tag{12}$$

Janssen

En el capítulo 4 sobre álgebra de tensores y transformaciones ortogonales, **Janssen** [1, p 67-79] aplica a los vectores los calificativos columna y fila para referirse a vector contravariante y covariante respectivamente e introduce la notación bra y ket para distinguirlos. El vector posición $|x\rangle$ expresado en función de los vectores de la base de un espacio vectorial N-dimensional $\{|e_i\rangle\}$ y usando el convenio de sumación de Einstein

$$|x\rangle = x^i\,|e_i\rangle \equiv \begin{pmatrix} x^1 \\ \vdots \\ x^N \end{pmatrix}. \tag{13}$$

[1] Un escalar tiene el mismo valor en cualquier sistema de coordenadas $\Psi(x') = \Psi(x(x'))$, si diferenciamos con respecto a las coordenadas x' y aplicamos la regla de la cadena obtenemos

$$\frac{\partial\Psi(x')}{\partial x'^i} = \frac{\partial\Psi(x)}{\partial x'^i} = \frac{\partial x^j}{\partial x'^i}\frac{\partial\Psi(x)}{\partial x^j}, \tag{11}$$

que indica cómo se transforman las componentes del gradiente del escalar, un vector covariante, en un cambio de coordenadas.

A continuación considera el espacio de las aplicaciones lineales $\langle y|$ que llevan los vectores del espacio vectorial $\mathbb{R}^N$ a los números reales $\mathbb{R}$: la aplicación $\langle y|$ actúa sobre el vector contravariante $|x\rangle$ de forma que $\langle y \mid x\rangle \in \mathbb{R}$. Este espacio de aplicaciones lineales tiene la estructura de un espacio vectorial, llamado espacio vectorial dual $^*\mathbb{R}^N$, así comenta la posibilidad de considerar los elementos $\langle y|$ como un tipo de vectores, distintos a los $|x\rangle$, y construir una base dual $\{\langle e^i|\}$ que permite descomponerlos

$$\langle y| = y_i \left\langle e^i\right| \equiv \left(y_1 \ldots y_N\right). \tag{14}$$

siendo estos un tipo de vectores covariantes o uno-formas, que implican que el producto escalar de los vectores de la base del espacio vectorial y del espacio dual cumplan la condición

$$\left\langle e^i \mid e_j\right\rangle = \delta^i{}_j \tag{15}$$

y así

$$\langle y \mid x\rangle = y_i \left\langle e^1\right| \cdot x^j \left|e_j\right\rangle = y_i \cdot x^j \left\langle e^1 \mid e_j\right\rangle = y_i \cdot x^j \cdot \delta^i{}_j = y_i \cdot x^i \in \mathbb{R} \tag{16}$$

Hooft

Al comienzo de su introducción a la teoría de la relatividad especial, **Hooft** [7] hace una referencia expresa a la notación, diciendo que "el uso intermitente de superíndices $(\{\}^\mu)$ y subíndices $(\{\}_\mu)$ no tiene importancia, pero la tendrá mas adelante". De esta manera expresa las transformaciones de Lorentz o "transformaciones de velocidad" como el conjunto de transformaciones lineales

$$(x^\mu)' = \sum_{\nu=1}^{4} L^\mu{}_\nu \; x^\nu \tag{17}$$

sujeto a la condición de que la cantidad σ permanezca invariante

$$\sigma^2 = \sum_{\mu=1}^{4} (x^\mu)^2 = |\vec{x}|^2 - c^2 t^2. \tag{18}$$

Hooft considera $L^\mu{}_\nu$[2] como una matriz de μ filas y ν columnas y $(x^\mu)'$ y (x^μ) matrices columna entonces

$$\sum_{\mu=1}^{4} (x^\mu)^2 = \begin{pmatrix} x^1 & \ldots & x^4 \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^4 \end{pmatrix} = [x^\mu]^T [x^\mu] \tag{19}$$

$$[x^\mu]^T [x^\mu] = \left[x^{\nu'}\right]^T \left[x^{\nu'}\right] = [x^\mu]^T \left[L^\nu{}_\mu\right]^T \left[L^\nu{}_\mu\right] [x^\mu] \tag{20}$$

$$\left[L^\nu{}_\mu\right]^T \left[L^\nu{}_\mu\right] = \delta^\nu_\mu \tag{21}$$

que demuestra que la condición impuesta a la transformación de Lorentz implica que la matriz de dicha transformación tiene que ser ortogonal, llamándose al conjunto de matrices ortogonales del grupo ortogonal $O(n, \mathbb{C})$ o $SO(n, \mathbb{C})$ a un subgrupo especial si el

[2] Según esta afirmación la forma de escribir la transformación podría ser $L_\mu{}^\nu$ o $L^{\mu\nu}$, pero siempre el primer índice se refiere a las filas y el segundo a las columnas.

determinante es ± 1. Pone de ejemplo como elemento del grupo de Lorentz una rotación en el espaciotiempo (n=4) de los ejes $x^3 \equiv z$ y $x^4 \equiv ict$

$$[L^\mu{}_\nu] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh\chi & -i\sinh\chi \\ 0 & 0 & i\sinh\chi & \cosh\chi \end{pmatrix} \tag{22}$$

que produce la siguiente rotación

$$x' = x \ ; \ y' = y \ ; z' = z \ \cosh\chi - ct \ \sinh\chi \ ; t' = -\frac{z}{c}\sinh\chi + t \ \cosh\chi \tag{23}$$

que es el movimiento del sistema con primas a lo largo del eje Z, con velocidad

$$v = c \ \tanh\chi \tag{24}$$

En una transformación de Lorentz hay magnitudes que son invariantes, como la velocidad de la luz o cualquier escalar, pero otras aun no siendo invariantes son covariantes ante esta transformación, por ejemplo el tetravector momento energía

$$p^\mu = \begin{pmatrix} p_x \\ p_y \\ p_z \\ iE \end{pmatrix} \ ; \ (p^\mu)' = L^\mu{}_\nu \ p^\nu . \tag{25}$$

Hooft define tensor como cantidades que transforman según la ecuación

$$(X^{\mu\nu\ldots}{}_{\alpha\beta\ldots})' = L^\mu{}_\kappa L^\nu{}_\lambda \ldots L^\alpha{}_\gamma L^\beta{}_\delta \ X^{\kappa\lambda\ldots}{}_{\gamma\delta} \tag{26}$$

Plebansky y Krasinski

Como hemos visto en los textos citados la transformación es la palabra clave en la definición de tensor, **Plebansky** y **Krasinski** [6] dicen que “los tensores son objetos definidos para que ningún sistema sea privilegiado” y comienzan dando una definición poco precisa “supongamos que cambiamos el sistema de coordenadas en un espacio n-dimensional de $\{x^\alpha\}, \alpha = 1, 2, \ldots, n$ a $\{x^{\alpha'}\}, \alpha' = 1, 2, \ldots, n$, un tensor es una colección de funciones en este espacio que cambia de una manera específica bajo dicha transformación de coordenadas”.

Pope

Aprovechamos que **Pope** [4] dice “de hecho el conjunto de los números reales es el prototipo de variedad; es un espacio topológico parametrizado por los puntos de la línea real” para definir una variedad N-dimensional M como un espacio que se parece localmente a $\mathbb{R}^N$, que está compuesta de parches N-dimensionales y allí donde los parches se solapan están pegados topológicamente, es decir se puede definir una función diferenciable

$$f : \ M \to \mathbb{R} \tag{27}$$

una aplicación que asigna un número real a cada punto de la variedad, si además en un parche de M se define un sistema de coordenadas

$$\phi : \ M \to \mathbb{R}^N \tag{28}$$

componemos ambas aplicaciones

$$f \circ \phi^{-1} : \mathbb{R}^N \to \mathbb{R} \tag{29}$$

si las coordenadas son x^i la aplicación $f \circ \phi^{-1}$ puedes escribirse como $f(x^i)$, es decir que $f(x^i)$ representa el valor de f en el punto de M especificado por las coordenadas x^i. Pope propone definir vector de una manera diferente al vector desplazamiento cuando se habla de espacios curvos, y lo relaciona con el límite infinitesimal de la línea que une el punto A con $A + \delta A$, de hecho esto significa que se piensa en el plano tangente a un punto en el espacio mientras que los vectores son desplazamientos infinitesimales en este plano. Considera un parche U en la variedad M, en el que se define un sistema de coordenadas locales x^i, una curva que cruce U puede ser descrita al especificar los valores de las coordenadas de los puntos por los que pasa, y para ello se introduce un parámetro λ que incrementa monótonamente a lo largo de la curva de forma que cada punto de M perteneciente a la curva se expresa así

$$x^i = x^i(\lambda) \tag{30}$$

si se tiene definida una función f en M, los valores que toma a lo largo de la curva viene dados por $f(x^i(\lambda))$, y al aplicar la regla de la cadena se obtiene

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \frac{\partial x^i}{\partial \lambda}\frac{\mathrm{d}f}{\mathrm{d}x^i} = \frac{\partial x^i}{\partial \lambda}\partial_i(f) = V\ f \tag{31}$$

donde vemos que el producto de las componentes de dos objetos geométricos conducen a un valor real

$$V = \frac{\partial x^i}{\partial \lambda}\ \partial_i\ ;\ w = \partial_i(f)\ \mathrm{d}x^i \tag{32}$$

el primer objeto es el vector tangente expresado en la base $\{\partial_i\}$ y el otro recibe el nombre de covector que pertenece a un espacio dual y que viene expresado en la base $\{\mathrm{d}x^i\}$. La condición para que dos espacios vectoriales sean duales viene dada por

$$\langle w \mid V\rangle \in \mathbb{R} \tag{33}$$

que obliga a que

$$\left\langle \mathrm{d}x^i \mid \partial_j \right\rangle = \delta^i{}_j \tag{34}$$

y así

$$\langle w \mid V\rangle = w_i\ V^i = \partial_i(f)\frac{\partial x^i}{\partial \lambda} = \frac{\mathrm{d}f}{\mathrm{d}\lambda}. \tag{35}$$

Vemos que el vector V lo hemos definido como un operador $\mathrm{d}/\mathrm{d}\lambda$ que identificamos con el vector tangente cuando definimos un sistema de coordenadas x^i, siendo sus componentes dependientes de la base, ante un cambio de sistema de coordenadas x'^i el vector no cambia pero si lo hacen sus componentes, para ello aplicamos la regla de la cadena al tener que $x'^i = x'^i(x^j)$

$$V = V^j\frac{\partial}{\partial x^j} = V^j\frac{\partial x'^i}{\partial x^j}\frac{\partial}{\partial x'^i} \equiv V'^i\frac{\partial}{\partial x'^i} \tag{36}$$

donde

$$V'^i = \frac{\partial x'^i}{\partial x^j}V^j \tag{37}$$

que se conoce como la regla de transformación de coordenadas general que debe cumplir un vector ante un cambio de coordenadas arbitrario.

Y los covectores transforman de la forma

$$w = w_i \; \mathrm{d}x^i = \partial_i f \; \mathrm{d}x^i = \partial_i f \; \frac{\partial x^i}{\partial x'^j} \mathrm{d}x'^j \equiv \partial_{'j} \; \mathrm{d}x'^j = w_{'j} \; \mathrm{d}x'^j \tag{38}$$

donde

$$w_{'j} = \frac{\partial x^i}{\partial x'^j} \; w_i \tag{39}$$

Una vez introducida la noción de vector y covector es mas fácil la generalización a tensores de rango arbitrario, como objetos geométricos que viven en un espacio producto tensorial ($\otimes$) que involucra r factores de espacios tangente $T_p(M)$ y t factores de espacios cotangente $T_p^*(M)$. Un tensor $Z^r{}_t$ se designa con el tipo (r,t) o $\binom{r}{t}$ y se dice que su rango es la suma de ambos[3]

$$Z = Z^{i_1 \ldots i_r}{}_{j_1 \ldots j_t} \; \partial_{i_1} \otimes \ldots \otimes \partial_{i_r} \otimes \mathrm{d}x^{j_1} \otimes \ldots \otimes \mathrm{d}x^{j_t} \tag{40}$$

este mismo tensor es independiente del sistema de coordenadas por lo que en otro sistema (indicado con primas) sería

$$Z = Z^{'i_1 \ldots 'i_r}{}_{'j_1 \ldots 'j_t} \; \partial_{'i_1} \otimes \ldots \otimes \partial_{'i_r} \otimes \mathrm{d}x'^{j_1} \otimes \ldots \otimes \mathrm{d}x'^{j_t} \tag{41}$$

y sus componentes transforman de acuerdo a la regla señalada anteriormente para vectores y covectores

$$Z^{'i_1 \ldots 'i_r}{}_{'j_1 \ldots 'j_t} = \frac{\partial x'^{i_1}}{\partial x^{k_1}} \cdots \frac{\partial x'^{i_r}}{\partial x^{k_r}} \frac{\partial x^{l_1}}{\partial x'^{j_1}} \cdots \frac{\partial x^{l_t}}{\partial x'^{j_t}} Z^{k_1 \ldots k_r}{}_{l_1 \ldots l_t} \tag{42}$$

por esto podemos decir que un tensor es del tipo (r,t) si en transformaciones de coordenadas generales si y solo si sus componentes se transforman como en (42) bajo transformaciones de coordenadas generales.

Schutz

Schutz [5] hace la siguiente definición "Un tensor del tipo $\binom{0}{N}$ es una función de de N vectores hacia los números reales, lineal en cada uno de sus N argumentos". El cuadrado del intervalo en un espaciotiempo bajo una transformación de Lorentz es un invariante, y por analogía la magnitud de tetravector $\vec{A}$ también lo es

$$\vec{A}^2 = -(A^0)^2 + (A^1)^2 + (A^2)^2 + (A^3)^2 = -(A^{\tilde{0}})^2 + (A^{\tilde{1}})^2 + (A^{\tilde{2}})^2 + (A^{\tilde{3}})^2 \tag{43}$$

que se puede ampliar al producto escalar de dos tetravectores $\vec{A}$ y $\vec{B}$

$$\vec{A} \cdot \vec{B} = -A^0 B^0 + A^1 B^1 + A^2 B^2 + A^3 B^3 \tag{44}$$

si reescribimos la anterior expresión usando matrices tenemos

$$\vec{A} \cdot \vec{B} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A^0 \\ A^1 \\ A^2 \\ A^3 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} = \begin{pmatrix} -A^0 & A^1 & A^2 & A^3 \end{pmatrix} \begin{pmatrix} B^0 \\ B^1 \\ B^2 \\ B^3 \end{pmatrix} \tag{45}$$

[3] Por ejemplo un tensor de rango 2 obtenido del producto tensorial de dos espacios vectoriales de la misma dimensión N, tendrá N^2 componentes, otro tensor de rango 4 que se obtenga del producto tensorial de tres espacios de N dimensiones con uno de M, tendrá $N^3 \times M$ componentes. Un tensor de rango 2 puede ser del tipo (0,2) $Z_{\alpha\beta}$, o del tipo (2,0) $Z^{\alpha\beta}$, o del tipo (1,1) $Z^\alpha{}_\beta$.

siendo la primera matriz el tensor métrico g en el espaciotiempo, que se puede considerar como un tensor del tipo $\binom{0}{2}$ que toma como argumentos dos tetravectores para obtener un escalar

$$g(\vec{A}, \vec{B}) := \vec{A} \cdot \vec{B} \tag{46}$$

siendo este tensor métrico

$$g = -1\vec{e}_0 \otimes \vec{e}_0 + 1\vec{e}_1 \otimes \vec{e}_1 + 1\vec{e}_2 \otimes \vec{e}_2 + 1\vec{e}_3 \otimes \vec{e}_3 \tag{47}$$

sólo cuatro de sus dieciséis componentes son diferentes de cero. Para resumen el producto escalar de dos vectores puede expresarse mediante el producto de las componentes de los tensores, en este caso uno del tipo $\binom{0}{2}$ y los otros dos del tipo $\binom{1}{0}$

$$\vec{A} \cdot \vec{B} = g_{\alpha\beta} A^\alpha B^\beta \tag{48}$$

Referencias

[1] Bert Janssen. *Teoría de la Relatividad General.* Departamento de Física Teórica y del Cosmos, Granada, 2013. http://www.ugr.es/ bjanssen/text/BertJanssen-RelatividadGeneral.pdf.

[2] A.A. Logunov. *Curso de la Teoría de la Relatividad y de la Gravitación.* URSS, Moscú, 1998.

[3] J.V. Narlikar. *Introduction to Cosmology.* Cambridge University Press, Cambridge, 1993.

[4] Christopher Pope. *Geometry and Group Theory.* Texas AM University, Texas, 2006.

[5] Bernard F. Schutz. *A First Course in General Relativity.* Cambridge University Press, Cambridge, 2009.

[6] Jerzy Plebański y Andrzej Krasiński. *An Introduction to General Relativity and Cosmology.* Cambridge University Press, Cambridge, 2006.

[7] Gerard 't Hooft. *Introduction to theory of relativity.* Institute for Theoretical Physics Utrecht University and Spinoza Institute, Utrecht, 2012. http://www.staff.science.uu.nl/ hooft101/lectures/gr.html.