Notas sobre Física, Matemática y Física-Matemática

Martín de la Rosa Díaz

16 de agosto de 2020 – Revisión **1.a**

Prólogo

Este documento está constituido por unidades temáticas que abarcan múltiples aspectos de la Física, la Matemática y la profunda conexión entre ambas. Dichas unidades no tienen por qué guardar relación entre sí, hasta el punto de que han podido redactarse con varios años de separación temporal. Es por ello que no se caracterizan por una única perspectiva físico-matemática, ni tampoco por un estilo o formato uniformes, sino que deben reflejar la evolución en la mentalidad y conocimientos del autor.

Se pretende que, gradualmente, este documento vaya llenándose de contenido. Futuras actualizaciones, además de añadir nuevos temas y secciones, corregirán o alterarán de manera más o menos sutil lo que ya formaba parte de versiones previas. El lector queda invitado a trasladar sus valoraciones, críticas o no, a través de correo electrónico o cualquier otro medio por el que pueda comunicarse con el autor.

16 de agosto de 2020

Índice

 ✉ f42rodim@uco.es

1. Un alegato en defensa de los tensores en Física

La Física es una disciplina tensorial.

Miguel Ángel Hernández Aláez

Antes de nada, notar que:

- Los **escalares** son objetos cuya representación es la misma, un 'número', en todas las bases. Los escalares son *tensores de orden cero*.
- Los **vectores** son objetos cuya representación en una base dada es un array unidimensional de 'números' (una matriz fila o columna). Los vectores son *tensores de orden uno*.
- Los **tensores de orden dos** son objetos cuya representación en una base dada es un array bidimensional (una matriz cuadrada).
- Y del mismo modo con tensores de orden tres y superior.

Los tensores son, en algún sentido, una generalización de los escalares y vectores, que, en efecto, son casos particulares de estos. Por eso, todas las consideraciones físicas generales sobre tensores se aplican a escalares y vectores.

¿Por qué son los tensores útiles en física? El objetivo último de los físicos y filósofos naturales es aprehender el universo a la escala más fundamental posible. Esto bien podría ser una meta fuera del alcance de nuestra especie, o incluso cabe la posibilidad de que la naturaleza sea fundamentalmente inaprehensible. Realmente, apenas si disponemos de modelos matemáticos que predicen el resultado de experimentos con cierto grado de aproximación, y que solo son válidas hasta escalas de energía finitas. Los físicos teóricos se jactan de que disponen de teorías fundamentales como la relatividad general o la teoría cuántica de campos. ¿Pero cómo podemos estar tan seguros de que nuestros modelos captan siquiera un ápice de la estructura fundamental del cosmos, si es que tal cosa existe? No hay manera de saberlo. Nadie antes del advenimiento de la relatividad especial habría imaginado que la mecánica newtoniana acabaría siendo remplazada por un constructo de la cual esta es solo un límite de bajas velocidades. Algunos atribuyen (erróneamente) a Lord Kelvin la creencia de que todo lo que quedaba por hacer en la física a finales del siglo XIX era computar predicciones con más cifras significativas y efectuar mediciones más precisas. Paralelamente, en la actualidad se piensa que gozamos de una descripción satisfactoria de tres de las cuatro interacciones fundamentales. ¿Pero son estas, en un sentido categórico, las interacciones fundamentales? Más preocupante resulta que en las últimas décadas no hemos progresado en el entendimiento esencial del mundo. La teoría de cuerdas, que en sus albores fue declarada la prometida gran teoría unificada de la gravedad y el resto de fuerzas, sigue sin dar fruto. El paradigma de conocimiento aceptado, en lo que a la física más teórica respecta, no ha sufrido apenas cambios en los últimos 50 años. Y, a pesar de ello, las nuevas generaciones siguen lanzándose al estudio de la teoría de cuerdas, la supersimetría y otras ideas especulativas absolutamente carentes de verificación experimental. ¿Estamos haciendo las cosas bien?

De este elongado discurso que poco o nada tiene que ver con el objetivo del presente curso, hemos de quedarnos con la idea de que es bueno construir nuestros modelos y descripciones de la naturaleza empleando objetos tan fundamentales como se pueda, desde un punto de vista físico, puesto que en general nos aproxima a ese objetivo último que hemos esbozado en el párrafo anterior y facilita la comprensión conceptual de la teoría, la torna más transparente. Sean A y B dos magnitudes físicas u observables, esto es, cosas que se prestan a la medición. Los distintos observadores o sistemas de referencia inerciales[1] están facultados para mesurar A y B de manera independiente. Se comprueba entonces que todos los observadores están de acuerdo en lo que vale A, mas no así con B. Concluiremos pues que, a diferencia de B, la mangitud A es *independiente del observador*. Parece sensato asignar una existencia física genuina solo a aquellos observables que son como A. ¿Por qué? Ya en 1632, el italiano Galileo Galilei afirmó que no es

[1] Nos limitamos a observadores inerciales en la presente discusión solo por simplicidad. Todo lo dicho aquí es extensible a observadores en cualquier estado de movimiento.

 ✉ f42rodim@uco.es

posible distinguir mediante experimentos de carácter mecánico el estado de movimiento de un sistema que se desplaza con velocidad constante. Y, algo menos de tres siglos más tarde, Einstein amplió este principio a toda clase de fenómenos físicos, en ausencia de gravedad, haciendo de este uno de los dos postulados de la teoría de la relatividad especial. En suma, *todos los observadores inerciales perciben la misma física*. Esta oración, concisa y rotunda, hace resonar el tímpano de los físicos teóricos, que vislumbran (o quieren ver) en ella el paradigma de principio fundamental de la naturaleza. Guiados por esta doctrina, nos sentimos compelidos a desterrar de nuestras teorías a aquellas magnitudes que son como B, ya que, en virtud del principio de relatividad, no encierran contenido físico alguno. Lo que solía ser enrevesado e indecoroso se vuelve elegante y claro al pasar a un modelo basado en objetos matemáticos como A que no dependen del observador. Pues bien, *estos objetos son tensores*.

A continuación, discutiremos algunos ejemplos bien conocidos de magnitudes que son como A o como B. Pensemos en la energía cinética de un cuerpo. Un observador en reposo con respecto al cuerpo asignaría a este una energía cinética nula, mientras que desde un sistema de referencia que se mueve en relación al cuerpo se diría que este posee una energía cinética distinta de cero. Por tanto, la energía cinética es como B. Sin embargo, todos los observadores, con independencia de su orientación o estado de movimiento, miden, en mecánica prerrelativista, las mismas altura o anchura del cuerpo. Más generalmente, la distancia entre dos puntos o la separación temporal entre dos eventos son observables del estilo de A. Este hecho, de apariencia trivial e inocente, deja de ser cierto cuando damos el salto a la relatividad especial. Distintos observadores inerciales perciben diferentes separaciones espaciales y temporales como consecuencia de las archiconocidas contracción de longitudes y dilatación de tiempos. Un observable que sí es independiente del observador tanto en mecánica newtoniana como einsteiniana es la masa[2] de una partícula. Todas estas magnitudes que no dependen del observador

[2]En algunos tratados antiguos sobre relatividad especial se define la *masa relativista*, que sí depende de la velocidad relativa entre partícula y observador. Este concepto obscurece más que ilumina y acabó por desecharse (aunque todavía hay quien se aferra a él). En la actualidad, la masa (que antes era llamada

vienen representadas por un único valor numérico, es decir, son *tensores de orden cero* o *escalares*. Así, masa, longitud y altura del cuerpo son escalares, mientras que la energía cinética no lo es, a pesar de que todas ellas toman la forma de números cuando son medidas por un observador.

Como es bien sabido, no todos los observables son escalares. La velocidad, por ejemplo, no es un escalar pues su representación en un sistema de referencia dado consta no de un único número sino de una ristra de números, i.e. una matriz fila o columna. Esto nos tienta a pensar que la velocidad podría tratarse de un vector. No obstante, la velocidad depende manifiestamente del estado del movimiento de quien la cuantifica. El ejemplo canónico de *vector* en física (o *tensor de orden uno*), de acuerdo con nuestras definiciones, es la fuerza ejercida sobre una partícula en un sistema de referencia inercial (o, equivalentemente, su aceleración). De hecho, el contenido explícito del principio de relatividad de Galileo no es sino que la segunda ley de Newton $F = ma$ es igual para todos los observadores inerciales. ¿Qué significa exactamente que F sea un vector? En el caso de las magnitudes escalares la respuesta era directa: Los escalares son aquellos observables que toman el mismo valor numérico en todos los sistemas de referencia. Es por ello que a los escalares se les conoce también como *invariantes*. Con los vectores no sucede exactamente lo mismo, puesto que la representación de un vector sí es formalmente dependiente del observador. Piénsese en una partícula moviéndose en el plano sobre la que actúa una fuerza en la dirección del eje x. Cierto observador tendría a bien asignar a la citada fuerza las coordenadas $(F, 0)$. Imaginemos un segundo sistema de referencia que está rotado 180° respecto del primero, de manera que la fuerza es vista apuntando en sentido contrario. Se sigue que aquí la fuerza vendría dada por $(-F, 0)$. Por tanto, en oposición a lo que acaecía con los escalares, un vector asume diferentes *representaciones coordenadas* según el observador, y por eso no es un invariante absoluto. Ahora bien, el vector, como *entidad geométrica*, es invariablemente el mismo, con completa independencia de quien lo observe. Lo único que está cambiando es la perspectiva del observador.

masa en reposo para diferenciarla de la relativista) se interpreta como una propiedad intrínseca de cada partícula.

Y esto se aplica igualmente a observadores que se mueven con velocidad constante. Por contra, la velocidad no es un vector. Prueba clara de ello es que (en mecánica prerrelativista) siempre es posible encontrar un observador respecto del cual la velocidad de una partícula concreta es cero, mientras que una magnitud tensorial, como la masa o la fuerza, o es nula en todos los sistemas de referencia o no lo es en ninguno. El que la representación en coordenadas se transforme de forma adecuada para preservar el vector como ser geométrico es lo que en física, en lugar de invariancia, se denomina *covariancia*.

¿Cómo entronca lo anterior con el tratamiento de escalares y vectores desde el punto de vista del álgebra lineal? Recordemos que los espacios vectoriales admiten múltiples bases. Con respecto de cada base, un mismo vector abstracto queda ligado a distintas representaciones coordenadas. Concretamente, sean $\{e_1, \ldots, e_n\}$ y $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$ dos bases de un espacio vectorial V. El vector $v \in V$ se escribe como combinación lineal de los elementos de las dos bases

$$v = \sum_{i=1}^{n} v^i e_i = \sum_{i=1}^{n} \tilde{v}^i \tilde{e}_i$$

Los números v^i y $\tilde{v}^i$ son las *coordenadas* de v con respecto a las bases $\{e_1, \ldots, e_n\}$ y $\{\tilde{e}_1, \ldots, \tilde{e}_n\}$, respectivamente. A pesar de que las coordenadas no coinciden, definen el mismo vector, pues este trasciende todo el asunto de bases y coordenadas. Los vectores en un espacio vectorial gozan de una existencia por completo independiente de que elijamos o no una base en la que representarlos. Como se verá, algo análogo ocurre con tensores de orden superior. Y esto nos recuerda inevitablemente a las magnitudes físicas que son como A. Ciertamente, *cada observador o sistema de referencia inercial lleva asociado de forma natural un conjunto de ejes, esto es, una base de un espacio vectorial que modeliza el espacio o espaciotiempo físicos*. Las magnitudes como A, que son independientes del observador, es decir, de la base, *están pidiendo a gritos ser encarnadas en el modelo matemático por escalares, vectores, tensores de orden superior y demás objetos geométricos*, que por definición incorporan la propiedad fundamental de existir más allá de los sistemas de referencia.

En un espacio vectorial, los elementos del cuerpo se comportan como escalares bajo cambios de base. ¿Qué números hay en un espacio vectorial que no permanecen invariantes al pasar de una base a otra? Claramente, las coordenadas (o componentes) de un vector cualquiera. Si se toma una de estas componentes por separado, mutará con el cambio de base. Estas consideraciones podrían arrojar luz sobre la naturaleza de las magnitudes físicas que son como B. ¿Tal vez la energía cinética, por ejemplo, sea en verdad la componente de un vector y por eso no se preserva al cambiar de observador? La respuesta es afirmativa. Energía y cantidad de movimiento constituyen en el ámbito de la relatividad especial un solo vector (de cuatro componentes). De igual modo tiene lugar con la longitud de los cuerpos o las separaciones temporales entre eventos. De los postulados de la relatividad se desprende que tanto las distancias espaciales como las temporales no son sino componentes de un objeto geométrico que combina espacio y tiempo, *el intervalo espaciotemporal*. Y de ahí que tengan lugar la contracción de longitudes y la dilatación de tiempos, que no son otra cosa que la transformación de las coordenadas de un vector (de cuatro componentes) cuando se modifica la base. Puede verse que esta transformación es tal que si la velocidad relativa de los observadores inerciales es despreciable frente a la de la luz, como es propio de nuestra vida cotidiana, longitudes y tiempos son esencialmente invariantes y se recupera la relatividad galileana.

Para cerrar esta introducción, presentaremos un caso especialmente relevante de tensor de segundo orden en física. Todo aquel que ha estudiado los rudimentos de electromagnetismo en el instituto o los primeros cursos del grado ha debido aceptar que el *campo eléctrico* y el *campo magnético* son magnitudes vectoriales. Ocasionalmente, el aprendizaje, sobre todo si requiere elevar el grado de abstracción, como acostumbra a pasar en matemáticas y física, procede por medio del olvido[3]. En verdad, en verdad os digo: ¡Los campos eléctrico y magnético no son vectores! Y lo pondremos de manifiesto a través de un sencillo *gedankenexperiment*[4], de esos que tanto gustaban a Einstein. Imaginemos una carga eléctrica que se desplaza con velocidad constante. Por poseer carga eléctrica, esta

[3]Decía Platón en su teoría de la reminiscencia que *aprender es recordar*. Aquí estamos sugiriendo que, al menos a veces, *aprender es olvidar*.

[4]Experimento mental.

produce en sus inmediaciones un campo eléctrico. Y por poseer carga eléctrica y además hallarse en movimiento, genera asimismo un campo magnético. ¡O eso experimenta un observador en las citadas condiciones! Invoquemos un segundo sistema de referencia inercial que se mueve a la misma velocidad que la carga. Con respecto de este, la carga está en reposo. Y así es como el fantasmagórico campo magnético se desvanece sin más que alterar nuestra cinemática. La conclusión es traumática, pero indubitable: *El campo magnético no contiene realidad física*. No hay *un* campo magnético si el experimentador de turno no lo mide con un teslámetro, exactamente de la misma forma que un vector no recibe *un* conjunto de coordenadas hasta que se fija una base. El campo eléctrico en esta situación tampoco se transforma como un vector, dado que el campo eléctrico de una carga en movimiento es *físicamente distinto* del de una carga en reposo. Mostrar que el campo eléctrico puede anularse mediante un cambio de observador es más arduo, dada la hipotética inexistencia de monopolos magnéticos. Este tipo de cavilaciones condujo a Einstein a su teoría de la relatividad especial, y es por eso que el artículo en el que la postuló se titula *Sobre la Electrodinámica de los Cuerpos en Movimiento*.

¡Mas algún atisbo de realidad física habrán de esconder los fenómenos electromagnéticos! En efecto, pero dicha realidad nace de la simbiosis de los campos eléctrico y magnético. Son dos caras de una única moneda, *el campo electromagnético*. Mencionamos anteriormente que la energía o la longitud eran componentes de una estructura geométrica mayor, en su caso sendos vectores. Se arguye que los campos eléctrico y magnético son igualmente las componentes de un *tensor de orden dos antisimétrico* que en coordenadas respecto de un observador-base adquiere el aspecto de una matriz cuadrada

$$\begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{pmatrix}$$

siendo E_i, B_i las componentes de los campos eléctrico y magnético, respectivamente.

2. Tensores en un espacio vectorial

2.1. Comentarios previos

Usaremos y abusaremos del convenio de suma de Einstein, que consiste en omitir todo sumatorio que se preste a omisión (hay quien va por ahí enseñando que uno de los principios no escritos de todo físico es que cuanto menos tinta haya que derramar, mejor). Porque sí, a veces se peca de omisión y otras veces, de redundancia y petulancia (como podría usted pensar de mí mientras lee este ampuloso y a veces inane texto). En fin, que, en el trajín tensorial, (casi) siempre que hay índices repetidos, uno arriba y otro abajo, se están sumando desde 1 hasta la dimensión del espacio vectorial. Solo con percibir esta disposición indicial basta para entender que hay una suma, y quitando la Σ nos ahorramos un poco de tinta y/o esfuerzo manual y otro tanto de fotones negros (o del color que a vd. más le guste) que salen del sumatorio e impactan contra nuestra retina (en realidad no, no existen los fotones negros). Escribiremos, pues

$$\sum_i v^i e_i \equiv v^i e_i$$

O, incluso

$$\sum_{i_1,\ldots,i_r,j_1,\ldots,j_s} T^{i_1\ldots i_r}{}_{j_1\ldots j_s}\, e_{i_1}\otimes\cdots\otimes e_{i_r}\otimes e^{j_1}\otimes\cdots\otimes e^{j_s} \equiv T^{i_1\ldots i_r}{}_{j_1\ldots j_s}\, e_{i_1}\otimes\cdots\otimes e_{i_r}\otimes e^{j_1}\otimes\cdots\otimes e^{j_s}$$

Nos limitaremos a lo largo del tema al estudio de tensores sobre espacios vectoriales reales, aunque todas las nociones presentadas pueden extenderse directamente a espacios vectoriales sobre otros cuerpos, como el de los números complejos.

Con la notación (g_{ij}), siendo g_{ij} una colección de números, aludiremos a la matriz cuya entrada $i, j-$ésima es g_{ij}.

2.2. Espacio vectorial dual

Antes de enunciar la definición general de *tensor*, consideraremos algunos casos particulares que ya nos son familiares. Todas las estructuras que se tratan a continuación están edificadas sobre un espacio vectorial abstracto V de dimensión finita y sobre su espacio dual. A pesar de que el estudio de estos conceptos pertenece a la asignatura Álgebra Lineal y Geometría I, damos aquí, en afán de hacer de esta una referencia completa, las definiciones y resultados que más adelante necesitaremos en la construcción de tensores.

Definición 2.2.1 *Sea V un espacio vectorial real. El espacio vectorial dual de V es el conjunto*

$$V^* \equiv \{\alpha : V \longrightarrow \mathbb{R} \text{ lineal}\}$$

dotado de las operaciones suma $(+)$ y producto por escalar $(\cdot)$ dadas por

$$(\alpha + \beta)(v) \equiv \alpha(v) + \beta(v) \qquad \alpha, \beta \in V^*, v \in V$$

$$(\lambda\alpha)(v) \equiv \lambda\alpha(v) \qquad \alpha, \in V^*, v \in V,\ \lambda \in \mathbb{R}$$

El espacio dual se compone, por tanto, de *aplicaciones lineales* que engullen vectores y excretan números reales[5][6]. Es fácil ver que la anterior estructura satisface todos los axiomas de la definición de espacio vectorial. De hecho, si $\dim V = n$, entonces $\dim V^* = n$, como probamos a continuación:

Proposición 2.2.1 *Sea $B = \{e_1, \ldots, e_n\}$ una base de V. Existe una (y solo una) base de V^*, $B^* = \{e^1, \ldots, e^n\}$, cuyos elementos verifican*

$$e^i(e_j) = \delta^i_{\ j}$$

[5] Más generalmente, los elementos del espacio dual son aplicaciones lineales de V en K, donde K es el cuerpo sobre el que se sustenta el espacio vectorial.

[6] Es habitual encontrar en libros y otros textos la aplicación

$$\langle \cdot , \cdot \rangle : V^* \times V \longrightarrow \mathbb{R}$$

$$\alpha, v \longmapsto \langle \alpha, v \rangle \equiv \alpha(v)$$

bajo el nombre de *natural pairing*. Además de ser lineal en sus dos entradas, es *no degenerada*, es decir, $\forall x \in V : \langle \alpha, x \rangle = 0 \Longleftrightarrow \alpha = 0$ y $\forall \beta \in V^* : \langle \beta, v \rangle = 0 \Longleftrightarrow v = 0$.

En consecuencia, $\dim V^* = \dim V$. B^* *se llama base dual de* B

Demostración: En general, una aplicación lineal $f : V \longrightarrow W$ queda completamente especificada cuando se da su imagen bajo los elementos de una base cualquiera de V. Así es precisamente como estamos definiendo los objetos e^i, lo que establece su existencia y unicidad.

Queda comprobar que B^* es una base. En primer lugar, veamos que es sistema de generadores. Tomamos $\alpha \in V^*$ y $v \in V$. En la base B, v se expande como

$$v = v^i e_i$$

entonces, aprovechando la linealidad de α

$$\alpha(v) = \alpha(v^i e_i) = v^i \alpha(e_i)$$

Llamando $\alpha_i \equiv \alpha(e_i)$, tenemos

$$\alpha(v) = \alpha_i v^i = \alpha_i e^i(v)$$

ya que

$$e^i(v) = e^i(v^j e_j) = v^j e^i(e_j) = v^j \delta^i{}_j = v^i$$

Como v, α son arbitrarios, hemos deducido que

$$\forall \alpha \in V^* : \alpha = \alpha_i e^i$$

lo que certifica que B^* es un sistema de generadores de V^*. Para finalizar, B^* es un sistema libre, i.e. está formado por vectores linealmente independientes. En efecto, supongamos que

$$\beta = \lambda_i e^i = 0 \qquad \lambda_i \in \mathbb{R}$$

Evaluando β en los elementos de la base B resulta $\beta(e_j) = 0 \Longrightarrow \lambda_j = 0 \; \forall j$ ■

A continuación, estudiaremos el cambio de base en el espacio dual. Recordemos que, dadas dos bases $B = \{e_i\}, \tilde{B} = \{\tilde{e}_j\}$ de V, las coordenadas de un mismo vector $x \in V$

respecto de las bases $B, \tilde{B}$ están relacionadas por medio de la *matriz de cambio de base* $\mathscr{M}(B, \tilde{B})$ como sigue

$$\begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix} = \mathscr{M}(B, \tilde{B}) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

o bien, llamando a^i_j a las entradas de la matriz $\mathscr{M}(B, \tilde{B})$

$$\tilde{x}^i = a^i_j x^j$$

Se tiene el resultado:

Proposición 2.2.2 *Sean $B, \tilde{B}$ dos bases de V y $B^*, \tilde{B}^*$, las correspondientes bases duales. La relación entre las coordenadas de una forma $\alpha \in V^*$ respecto de sendas bases viene dada por*

$$\alpha_i = a^j_i \tilde{\alpha}_j$$

donde a^j_i son las entradas de $\mathscr{M}(B, \tilde{B})^t$

Demostración: Denotemos $B^* = \{e^i\}_i$, $\tilde{B}^* = \{\tilde{e}^j\}_j$. La igualdad

$$\tilde{x}^j = a^j_i x^i$$

es equivalente a (véase la demostración de la proposición anterior)

$$\tilde{e}^j(x) = a^j_i e^i(x)$$

Dado que esta relación es cierta $\forall x \in V$, se sigue la igualdad entre aplicaciones

$$\tilde{e}^j = a^j_i e^i$$

Tomamos ahora una forma $\alpha \in V^*$. Como buen objeto geométrico, ha de ser independiente de las coordenadas elegidas para expresarlo. Por tanto

$$\alpha = \alpha_i e^i = \tilde{\alpha}_j \tilde{e}^j$$

Utilizando la relación entre e^i y $\tilde{e}^j$

$$\alpha_i e^i = \tilde{\alpha}_j \tilde{e}^j = \tilde{\alpha}_j a^j_i e^i$$

luego

$$(\alpha_i - \tilde{\alpha}_j a_i^j)e^i = 0$$

Ahora bien, los n vectores e^i son linealmente independientes, con lo cual

$$\alpha_i - \tilde{\alpha}_j a_i^j = 0 \Longrightarrow \alpha_i = a_i^j \tilde{\alpha}_j$$

■

A priori, los constituyentes de V y V^* no parecen tener demasiado en común. Es cierto que los dos objetos son vectores en el sentido abstracto. Sin embargo, las formas son aplicaciones lineales y los elementos de V, no. ¿O quizá sí? Intuitivamente, es fácil ver que un vector (de V) adquiere de manera natural el carácter de aplicación lineal en presencia del espacio dual. Sean $x \in V, \alpha \in V^*$. Definimos

$$x(\alpha) \equiv \alpha(x) \in \mathbb{R}$$

Así, hemos construido una aplicación $x : V^* \longrightarrow \mathbb{R}$ bien definida y manifiestamente lineal, ya que, si $\alpha, \beta \in V^*,\ a, b \in \mathbb{R}$

$$x(a\alpha + b\beta) = (a\alpha + b\beta)(x) = a\alpha(x) + b\beta(x) = ax(\alpha) + bx(\beta)$$

Siendo algo más rigurosos, lo que acabamos de hacer no es sino identificar $x \in V$ con un elemento del dual del espacio dual, conocido como *espacio bidual*, $x \in V^{**}$, quien, por definición, es una aplicación lineal $x : V^* \longrightarrow \mathbb{R}$. Esto es posible gracias a la existencia de un *isomorfismo lineal* entre V y V^{**} que, además, resulta ser *canónico* o *natural*, es decir, su construcción no requiere especificar bases y coordenadas (a diferencia del isomorfismo lineal *cartesiano* entre dos espacios vectoriales de igual dimensión). V y V^{**} son, a nivel de espacios vectoriales, estructuralmente idénticos, lo que legitima hablar indistintamente de vectores o formas lineales sobre el espacio dual.

Proposición 2.2.3 *(Teorema de reflexividad) La aplicación $\Phi : V \longrightarrow V^{**}$ dada por*

$$[\Phi(x)](\alpha) \equiv \alpha(x)$$

es un isomorfismo lineal.

Demostración:

1. ¿Φ *está bien definida?* ¿Hace corresponder un elemento de V a un solo objeto del bidual? Afirmativo, pues la imagen de $\Phi(x) \in V^{**}$ sobre cualquier $\alpha \in V^*$ queda unívocamente especificada.

2. ¿Φ *es lineal?* Afirmativo. Véase la discusión previa al enunciado del teorema.

3. ¿Φ *es inyectiva?* Afirmativo, pues

$$\Phi(x) = 0 \iff \forall\alpha : [\Phi(x)](\alpha) = 0 \iff \forall\alpha : \alpha(x) = 0 \iff x = 0$$

se sigue que $\ker\Phi = \{0\}$, lo que, para aplicaciones lineales, equivale a la inyectividad.

4. ¿Φ *es sobreyectiva?* Afirmativo. En virtud de la fórmula de las dimensiones

$$\dim V = \dim\ker\Phi + \dim \mathrm{Im}\Phi$$

Del tercer punto se desprende que $\dim\ker\Phi = 0$. Además, la dimensión de un espacio vectorial finitamente generado coincide con la de su despacio dual (proposición 2.2.1), y también con la de su espacio bidual, ya que este es el dual del dual del susodicho espacio. De este modo

$$\dim V^{**} = \dim \mathrm{Im}\Phi$$

Como quiera que $\mathrm{Im}\Phi$ es un subespacio vectorial de V^{**} con la misma dimensión que el bidual, ha de ser $\mathrm{Im}\Phi = V^{**}$, esto es, la aplicación es sobreyectiva.

En suma, Φ es un isomorfismo lineal, y además es canónico pues su construcción no supone especificar una base ■

Del teorema de reflexividad se desprende que todo vector $v \in V$ es *naturalmente equivalente* a $\Phi(v) \in V^{**}$. En lo que sigue, abusaremos de la notación remplazando $\Phi(v)$ por v.

2.3. Concepto abstracto de tensor

A la luz de las elucubraciones expuestas en la sección anterior, sabemos de la existencia de dos tipos de entes fundamentales:

- Los elementos del espacio vectorial o *vectores*, que por medio del isomorfismo de dualidad (proposición 2.2.3) se identifican con aplicaciones lineales $v : V^* \longrightarrow \mathbb{R}$.
- Los elementos del espacio dual, conocidos como *formas lineales* o *covectores*[7], es decir, aplicaciones lineales $\alpha : V \longrightarrow \mathbb{R}$.

Uno de los recursos por el que la matemática (amén de la física) prospera es la unificación de varias nociones aparentemente dispares en una sola. En lugar de construir sendas teorías matemáticas para vectores y covectores, resulta mucho más económico e iluminador percatarse de que estos son casos particulares de una estructura más general, los *tensores*, cuya definición conoceremos en breve. Es pues deseable elaborar una teoría que abarque todas las clases de tensores, ganando en potencia y generalidad y ahorrando en papel y tinta.

Muy resumidamente, el proceso intelectivo por el cual se da el salto de vectores y covectores a tensores es el siguiente. Cambalacheando con los espacios vectoriales hemos hallado aplicaciones lineales[8] que mandan elementos de V y de V^* a números reales. ¿Por qué no considerar entonces aplicaciones lineales que toman *varios* elementos de V y devuelven un real, i.e. cuyo dominio es el producto cartesiano $V \times \cdots \times V$? ¿O bien, aplicaciones lineales que se evalúan en cierto número de covectores para dar un número, o sea, con dominio $V^* \times \cdots \times V^*$? Y ya que estamos en ello, ¿por qué no combinar estos dos casos y considerar dominios dados por productos de unos pocos V y otros tantos V^*? He aquí en qué consiste un tensor: no es sino un mapa que se abastece de s vectores y r covectores para producir un número real y que es lineal en cada una de las entradas.

[7]No olvidemos que *los covectores son también vectores* en tanto que pertenecen a un espacio vectorial, el espacio dual.

[8]Más técnicamente, se habla de *aplicaciones multilineales*, aludiendo a aplicaciones con más de una entrada que son lineales en todas ellas.

 ✉ f42rodim@uco.es

Específicamente, tal objeto constituye un tensor $r-$contravariante y $s-$covariante, o, de forma abreviada, un tensor de tipo (r,s).

Evidentemente, lo que motivó la introducción de los tensores desde un punto de vista histórico es muchísimo más complejo que la cadena argumentativa del anterior párrafo. Antes de enunciar la definición general de tensor ya habían aparecido en matemáticas otros ejemplos de tensores, siendo especialmente prominente el formidable tensor 2-covariante que es el *producto escalar*, el cual presentaremos posteriormente, así como varios objetos físicos semejantes a los tensores matemáticos como el tensor de esfuerzos o el tensor de inercia[9]. Si bien la historia detrás de la formalización de los tensores en matemáticas y su ingente aplicación a la física es compleja e interesante, no nos detendremos aquí a tratarla so pena de desviarnos sensiblemente de nuestra meta, que es definirlos y estudiarlos:

Definición 2.3.1 *Sea V un espacio vectorial real. Un tensor $r-$contravariante, $s-$covariante es una aplicación*

$$T : V^* \times \ldots^{(r)} \times V^* \times V \times \ldots^{(s)} \times V \longrightarrow \mathbb{R}$$

$$\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s \longmapsto T(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \in \mathbb{R}$$

que es lineal en todos sus argumentos, esto es, dados $\alpha, \beta \in V^*$, $\lambda, \mu \in \mathbb{R}$, para cualquier argumento contravariante

$$T(\alpha^1, \ldots, \lambda\alpha + \mu\beta, \ldots, \alpha^r, v_1, \ldots, v_s) = \lambda T(\alpha^1, \ldots, \alpha, \ldots, \alpha^r, v_1, \ldots, v_s) + $$
$$+ \mu T(\alpha^1, \ldots, \beta, \ldots, \alpha^r, v_1, \ldots, v_s)$$

y, dados $v, w \in V$, para cualquier argumento covariante

$$T(\alpha^1, \ldots, \alpha^r, v_1, \ldots, \lambda v + \mu w, \ldots v_s) = \lambda T(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v, \ldots v_s) + $$
$$+ \mu T(\alpha^1, \ldots, \alpha^r, v_1, \ldots, w, \ldots, v_s)$$

[9]Nótese que se ha empleado el sintagma adjetival *semejantes a los tensores matemáticos* para describir al tensor de esfuerzos y al tensor de inercia, que, a pesar de incluir el término *tensor* en sus títulos, no son, en un sentido estricto, tensores.

Por muy abstracta que pueda parecer la definición 2.3.1, lo cierto es que hemos vivido rodeados de tensores sin apenas si ser conscientes de ello. Enumeramos algunos ejemplos de tensores cotidianos:

- Las formas lineales o covectores, al ser aplicaciones lineales $\alpha : V \longrightarrow \mathbb{R}$, se corresponden con tensores $1-$covariantes (y $0-$contravariantes) o de tipo $(0, 1)$, pues entre sus argumentos se cuentan un vector y cero covectores.

- Los vectores, al ser aplicaciones lineales $\alpha : V^* \longrightarrow \mathbb{R}$, se corresponden con tensores $1-$contravariantes (y $0-$covariantes) o de tipo $(1, 0)$, pues entre sus argumentos se cuentan un covector y cero vectores.

- Los escalares o elementos del cuerpo $\mathbb{R}$ son tensores de tipo $(0, 0)$, ya que no necesitan ser evaluados en ningún vector o covector para producir un número real. Al menos, se toma así por convenio.

- Vamos con ejemplos algo más elaborados. En el contexto del álgebra y geometría lineales es habitual enfrentarse a *formas bilineales*, que son aplicaciones

$$g : V \times V \longrightarrow \mathbb{R}$$

  respetando linealidad en sus dos entradas. De la definición 2.3.1 tenemos que las formas bilineales son exactamente lo mismo que los tensores 2-covariantes o de tipo (0,2). Entre todas las formas bilineales destacan los *productos escalares*, tensores 2-covariantes simétricos y no degenerados[10]. Daremos buena cuenta de los productos escalares más adelante. Sí nos detenemos a mencionar que un tensor se dice *simétrico* en dos argumentos dados del mismo tipo (dos lugares ocupados por vectores o por formas lineales)[11] si la imagen no se ve afectada por intercambiar los inputs alojados en dichas entradas. Por ejemplo, un tensor 2-covariante (o forma

[10]La definición de producto escalar que proporcionamos aquí es propia del ámbito de la geometría. En análisis, entre otros, se acostumbra a pedir que g sea además definido positivo.

[11]Evidentemente, no tiene sentido hablar de simetría entre una entrada vectorial y una entrada covectorial. Así, un tensor de tipo $(1, 1)$, que es una aplicación $T : V^* \times V \longrightarrow \mathbb{R}$, nunca podrá ser simétrico porque $T(\alpha, v)$ tiene sentido pero $T(v, \alpha)$, no, siendo $v \in V, \alpha \in V^*$.

bilineal) T es simétrico (en las dos únicas entradas que tiene) si $T(v,w) = T(w,v)$ para cualesquiera vectores v, w.

- El *determinante* de k vectores, $\det(v_1, \ldots, v_k)$, conocido desde los estudios de bachillerato, es un tensor $k-$covariante al alimentarse de k vectores y ser lineal en todas sus entradas. Más aún, se trata de un tensor *completamente antisimétrico*. Para empezar, un tensor se dice *antisimétrico* en dos entradas de igual tipo si siempre que se intercambian sus elementos de entrada el resultado sufre un cambio de signo. Para el caso particular del tensor 2-covariante T, esto se traduce en que $T(v,w) = -T(w,v)$ para cualesquiera vectores v, w. Si un tensor cuyas entradas son de igual tipo es antisimétrico en todas las parejas de entradas, se le llama *completamente antisimétrico*, y esta es la propiedad que caracteriza al determinante. Estudiaremos asimismo los tensores $k-$covariantes[12] completamente antisimétricos con mayor fruición en un apartado futuro.

- Aportamos, para terminar, un ejemplo de tensor *mixto*, i.e. que tiene al menos una entrada covariante y otra contravariante. Sea $f : V \longrightarrow V$ un endomorfismo (una aplicación lineal cuyos dominio y codominio coinciden). Se define

$$T_f : V^* \times V \longrightarrow \mathbb{R}$$

$$T_f(\alpha, v) \equiv \alpha[f(v)]$$

  y se comprueba fácilmente que T_f es lineal en sus dos argumentos. Por consiguiente, T_f es un tensor de tipo $(1,1)$. Más aún, se demuestra que el mapa que asigna f a T_f es un isomorfismo lineal (canónico), por lo que especificar un endomorfismo de V es equivalente a dar un tensor $(1,1)$ sobre V.

Por otra parte, existen cuantiosos tensores relacionados con nuestra modelización de la naturaleza. Nombramos unos pocos para abrir boca al lector de inclinación física. Adelantamos que, en un contexto físico, en muchos casos (mas no siempre) el espacio

[12] Se podría efectuar un estudio absolutamente análogo para los tensores $k-$contravariantes antisimétricos, pero, por costumbre (y porque el determinante, que es un objeto con mucha presencia, es covariante), se propende a trabajar con el otro caso.

vectorial está dotado de un producto escalar, y, en presencia de este, surgen isomorfismos lineales canónicos que facultan la identificación de tensores de distinto tipo pero igual *orden*, que será como nos referiremos al número de argumentos de un tensor. Verbigracia, un tensor de tipo $(2,3)$ es de orden $2+3=5$. Como decíamos, a través de estos isomorfismos, a cada tensor de tipo $(0,2)$ se asocia uno de tipo $(1,1)$ y otro de tipo $(2,0)$, puesto que todos ellos poseen el mismo orden, y así más generalmente. Por este motivo, en física se suele establecer el orden de un tensor pero no su tipo.

- El *tensor de Faraday* o *tensor de campo electromagéntico* $\mathscr{F}$ es un tensor de orden 2 que unifica los campos eléctrico y magnético (entes no tensoriales, no físicos). Uno puede expresar todo el electromagnetismo clásico (en particular, las ecuaciones de Maxwell-Hertz) en clave de $\mathscr{F}$, construyendo de este modo la *formulación (manifiestamente) covariante* del electromagnetismo.

- Es sobradamente sabido que la curvatura del espaciotiempo es el trasunto de la interacción gravitacional en relatividad general. Esta curvatura viene codificada por el *tensor de Riemann* (o *tensor de curvatura*), que resulta ser un tensor de orden 4 y que típicamente se encuentra en sus formas $(1,3)$ o $(0,4)$.

- A partir del tensor de Riemann, mediante un proceso conocido como *contracción* que definiremos posteriormente, se construye el *tensor de Ricci* Ric, un tensor de orden 2, en principio de tipo $(0,2)$.

- Combinando el tensor de Ricci con un segundo tensor $2-$covariante, el producto escalar g, definimos
$$\mathscr{G} = \mathrm{Ric} - \frac{1}{2}Sg$$
donde S es el *escalar de curvatura*, resultante de aplicar la contracción al tensor de Ricci. $\mathscr{G}$ recibe el nombre del tensor de Einstein, la pieza fundamental en sus ecuaciones de campo de la relatividad general.

- Si el tensor de Einstein solo contiene información relativa a la geometría del espaciotiempo, el otro miembro de las ecuaciones de Einstein, el *tensor de energía-impulso* $\mathscr{T}$, es también un tensor de orden 2 y depende exclusivamente del con-

 ✉ f42rodim@uco.es

tenido energético (materia, presiones, etc.) del espaciotiempo en cada punto. La ecuación de Einstein es una *ecuación de fuentes* para la geometría del espaciotiempo, siendo las fuentes estos fenómenos energéticos. En términos más mundanos, la existencia de cualquier cosa con energía en un punto altera la geometría en torno a ese punto. Esta ecuación se escribe

$$\mathscr{G} = \kappa \mathscr{T}$$

siendo κ una constante de proporcionalidad que depende de la constante de gravedad de Newton y de la velocidad de la luz.

2.4. Bases y coordenadas en espacios de tensores

Notaremos $\mathscr{T}_s^r(V)$ al conjunto de tensores de tipo (r,s) sobre V. Tratándose de casos particulares de espacios de aplicaciones lineales, los elementos de un $\mathscr{T}_s^r(V)$ pueden sumarse y multiplicarse por elementos del cuerpo $\mathbb{R}$, y estas operaciones respetan todas las operaciones propias de un espacio vectorial. Por tanto:

Proposición 2.4.1 *Para cada $r, s \geq 0$, la estructura $(\mathscr{T}_s^r(V), \mathbb{R}, +, \cdot)$, siendo $+$ la suma de tensores y $\cdot$ el producto de un tensor por un elemento del cuerpo, es un espacio vectorial.*

Sabido esto[13], lo natural a continuación es preguntarse cuál es la dimensión de estos espacios vectoriales y si podemos proporcionar una base explícita para los mismos. Antes de proceder con este ejercicio, introduciremos la operación básica entre tensores generales:

Definición 2.4.1 *Sean $T \in \mathscr{T}_s^r(V), S \in \mathscr{T}_q^p(V)$. El producto tensorial de T y S es el tensor $T \otimes S \in \mathscr{T}_{s+q}^{r+p}(V)$ tal que*

$$(T \otimes S)(\alpha^1, \ldots, \alpha^r, \alpha^{r+1}, \ldots, \alpha^{r+p}, v_1, \ldots, v_s, v_{s+1}, \ldots, v_{s+q}) = T(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) \cdot$$
$$\cdot S(\alpha^{r+1}, \ldots, \alpha^{r+p}, v_{s+1}, \ldots, v_{s+q})$$

[13] De nuevo, los elementos de cada $(\mathscr{T}_s^r(V), \mathbb{R}, +, \cdot)$ (por definición, tensores de tipo (r,s)) son al mismo tiempo vectores abstractos por componer un espacio vectorial.

 ✉ f42rodim@uco.es

La definición puede parecer engorrosa, pero en el fondo es muy sencilla. Vamos a particularizarla al producto tensorial de dos covectores $\alpha, \beta \in V^* \equiv \mathcal{T}_1^0(V)$, que es el tensor 2−covariante $\alpha \otimes \beta \in \mathcal{T}_2^0(V)$. Dados dos vectores $v, w \in V$, se tiene

$$(\alpha \otimes \beta)(v, w) = \alpha(v) \cdot \beta(w)$$

Dicho en palabras, el producto tensorial de dos covectores evaluado en dos vectores da como resultado el producto del primer covector evaluado en el primer vector (que es un número real) por el segundo covector evaluado en el segundo vector (que es otro número real). Y esta misma idea se traslada a tensores de cualquier tipo. Sería menester demostrar que, tal y como afirmamos, esta operación establece un tensor como se dice en (2.4.1). Ahora bien, la linealidad de $\alpha \otimes \beta$ en cualquiera de sus dos entradas se desprende directamente de la definición, pues

$$(\alpha \otimes \beta)(\lambda_1 v_1 + \lambda_2 v_2, w) = \alpha(\lambda_1 v_1 + \lambda_2 v_2) \cdot \beta(w) = \lambda_1 \alpha(v_1) \cdot \beta(w) + \lambda_2 \alpha(v_2) \cdot \beta(w) =$$
$$= \lambda_1 (\alpha \otimes \beta)(v_1, w) + \lambda_2 (\alpha \otimes \beta)(v_2, w)$$

siendo $v_1, v_2 \in V$ y $\lambda_1, \lambda_2 \in \mathbb{R}$, y del mismo modo para el segundo argumento. Queda probado que $\alpha \otimes \beta \in \mathcal{T}_2^0(V)$.

Ejemplo .1 No conmutatividad del producto tensorial

Tómense $V = \mathbb{R}^2$, una base cualquiera de este espacio vectorial $B = \{e_1, e_2\}$ así como su base dual $B^* = \{e^1, e^2\}$. Entonces

$$(e^1 \otimes e^2)(e_1, e_2) = e^1(e_1) \cdot e^2(e_2) = 1 \cdot 1 = 1$$

$$(e^1 \otimes e^2)(e_2, e_1) = e^1(e_2) \cdot e^2(e_1) = 0 \cdot 0 = 0$$

De este cálculo se sigue que *el producto tensorial no es conmutativo.*

Proposición 2.4.2 *Sean $B = \{e_1, \ldots, e_n\}$ una base de V y $B^* = \{e^1, \ldots, e^n\}$, su correspondiente base dual. El conjunto $B_{\mathcal{T}_2^0(V)} = \{e^i \otimes e^j\}_{i,j=1}^n$ es una base de $\mathcal{T}_2^0(V)$. Por tanto,* $\dim \mathcal{T}_2^0(V) = n^2$.

Demostración: Tomemos $T \in \mathscr{T}_2^0(V)$ y dos vectores $v, w \in V$ cuya expresión en coordenadas respecto de B es $v = v^i e_i, w = w^j e_j$. Se tiene

$$T(v,w) = T(v^i e_i, w^j e_j) = v^i w^j T(e_i, e_j) = T(e_i, e_j) e^i(v) e^j(w) = T(e_i, e_j)\,(e^i \otimes e^j)(v_i, w_j)$$

De este modo, cualquier tensor 2-covariante T es combinación lineal de los elementos de $B_{\mathscr{T}_2^0(V)}$. Denotando $T(e_i, e_j) \equiv T_{ij}$

$$T = T_{ij} e^i \otimes e^j$$

Queda probar que $B_{\mathscr{T}_2^0(V)}$, además de sistema generador, es libre. Para ello, consideremos una combinación lineal genérica de los elementos de dicho conjunto que haremos igual a cero

$$\lambda_{ij} e^i \otimes e^j = 0$$

Esto quiere decir que la aplicación del miembro izquierdo da cero cuando se evalúa en cualquiesquiera elementos de entrada. En particular, si introducimos e_k, e_l, tenemos

$$\lambda_{ij}(e^i \otimes e^j)(e_k, e_l) = \lambda_{ij} \delta^i_k \delta^j_l = \lambda_{kl} = 0$$

Y dado que esto es cierto $\forall k, l \in \{1, \ldots, n\}$, todos los coeficientes de la combinación lineal han de ser nulos, constatando la independencia lineal de los $e^i \otimes e^j$ ■

La anterior prueba se extiende directamente a tensores de tipo arbitrario[14]. Enunciamos

[14] Dejamos aquí la demostración explícita para el lector ávido por hacer malabarismos con índices (no tiene más interés). Tomemos un $T \in \mathscr{T}_s^r(V)$, vectores $v_k = v_k^{j_k} e_{j_k}$ con $k = 1, \ldots, s$ y covectores $\alpha^l = \alpha^l_{i_l} e^{i_l}$ con $l = 1, \ldots, r$. Vemos que

$$T(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s) = T(\alpha^1_{i_1} e^{i_1}, \ldots, \alpha^r_{i_r} e^{i_r}, \ldots, v_1^{j_1} e_{j_1}, \ldots, v_s^{j_s} e_{j_s}) = \alpha^1_{i_1} \ldots \alpha^r_{i_r} v_1^{j_1} \ldots v_s^{j_s}\, T(e^{i_1}, \ldots, e^{i_r}, \ldots, e_{j_1}, \ldots, e_{j_s}) =$$
$$= T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}\,(e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s})(\alpha^1, \ldots, \alpha^r, v_1, \ldots, v_s)$$

por lo que $B_{\mathscr{T}_s^r(V)}$ es sistema generador de $T \in \mathscr{T}_s^r(V)$. En segundo lugar, sea una combinación lineal arbitraria Λ de elementos de $B_{\mathscr{T}_s^r(V)}$, que hacemos nula

$$\Lambda = \lambda^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}\, e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s} = 0$$

Entonces

$$\Lambda(e^{k_1}, \ldots, e^{k_r}, e_{l_1}, \ldots, e_{l_s}) = \lambda^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}\,(e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s})(e^{k_1}, \ldots, e^{k_r}, e_{l_1}, \ldots, e_{l_s}) =$$
$$= \lambda^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}\, e_{i_1}(e^{k_1}) \ldots e_{i_r}(e^{k_r}) e^{j_1}(e_{l_1}) \ldots e^{j_s}(e_{l_s}) = \lambda^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \delta^{k_1}_{i_1} \ldots \delta^{k_r}_{i_r} \delta^{j_1}_{l_1} \ldots \delta^{j_s}_{l_s} = \lambda^{k_1 \ldots k_r}{}_{l_1 \ldots l_s} = 0$$

el resultado general

Proposición 2.4.3 *Sean* $B = \{e_1, \ldots, e_n\}$ *una base de* V *y* $B^* = \{e^1, \ldots, e^n\}$, *su correspondiente base dual. El conjunto*

$$B_{\mathscr{T}_s^r(V)} = \{e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}\}_{i_1,\ldots,i_s,j_1,\ldots,j_r=1}^{n}$$

es una base de $\mathscr{T}_s^r(V)$, *esto es, un tensor de tipo* (r,s), $T \in \mathscr{T}_s^r(V)$, *se expresa como*

$$T = T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}\, e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes e^{j_1} \otimes \cdots \otimes e^{j_s}$$

de manera única, siendo $T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} \equiv T(e^{i_1}, \ldots, e^{i_r}, e_{j_1}, \ldots, e_{j_s})$. *Por tanto,* $\dim \mathscr{T}_s^r(V) = n^{r+s}$.

¿Qué denominación debemos otorgar a los coeficientes $T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}$? Si nos restringimos al caso de tensores $(1,0)$[15], que hemos venido llamando tradicionalmente vectores v, la anterior proposición establece que $v = v^i e_i$. Los v^i no son sino las coordenadas de v en la base $B = \{e_1, \ldots, e_n\}$, y, siendo que cumplen la misma función, adoptaremos la misma nomenclatura para el resto de tensores:

Definición 2.4.2 *Sean* $T \in \mathscr{T}_s^r(V)$ *y* $B = \{e_1, \ldots, e_n\}$, *una base de* V. *Los números* $T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s} = T(e^{i_1}, \ldots, e^{i_r}, e_{j_1}, \ldots, e_{j_s})$ *se denominan coordenadas de* T *en la base* B.

Tal y como se anticipaba en el párrafo previo, no hay diferencia conceptual alguna entre las coordenadas de un vector (idea bien digerida en Álgebra Lineal y Geometría I) y de un tensor. Fijada una base, un tensor de orden (r,s) queda asociado a una colección de n^{r+s} números, sus coordenadas $T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}$, que pueden emplearse equivalentemente para efectuar operaciones en sustitución del tensor abstracto, **siendo conscientes de que la representación en coordenadas de un tensor va inextricablemente ligada a la elección de una base**. En analogía a lo que sucedía con los vectores, si se toman dos bases distintas del espacio vectorial resultan dos conjuntos de coordenadas no coincidentes para un

y como esto se cumple para cualesquiera $k_1, \ldots, k_r, l_1, \ldots, l_s = 1, \ldots, n$, encontramos que $\Lambda = 0$, es decir, $B_{\mathscr{T}_s^r(V)}$ es un sistema libre ■

[15] La proposición 2.4.3 no aporta información nueva cuando se aplica a tensores $(1,0)$, ya que la condición de que todo vector $v = v^i e_i$ de manera única no deja de ser la definición de base.

mismo tensor. Si no se especifica una base, no procede hablar de coordenadas, a pesar de que el tensor en sí no precisa de bases. Por eso, la existencia de la que gozan las coordenadas en el universo algebraico es un tanto ficticia, consecuencia de una descripción que introducimos externamente, al igual que en nuestra vida cotidiana hablamos de norte, sur, este y oeste, con el propósito de hacer indicaciones respecto de unas direcciones de referencia concretas, si bien la naturaleza carece de un norte o un este bien definidos: todas las direcciones son equivalentes. Las coordenadas emanan de nuestra injerencia, imperfecta y artificial. Son la fruta prohibida del Edén platónico, ese proyecto de Idea desterrado a la caverna que tendemos a confundir con los genuinos tensores cuando no son otra cosa que las veleidosas sombras que estos proyectan sobre la piedra. Pero como todo pecado, tentador y placentero, no hay físico que no sucumba al uso frecuente de coordenadas y sistemas de referencia so pena de ser incapaz de resolver los más sencillos problemas del Tipler. El lenguaje abstracto e independiente de bases, idioma natural de cualquier vaina tensorial, la *lingua franca* que conecta a los que hablan en la lengua de diferentes bases, en cuyo seno se gestan las definiciones y se demuestran los más bellos teoremas, no es el más apto para realizar cálculos en situaciones concretas. En esto, las coordenadas, pese a quien le pese, ostentan la hegemonía, y por eso a físicos e ingenieros se les inculca desde la más temprana infancia una cosmovisión fundamentada en coordenadas, y que los vectores son filas o columnas de números, y que los tensores son matrices, entre otras blasfemias. No es para nada execrable que los matemáticos se mofen, o más bien, sientan compasión por estas criaturas y su ciega fe en el dogma de las coordenadas y las bases, pues viven en un engaño sin saberlo.

Ejemplo .2 Un tensor en dos bases

Tómense $V = \mathbb{R}^2$ y una base cualquiera de este espacio vectorial $B = \{e_1, e_2\}$. Sea el tensor 2-contravariante

$$T = e_1 \otimes e_2$$

La expresión general de un tensor 2-contravariante en coordenadas con respecto a

✉ f42rodim@uco.es

B es

$$T = T^{ij} e_i \otimes e_j = T^{11} e_1 \otimes e_1 + T^{12} e_1 \otimes e_2 + T^{21} e_2 \otimes e_1 + T^{22} e_2 \otimes e_2$$

Comparando, deducimos las coordenadas de nuestro T en esta base

$$T^{12} = 1 \qquad\qquad T^{11} = T^{21} = T^{22} = 0$$

Sea $\tilde{B} = \{\tilde{e}_1, \tilde{e}_2\}$ una segunda base tal que

$$e_1 = \tilde{e}_1 + \tilde{e}_2 \qquad\qquad e_2 = \tilde{e}_1 - \tilde{e}_2$$

Estas últimas ecuaciones permiten encontrar las coordenadas de T respecto de $\tilde{B}$. Ciertamente, sustituyendo

$$T = e_1 \otimes e_2 = (\tilde{e}_1 + \tilde{e}_2) \otimes (\tilde{e}_1 - \tilde{e}_2) = \tilde{e}_1 \otimes \tilde{e}_1 - \tilde{e}_1 \otimes \tilde{e}_2 - \tilde{e}_2 \otimes \tilde{e}_1 + \tilde{e}_2 \otimes \tilde{e}_1$$

Y poniendo $T = \tilde{T}^{ij} \tilde{e}_i \otimes \tilde{e}_j$, llegamos a

$$\tilde{T}^{11} = \tilde{T}^{22} = 1 \qquad\qquad \tilde{T}^{12} = \tilde{T}^{21} = -1$$

Podemos encontrar las ecuaciones generales que relacionan las coordenadas de un tensor en dos bases. Sean, pues, dos bases $B = \{e_1, \dots, e_n\}$, $\tilde{B} = \{\tilde{e}_1, \dots, \tilde{e}_n\}$. Un vector $v \in V$ admite las expresiones

$$v = v^i e_i = \tilde{v}^i \tilde{e}_i$$

donde las coordenadas están relacionadas por la matriz de cambio de base $\mathscr{M}(B, \tilde{B})$, cuyas entradas denotamos $a^i{}_j$

$$\tilde{v}^i = a^i{}_j v^j$$

Equivalentemente, los elementos de las bases satisfacen la correspondencia *inversa*

$$e_j = a^i{}_j \tilde{e}_i \qquad\qquad \tilde{e}^i = a^j{}_i e^j$$

Tomemos ahora, como es costumbre para iniciarnos con un caso sencillo, un tensor 2-contravariante $S \in \mathscr{T}_0^2(V)$ (lo tomamos contravariante y no covariante porque estamos más habituados a la ley de transformación de un vector, o tensor 1-contravariante). Por definición, las coordenadas de S respecto de B son los números $S^{ij} = S(e^i, e^j)$, mientras que sus coordenadas con respecto de $\tilde{B}$ vienen dadas por $\tilde{S}^{ij} = S(\tilde{e}^i, \tilde{e}^j)$. De este modo

$$S(\tilde{e}^i, \tilde{e}^j) = S(a^i{}_k e^k, a^j{}_l e^l) = a^i{}_k a^j{}_l S(e^k, e^l)$$

que es lo mismo que

$$\tilde{S}^{ij} = a^i{}_k a^j{}_l S^{kl}$$

Vemos que la regla de transformación es semejante a la de los vectores, mas aparecen dos matrices cambio de base en lugar de una. Más generalmente, es inmediato ver que *habremos de introducir un factor $a^i{}_j$ (con los índices apropiados) por cada índice contravariante del que conste nuestro tensor*. Efectuando el mismo análisis para tensores covariantes, encontramos:

Proposición 2.4.4 *(Cambio de base) Sean $T \in \mathscr{T}_s^r(V)$ y $a^i{}_j$, $(a^i{}_j)^{-1}$, las componentes de las matrices $\mathscr{M}(B, \tilde{B})$, $\mathscr{M}(B, \tilde{B})^{-1}$, respectivamente. Se tiene*

$$\tilde{T}^{i_1 \dots i_r}{}_{j_1 \dots j_s} = a^{i_1}{}_{k_1} \dots a^{i_r}{}_{k_r} (a^{l_1}{}_{j_1})^{-1} \dots (a^{l_s}{}_{j_s})^{-1} T^{k_1 \dots k_r}{}_{l_1 \dots l_s} \tag{1}$$

Los resultados precedentes ponen de manifiesto la equivalencia entre la noción abstracta, matemática de tensor (a la que nos hemos adherido en el presente compuscrito) y la que se suele conocer como definición *física* o *clásica* de tensor, que típicamente figura en manuales tradicionales de mecánica clásica, teoría de la relatividad u otras disciplinas con elevado contenido tensorial explícito[16]. Esta versión alternativa, igual de plausible pero infinitamente menos elegante, ha venido satisfaciendo a los científicos de insuficiente inclinación matemática, que huyen despavoridos tras el mínimo contacto con un espacio vectorial o una aplicación lineal, como un químico espantado ante el descubrimiento de que la física y química cuánticas precisan del uso de números complejos[17]. La definición clásica se ampara única y exclusivamente en la particular transformación

[16] *Todas* las ramas de la física admiten una formulación tensorial y/o geométrica. Excluir una cierta teoría de esta caracterización supondría negar cualquier atisbo de contenido genuinamente físico en la misma.

[17] El autor da fe de haber sido testigo de cómo un conjunto de unos 50 alumnos de segundo de química se echó las manos a la cabeza al vislumbrar una unidad imaginaria en las transparencias de una asignatura de química cuántica.

que experimentan las componentes de un tensor bajo un cambio de base, sin indagar en su naturaleza algebraica. Además, este otro enfoque se presta a una interpretación más física, identificando *bases* con *sistemas de referencia* (que, en función del contexto, incluyen o no una dirección temporal junto con las direcciones espaciales necesarias). Podemos enunciarla tal que así:

Definición 2.4.3 *Un tensor $n-$dimensional, $r-$contravariante, $s-$covariante es una asignación de n^{r+s} números $T^{i_1\dots i_r}{}_{j_1\dots j_s}$ a cada sistema de referencia tales que la relación entre los conjuntos de números asociados a dos sistemas de referencia cualesquiera viene dada por (1).*

Una ventaja que ofrece esta reinterpretación, astutamente empleada en múltiples ramas de la física, es que preconiza la introducción de unos objetos más generales que los tensores: asignaciones de n^{r+s} números a cada base que no se corresponden con ningún tensor, porque no obedecen la regla (1) en cualquier circunstancia, pero que sí se comportan tensorialmente bajo un conjunto selecto de cambios de base o transformaciones entre sistemas de referencia. El ejemplo estándar es la construcción a la que en muchas referencias se alude como *tensor ortogonal* (o, en una rebelde economización del lenguaje que ha costado terriblemente cara, *tensor* asecas), y que no es otra cosa que una vaina como la de (2.4.3) con la salvedad de que apenas satisface (1) cuando la transformación que liga las dos bases o sistemas de referencia es *ortogonal*[18]. Es más, ni siquiera tiene por qué estar definida la asignación para bases no ortonormales (¡qué horror!). El archiconocido *tensor de inercia* se encuentra entre estos engendros, abominables mutaciones de los tensores platónicos con talento suficiente para engañar a físicos e ingenieros despistados. Lamentablemente, no es posible aprender y aprehender la física en su concepción actual más extendida sin asimilar y amaestrar el manejo de estos bichos, parásitos que han infectado sin remedio la relatividad especial y teoría cuántica de campos. Dejamos aquí su definición genérica, por cortesía[19]:

[18]Esto significa que la transformación preserva el producto escalar (del que se asume dotado el espacio físico) o, lo que es lo mismo, las longitudes (normas) de los vectores.

[19]La denominación $G-$*tensor* bien podría ser un invento del autor.

Definición 2.4.4 *Un G−tensor n−dimensional, r−contravariante, s−covariante es una asignación de n^{r+s} números $T^{i_1 \ldots i_r}{}_{j_1 \ldots j_s}$ a cada sistema de referencia tales que la relación entre los conjuntos de números asociados a dos sistemas de referencia cualesquiera viene dada por (1) siempre y cuando la transformación entre estos sistemas de referencia sea un elemento del grupo G.*

Así, en tres dimensiones, el (mal llamado) tensor de inercia es un $SO(3)-$tensor. La relatividad especial o restringida, bien entendida, confiere el protagonismo a los llamados *tensores de Lorentz*, o $SO(3,1)-$tensores.

2.5. Contracción de tensores

Motivamos las definiciones centrales de esta sección por medio de un listado aleatorio de *facts*:

- Dada una matriz cuadrada, su *traza* es la suma de sus elementos diagonales. ¿Por qué esta operación es más interesante que otras fórmulas de semejante semblanza, como podría ser la suma de todos los elementos de la matriz?

- Un endomorfismo, como se esgrimió anteriormente, es esencialmente un tensor $(1,1)$. Sus coordenadas con respecto a una base, $T^i{}_j$, acostumbran a presentarse en forma de matriz cuadrada, de la cual es expedito determinar su traza, $T^i{}_i$. Nos gustaría otorgar a esta cuantía el título *traza del endomorfismo*, pero esto daría a entender que no depende de la base, cuando su cálculo involucra explícitamente las componentes del tensor en una base.

- Si analizamos el quilombo con más alumbramiento de mente, nos percataremos de que, en efecto, cojas la base que cojas te va a salir lo mismo. Ya que, con las notaciones sobradamente utilizadas en lo que va de capítulo

$$\tilde{T}^i{}_i = T(\tilde{e}^i, \tilde{e}_i) = T(a^i{}_j e^j, (a^k{}_i)^{-1} e_k) = (a^k{}_i)^{-1} a^i{}_j T^j{}_k = \delta^k{}_j T^j{}_k = T^j{}_j$$

  y de ahí la bondad de la traza de un endomorfismo.

 ✉ f42rodim@uco.es

- Obsérvese que la operación consistente en tomar la traza de un endomorfismo o tensor $(1,1)$ produce un escalar o tensor $(0,0)$. Esta operación recibe el nombre de *contracción*[20]. Se trata, es más, de una aplicación lineal, siendo que la traza de la suma de matrices es la suma de las trazas, y la traza de una constante por una matriz es la constante por la traza.

- ¿Cómo lo apañamos para extender esta contracción a tensores de tipo arbitrario? La idea es un tanto ñoña. Sabemos contraer un tensor $(1,1)$, luego habríamos dado jaque mate a la problemática si de un tensor de orden superior extraemos uno (o varios) tensores $(1,1)$. Pues bien, si $f(x,y)$ es una función de dos variables y fijamos una de ellas, o sea, evaluamos f en su primer argumento para, digamos, $x = x_0$, entonces $f(x_0,y)$ es una función de una sola variable. Análogamente, si tomo un tensor (r,s) y *congelo* exactamente $r-1$ argumentos contravariantes y $s-1$ argumentos covariantes, resulta un tensor $(1,1)$, cuya contracción arroja un número. Si hago esto para todos los valores de los argumentos congelados, me queda un tensor $(r-1,s-1)$. Este proceso queda encapsulado en la definición:

Definición 2.5.1 *Sean $T \in \mathscr{T}^r_s(V)$, $B = \{e_1,\dots,e_n\}$, una base de V y $B^* = \{e^1,\dots,e^n\}$, su base dual. La contracción de T respecto de su $k-$ésimo argumento contravariante y su $l-$ésimo argumento covariante es el tensor $\mathscr{C}^k_l T \in \mathscr{T}^{r-1}_{s-1}(V)$ cuya evaluación en los covectores $\alpha^1,\dots,\alpha^{r-1} \in V^*$ y los vectores $v_1,\dots,v_{s-1} \in V$ da*

$$\begin{aligned}&\mathscr{C}^k_l T(\alpha^1,\dots,\alpha^{r-1},v_1,\dots,v_{s-1})\\&= T(\alpha^1,\dots,\alpha^{k-1},e^i,\alpha^{k+1},\dots,\alpha^{r-1},v_1,\dots,v_{l-1},e_i,v_{l+1},\dots,v_{s-1})\end{aligned}$$

Ejemplo .3 El tensor de Riemann

El *tensor de curvatura de Riemann* R es naturalmente un tensor $(1,3)$, por lo que posee componentes $R^i{}_{jkl}$. Se define el *tensor de Ricci* Ric como el tensor $(0,2)$ fruto de la contracción

$$\mathrm{Ric} := \mathscr{C}^1_1 R$$

[20]Ruego disculpen la rima interna al puro estilo 'el spinozismo es un racionalismo'.

 ✉ f42rodim@uco.es

En consecuencia, las componentes de Ric se hallan según

$$\mathrm{Ric}_{ij} = R^{k}{}_{kij}$$

Recalcamos que solo los tensores mixtos, es decir, aquellos que gozan de al menos un índice de cada tipo, son susceptibles de sufrir contracción. Podría creerse que, como las componentes de un tensor 2-covariante o 2-contravariante se organizan igualmente en forma de matriz cuadrada, les queda asociada a estos una traza o contracción. Sin embargo, esta traza *no es invariante* ante un cambio de base y no establece un tensor.

2.6. Producto escalar e isomorfismos musicales

Así como Dios creó a Adán en la máxima desnudez, inerme y sin harapos, y hasta le confiscó una costilla para alumbrar a Eva, los espacios vectoriales fueron engendrados sin estructura adicional alguna. Es imposible enunciar o responder en un espacio vectorial abstracto y desvestido cuestiones como '¿son estos dos vectores perpendiculares?' o '¿cuánto mide tal y cual vector?'. Sería tan cruel e inútil como preguntar a un ciego que de qué color es mi mascarilla quirúrgica. Por fortuna, la ceguera que padece un espacio vectorial con respecto a estos asuntos métricos es sanable sin recurrir a milagros médicos o divinos. Es suficiente con proveer al espacio vectorial de un *producto escalar*, en lugar de dejarlo en cueros campando a sus anchas por el Edén geométrico sin tan siquiera la capacidad de medirse sus partes nobles.

Los pensamientos que condujeron al concepto pulido de producto escalar, sin perdernos en el inextricable dédalo de la historia de la matemática, lo mismo fueron *grosso modo* los siguientes[21]. En virtud del teorema de Pitágoras, cuyo descubridor muy probablemente no fue Pitágoras, la *longitud* de un segmento inmerso en el espacio tridimensional físico cuyas proyecciones o componentes son $A = (x, y, z)$ tiene por valor $\sqrt{x^2 + y^2 + z^2}$. La aplicación que a cada segmento asigna el cuadrado de esta cantidad, es decir

$$(x, y, z) \longmapsto x^2 + y^2 + z^2$$

[21]Seguro que en verdad no tuvieron nada que ver con la milonga que estoy a punto de contar.

es un caso particular de *forma cuadrática*, como lo sería cualquier polinomio de grado 2 en las variables x, y, z. La magnitud $|A| := \sqrt{x^2 + y^2 + z^2}$ es la *norma* o *módulo* del vector A. Ahora bien, el conocimiento de las normas no basta para resolver los problemas de triángulos y otros polígonos que tanto entretenían a nuestros antepasados helenos. Y el teorema de Pitágoras apenas es válido para triángulos rectángulos, ¿qué hay del resto? Para abarcarlos en nuestro formalismo, debemos generalizar lo anterior como sigue. Como la forma cuadrática introducida es $|A|^2$, consideramos, dados ahora dos vectores A, B, el producto de la longitud de uno por la longitud de la *proyección* del otro sobre el primero, esto es, $|A||B|\cos\theta$, donde θ es el ángulo que comprenden. Cuando $A = B$, recuperamos la forma cuadrática original. Aunque esto parezca una invención de lo más alocada, hay motivos para idear esta combinación si uno se pone a darle vueltas a los acertijos de la geometría clásica. Como quiera que, si $A = (x_1, y_1, z_1)$ y $B = (x_2, y_2, z_2)$, se tiene[22]

$$|A||B|\cos\theta = x_1x_2 + y_1y_2 + z_1z_2 := g(A, B)$$

todos estos preludios y meandros desembocan en una aplicación $g : V \times V \longrightarrow \mathbb{R}$ lineal en sus dos argumentos (lo que viene siendo un tensor 2-covariante) que además es *simétrica*, es decir, $g(A, B) = g(B, A)\ \forall A, B$, y *definida positiva*: $g(A, A) \geq 0\ \forall A$, alcanzándose $g(A, A) = 0 \iff A = 0$. Esto garantiza la positividad de las longitudes, requerimiento que se nos antoja de lo más imperativo. Vemos, por tanto, que la presencia de esta estructura matemática, que llamaremos *producto escalar euclídeo estándar*, es la que posibilita calcular (y, más fundamentalmente, la que dota de significado a) longitudes y ángulos.

La física del siglo XX nos ha revelado que repudiar los productos escalares no definidos positivos acarrearía renunciar al modelado del espaciotiempo en su concepción relativista. Además, desde un punto de vista matemático, el afán de generalidad compele a aliviar esta condición, harto restrictiva, sustituyéndola por la de no degeneración. Una

[22] No mentiríamos si clamamos que $\cos\theta$ se define así con el propósito de que la igualdad se cumpla.

 ✉ f42rodim@uco.es

aplicación bilineal de estas se dice *no degenerada* si

$$\forall B \in V : g(A, B) = 0 \Longleftrightarrow A = 0$$

esto es, el único vector ortogonal[23] a todos los elementos del espacio vectorial es el nulo.

Definición 2.6.1 *Un producto escalar es un tensor $g \in \mathcal{T}_2^0(V)$ simétrico y no degenerado.*

No nos dedicaremos a recitar y comprobar las propiedades de los productos escalares, sino que, como resultado primordial de la actual sección, argüiremos que el producto escalar induce un isomorfismo canónico entre el espacio vectorial y su dual, criatura inexistente en ausencia de esta estructura[24].

Proposición 2.6.1 *Sea (V, g) un espacio vectorial dotado de un producto escalar. La aplicación*

$$\flat : V \longrightarrow V^*$$

$$v \longmapsto v^\flat := g(v, \cdot)$$

es un isomorfismo lineal.

Demostración: Notemos que $v^\flat$, que es la imagen del presunto isomorfismo $\flat$ bajo un vector $v \in V$, es una aplicación $v^\flat : V \longrightarrow \mathbb{R}$. Tenemos:

[23] Dos vectores $A, B \in V$ son *ortogonales* con respecto al producto escalar g si $g(A, B) = 0$

[24] No hay, al menos en conocimiento del autor, una frase para designar genéricamente un espacio vectorial provisto de un producto escalar según (2.6.1). Cuando el producto escalar es definido positivo y el cuerpo es $\mathbb{R}$, se llama *espacio vectorial euclídeo* al par (V, g). En caso de que el cuerpo pueda ser $\mathbb{C}$, el producto escalar se dice *hermítico* si verifica:

- Simetría bajo conjugación: $\forall A, B \in V : g(A, B) = \overline{g(B, A)}$
- Linealidad en el primer argumento: $\forall A_1, A_2, B \in V,\ \lambda_1, \lambda_2 \in \mathbb{C} : g(\lambda_1 A_1 + \lambda_2 A_2, B) = \lambda_1 g(A_1, B) + \lambda_2 g(A_2, B)$
- Definición positiva: $\forall A \in V : g(A, A) \geq 0$ con igualdad si y solo si $A = 0$

El par (V, g) se denomina entonces *espacio prehilbertiano*. Un espacio prehilbertiano sobre $\mathbb{R}$ es un espacio vectorial euclídeo.

Cuando el producto escalar g es tal que en una base ortonormal su representación en coordenadas tiene un -1 en la diagonal (o $n - 1$, siendo n la dimensión), la dupla (V, g) es un *espacio vectorial lorentziano*.

- $v^\flat$ es lineal

$$v^\flat(\lambda_1 x + \lambda_2 y) = g(v, \lambda_1 x + \lambda_2 y) = \lambda_1 g(v, x) + \lambda_2 g(v, y) = \lambda_1 v^\flat(x) + \lambda_2 v^\flat(y)$$

siendo $v, x, y \in V,\ \lambda_1, \lambda_2 \in \mathbb{R}$. Por ende, $v^\flat \in V^*$.

- $\flat$ es lineal

$$(\lambda_1 v + \lambda_2 w)^\flat(x) = g(\lambda_1 v + \lambda_2 w, x) = \lambda_1 g(v, x) + \lambda_2 g(v, y) = \lambda_1 v^\flat(x) + \lambda_2 w^\flat(x)$$

- $\flat$ es inyectiva. En efecto, supongamos que $v^\flat = 0$, es decir

$$\forall x \in V : v^\flat(x) = 0 \Longrightarrow g(v, x) = 0\ \forall x \in V \Longrightarrow v = 0$$

donde en el último paso se ha usado que g es no degenerado. En consecuencia, $\ker \flat = \{0\}$, que, para aplicaciones lineales, equivale a la inyectividad.

- $\flat$ es sobreyectiva, por los mismos motivos que el isomorfismo de reflexividad (2.2.3) es sobreyectivo[25].

■

[25] Damos una justificación alternativa. La sobreyectividad quiere decir que es posible encontrar para cualquier $\alpha \in V^*$ un $v \in V$ tal que $v^\flat = \alpha$, o sea

$$\forall x \in V : \alpha(x) = v^\flat(x) = g(v, x)$$

Particularicemos esta igualdad a los elementos de una base $\{e_1, \ldots, e_n\}$, sobre los que $v^\flat$ queda completamente especificada

$$\alpha(e_i) = g(v, e_i) \Longrightarrow \alpha_i = g_{ij} v^j$$

donde

$$\alpha = \alpha_i e^i \qquad\qquad g = g_{ij} e^i \otimes e^j$$

Las expresiones $g_{ij} v^j = \alpha_i$ constituyen un sistema lineal de n ecuaciones y n incógnitas v^j. Por el teorema de Rouché-Frobenius, posee solución única, puesto que la matriz (g_{ij}) es no singular (esto se desprende de que g es no degenerado).

Definición 2.6.2 *El isomorfismo*

$$\flat : V \longrightarrow V^*$$

$$v \longmapsto v^\flat := g(v, \cdot)$$

se denomina bemol. El isomorfismo inverso

$$\sharp : V^* \longrightarrow V$$

$$\alpha \longmapsto \alpha^\sharp := \flat^{-1}(\alpha)$$

es el sostenido. Colectivamente, se les llama aplicaciones musicales.

Vemos que en un conglomerado (V, g) se erige una correspondencia biunívoca entre vectores y covectores que es manifiestamente independiente de bases. Deduzcamos a continuación las expresiones coordenadas de estas aplicaciones. Sea un vector $v = v^i e_i \in V$, ¿cuál es su bemol? Pues, por la definición, para algún $x = x^i e_i \in V$

$$v^\flat(x) = g(v, x) = g(v^i e_i, x^j e_j) = v^i g_{ij} x^j = v^i g_{ij}\, e^j(x)$$

de donde

$$v^\flat = v^i g_{ij}\, e^j$$

En un abuso de notación, se acostumbra a notar $v_j = v^i g_{ij}$ las componentes de $v^\flat$. Es por ello que se habla coloquialmente de que la acción de $\flat$ comporta una *bajada de índice*. En lo que concierne al sostenido, podemos emplear que

$$(v^\flat)^\sharp = v \Longrightarrow A^{ij} g_{jk} v^k = v^i$$

siendo A^{ij} la matriz de la aplicación $\sharp$ en esa base. Haciendo $A^{ij} = g^{ij}$, donde g^{ij} son las entradas de la matriz inversa a (g_{ij}), que existe por ser g no degenerada, encontramos que

$$A^{ij} g_{jk} v^k = g^{ij} g_{jk} v^k = \delta^i{}_k v^k = v^i$$

por lo que nuestra elección es adecuada, y, si $\alpha = \alpha_i e^i \in V^*$

$$\alpha^\sharp = \alpha_i g^{ij}\, e_j$$

✉ f42rodim@uco.es

Nuevamente, suele escribirse $\alpha^j = \alpha_i g^{ij}$ para las componentes del sostenido, que actúa *subiendo el índice*, como se viene escuchando en la parlanza física. Esta trapisonda de multiplicar por g_{ij} o g^{ij} con objeto de subir y bajar índices se extiende fácilmente a tensores de orden superior. Tomemos así a chorón un $T \in \mathscr{T}_3^1(V)$, con coordenadas $T^i{}_{jkl}$ en alguna base. Por la lógica anterior, construimos un tensor de componentes T_{ijkl} según

$$T_{ijkl} := T^p{}_{jkl}\, g_{pi}$$

que, por sus índices, habrían de corresponder a un tensor $(0,4)$. O, por qué no, somos asimismo libres de producir unas coordenadas $T^i{}_j{}^k{}_l$ tales que

$$T^i{}_j{}^k{}_l := T^i{}_{ipl}\, g^{pk}$$

que tendrían que representar un tensor $(2,2)$, en principio distinto de

$$T^i{}_{jk}{}^l := T^i{}_{ikp}\, g^{pl}$$

a pesar de ser ambos tensores $(2,2)$ procedentes de un mismo germen. Y ya que andamos inmersos en el galimatías este, todavía no ahítos de tanta subida y bajada indicial, nos podemos coronar con

$$T^{ijkl} := T^i{}_{pqr}\, g^{pj} g^{qk} g^{rl}$$

que es un tensor $(4,0)$. Concluimos, al menos heurísticamente, que, de tal suerte, las aplicaciones musicales dan lugar a toda una hueste de isomorfismos entre espacios de tensores de igual *orden*, que es su número total de índices. En la comunidad física, el abuso del lenguaje se lleva al extremo de afirmar, sin ningún remordimiento, que los tensores resultantes de subir y bajar índices a partir de uno ofrendado son la encarnación de un único *tensor físico* (no se puede negar que, matemáticamente, la aseveración es ridícula, ¡si ni siquiera pertenecen al mismo espacio!).

Nos encantaría cimentar estas ideas sin que haya coordenadas de por medio. Nos inspiraremos en el caso simple de un tensor $(0,2)$, $S = S_{ij}\, e^i \otimes e^j$. Subiéndole un índice, construimos el tensor $(1,1)$

$$\tilde{S} = S^i{}_j\, e_i \otimes e^j \qquad\qquad S^i{}_j = S_{kj} g^{ki}$$

Evaluemos $\tilde{S}$ en dos argumentos aleatorios

$$\tilde{S}(\alpha, v) = S^i{}_j \alpha_i v^j = S_{kj} g^{ki} \alpha_i v^j$$

Ahora bien, $g^{ki}\alpha_i = \alpha^k$, por lo que

$$\tilde{S}(\alpha, v) = S^i{}_j \alpha_i v^j = S_{kj} \alpha^k v^j = S(\alpha^\sharp, v)$$

Y esta última ecuación nos da la definición independiente de coordenadas que buscábamos:

Definición 2.6.3 *Sea $T \in \mathscr{T}^r_s(V)$. De subir su k−ésimo índice covariante se obtiene el tensor $(\uparrow^k T) \in \mathscr{T}^{r+1}_{s-1}(V)$ cuya evaluación en el nuevo argumento contravariante da*

$$(\uparrow^k T)(\dots, \alpha^k, \dots) := T(\dots, (\alpha^k)^\sharp, \dots)$$

mientras que sobre el resto de entradas opera como T. De bajar su k−ésimo índice contravariante se obtiene el tensor $(\downarrow^k T) \in \mathscr{T}^{r-1}_{s+1}(V)$ cuya evaluación en el nuevo argumento covariante da

$$(\downarrow^k T)(\dots, v_k, \dots) := T(\dots, (v_k)^\flat, \dots)$$

y las demás entradas quedan como estaban.

Ejemplo .4 Un caso práctico

Partimos del espacio vectorial con producto escalar (V, g), donde, en alguna base

$$g = e^1 \otimes e^1 + 2e^1 \otimes e^2 + 2e^2 \otimes e^1 + 3e^2 \otimes e^2$$

a) *Calcúlense $e_1^\flat, e_2^\flat$.* Usando la definición de bemol y la forma explícita de g, tenemos

$$e_1^\flat = g(e_1, \cdot) = e^1(e_1) \otimes e^1 + 2e^1(e_1) \otimes e^2 + 2e^2(e_1) \otimes e^1 + 3e^2(e_1) \otimes e^2 = e^1 + 2e^2$$

$$e_2^\flat = g(e_2, \cdot) = e^1(e_2) \otimes e^1 + 2e^1(e_2) \otimes e^2 + 2e^2(e_2) \otimes e^1 + 3e^2(e_2) \otimes e^2 = 2e^1 + 3e^2$$

donde se ha tenido en cuenta que $e^i(e_j) = \delta^i{}_j$.

b) *Obténgase $e^{1\sharp}$.* Sabemos que

$$g(e^{1\sharp}, v) = e^1(v) \quad \forall v \in V$$

 ✉ f42rodim@uco.es

En particular

$$g(e^{1\sharp}, e_1) = 1 \qquad\qquad g(e^{1\sharp}, e_2) = 0$$

Llamando $e^{1\sharp} = x^1 e_1 + x^2 e_2$, esto da el sistema de ecuaciones lineales

$$\begin{cases} x^1 + 2x^2 = 1 \\ 2x^1 + 3x^2 = 0 \end{cases}$$

La resolución arroja $x^1 = -3, x^2 = 2$, por lo que $e^{1\sharp} = -3e_1 + 2e_2$

c) *Sea* $T = e_1 \otimes e_2 + 5e_2 \otimes e_1$. *Determínese* $(\downarrow^1 T)$. Por definición

$$(\downarrow^1 T)(v, \alpha) = T(v^\flat, \alpha) \quad \forall v \in V, \alpha \in V^*$$

Ahora bien, notemos que $v(w^\flat) = w^\flat(v) = g(w, v) = g(v, w) = v^\flat(w)\ \forall v, w \in V$, lo que implica

$$(\downarrow^1 T) = e_1^\flat \otimes e_2 + 5e_2^\flat \otimes e_1$$

Sustituyendo los resultados previos

$$(\downarrow^1 T) = (e^1 + 2e^2) \otimes e_2 + 5(2e^1 + 3e^2) \otimes e_1 = e^1 \otimes e^2 + 2e^2 \otimes e^2 + 10e^1 \otimes e^1 + 15e^2 \otimes e^1$$

d) Podría plantearse un ejercicio análogo para, por ejemplo, un tensor $S \in \mathscr{T}_2^0(V)$. Sería

$$(\uparrow^1 S)(\alpha, v) = S(\alpha^\sharp, v)$$

y como $\beta(\alpha^\sharp) = (\beta^\sharp)^\flat(\alpha^\sharp) = g(\beta^\sharp, \alpha^\sharp) = g(\alpha^\sharp, \beta^\sharp) = (\alpha^\sharp)^\flat(\beta^\sharp) = \alpha(\beta^\sharp) = \beta^\sharp(\alpha)$, el cálculo procedería, como en el apartado anterior, aplicando el sostenido a los primeros factores, esto es, si

$$S = S_{ij} e^i \otimes e^j$$

entonces

$$(\downarrow^1 S) = S_{ij} e^{i\sharp} \otimes e^j$$

✉ f42rodim@uco.es

Ejemplo .5 Producto escalar de covectores

En un espacio vectorial con producto escalar (V, g), los isomorfismos musicales permiten conferir un producto escalar natural al espacio dual, $G : V^* \times V^* \longrightarrow \mathbb{R}$, según

$$G(\alpha, \beta) := g(\alpha^\sharp, \beta^\sharp)$$

En otras palabras

$$G = \uparrow^1 \uparrow^1 g$$

Si $\alpha = \alpha_i e^i, \beta = \beta_i e^i$, entonces

$$G(\alpha, \beta) := g(\alpha_i g^{ij} e_j, \beta_k g^{kl} e_l) = \alpha_i \beta_k g^{ij} g^{kl} g(e_j, e_l) = \alpha_i \beta_k g^{ij} g_{jl} g^{kl} = \alpha_i \beta_k \delta^i{}_l g^{kl} = \alpha_i \beta_j g^{ij}$$

Por tanto, la matriz de este producto escalar es (g^{ij}).

La imposibilidad de contraer dos índices de igual tipo desaparece gracias a las aplicaciones musicales:

Definición 2.6.4 *La contracción métrica de dos índices covariantes (resp. contravariantes) se define como la subida (resp. bajada) de uno de los índices seguida de su contracción.*

Ejemplo .6 Escalar de curvatura

El escalar de curvatura S se construye a partir del tensor de Ricci contrayendo métricamente sus dos índices, esto es

$$S := \mathscr{C}_1^1 \uparrow^1 \mathrm{Ric} = \mathrm{Ric}_{ij} g^{ij}$$

En el contexto de la relatividad general, este objeto es el ingrediente esencial de la acción de Einstein-Hilbert

$$S_{EH} \sim \int \sqrt{-\det(g_{ij})}\, S\, d^4x$$

cuya extremización conduce a las ecuaciones de campo de Einstein.

✉ f42rodim@uco.es

2.7. Tensores antisimétricos

De entre todos los tensores con alguna característica peculiar, los que se llaman *antisimétricos* gozan de especial prominencia, tanto por su interés matemático (e.g. el determinante de vectores, la teoría de formas diferenciales e integración) como físico (todo el electromagnetismo puede entenderse como la historia de un determinado tensor antisimétrico, el tensor de Faraday o campo electromagnético). Es por ello que merece la pena invertir algo de esfuerzo (tampoco en demasía) en su estudio pormenorizado[26].

Definición 2.7.1 *Un tensor $s-$covariante $T \in \mathscr{T}_s^0(V)$ es (totalmente) antisimétrico, o una s-forma, si, para cualquier par de argumentos*

$$\forall v, w \in V : T(\dots, v, \dots, w, \dots) = -T(\dots, w, \dots, v, \dots)$$

Es decir, si se permutan dos entradas, apenas cambia el signo del resultado. Cuando se practica una permutación general[27] $\sigma \in S_s$, sucede que

$$\forall v_1, \dots, v_s \in V : T(v_{\sigma(1)}, \dots, v_{\sigma(s)}) = \operatorname{sgn}(\sigma)\, T(v_1, \dots, v_s)$$

Ofrecemos una caracterización equivalente:

Proposición 2.7.1 *Un tensor $s-$covariante $T \in \mathscr{T}_s^0(V)$ es antisimétrico si y solo si para cualquier par de argumentos*

$$\forall v \in V : T(\dots, v, \dots, v, \dots) = 0$$

Demostración: Supongamos que T es antisimétrico. Entonces

$$T(\dots, v, \dots, v, \dots) = -T(\dots, v, \dots, v, \dots)$$

[26] Lo haremos para el caso covariante por darse con más frecuencia, pero, como es o será evidente, puede desarrollarse un formalismo idéntico para tensores contravariantes. Lo que sí es claro es que todos los índices deben ser del mismo tipo para que siempre tenga sentido permutar dos argumentos cualesquiera.

[27] Una *permutación* de s elementos es una biyección $\sigma : \{1, \dots, s\} \longrightarrow \{1, \dots, s\}$. El conjunto de todas las permutaciones de s elementos se denota S_s. La *signatura* o *signo* de una permutación σ es $\operatorname{sgn}(\sigma) := (-1)^r$, siendo r un número de trasposiciones o permutaciones de pares necesario para ejecutar σ.

como se ve de permutar las dos entradas indicadas y aplicar la definición. Se sigue de esto que $T(\ldots, v, \ldots, v, \ldots) = 0$. Recíprocamente, si $\forall v \in V : T(\ldots, v, \ldots, v, \ldots) = 0$, se tendrá también

$$\forall v, w \in V : T(\ldots, v + w, \ldots, v + w, \ldots) = 0$$

Por linealidad, esto es equivalente a

$$T(\ldots, v, \ldots, v, \ldots) + T(\ldots, w, \ldots, w, \ldots) + T(\ldots, v, \ldots, w, \ldots) + T(\ldots, w, \ldots, v, \ldots) = 0$$

Los dos primeros sumandos son cero por la hipótesis, con lo cual

$$T(\ldots, v, \ldots, w, \ldots) + T(\ldots, w, \ldots, v, \ldots) = 0 \Longrightarrow T(\ldots, v, \ldots, w, \ldots) = -T(\ldots, w, \ldots, v, \ldots)$$

■

Llamaremos $\Lambda_s(V)$ al conjunto de las $s-$formas sobre V. El anterior resultado conlleva que $\Lambda_s(V) = \{0\}$ cuando s es mayor que la dimensión de V.

Proposición 2.7.2 *El conjunto $\Lambda_s(V)$ provisto de la suma de tensores y el producto por escalares es un espacio vectorial sobre $\mathbb{R}$.*

Demostración: Evidentemente, tanto la suma de dos tensores antisimétricos $s-$covariantes como el producto de un escalar por un tensor antisimétrico $s-$covariante da sendos tensores antisimétricos $-s$ covariantes. En consecuencia, $\Lambda_s(V)$ es un subespacio vectorial de $\mathscr{T}_s^0(V)$ y, particularmente, un espacio vectorial *per se* ■

La siguiente transformación asigna una $s-$forma a un tensor $s-$covariante dado:

Definición 2.7.2 *Se llama antisimetrizador de orden s a la aplicación*

$$h^s : \mathscr{T}_s^0(V) \longrightarrow \Lambda_s(V)$$

definida para un $T \in \mathscr{T}_s^0(V)$ como

$$\forall v_1, \ldots, v_s \in V : h^s(T)(v_1, \ldots, v_s) := \sum_{\sigma \in S_s} \operatorname{sgn}(\sigma)\, T(v_{\sigma(1)}, \ldots, v_{\sigma(s)})$$

✉ f42rodim@uco.es

Ejemplo .7 Antisimetrizador de orden 3

Puede comprobarse que, si $T \in \mathscr{T}_3^0(V)$

$$\forall v_1, v_2, v_3 \in V : h^3(T)(v_1, v_2, v_3) = T(v_1, v_2, v_3) + T(v_2, v_3, v_1) + T(v_3, v_1, v_2) -$$
$$- T(v_2, v_1, v_3) - T(v_1, v_3, v_2) - T(v_2, v_1, v_3)$$

Ejemplo .8 Tensores simétricos

En contraposición a la definición 2.7.1, un tensor $T \in \mathscr{T}_s^0(V)$ se dice *(totalmente) simétrico* si

$$\forall v_1, \ldots, v_s \in V : T(v_{\sigma(1)}, \ldots, v_{\sigma(s)}) = T(v_1, \ldots, v_s)$$

que es como decir que la imagen no cambia al transponer cualquier par de elementos de entrada. Sea $\Sigma_s(V)$ el conjunto de los tensores $s-$covariantes simétricos sobre V. Por los mismos motivos que en el escenario antisimétrico, con las operaciones adecuadas, $\Sigma_s(V)$ es un subespacio vectorial de V. Es más, se verifica

$$\mathscr{T}_s^0(V) = \Sigma_s(V) \oplus \Lambda_s(V)$$

Por ejemplo, todo $T \in \mathscr{T}_2^0(V)$ admite la escritura

$$\forall v, w \in V : T(v, w) = S(v, w) + A(v, w)$$
$$S(v, w) := \frac{1}{2}[T(v, w) + T(w, v)]$$
$$A(v, w) := \frac{1}{2}[T(v, w) - T(w, v)]$$

donde $S \in \Sigma_2(V)$, $A \in \Lambda_2(V)$. Más generalmente, para un $T \in \mathscr{T}_s^0(V)$, se comprueba que

$$T = \frac{1}{s!}[H^s(T) + h^s(T)]$$

donde $H^s(T) \in \Sigma_s(V)$ es el *simetrizador*

$$\forall v_1, \ldots, v_s \in V : H^s(T)(v_1, \ldots, v_s) := \sum_{\sigma \in S_s} T(v_{\sigma(1)}, \ldots, v_{\sigma(s)})$$

Ansiamos edificar una base para $\Lambda_s(V)$ a partir de una base de V^*. Ello precisará, al igual que cuando dimos una base de $\mathscr{T}_s^r(V)$, una operación que combine tensores antisimétricos para dar otros de mayor orden. El producto tensorial no nos sirve, puesto que, salvo excepcionalmente, $T_1 \otimes T_2 \notin \Lambda_s(V)$ aunque $T_1, T_2 \in \Lambda_s(V)$. La cosa cambia si *antisimetrizamos el producto*, en el siguiente sentido. Consideremos dos covectores, o 1-formas, $\alpha, \beta \in V^* \equiv \Lambda_1(V)$. Es claro que[28]

$$\alpha \wedge \beta := \alpha \otimes \beta - \beta \otimes \alpha \in \Lambda_2(V)$$

La idea se extiende directamente a cualquier número de factores:

Definición 2.7.3 *Sean $\alpha^1, \ldots, \alpha^s \in \Lambda_1(V)$. Su producto exterior (o producto antisimétrico, wedge, Grassmann) es la $s-$forma $\alpha^1 \wedge \cdots \wedge \alpha^s \in \Lambda_s(V)$ dada por*

$$\alpha^1 \wedge \cdots \wedge \alpha^s := \sum_{\sigma \in S_s} \alpha^{\sigma(1)} \otimes \cdots \otimes \alpha^{\sigma(s)}$$

Obsérvese que

$$\alpha^1 \wedge \cdots \wedge \alpha^s = h^s(\alpha^1 \otimes \cdots \otimes \alpha^s)$$

El producto exterior de 1-formas, además de asociativo y distributivo con respecto a la suma, es naturalmente antisimétrico

$$\forall \alpha, \beta \in \Lambda_1(V) : \alpha \wedge \beta = -\beta \wedge \alpha$$

Como consecuencia, se deduce que $\forall \alpha \in \Lambda_1(V) : \alpha \wedge \alpha = 0$. Para la situación más general, véase la proposición 2.7.4, que probaremos más adelante.

Proposición 2.7.3 *Sean $\{e_1, \ldots, e_n\}$ una base de V y $\{e^1, \ldots, e^n\}$, su base dual. El conjunto $\{e^{j_1} \wedge \cdots \wedge e^{j_s}\}_{1 \leq j_1 < \ldots < j_s \leq n}$ es una base de $\Lambda_s(V)$, con $s \leq n$.*

Demostración: Tomemos $T \in \Lambda_s(V)$ y vectores arbitrarios $v_1, \ldots, v_s \in V$. En coordenadas

$$v_i = v_i^{j_i} e_{j_i}$$

[28] Algunos autores incluyen un factor $1/2$ en esta definición. Es irrelevante exactamente en qué punto aparezcan los factores del tipo $1/s!$, lo que importa es que figuren en alguna parte.

✉ f42rodim@uco.es

Por tanto

$$T(v_1,\dots,v_s) = T(v_1^{j_1}e_{j_1},\dots,v_s^{j_s}e_{j_s}) = v_1^{j_1}\dots v_s^{j_s}\,T(e_{j_1},\dots,e_{j_s}) = T_{j_1\dots j_s}e^{j_1}(v_1)\dots e^{j_s}(v_s)$$

Aquí finalizaría la prueba si quisiéramos concluir que $T = T_{j_1\dots j_s}e^{j_1}\otimes\cdots\otimes e^{j_s}$. No obstante, como T es antisimétrico, podemos ir más lejos. Como quiera que T se anula cuando dos o más entradas son iguales, todos los sumandos distintos de cero de la última igualdad son tales que $\{j_1,\dots,j_s\}$ se corresponden con permutaciones de $\{1,\dots,s\}$, por lo que

$$T_{j_1\dots j_s}e^{j_1}(v_1)\dots e^{j_s}(v_s) = \sum_{1\,\leq\, j_1\,<\,\dots\,<\,j_s\,\leq\, n}\left[\sum_{\sigma\in S_s} T_{\sigma(j_1)\dots\sigma(j_s)}e^{\sigma(j_1)}(v_1)\dots e^{\sigma(j_s)}(v_s)\right]$$

donde se ha añadido la condición $1 \leq j_1 < \dots < j_s \leq n$ para evitar términos que, si bien asociados a permutaciones disitntas, son iguales[29]. De la definición 2.7.1 se desprende que, para una permutación σ

$$T_{j_1\dots j_s} = \mathrm{sgn}(\sigma)\,T_{\sigma(j_1)\dots\sigma(j_s)}$$

Así

$$T(v_1,\dots,v_s) = \sum_{1\,\leq\, j_1\,<\,\dots\,<\,j_s\,\leq\, n} T_{j_1\dots j_s}\left[\sum_{\sigma\in S_s}\mathrm{sgn}(\sigma)\,e^{\sigma(j_1)}(v_1)\dots e^{\sigma(j_s)}(v_s)\right]$$

A la luz de la definición 2.7.3, esto es equivalente a

$$T(v_1,\dots,v_s) = \left[\sum_{1\,\leq\, j_1\,<\,\dots\,<\,j_s\,\leq\, n} T_{j_1\dots j_s}\,e^{j_1}\wedge\cdots\wedge e^{j_s}\right](v_1,\dots,v_s)$$

lo que muestra que $\{e^{j_1}\wedge\cdots\wedge e^{j_s}\}_{1\,\leq\, j_1\,<\,\dots\,<\,j_s\,\leq\, n}$ es un sistema de generadores de $\Lambda_s(V)$. A continuación, escojamos una combinación lineal de estos objetos y hagámosla cero

$$\sum_{1\,\leq\, j_1\,<\,\dots\,<\,j_s\,\leq\, n} \lambda_{j_1\dots j_s}e^{j_1}\wedge\cdots\wedge e^{j_s} = 0$$

En particular

$$\left[\sum_{1\,\leq\, j_1\,<\,\dots\,<\,j_s\,\leq\, n} \lambda_{j_1\dots j_s}e^{j_1}\wedge\cdots\wedge e^{j_s}\right](e_{k_1},\dots,e_{k_s}) = 0$$

[29]Otra alternativa igual de válida consistiría en no tocar la suma sobre los índices e incorporar un factor $1/s!$

✉ f42rodim@uco.es

O bien

$$\sum_{1 \le j_1 < ... < j_s \le n} \lambda_{j_1 \ldots j_s} \sum_{\sigma \in S_s} \mathrm{sgn}(\sigma)\, \delta^{\sigma(j_1)}{}_{k_1} \ldots \delta^{\sigma(j_s)}{}_{k_s} = 0$$

Esto es un sistema de ecuaciones lineales homogéneas. Hay tantas ecuaciones como incógnitas independientes $\lambda_{j_1 \ldots j_s}$ y la matriz de coeficientes del sistema tiene rango máximo. Por ende, el sistema presenta solución única, i.e. la trivial

$$\forall j_1, \ldots, j_s : \lambda_{j_1 \ldots j_s} = 0$$

Esto evidencia que $\{e^{j_1} \wedge \cdots \wedge e^{j_s}\}_{1 \le j_1 < ... < j_s \le n}$ es un sistema libre ■

Proposición 2.7.4 *Sean $\alpha^1, \ldots, \alpha^s \in \Lambda_1(V)$. Se cumple*

$$\alpha^1 \wedge \cdots \wedge \alpha^s = 0$$

si y solo si $\alpha^1, \ldots, \alpha^s$ son linealmente dependientes.

Demostración: Si $\alpha^1, \ldots, \alpha^s$, con $s \le n$, son linealmente independientes, existe una base de V tal que estas formas lineales coinciden con los primeros elementos de su base dual: $\alpha^1 = e^1, \ldots, \alpha^s = e^s$. Es claro entonces que $\alpha^1 \wedge \cdots \wedge \alpha^s = e^1 \wedge \cdots \wedge e^s \neq 0$. Verbalmente, si las 1-formas son independientes, su producto exterior es distinto de cero. Ergo, si el producto exterior de un conjunto de formas es cero, han de ser linealmente dependientes[30]. Cualquier producto de s formas lineales, con $s > 0$, es cero al no haber $s-$formas no nulas.

Para el recíproco, digamos que $\alpha^1, \ldots, \alpha^s$ son dependientes, lo que acarrea

$$\alpha^1 = \lambda_2 \alpha^2 + \cdots + \lambda_s \alpha^s$$

para ciertos coeficientes $\lambda_2, \ldots, \lambda_s \in \mathbb{R}$, de modo que

$$\alpha^1 \wedge \cdots \wedge \alpha^s = (\lambda_2 \alpha^2 + \cdots + \lambda_s \alpha^s) \wedge \cdots \wedge \alpha^s$$

Todos los términos resultantes de aplicar la propiedad distributiva al primer producto exterior constan de dos factores iguales y, por antisimetría, se cancelan, hallando

$$\alpha^1 \wedge \cdots \wedge \alpha^s = 0$$

[30] Ley de la contraposición lógica: $(A \Longrightarrow B) \Longleftrightarrow (\neg B \Longrightarrow \neg A)$

 ✉ f42rodim@uco.es

■

Por último, definimos el producto *wedge* de dos tensores antisimétricos de cualquier orden:

Definición 2.7.4 *Sean $T \in \Lambda_s(V)$, $S \in \Lambda_l(V)$. Su producto exterior $T \wedge S \in \Lambda_{s+l}(V)$ viene dado por*

$$T \wedge S := \frac{1}{s!l!} h^{s+l}(T \otimes S)$$

✉ f42rodim@uco.es

3. Una introducción sencilla a la Relatividad Especial

3.1. Antecedentes

3.1.1. La teoría del éter

En 1864, James Clerk Maxwell publica el artículo *A Dynamical Theory of the Electromagnetic Field* en el que se postula por primera vez el conjunto de ecuaciones que en la actualidad lleva su nombre y que describe por completo la dinámica de los campos eléctricos y magnético tal y como son concebidos clásicamente. Valiéndose de estas expresiones, el físico escocés derivó sendas **ecuaciones de onda** para los citados campos $(\mathbf{E}, \mathbf{B})$. En notación moderna, se escriben

$$\nabla^2\mathbf{E} - \varepsilon\mu\frac{\partial^2\mathbf{E}}{\partial t^2} = 0 \qquad \nabla^2\mathbf{B} - \varepsilon\mu\frac{\partial^2\mathbf{B}}{\partial t^2} = 0$$

donde ε y μ son la permitividad eléctrica y la permeabilidad magnética del medio, respectivamente. Maxwell computó el valor del producto $\varepsilon_0\mu_0$ (particularización de estas magnitudes para el vacío o *éter lumínífero*) y, percatándose de su concordancia con $1/c^2$, siendo c la velocidad de la luz en el vacío (de la que ya se tenían algunos datos experimentales relativamente exactos), concluyó que **la luz**[31] **es una onda electromagnética**, hipótesis que sería confirmada años más tarde en el laboratorio por Heinrich Rudolf Hertz. Esto constituyó la unificación del Electromagnetismo y la Óptica.

Los físicos del siglo S-XIX no estaban preparados para admitir que las ondas electromagnéticas se propagan en el vacío. Por ello, asumieron que existía una suerte de fluido denominado **éter** presente en todas partes cuyas perturbaciones mecánicas se corresponden con los campos eléctrico y magnético, del mismo modo que el sonido es una vibración en un medio material. Nunca se llegó a disponer de evidencia experimental de su existencia. Con el advenimiento de las ecuaciones de Maxwell, la creencia ciega de los científicos en el éter empezó a remitir. Esto es debido a que en la formulación de estas leyes no se advierte ningún parámetro relacionado con el éter. Más aún, *las ecuaciones de onda del campo electromagnético son independientes del estado de movimiento del mismo*,

[31] Por *luz* entendemos radiación electromagnética de cualquier frecuencia.

lo cual está en contradicción con la fenomenología asociada a las ondas mecánicas. Por otra parte, se encontró que las *transformaciones de Galileo* entre dos sistemas de referencia inerciales (entendiendo por *inercial* que en él se cumple la segunda ley de Newton)

$$x' = x - vt \qquad y' = y \qquad z' = z \qquad t' = t$$

no dejan invariantes las ecuaciones de Maxwell, esto es, la forma de las leyes del Electromagnetismo dependerían del SRI escogido. Esto sugiere que existe un sistema de referencia privilegiado en el que las ecuaciones de Maxwell se expresan de la manera más sencilla: aquel que se desplaza con el éter.

En las últimas décadas del siglo XIX se efectuaron experimentos de complejidad creciente para verificar de una vez por todas la hipótesis del éter, entre ellos el archiconocido *experimento de Michelson-Morley* (1887), todos ellos desfavorables a esta creencia. Contrariados pero todavía no dispuestos a desechar la teoría del éter, los físicos George FitzGerald y Hendrik Antoon Lorentz introdujeron la noción de **contracción de longitudes**, que permitía explicar la ausencia de pruebas experimentales favorables al éter. A pesar de ser funcional, muchos objetaron que la contracción de longitudes era una conjetura *ad hoc* y que la formulación matemática de la misma requería la introducción arbitraria de un gran número de parámetros.

En el desarrollo de su teoría de la contracción de longitudes, Lorentz arguyó que un fenómeno semejante afectaba al tiempo (más tarde se bautizaría bajo el nombre **dilatación del tiempo**). En el proceso, derivó una transformación de coordenadas que sí debaja invariantes las ecuaciones de Maxwell: la **transformación de Lorentz**

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}} \qquad y' = y \qquad z' = z \qquad t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}}$$

A partir de estas transformaciones se deducen trivialmente tanto la contracción de longitudes como la dilatación de tiempos.

 ✉ f42rodim@uco.es

3.1.2. La Relatividad Especial

En su artículo de 1905 titulado *Sobre la Electrodinámica de los cuerpos en movimiento*, el físico alemán Albert Einstein obtuvo las transformaciones de Lorentz sin asumir la existencia del éter y produjo, por consiguiente, una teoría equivalente a la de FitzGerald y Lorentz que, además de partir de un menor número de premisas, es mucho más natural. Así, el dogma del éter terminó por disiparse y se impuso la **Teoría de la Relatividad** de Einstein.

Los dos postulados de la Relatividad Especial eran conocidos con anterioridad a la publicación del artículo de 1905. En concreto, ambos fueron enunciados o anticipados de una manera u otra por el matemático francés Henri Poincaré. El primero es una generalización del *principio de relatividad de Galileo* y reza así

Primer postulado: La forma de las leyes de la Física es la misma en todos los sistemas de referencia inerciales.

En consecuencia, no existen sistemas de referencia privilegiados, dado que son todos equivalentes. Por tanto, quedan abolidas tanto la teoría del éter como la hipótesis del *espacio absoluto* consagrada por Isaac Newton. Sorprendentemente, las ecuaciones de Maxwell están en total consonancia con este postulado, dado que no dependen de los estados de movimiento del emisor o del receptor de ondas electromagnéticas.

La aceptación de este postulado nos llevaría además a sustituir la transformación de Galileo por otra que preserve la forma de las ecuaciones de Maxwell (que no es sino la transformación de Lorentz). Sin embargo, las transformaciones de Galileo se conocen desde hace siglos y su bondad está más que comprobada experimentalmente. ¿Cómo es entonces que no son correctas? De momento, obviaremos este hecho.

El objetivo de Einstein era seleccionar un conjunto de postulados que fijaran unívocamente la transformación de coordenadas físicamente aceptable entre sistemas de referencia inerciales. Guiado por principios de economía y naturalidad, probó que bastaba con añadir el siguiente segundo postulado

Segundo postulado: La velocidad de la luz en el vacío es la misma en todos los sistemas de referencia inerciales.

Para un científico que no supiese de las leyes de Maxwell y sus consecuencias, este postulado sería completamente absurdo. No obstante, de acuerdo con las ecuaciones de ondas electromagnéticas que se siguen de las anteriores leyes, la luz en el vacío se propaga a velocidad constante $c = 1/\sqrt{\varepsilon_0 \mu_0}$. Podría entenderse que el segundo postulado es una consecuencia del primero, pues la ecuación de ondas es una ley física *per se*.

Cabe destacar además que la constancia de la velocidad de la luz es incompatible con la transformación de Galileo. En efecto, sean $\mathtt{S}$ y $\mathtt{S}'$ dos SRI (= sistemas de referencia inerciales) tales que $\mathtt{S}'$ se aleja de $\mathtt{S}$ con velocidad relativa v. Un dispositivo láser en reposo con respecto a $\mathtt{S}$ emite un rayo de luz (velocidad c visto desde $\mathtt{S}$). En virtud de la transformación de Galileo, un observador en $\mathtt{S}'$ percibiría que el rayo de luz se propaga con velocidad

$$x' = x + vt \Longrightarrow \frac{dx'}{dt} = \frac{dx}{dt} + v \Longrightarrow c' = c + v \neq c$$

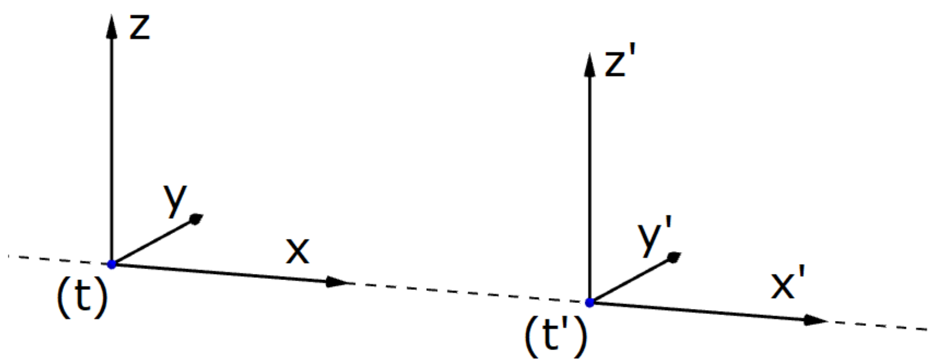
Por el contrario, la transformación de Lorentz sí respeta este postulado.

3.2. La transformación de Lorentz

3.2.1. Una derivación

En esta sección, obtendremos la transformación de Lorentz partiendo exclusivamente de los dos postulados de la Teoría de la Relatividad. Ha de hacerse notar que esta derivación no es la que Einstein incluyó en su artículo original de 1905. Existen numerosas maneras de proceder, y la que presentamos aquí es una de las más sencillas, sin recurrir a herramientas matemáticas sofisticadas.

Consideremos dos SRI, $\mathtt{S}$ y $\mathtt{S}'$, desplazándose uno con respecto al otro a una velocidad v. Dotaremos a ambos sistemas de referencia de los tres ejes cartesianos (x, y, z) y (x', y', z') y de sendas coordenadas temporales no necesariamente idénticas t y t'. Para facilitar las cosas, tomaremos los triedros cartesianos con las mismas direcciones en los dos SRI (i.e. el eje x paralelo al x', etc.) y supondremos que la velocidad relativa v tiene la dirección del eje x, de forma que las otras dos coordenadas espaciales no estarán involucradas en el razonamiento matemático posterior. Los relojes asociados a cada SRI se sincronizarán en el instante inicial $t = t' = 0$, y sus orígenes coincidirán en dicho instante. Nuestra pretensión es hallar unas ecuaciones de transformación $x' = F(x, t), t' = G(x, t)$ que satisfagan los postulados.



Disposición de los ejes de los SRI en movimiento relativo.

En primer lugar, limitaremos nuestra búsqueda a transformaciones *lineales*, esto es, de la forma $x' = \gamma x + \alpha t, t' = \beta x + \phi t$, siendo $\gamma, \alpha, \beta, \phi$ parámetros a determinar. Para convencerse de que ninguna transformación no lineal es físicamente aceptable, disponemos de los dos razonamientos que siguen:

I) Si las transformaciones de coordenadas del sistema $\mathtt{S}$ al $\mathtt{S}'$ vienen dadas por dos funciones $F(x, t), G(x, t)$, los cambios de coordenadas de $\mathtt{S}'$ a $\mathtt{S}$ se corresponderán con sus inversas, $F^{-1}(x', t'), G^{-1}(x', t')$. En general, la forma de una función y la de su inversa no son la misma. Verbigracia, sea $f(x) = x^3$, entonces $f^{-1}(x) = x^{1/3}$.

 ✉ f42rodim@uco.es

Así, una transfomación de coordenadas arbitraria no respetaría el primer postulado, puesto que la transformación de Lorentz es una ley física y su forma ha de ser independiente del SRI. Las únicas funciones que tienen la misma apariencia matemática que sus inversas son las funciones lineales, $f(x) = ax$.

II) La condición de *homogeneidad del espacio* (= carece de puntos distinguidos) obliga igualmente a pedir linealidad en la transformación de coordenadas. Esto es coherente con nuestra intuición. Asumiendo que la transformación es también *diferenciable*, se puede construir un argumento riguroso de que homogeneidad implica linealidad. Lo omitiremos en esta derivación.

Una partícula en reposo respecto de $\mathtt{S}'$ y situada en $x' = 0$ es vista moviéndose con velocidad constante v desde $\mathtt{S}$. Dicho de otro modo, $x = vt$ debe implicar $x' = 0$. Sustituyendo en la ecuación genérica $x' = \gamma x + \alpha t$ se tiene que $\alpha = -v\gamma$, lo que restringe la transformación a $x' = \gamma(x - vt)$.

Merced del primer postulado, la transformación de coordenadas inversa, $x = H(x', t')$, tendrá la misma forma, excepto por el signo de la velocidad (mientras que $\mathtt{S}$ percibe que $\mathtt{S}'$ se aleja con velocidad v, $\mathtt{S}'$ ve cómo $\mathtt{S}$ va distanciándose del primero en sentido contrario, a velocidad $-v$): $x = \gamma(x' + vt')$.

El resto es dejarse llevar por la Matemática. Tomamos diferenciales en las transformaciones directa e inversa y las multiplicamos

$$\begin{cases} dx' = \gamma(dx - vdt) \\ dx = \gamma(dx' + vdt') \end{cases} \Longrightarrow dxdx' = \gamma^2(dx - vdt)(dx' + vdt')$$

Equivalentemente

$$\frac{dx}{dx - vdt} = \gamma^2 \frac{dx' + vdt'}{dx'} \Longrightarrow \frac{dx/dt}{dx/dt - v} = \gamma^2 \frac{dx'/dt' + v}{dx'/dt'}$$

donde hemos dividido numerador y denominador entre dt o dt', dependiendo del miembro de la igualdad. Admitiendo que las coordenadas x, x' van referidas a un rayo de luz propagándose en el vacío, el segundo postulado implica que su velocidad es c en los dos SRI: $dx/dt = dx'/dt' = c$, con lo cual

$$\frac{c}{c-v} = \gamma^2 \frac{c+v}{c} \Longrightarrow \gamma = \frac{1}{\sqrt{1-v^2/c^2}}$$

El parámetro γ se conoce como **factor de Lorentz**. Para abreviar la notación, es frecuencte escribir $\beta \equiv v/c$ de modo que $\gamma = 1/\sqrt{1-\beta^2}$. Llegados a este punto, conocemos

$$x' = \frac{x - vt}{\sqrt{1-\beta^2}} \qquad\qquad x = \frac{x + vt}{\sqrt{1-\beta^2}}$$

Combinando las dos ecuaciones, despejamos t'

$$t' = \frac{1}{v}(x\sqrt{1-\beta^2} - x') = \frac{1}{v}\left(x\sqrt{1-\beta^2} - \frac{x-vt}{\sqrt{1-\beta^2}}\right) = \frac{1 - vx/c^2}{\sqrt{1-\beta^2}}$$

Hemos finalizado:

TRANSFORMACIÓN DE LORENTZ PURA (*BOOST*)

$$x' = \gamma(x - vt) \qquad y' = y \qquad z' = z \qquad t' = \gamma\left(t - \frac{vx}{c^2}\right) \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

Para velocidades inferiores a la de la luz, $v < c$, el factor de Lorentz es estrictamente mayor que 1. En particular, cuando las velocidades relativas son muy inferiores a la de la luz, $\gamma \cong 1$ y recuperamos la transformación de Galileo. Esto hace presagiar que *la Mecánica Newtoniana es el límite de velocidades muy pequeñas de la Relatividad Especial* . Debemos entender, por ende, que la Mecánica Newtoniana constituye un marco adecuado para estudiar fenómenos físicos compatibles con la condición $v \ll c$, que se cumple siempre en situaciones macroscópicas en las que el campo electromagnético no está presente. Aunque resulte turbador, la Física no deja de ser un conjunto de aproximaciones y modelos. La realidad última de las cosas está con casi toda probabilidad fuera de nuestro alcance.

3.2.2. Algunas consecuencias

Definimos **suceso** como una colección de coordenadas (x, y, z, t) dadas con respecto a un sistema de referencia inercial determinado. A la luz de la transformación de Lorentz, las coordenadas de un suceso (incluida la temporal) cambian al pasar de un SRI a otro. Es por esto que se suele decir que *espacio y tiempo son relativos*. Sin embargo, el suceso posee una realidad física que trasciende toda elección de coordenadas o sistemas de referencia.

Relatividad de la simultaneidad: Dado un SRI $\mathtt{S}$ con coordenadas asociadas (x, y, z, t), la coordenada t se interpreta como el tiempo marcado por un reloj en reposo respecto de este. Consideremos dos sucesos S_1, S_2 cuyas coordenadas en $\mathtt{S}$ son (x_1, y_1, z_1, t) y (x_2, y_2, z_2, t). Para un observador en $\mathtt{S}$, S_1 y S_2 son simultáneos. Por el contrario, desde un segundo SRI $\mathtt{S}'$, los sucesos S_1, S_2 tienen coordenadas temporales distintas

$$t_1' = \gamma\left(t - \frac{vx_1}{c^2}\right) \qquad\qquad t_2' = \gamma\left(t - \frac{vx_2}{c^2}\right)$$

Más generalmente

Relatividad de la simultaneidad: Dos sucesos simultaneos en un SRI no tienen por qué serlo en otro SRI.

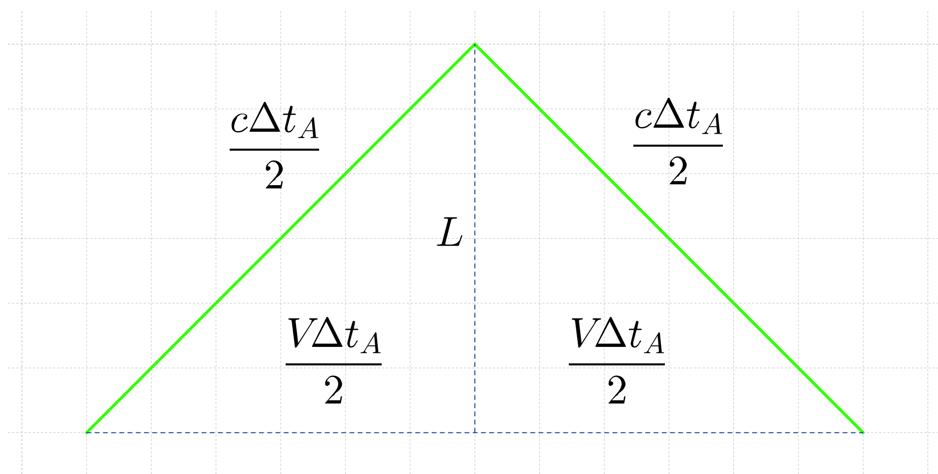
Dilatación de tiempos: El resultado anterior pone de manifiesto que, usualmente, los tiempos medidos por dos SRI en movimiento relativo difieren. A continuación, hallaremos la expresión analítica que los relaciona. Antes de echar mano de la transformación de Lorentz, expondremos un experimento mental del cual se sigue que la constancia de la velocidad de la luz compele a renunciar a la idea de *tiempo absoluto*.

Beny es uno de los alumnos más pendencieros de su escuela. Tras sucesivas participaciones en el juego de los patitos de la feria, fue recompensado con un puntero láser. A bordo de un tren, extrajo el láser y lo dispuso verticalmente, apuntando al techo del vagón. Tan pronto como presionó el interruptor vio un punto verde luminoso en la techumbre, amén de su reflejo incandescente sobre el propio puntero. Casi igual de instantánea fue la reprimenda que recibió por parte de su madre, que vislumbraba a su hijo marchar desde el andén: *¡Beny, desconecta de una vez esa fuente coherente de fotones!* Los bramidos se advirtieron en toda la estación.

Asignaremos sendos sistemas de referencia inerciales `T` (tren) y `A` (andén) a los dos observadores involucrados en esta peculiar narración. Notaremos V a la celeridad del convoy referida al andén y L a la separación entre el suelo y el emisor láser. Beny, cuyas observaciones se rigen por el sistema de referencia `T`, encontrará que el rayo de luz va al techo y regresa describiendo segmentos perpendiculares al mismo de longitud L. Por tanto, el tiempo que consume en cubrir el trayecto completo es

$$\Delta t_T = \frac{2L}{c}$$

¿Qué hay del observador que opta por experimentar el fenómeno desde la óptica del sistema de referencia `A`? La irritada parienta, con respecto a la cual su hijo y el puntero son objeto de una traslación de velocidad uniforme V, asignará al rayo de luz un recorrido semejante al que se muestra en el siguiente esquema



Trayectoria del rayo de luz en el sistema de referencia `A`.

donde hemos asumido que la separación vertical entre el disparador y el techo es la misma en los dos sistemas de referencia (L). Aplicando entonces el teorema de Pitágoras

$$\left(\frac{c\Delta t_A}{2}\right)^2 = L^2 + \left(\frac{V\Delta t_A}{2}\right)^2 \Longrightarrow \Delta t_A = \frac{2L}{c}\frac{1}{\sqrt{1-V^2/c^2}} = \frac{2L}{c}\frac{1}{\sqrt{1-\beta^2}}$$

✉ f42rodim@uco.es

Como c es la misma en los dos SRI, llegamos a que $\Delta t_T \neq \Delta t_A$. Más específicamente

$$\Delta t_A = \frac{\Delta t_T}{\sqrt{1-\beta^2}} = \gamma \Delta t_T \Longrightarrow \Delta t_A > \Delta t_T$$

Esto significa que, si madre e hijo portaran relojes funcionales, la madre observaría que su reloj se adelanta con respecto al de su vástago, de ahí la denominación *dilatación de tiempos*.

Este misma conclusión puede derivarse directamente de la transformación de Lorentz. Sean S_1 y S_2 sucesos cuyas coordenadas referidas a un sistema de referencia móvil $\mathtt{S}'$ son (x', y', z', t') y $(x', y', z', t' + \Delta t')$, respectivamente. En clave del sistema de referencia fijo $\mathtt{S}$, los sucesos tienen las siguientes coordenadas temporales

$$t + \Delta t = \gamma(t' + \Delta t' + vx'/c^2) \qquad\qquad t = \gamma(t' + vx'/c^2)$$

Calculando su diferencia, resulta

> Dilatación de tiempos
>
> $$\Delta t = \gamma \Delta t'$$

Este fenómeno no debe interpretarse como un mal funcionamiento de los relojes que se retrasan o adelantan. Al contrario, los cronómetros cuya lectura temporal se distorsiona según $\Delta t = \gamma \Delta t'$ son los que marchan adecuadamente. Evidentemente, con *reloj* nos referimos a cualquier entidad que se vea afectada por el paso del tiempo, sin excluir el tiempo biológico (la dilatación del tiempo es el fundamento de la versión simplificada de la *paradoja de los gemelos*).

La dilatación de tiempos es una suerte de ilusión que incumbe a todo reloj ajeno al propio observador. Desde su perspectiva, el cronómetro que porta consigo jamás se retrasa o adelanta, incluso si se desplaza a velocidades próximas a c. El tiempo marcado por este reloj se conoce como **tiempo propio** del observador, y se suele simbolizar por τ.

Ejemplo: El *muón* es una partícula subatómica que comparte todas sus características con el afamado electrón excepto su masa. La interacción de rayos cósmicos con las capas externas de la atmósfera genera un alto número de muones. Se sabe que la vida media del muón es de unos $1{,}56 \cdot 10^{-6}$ s (es decir, en promedio, el muón se desintegrará transcurrido ese tiempo desde su producción). La velocidad del muón es elevada, pero no lo suficiente como para alcanzar la superficie terrestre en ese tiempo. Los datos experimentales nos hacen dudar de esta reflexión, pues se detectan muchos más muones de los que esta sugiere. La Relatividad Especial y, concretamente, la dilatación de tiempos, resuelven el dilema. Ciertamente, tomando como velocidad del muón $v = 0{,}98c$ respecto del suelo terrestre, el factor de Lorentz vale $\gamma \cong 5$. De esta forma, la vida media del muón, medida desde la superficie de la Tierra, es aproximadamente $5 \cdot 1{,}56 \cdot 10^{-6}$ s, lo que le permite recorrer una distancia cinco veces mayor que la predicha por la Física Prerrelativista.

Contracción de longitudes: Tómense ahora dos sucesos S_1, S_2 cuyas coordenadas en el sistema fijo S son (x, y, z, t) y $(x + \Delta x, y, z, t)$, respectivamente. Usamos la transformación de Lorentz para obtener las coordenadas horizontales de estos sucesos vistos desde el sistema móvil S′

$$x' + \Delta x' = \gamma(x + \Delta x - vt) \qquad x' = \gamma(x - vt)$$

Sustrayéndolas, se encuentra la relación

Contracción de longitudes

$$\Delta x = \frac{\Delta x'}{\gamma}$$

Dado que $\Delta x < \Delta x'$, concluimos que todo cuerpo aparenta contraerse en la dirección de su movimiento en un factor γ. A veces, se dice que $\Delta x'$, la extensión del cuerpo en el SRI respecto del que se halla en reposo, es su *longitud propia*. Tanto la dilatación de tiempos como la contracción de longitudes se producen solo en esta dirección, puesto que $y' = y, z' = z$.

Ejemplo: En el apartado anterior, lidiamos exitosamente con la discrepancia entre los datos teóricos y experimentales de la desintegración del muón, *analizando el problema desde un sistema de referencia en reposo con respecto a la superficie de nuestro planeta*. ¡Pero el muón no ve su vida media dilatada desde su propio SRI! Lo que sí experimenta el leptón es una contracción del tamaño de las dimensiones de la atmósfera en la dirección del movimiento, en un factor $\gamma \cong 5$. Su evolución temporal no se ralentiza, mas ha de recorrer una distancia aparente menor hasta alcanzar la corteza terrestre.

3.2.3. Adición relativista de velocidades

Según Galileo y otros científicos 'absolutistas', las coordenadas de dos sistemas de referencia inerciales están relacionadas según $x' = x - vt$. ¿Cómo están ligadas las velocidades? Basta con derivar ambos miembros de la igualdad respecto al único tiempo del que tiene sentido hablar en Física Prerrelativista, el *tiempo absoluto* t, arrojando $V' = V - v$.

Es una verdadera desgracia que las transformaciones *buenas*, las de Lorentz, son algo más complejas y, en particular, no contemplan un tiempo absoluto. Por lo tanto, la cuestión antes formulada en terreno relativista carece de una respuesta tan directa como lo era *derivar con respecto al tiempo*. *¿Qué tiempo?* Cada observador empleará el tiempo asociado a su SRI.

Explícitamente, dado un cuerpo móvil de coordenadas espaciales (x', y', z') en $\mathtt{S}'$, les serán asignadas las siguientes componentes a su vector velocidad

$$u'_x = \frac{dx'}{dt'} \qquad u'_y = \frac{dy'}{dt'} \qquad u'_z = \frac{dz'}{dt'}$$

E igualmente, *mutatis mutandis*, en $\mathtt{S}$. Partamos, pues, de la transformación de Lorentz de $\mathtt{S}'$ a $\mathtt{S}$

$$x = \gamma(x' + vt') \qquad y = y' \qquad z = z' \qquad t = \gamma\left(t' + \frac{vx'}{c^2}\right)$$

✉ f42rodim@uco.es

Diferenciamos todas estas expresiones

$$dx = \gamma(dx' + vdt') \qquad dy = dy' \qquad dz = dz' \qquad dt = \gamma\left(dt' + \frac{vdx'}{c^2}\right)$$

De esta forma, se tienen los cocientes

$$\frac{dx}{dt} = \frac{dx' + vdt'}{dt' + vdx'/c^2} \qquad \frac{dy}{dt} = \frac{dy'}{\gamma(dt' + vdx'/c^2)} \qquad \frac{dz}{dt} = \frac{dz'}{\gamma(dt' + vdx'/c^2)}$$

Los miembros izquierdos de estas ecuaciones son las componentes de la velocidad en el sistema S, (u_x, u_y, u_z). Para terminar, dividimos los numeradores y los denominadores de los miembros derechos entre dt', obteniendo

Transformación de velocidades

$$u_x = \frac{u'_x + v}{1 + vu'_x/c^2} \qquad u_y = \frac{u'_y}{\gamma(1 + vu'_x/c^2)} \qquad u_z = \frac{u'_z}{\gamma(1 + vu'_x/c^2)}$$

Es menester cerciorarse de que la transformación encontrada cumple:

I) *Se reduce a la transformación de Galileo en el límite $v \ll c$.* Afirmativo, puesto que, en el citado límite, $\gamma \cong 1$ y el sumando vu'_x/c^2 es perfectamente despreciable. Así

$$u_x = u'_x + v \qquad u_y = u'_y \qquad u_z = u'_z$$

II) *Respeta el postulado de la constancia de la velocidad de la luz.* Afirmativo. Digamos que $u'_x = c, u'_y = 0, u'_z = 0$, entonces

$$u_x = \frac{c + v}{1 + v/c} = c\frac{c + v}{c + v} = c \qquad u_y = 0 \qquad u_z = 0$$

Textualmente, un rayo de luz que se propaga en la dirección del eje x lo hace a velocidad c en todos los sistemas de referencia inerciales no rotados entre sí. ¿Qué sucede si el rayo se mueve en otra dirección, por ejemplo $u'_x = 0, u'_y = c, u'_z = 0$? La transformación de velocidades devuelve

$$u_x = v \qquad u_y = \frac{c}{\gamma} = c\sqrt{1 - \beta^2} \qquad u_z = 0$$

 ✉ f42rodim@uco.es

Resulta una velocidad con más de una componente. Empero, su *módulo*, en conformidad con nuestro estimado segundo postulado, habría de seguir siendo c. Pues bien

$$\|\mathbf{u}\| = \sqrt{v^2 + c^2(1-\beta^2)} = \sqrt{v^2 + c^2\left(1+\frac{v^2}{c^2}\right)} = c$$

3.3. Dinámica Relativista

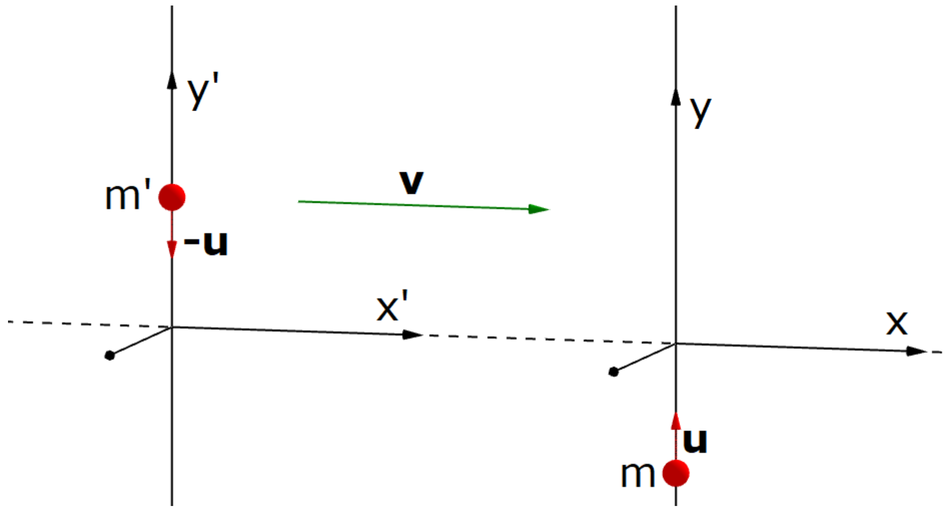
3.3.1. ¿Masa relativista?

Hasta ahora, analizamos cómo los postulados de la Relatividad condicionan el movimiento de partículas libres (i.e. no aceleradas) sin preocuparnos por las causas motrices (*Cinemática Relativista*). En el resto de esta introducción elemental a la teoría de Einstein *et al*, introduciremos algunas nociones de Dinámica Relativista. Esperamos que el lector no se sienta decepcionado al no encontrar una generalización satisfactoria de *fuerza* a este inverosímil marco teórico. Sirva de consuelo que, al menos, sí hallará cantidades de movimiento y energías relativistas.

¿A qué vienen las interrogaciones en el título de esta sección? Si bien fue algo completamente estándar durante las décadas posteriores al nacimiento de la Relatividad, el concepto de **masa relativista** ha sido vilipendiado por cuantiosos doctos y doctas en la teoría einsteiniana a causa de su naturaleza abstrusa. El autor de estos apuntes pertenece a este colectivo (quizá no tanto por docto como por detractor). Por desfortuna para él, ante la imposibilidad de colegir argumentos que justifiquen su encono para con la masa relativista sin que el lector disponga de conocimientos acerca de esta vaina, mantendrá una ideología neutral que cesará abruptamente unas páginas más adelante.

Iniciémonos en el insólito mundo de la Dinámica Relativista por medio de un *gedankenexperiment*[32]. Se tienen dos SRI $\mathtt{S}$ y $\mathtt{S}'$ de ejes horizontales (x) coincidentes, como es costumbre, y velocidad relativa v. En el eje vertical (y) de cada SRI hay ensartada una pe-

[32] Experimento mental. Término acuñado por el físico y pensador austriaco Ernst Mach, una de las fuentes de inspiración de Albert Einstein.



Disposición de los SRI y las esferas en un instante anterior a la colisión.

queña esfera de masa *en reposo* m_0 (si solo es posible medir la masa de un cuerpo cuando se encuentra en reposo con respecto al instrumento de medición, ¿por qué empeñarnos en utilizar la denominación *en reposo*?), una en el semieje positivo y la otra, en el negativo. Ambas están dotadas de una velocidad constante u dirigida hacia el origen de sus respectivos SRI. Cuadramos la cosa para que los orígenes de los SRI se superpongan en el instante en el que las dos esferas los alcanzan, originándose una *colisión elástica* entre los bolondros (no tendría por qué ser elástica, es algo que imponemos nosotros con el objeto de simplificar la situación).

Como parece que el *leitmotiv* de la Relatividad reza *todo es relativo* (¿no decían que la Física era objetiva?), convendremos en adjudicar masas m y m' a las esferas, percibidas en movimiento desde el sistema de referencia inercial en reposo (¡también relativo!) S. Tirando de la definición newtoniana, las **cantidades de movimiento** $\mathbf{p} = m\mathbf{v}$ de las esferas, antes y después del chosquín, vienen dadas por

$$\mathbf{p}_{\text{before}} = (0, mu, 0) \qquad \mathbf{p}'_{\text{before}} = (m'v, -m'u/\gamma, 0)$$

$$\mathbf{p}_{\text{after}} = (0, -mu, 0) \qquad \mathbf{p}'_{\text{after}} = (m'v, m'u/\gamma, 0)$$

donde se ha tenido en cuenta que, al ser la colisión elástica y unidimensional, las esferas salen despedidas con la misma velocidad modular de la que estaban animadas inicialmente. En ausencia de fuerzas, la cantidad del movimiento total es una constante del movimiento. De ser así, igualando las componentes verticales

$$mu - m'u/\gamma = -mu + m'u/\gamma \Longrightarrow m' = \gamma m$$

El resultado es independiente de la velocidad con la que las esferas se aproximan al origen, u. En particular, en el caso $u = 0$, $m = m_0$ por encontrarse esta esfera en reposo respecto de $\mathtt{S}$. Esto implica que un observador afincado en $\mathtt{S}$ vería la masa de la esfera en movimiento incrementada en un factor γ (*variación relativista de la masa*). Por tanto, su cantidad de movimiento es

Cantidad de movimiento relativista

$$\mathbf{p} = \gamma m \mathbf{v} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}}$$

Concluido el experimento mental, nos preguntamos: ¿Es estrictamente necesario juzgar que la masa de las esferas depende de su estado de movimiento? En absoluto. Resulta que, a efectos formales, es indiferente considerar la expresión de la masa relativista tal y como se acaba de derivar o aceptar que la masa es invariante y modificar la definición de cantidad de movimiento a $\mathbf{p} = \gamma m \mathbf{v}$ (recuperándose la fórmula newtoniana para $v \ll c$). En definitiva, estamos ante una cuestión de preferencias. En lo que sigue se esgrimen algunos de los motivos por los que la mayoría de los físicos actuales han repudiado la idea de masa relativista:

I) So pena de que tanta relatividad nos produzca una indigestión, es oportuno aprovechar que se nos ha concedido la oportunidad de preservar sin rectificación esencial uno de los conceptos fundamentales de la Mecánica Newtoniana: la *masa inercial*. Porque si admitimos que los cuerpos carecen de una masa inherente a su constitución e independiente del movimiento, ¿entonces qué nos queda? ¿Son todas las magnitudes físicas conocidas hasta ahora vanas quimeras? Estamos a tiempo de conservar algo del sabor parmenídeo del que gozaba la Física de toda la vida (posteriormente revelaremos que ni siquiera esta masa m_0 es inmutable, y que las leyes de conservación precisarán ser revisadas).

II) El que la masa varíe con el estado de movimiento del cuerpo puede inducir a pensar que también lo hace su constitución interna. Esto es, como poco, desacertado. Si nos limitamos a actualizar la definición de cantidad de movimiento, que ha sido una función del movimiento desde que fue concebida por Newton, no se genera tal confusión.

III) Uno de los objetivos de aquellos que utilizan la masa relativista m_{rel} es mantener la forma de las ecuaciones de la Mecánica Newtoniana, e.g. $\mathbf{p} = m_{\text{rel}}\mathbf{v}$. No obstante, no es cierta la regla 'lo que era verdadero en Física Newtoniana lo sigue siendo en Física Einsteiniana salvo intercambio de m por m_{rel}', e.g. $\mathbf{F} \neq m_{\text{rel}}\mathbf{a}$ por lo general. Esto hace que adherirse a la masa relativista sea algo arriesgado, sobre todo si no se domina conceptualmente la Teoría de la Relatividad.

En base a las razones anteriores y a otras que se enunciarán más adelante, abandonaremos la *masa relativista* y trabajaremos únicamente con la *masa en reposo* o, sencillamente, *masa* (en Relatividad Especial, no nos preocuparemos por hacer la distinción entre *masa inercial* y *masa gravitatoria*. Como la gravedad está ausente, se sobreentenderá que nos referimos siempre a la primera).

3.3.2. Energía relativista

En el contexto de la Mecánica Newtoniana, existen dos procedimientos para abordar la resolución de un problema: ora se listan todas las fuerzas actuantes y se aplica la segunda ley de Newton, ora se echa mano de la conservación de la energía. Matemáticamente, la segunda vía tiende a ser la más viable. En *Mecánica Analítica*, un formalismo posterior a la Mecánica de Newton pero físicamente equivalente a la misma, se estudia la evolución temporal de los sistemas mecánicos a partir de las funciones *lagrangiana* y *hamiltoniana*, que dependen directamente de las energías cinética y (en la mayoría de los casos de interés) potencial de estos. La *Mecánica Cuántica*, que debe buena parte de su formulación a la de Hamilton y Lagrange, es una teoría eminentemente energética. Algo parecido acaece en Relatividad: conforme se progresa en su estudio, más se ocultan las fuerzas y más se dejan ver las energías.

Con el fin de encontrar la expresión de la energía relativista de una partícula de masa m inicialmente en reposo, rescataremos de la Física de Newton la siguiente fórmula para determinar el trabajo elemental dW efectuado por la fuerza $\mathbf{F}$ al producir sobre la partícula un desplazamiento $d\mathbf{r}$

$$dW = \mathbf{F} \cdot d\mathbf{r} = \frac{d\mathbf{p}}{dt} \cdot d\mathbf{r} = d\mathbf{p} \cdot \frac{d\mathbf{r}}{dt} = d\mathbf{p} \cdot \mathbf{v}$$

En virtud de la regla del producto o de Leibniz de la derivada

$$d(\mathbf{p} \cdot \mathbf{v}) = d\mathbf{p} \cdot \mathbf{v} + \mathbf{p} \cdot d\mathbf{v} \Longrightarrow d\mathbf{p} \cdot \mathbf{v} = d(\mathbf{p} \cdot \mathbf{v}) - \mathbf{p} \cdot d\mathbf{v}$$

El momento relativista guarda relación con la velocidad de la partícula

$$\mathbf{p} = \frac{m\mathbf{v}}{\sqrt{1-\beta^2}}$$

Por tanto, usando que $\beta = v/c$

$$dW = d\left(\frac{mv^2}{\sqrt{1-\beta^2}}\right) - \frac{mv}{\sqrt{1-\beta^2}}\,dv = d\left(\frac{mc^2\beta^2}{\sqrt{1-\beta^2}}\right) - \frac{mc^2\beta}{\sqrt{1-\beta^2}}\,d\beta$$

Finalmente, integramos entre 0 y cierto valor de β (esto es, entre las velocidades inicial y final). La integración que queda es inmediata

$$\begin{aligned} W &= \int_0^\beta \left[d\left(\frac{mc^2\tilde{\beta}^2}{\sqrt{1-\tilde{\beta}^2}}\right) - \frac{mc^2\tilde{\beta}}{\sqrt{1-\tilde{\beta}^2}}\,d\tilde{\beta}\right] = mc^2\left(\frac{\beta^2}{\sqrt{1-\beta^2}} + \sqrt{1-\beta^2} - 1\right) \\ &= mc^2\left(\frac{1}{\sqrt{1-\beta^2}} - 1\right) = mc^2(\gamma - 1) \end{aligned}$$

Así, la energía cinética que adquiere la partícula al pasar de velocidad 0 a β (*teorema de las fuerzas vivas*, $W = \Delta E_c \equiv \Delta T$) es

✉ f42rodim@uco.es

Energía cinética relativista

$$T = mc^2(\gamma - 1)$$

A diferencia de la energía cinética prerrelativista, $T = mc^2(\gamma - 1) = \gamma mc^2 - mc^2$ involucra dos sumandos, uno de los cuales es independiente de la velocidad de la partícula, i.e. mc^2. A priori, esto no parece tener mucho sentido. Con la pretensión de asegurarnos de que el resultado alcanzado es bondadoso, supondremos que $v \ll c$, buscando recobrar $T \cong mv^2/2$. La burda aproximación $\gamma = 1$ arroja $T = 0$ y no es funcional en este contexto. Hemos de ser algo más cuidadosos: Desarrollaremos $T = T(\beta)$ en *serie de Taylor*[33] en torno a $\beta = 0$

$$T = \left(mc^2 + \frac{1}{2}mv^2 + \frac{3}{8}mv^4 + \frac{5}{16}mv^6 + \dots\right) - mc^2$$

Cuando $v \ll c$, los términos de orden 4 y superior son despreciables. Podemos, de este modo, suspirar aliviados al ver que $T \cong mv^2/2$.

Siguiente reto: dotar de una interpretación placentera al término constante mc^2. Es claro que tiene dimensiones de energía y que es una cualidad inherente a la partícula, exenta de *relatividad*. Aunque se trata de una hipótesis impensable fuera de este marco teórico, lo más natural es postular que mc^2 es una energía que la partícula posee con independencia de su estado de movimiento. En tal caso, su energía total vendría dada por $E = T + mc^2 = \gamma mc^2$, donde ya no figura ningún sumando constante.

Energía total relativista

$$E = \gamma mc^2$$

[33]Dada una función real de una variable $f(x)$, su *serie de Taylor* asociada centrada en x_0 es la suma infinita $\sum_{n=0}^{\infty} \frac{f^{n)}(x_0)}{n!}(x - x_0)^n$, siendo $f^{n)}$ la derivada n-ésima de f. Si la función es lo suficientemente bien comportada, su serie de Taylor converge a esta en cierto entorno de x_0.

 ✉ f42rodim@uco.es

El asignar cierta energía a un cuerpo por el simple hecho de ser masivo es un atrevimiento osado con serias implicaciones prácticas. Para empezar, cuesta creerlo, ya que en nuestra vida rutinaria no somos sorprendidos por desapariciones y apariciones espontáneas de cuerpos extensos. ¿Se ha observado experimentalmente la conversión de masa en otras clases de energía, e.g. energía cinética o radiación? No solo la respuesta es un rotundo SÍ, sino que, además, algo más del 70 % de la energía consumida en Francia se genera gracias a este fenómeno inédito. El siguiente ejemplo evidencia su descomunal eficiencia:

Ejemplo: El

4. Termodinámica del equilibrio y geometría de contacto - Motivación

4.1. Formas diferenciales y termodinámica

Podría decirse que la física es una colección de modelos que tratan de describir, siempre con cierto grado de idealización, determinadas parcelas del universo sensible. Como buena teoría física, este es el caso de la termodinámica del equilibrio, que gira en torno a una de las más fantasiosas idealizaciones. En efecto, en esta disciplina, todo sistema se postula permanentemente en equilibrio termodinámico, por lo que dicho estado, diríamos, habría de permanecer eternamente inalterado. Sin embargo, de forma aparentemente paradójica, se resuelven problemas de gases que se expanden o medios que se magnetizan. Para ello, se admite que la evolución del sistema procede siempre por medio de estados de equilibrio termodinámico, requiriendo por tanto un tiempo infinito para completar el proceso, que denominamos *cuasiestático*. De esta manera, los parámetros termodinámicos varían de manera suave y están bien definidas sus derivadas con respecto a algún parámetro de evolución (que no tiene por qué tomarse igual al tiempo). Además, tal variación ocurre tan lentamente que tiene sentido considerarla infinitesimal. Es por eso que las ecuaciones de la termodinámica vienen generalmente expresadas en términos de diferenciales (lo que tradicionalmente se conoce como *ecuaciones pfaffianas*).

Por ejemplo, el primer principio suele escribirse

$$dU = T\,dS - \sum_i X_i\,dy_i + \sum_i \mu_i\,dN_i$$

Las 'cantidades' dS, dy_i, dN_i pueden entenderse meramente como infinitésimos cuya integración produce un valor finito o, desde una perspectiva más sofisticada e iluminadora, como *formas diferenciales*[34]. Muchos conceptos y resultados en termodinámica adquieren un sentido natural cuando se utiliza el segundo enfoque. En los cursos estándares de esta materia se establece una diferencia fundamental entre magnitudes como la entropía y la energía interna, por una parte, y el calor o el trabajo, por otra: solo las primeras son *funciones de estado*, esto es, dependen exclusivamente del estado del sistema en el momento de su medición. Dicho de otra manera, la variación de energía interna o de entropía no es función de la trayectoria que la evolución del sistema realiza en el espacio de parámetros termodinámicos, apenas importan sus extremos. Esto se traduce en que, infinitesimalmente, energía interna y entropía son formas diferenciales *exactas*, no así el calor o el trabajo, cuya integral entre dos puntos depende en general del camino escogido para conectarlos.

Puede ocurrir que, no siendo una forma diferencial exacta, el producto de esta por una función, denominada *factor integrante*, lo sea. En verdad, lo que se requiere es que esta forma diferencial sea *cerrada*. Ahora bien, en lo que sigue supondremos siempre que las variedades diferenciables sobre las que definimos todo el aparataje son simplemente conexas, de manera que las formas cerradas y las exactas coinciden. Esto sucede con el calor, pues de multiplicar δQ por el inverso de la temperatura $1/T$ resulta exactamente la entropía dS. El *teorema de Carathéodory*, a veces abstruso cuando se estudia desde el *approach* clásico, se desprende fácilmente en clave geométrica. Este afirma que en todo entorno de un estado termodinámico existen estados inaccesibles por medio de un proceso adiabático cuasiestático. Geométricamente, un proceso es una curva $\gamma : I \subseteq \mathbb{R} \longrightarrow \mathscr{M}$, donde $\mathscr{M}$ es el espacio de estados termodinámicos al que hemos dotado

[34] De hecho, esta es típicamente considerada la forma más rigurosamente correcta de interpretar todas las cantidades diferenciales en física y matemáticas.

 ✉ f42rodim@uco.es

de estructura de variedad diferenciable[35]. Si $\delta Q \in \Lambda^1(\mathscr{M})$[36], la curva γ representa un proceso adiabático cuando $\delta Q(\gamma') = 0$ para todos sus puntos. Invocamos ahora que

$$dS = \frac{\delta Q}{T} \Longrightarrow \delta Q = T\, dS$$

Así

$$\delta Q(\gamma') = T\, dS(\gamma') = T\, \gamma'(S) = T\, \frac{d}{d\lambda}(\gamma \circ S) = 0$$

Por tanto, no es posible ir entre estados de diferentes entropías a través de un proceso adiabático cuasiestático.

4.2. Primer principio y forma de contacto

Consideremos, por simplificar, un sistema termodinámico simple que satisface el primer principio en la forma

$$dU = T\, dS - p\, dV \tag{2}$$

Equivalentemente, estamos postulando la nulidad de la 1-forma[37]

$$\eta \equiv dU - T\, dS + p\, dV = 0$$

No obstante, η no es idénticamente nula en $\mathscr{M}$, solo lo es a lo largo de las curvas que implementan procesos cuasiestáticos de algún sistema termodinámico. Vamos a probar que cierto producto de η y su derivada exterior $d\eta$ sí se anula. Tenemos

$$d\eta = dp \wedge dV - dT \wedge dS$$

$$\eta \wedge d\eta = dU \wedge dp \wedge dV - dU \wedge dT \wedge dS - T\, dS \wedge dp \wedge dV - p\, dV \wedge dT \wedge dS$$

$$d\eta \wedge d\eta = -2dp \wedge dV \wedge dT \wedge dS$$

[35] En las referencias de termodinámica geométrica se da a $\mathscr{M}$ el nombre de *espacio de fases termodinámico*, en similitud con el espacio de fases en mecánica clásica. Se verá a medida que avancemos en nuestro estudio que hay cuantiosos puntos en común entre los formalismos geométricos que subyacen la mecánica y la termodinámica, hasta el extremo de que el segundo (la geometría de contacto) se describe como la extensión del primero (la geometría simpléctica) a variedades de dimensión impar.

[36] Igualmente, podríamos haber escrito $\delta Q \in \Gamma(T\mathscr{M}^*)$

[37] Algunos autores llaman a η la *forma de Gibbs*.

$$\eta \wedge (d\eta)^2 = -2dU \wedge dp \wedge dV \wedge dT \wedge dS$$

Vemos que $\eta \wedge (d\eta)^2 \neq 0$, pero que cualquier producto de η o $d\eta$ con este es cero. Las 1-formas con la citada propiedad se conocen como *formas de contacto*. De este modo, la primera ley de la termodinámica dota al espacio de parámetros termodinámicos de una *estructura de contacto* natural, esto es, la dupla $(\mathscr{M}, \eta)$. Podría pensarse que introducir la forma de contacto supone una complicación innecesaria, pero, como veremos, la termodinámica del equilibrio puede formularse de manera sencilla y elegante por medio de esta estructura.

¿Son la estructura de contacto y la forma precisa del primer principio (2) genuinamente equivalentes? La respuesta es afirmativa y se sustenta en el celebrado *teorema de Darboux*, en virtud del cual cualquier 1-forma de contacto α (en el caso 5-dimensional) se escribe (al menos localmente) como $\alpha = dU - T\,dS + p\,dV$ para algunas funciones U, T, S, p, V a las que se dota entonces de su significado físico habitual y se tratan como las coordenadas del espacio termodinámico, motivo por el que se les llama *coordenadas de Darboux*.

4.3. Sistemas termodinámicos y subvariedades de Legendre

Tradicionalmente, un sistema termodinámico concreto (en equilibrio) se define a través de un conjunto de *ecuaciones de estado* que ligan las variables termodinámicas, exigiendo el cumplimiento de la primera ley de la termodinámica (2). Por ejemplo, para los sistemas termodinámicos simples de la sección anterior, lo habitual es introducir dos ligaduras $f(p, V, T) = 0$ y $g(U, V, T) = 0$, denominadas ecuaciones térmica y energética de estado, respectivamente, e imponer el primer principio, que es equivalente a establecer una tercera ecuación $h(S, V, T) = 0$ para la entropía. ¿Cuál es el significado geométrico de los anteriores trapicheos? El espacio de parámetros termodinámicos sin constreñir, que antes habíamos simbolizado por $\mathscr{M}$, es una variedad diferenciable de cinco dimensiones. A primera vista, no hay motivos para no tomar $\mathscr{M} = \mathbb{R}^5$ y llamar (U, S, p, V, T) a sus funciones coordenadas. Cada una de las relaciones $f = 0, g = 0, h = 0$ define, si es lo suficientemente bien comportada, una *subvariedad* de $\mathscr{M}$, una región cuya dimensión es la total menos una ($5 - 1 = 4$), porque una de las variables termodinámicas queda expresada como función del resto. Si requerimos las tres condiciones simultáneamente, estamos limitándonos a la región que resulta de intersecar tres subvariedades de dimensión 4 y que tendrá, en general, dimensión 2. Por eso, un sistema termodinámico simple se corresponderá con una subvariedad bidimensional $\mathscr{L} \subset \mathscr{M}$.

Yendo hacia una formulación más geométrica y abstracta de la termodinámica del equilibrio, ¿bastaría con postular que un sistema termodinámico *es* una subvariedad bidimensional? No, ya que esto no garantiza la primera ley de la termodinámica. Recordemos que en la sección previa codificamos dicha ley en la 1-forma de contacto η, diciendo que (2) se satisface si η se hace cero. Así pues, un sistema termodinámico vendría dado por una subvariedad sobre la que la 1-forma de contacto es cero. [Pero esta última afirmación no es del todo rigurosa, ya que η está definida sobre todo $\mathscr{M}$, mientras que una definición geométricamente satisfactoria de sistema termodinámico debe ser lo más *intrínseca* posible]. Ahora bien, η induce una 1-forma sobre la subvariedad $\eta_{\mathscr{L}}$ dada por su *pullback* bajo la aplicación inclusión $i : \mathscr{L} \hookrightarrow \mathscr{M}$, es decir, $\eta_{\mathscr{L}} \equiv i^*\eta$. La condición para que $\mathscr{L}$ represente un sistema termodinámico en equilibrio es, pues, que $\eta_{\mathscr{L}} \equiv 0$. Puede verse que de $\eta \wedge (d\eta)^2 \neq 0$ se sigue que toda subvariedad sobre la que $\eta_{\mathscr{L}} \equiv 0$ tiene dimensión máxima 2. En el contexto matemático de la geometría de contacto, estas subvariedades se conocen como *subvariedades de Legendre*.

4.4. Potenciales termodinámicos y transformada de Legendre

Si bien antes mencionamos que son necesarias, normalmente, varias ecuaciones para especificar por completo un sistema termodinámico, a veces basta con solo una, sujeta al primer principio. Ciertamente, dada la energía interna en función de S y de V, esto es, $U(S, V)$, su diferencial es

$$dU = \frac{\partial U}{\partial S}\, dS + \frac{\partial U}{\partial V}\, dV$$

Comparando con (2), deducimos que

$$T = \frac{\partial U}{\partial S} \qquad\qquad p = -\frac{\partial U}{\partial V}$$

Estas dos ecuaciones de estado, combinadas con la primera ley, definen unívocamente una subvariedad de Legendre, y han sido engendradas a partir de $U(S,V)$. Acostumbra a decirse que, en conteniendo toda la información relativa al sistema termodinámico, $U(S,V)$ es una *ecuación fundamental* y que S,V son las *variables naturales* de la energía interna, ya que solo en función de estas cualifica como ecuación fundamental.

La fórmula $U(S,V)$ no es la única ecuación fundamental, y en los textos de termodinámica figura toda una hueste de funciones que, expresadas en términos de sus correspondientes variables naturales, son equivalentes al conocimiento de esta primera. Las más sonadas son la *entalpía* $H(S,p)$, la función de Helmholtz $F(V,T)$ y la *función de Gibbs* $G(p,T)$. Nos referiremos a estas, a la energía interna y a otras que hemos omitido como *potenciales termodinámicos*. El motivo por el que todas almacenan la misma información es que no se trata de funciones independientes, sino que se obtienen las unas de las otras por medio de *transformadas de Legendre*

$$H = U + pV \qquad G = H - TS \qquad F = U - TS$$

¿Cómo entroncan estas consideraciones con la descripción basada en geometría de contacto? Demostraremos que toda subvariedad de Legendre es generada por una función de dos variables $f(S,V), f(S,p), f(V,T), f(p,T)$ de una manera concreta (y que se entenderían como la energía interna, entalpía, función de Hemholtz o función de Gibbs del sistema, respectivamente). Denotando[38] $q^1 = S, q^2 = V, p_1 = T, p_2 = -p$, entonces cualquiera de estas funciones presenta la dependencia $f(p_i, q^j)$ donde $i \in I, j \in J$ e $I, J \subseteq \{1,2\}$ constituyen una partición[39] de $\{1,2\}$. En condiciones propicias, f permite construir, al menos localmente, la subvariedad bidimensional que es lugar geométrico de las ecuaciones

$$q^i = -\frac{\partial f}{\partial p_i} \qquad p_j = \frac{\partial f}{\partial q^j} \qquad U = f - \sum_{i\in I} p_i \frac{\partial f}{\partial p_i}$$

Obsérvese que la postrera ecuación es la versión general de la transformada de Legendre tal y como se estudia en los cursos de termodinámica. Las dos primeras fórmulas son

[38] Esta notación es también propia del ámbito de la mecánica teórica.

[39] $I \cup J = \{1,2\},\ I \cap J = \emptyset$

ecuaciones de estado para los dos parámetros termodinámicos restantes[40]. Así, según escogamos un potencial termodinámico u otro, tendremos distintas parametrizaciones de una misma subvariedad de Legendre. La multiplicidad de potenciales termodinámicos es por tanto análoga a la libertad en la elección de coordenadas o de bases. La subvariedad de Legendre, i.e. el sistema termodinámico, desde esta perspectiva geometrizada, existe más allá de que establezcamos una de sus ecuaciones fundamentales.

Siendo rigurosos, habríamos de probar que todas las ideas del párrafo previo son consistentes entre sí. El teorema asevera que a toda subvariedad de Legendre corresponde, al menos, una función f que la genera, pero cabe preguntarse en qué sentido dos funciones (posiblemente relacionadas por una transformada de Legendre) producen una misma subvariedad. Argüiremos que si f y $\tilde{f}$ son transformadas de Legendre recíprocas, sus subvariedades de Legendre son difeomorfas, lo que implica que representan a nuestros efectos el mismo sistema. Esto es consecuencia de que las transformadas de Legendre se incluyen en una amplia clase de transformaciones que preservan la estructura de contacto, a veces denominadas *contactomorfismos*.

4.5. Procesos termodinámicos y campos hamitlonianos de contacto

Sea $X \in \Gamma(T\mathscr{M})$ un campo vectorial sobre el espacio termodinámico $\mathscr{M}$. Las curvas integrales de X, recogidas colectivamente en su aplicación flujo, definen la evolución de los parámetros termodinámicos a medida que se incrementa el parámetros de la curva τ, que no debe interpretarse como el tiempo físico por la naturaleza cuasiestática del fenómeno. Es a través de estas curvas que abstraemos los procesos que llevan a un sistema termodinámico en equilibrio de un estado inicial a un estado final, mas no todas las curvas cumplen con el propósito. Si aspiramos a analizar los devenires de un sistema particular, lo que geométricamente se traduce en restringirnos a su subvariedad de Legendre $\mathscr{L}$, una curva γ podrá codificar una cierta evolución del sistema si está contenida en $\mathscr{L}$. De lo contrario, γ constaría de puntos que no representan ningún estado de

[40] Y recuerdan a las ecuaciones de Hamilton.

nuestro sistema. Para que las curvas integrales de X no abandonen la subvariedad, es preciso que X sea tangente a $\mathscr{L}$. Por ende, se tendrá $\eta(X) = 0$ para todos los puntos de $\mathscr{L}$, en cuyo caso cada (segmento de) curva integral de X se corresponde con un proceso del sistema.

Solo las curvas contenidas en una subvariedad de Legendre incorporan procesos termodinámicos físicos, pero algo no tiene por qué ser físico para despertar nuestro interés. En el mundo de la termodinámica geométrica no es extraño contemplar campos cuyas curvas integrales intersecan una sola vez a cada elemento de una familia de subvariedades de Legendre. Estas sirven pues para definir *familias uniparamétricas* de sistemas termodinámicos o, dicho de otra manera, cómo un sistema termodinámico se deforma en otro de manera suave. El atractivo de esta construcción reside en que existen curvas que ponen de manifiesto relaciones muy sutiles entre sistemas en apariencia por completo independientes, como sucede con el gas ideal y los agujeros negros cargados.

En mecánica hamiltoniana, a cada función sobre el espacio de fases se hace corresponder un campo vectorial cuyas curvas integrales producen una dinámica en la que dicha función es una cantidad conservada y que obedece las ecuaciones de Hamilton, siendo el hamiltoniano la función. En termodinámica geométrica encontramos la noción análoga de *campo hamiltoniano de contacto* asociado a una función por medio de la estructura de contacto, cuya interpretación física no parece ser tan clara como en el universo simpléctico. Además, la dinámica del campo no tiene por qué preservar la función. A pesar de todo, los campos hamiltonianos de contacto posibilitan la construcción de procesos termodinámicos de manera sencilla, ya que se demuestra que un campo hamiltoniano de contacto es tangente a una subvariedad de Legendre si y solo si la función a la que va asociado se hace cero en la subvariedad.

 ✉ f42rodim@uco.es

Historial

Revisión 1.a (16/08/20)

- Correcciones menores.

1º revisión (16/08/20)

- Creación del documento.
- Adición de las unidades:
  - Un alegato en defensa de los tensores en Física *(redactado entre febrero y marzo de 2020)*
  - Tensores en un espacio vectorial *(redactado entre agosto de 2018 y agosto de 2020)*
  - Una introducción sencilla a la Relatividad Especial *(redactado en verano de 2017)*
  - Termodinámica del equilibrio y geometría de contacto – Motivación *(redactado entre junio y julio de 2020)*