Capítulo 1

Álgebra tensorial

1.1. Tensores en $\mathbb{R}^n$

No es difícil tener prejuicios contra los tensores: al abrir cualquier libro con esa palabra en su título nos saltan a la vista un montón de subíndices y superíndices adornados con comas, puntos y comas y otros símbolos cuando se pasa al cálculo tensorial. Todo ello da una sensación misteriosa de taquigrafía impenetrable. Para los matemáticos, que muchas veces tienen otros prejucios contra la Física, los temores se agravan por la conocida aplicación de los tensores en la teoría de la relatividad o en una parte de la mecánica.

El propósito de esta primera sección constituida por definiciones y ejemplos es perder ese miedo. El álgebra tensorial no es más que una extensión natural del álgebra lineal. El cálculo tensorial, que se verá más adelante, sigue las mismas directrices que el cálculo en subvariedades del segundo curso: todo es lineal en entornos pequeños. Como no hay sistemas naturales de coordenadas las cosas se complican pero conceptualmente ni el álgebra ni el cálculo tensorial son especialmente difíciles[1].

En el curso de primero se estudió álgebra lineal de una variable vectorial, pero nada impide considerar dos, tres o más variables; lo cual lleva directamente a la noción de tensor sobre un espacio vectorial.

Antes de dar la definición precisa de tensor, veamos algunos ejemplos sencillos que la motivan.

Sea V un espacio vectorial de dimensión finita sobre $\mathbb{R}$. Los libros dicen que cualquiera de estos espacios es isomorfo a algún $\mathbb{R}^m$, así que para fijar ideas podemos suponer que de hecho $V = \mathbb{R}^m$. Con una tipografía difícil de mantener, todas las aplicaciones lineales

[1] Si se permite una comparación extravagante, es como si alguien dijera que es muy complicado jugar al mus. Evidentemente es muy difícil aprender a jugar sólo oyendo "sí, no, paso, órdago" pero las reglas las podría aprender cualquier tonto (y encima ganarnos). Tampoco en Matemáticas puede uno limitarse a mirar u oír.

$f : V \longrightarrow \mathbb{R}$ se expresan en coordenadas (lo cual requiere fijar una base) como

$$f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = (a_1, a_2, \dots, a_m)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = a_1x_1 + a_2x_2 + \cdots + a_mx_m.$$

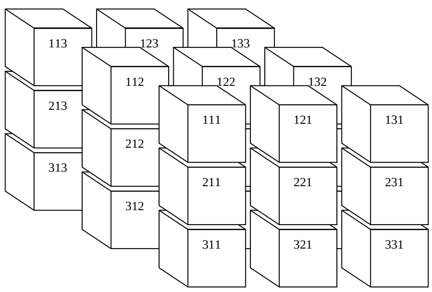
Es decir, cada una de ellas está determinada por una matriz de coeficientes $1 \times m$ o lo que es lo mismo un vector horizontal $(a_1, a_2, \dots, a_m)$. Recuérdese que al conjunto de estas aplicaciones lineales se le llama *espacio dual* y se denota con V^*. Como es sólo una cuestión estética escribir vectores en vertical o en horizontal (de hecho por razones tipográficas pocas veces se escriben en vertical), V y V^* son lo mismo; o dicho matemáticamente, isomorfos. Recuérdese que a una base de V, $\{\vec{e}_1, \dots, \vec{e}_m\}$, se le puede asignar una base de V^*, llamada la *base dual*, $\{\widetilde{\varphi}^1, \dots, \widetilde{\varphi}^m\}$, de manera que $\widetilde{\varphi}^i(\vec{e}_j) = 0$ si $i \neq j$ y $\widetilde{\varphi}^i(\vec{e}_i) = 1$.

Consideremos ahora una aplicación bilineal, esto es, lineal en dos variables:

$$\begin{array}{ll} a_1)\ f(\lambda\vec{x}, \vec{y}) = \lambda f(\vec{x}, \vec{y}) & b_1)\ f(\vec{x} + \vec{x}\,', \vec{y}) = f(\vec{x}, \vec{y}) + f(\vec{x}\,', \vec{y}) \\ a_2)\ f(\vec{x}, \lambda\vec{y}) = \lambda f(\vec{x}, \vec{y}) & b_2)\ f(\vec{x}, \vec{y} + \vec{y}\,') = f(\vec{x}, \vec{y}) + f(\vec{x}, \vec{y}\,') \end{array}$$

No es difícil comprobar que todas las funciones bilineales de $V \times V$ en $\mathbb{R}$ son de la forma

$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}\right) = (x_1, \dots, x_m)\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mm} \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$



Si ahora considerásemos una aplicación *trilineal* necesitaríamos una matriz tridimensional para colocar los vectores lo cual no es muy operativo, por ejemplo, para $m = 3$ tendríamos que desguazar un cubo de Rubik y poner un número en cada trozo; y en el caso $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ utilizar un cubo de Rubik cuatridimensional (parece que los hay virtuales en la red). Lo bueno de la abstracción matemática es que uno puede definir objetos sin necesidad de dibujarlos ni de que existan, y nadie protesta (demasiado). Así que consideremos las aplicaciones lineales "a lo grande".

Definición: Se dice que $f : V_1 \times V_2 \times \cdots \times V_n \longrightarrow W$, donde $V_1, V_2, \dots, V_n, W$ son espacios vectoriales, es una *aplicación multilineal* si para todo $1 \le i \le n$

a) $\quad f(\vec{v}_1, \dots, \lambda\vec{v}_i, \dots, \vec{v}_n) = \lambda f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_n) \quad$ con $\lambda \in \mathbb{R}$

b) $\quad f(\vec{v}_1, \dots, \vec{v}_i + \vec{v}_i\,', \dots, \vec{v}_n) = f(\vec{v}_1, \dots, \vec{v}_i, \dots, \vec{v}_n) + f(\vec{v}_1, \dots, \vec{v}_i\,', \dots, \vec{v}_n).$

Es habitual que las variables de una aplicación multilineal tengan todas la misma naturaleza y por tanto $V_1 = V_2 = \cdots = V_n$. Daremos un nombre a esta situación en el caso simple en que $W = \mathbb{R}$.

Definición: Se llama *tensor n veces covariante* a cualquier aplicación multilineal de la forma $T : V \times \overset{n\ \text{veces}}{\cdots} \times V \longrightarrow \mathbb{R}$.

Ejemplo: El producto escalar usual en $\mathbb{R}^m$ define un tensor dos veces covariante.

Ejemplo: El determinante aplicado a m vectores de $\mathbb{R}^m$ define un tensor m veces covariante.

Ejemplo: La función que asigna a n vectores de $\mathbb{R}^m$ el producto de sus primeras coordenadas (en la base canónica) es un tensor n veces covariante.

Al igual que en cálculo de varias variables se consideran funciones vectoriales, también podríamos definir algo así como tensores vectoriales, de la forma $f : V \times \cdots \times V \longrightarrow V$ o incluso complicar más las cosas permitiendo $f : V \times \cdots \times V \longrightarrow V \times V$, etc. Cada vector "vertical" puede pasarse a un número real (pre-)multiplicando por un vector "horizontal", así que a cada $f : V \times \cdots \times V \longrightarrow V$ se le puede asociar $T : V^* \times V \times \cdots \times V \longrightarrow \mathbb{R}$ dada por $T(\widetilde{\varphi}, \vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}) = \widetilde{\varphi}\big(f(\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n})\big)$ para cada $\widetilde{\varphi} \in V^*$. Además por el isomorfismo entre V y V^*, esta correspondencia es biyectiva[2].

En definitiva, da igual considerar los hipotéticos tensores vectoriales, sin gancho entre los matemáticos, que considerar los tensores antes definidos pero permitiendo sustituir algunos de los factores V por V^*. Lo más breve es generalizar de esta forma la definición anterior.

Definición: Se llama *tensor r veces contravariante y s veces covariante* o *tensor de tipo* (r, s) a una aplicación multilineal $T : V^* \times \overset{r\ \text{veces}}{\cdots} \times V^* \times V \times \overset{s\ \text{veces}}{\cdots} \times V \longrightarrow \mathbb{R}$.

Comparando con la definición previa, un tensor n veces covariante es un tensor de tipo $(0, n)$. Por otro lado los tensores de tipo $(n, 0)$ se dice que son n veces contravariantes. Por convenio además diremos que *una constante es un tensor de tipo* $(0, 0)$. Obsérvese que hay cierta lógica en esta notación porque una constante no depende de ningún vector.

Como ejemplo, nótese que un endomorfismo $f : V \longrightarrow V$ asigna a cada vector otro vector, y según la identificación anterior da lugar a un tensor de tipo $(1, 1)$. En coordenadas, si representamos el endomorfismo como $f(\vec{v}) = A\vec{v}$ para cierta matriz cuadrada A y un elemento $\widetilde{\varphi} \in V^*$ como un vector horizontal, el tensor correspondiente es $T(\widetilde{\varphi}, \vec{v}) = \widetilde{\varphi}(A\vec{v})$.

Al igual que hablamos de las componentes (o entradas o coeficientes) de una matriz en cierta base, nos podemos referir a las componentes de un tensor (excluiremos implícitamente el caso $r = s = 0$).

[2] Todo este párrafo se resume en lo siguiente: si tienes un vector y quieres un número, haz el producto escalar con otro vector arbitrario y si además quieres quedar bien, di que ésa es la acción de V^* sobre V.

Definición: Supongamos que $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \ldots, \vec{e_m}\}$ es una base de V y la base dual es $\mathcal{B}^* = \{\widetilde{\varphi}^1, \widetilde{\varphi}^2, \ldots, \widetilde{\varphi}^m\} \subset V^*$. Se llaman *componentes de un tensor*, T, de tipo (r,s), en estas bases a los números reales

$$T^{i_1 i_2 \ldots i_r}_{j_1 j_2 \ldots j_s} = T(\widetilde{\varphi}^{i_1}, \widetilde{\varphi}^{i_2}, \ldots, \widetilde{\varphi}^{i_r}, \vec{e_{j_1}}, \vec{e_{j_2}}, \ldots, \vec{e_{j_s}}).$$

A partir de ahora pondremos especial atención en enumerar los elementos de V (los vectores) con subíndices y los de V^* (a veces llamados *contravectores*) con superíndices para que sea más claro de dónde viene cada componente de un tensor

Ejemplo: Calcular las componentes del tensor D definido por el determinante en $\mathbb{R}^2$ con la base usual.

Claramente $D(\vec{e}_1, \vec{e}_1) = D(\vec{e}_2, \vec{e}_2) = 0$ y $D(\vec{e}_1, \vec{e}_2) = -D(\vec{e}_2, \vec{e}_1) = 1$, por lo que sus componentes son $D_{11} = D_{22} = 0$, $D_{12} = -D_{21} = 1$. Esto está estrechamente relacionado con la igualdad (inútil)

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix}.$$

Nótese que una igualdad similar para el determinante en $\mathbb{R}^m$ requeriría algo así como "matrices $m-$dimensionales" cuyos elementos serían las componentes del tensor.

Ejemplo: Escribir un tensor $(2,1)$, $S : (\mathbb{R}^2)^* \times (\mathbb{R}^2)^* \times \mathbb{R}^2 \longrightarrow \mathbb{R}$, tal que, empleando la base canónica, tenga $S^{12}_2 = 1$ como única componente no nula.

Basta tomar el tensor definido por

$$S\left(\begin{pmatrix} a & b \end{pmatrix}, \begin{pmatrix} c & d \end{pmatrix}, \begin{pmatrix} e \\ f \end{pmatrix} \right) = adf.$$

Está claro que una vez fijada una base un tensor está determinado por sus componentes. Por ejemplo, el tensor T de tipo $(1,1)$ correspondiente a un endomorfismo tiene como componente T^i_j el elemento ij de la matriz que lo define en cierta base. Para el endomorfismo identidad las componentes se suelen denotar con el símbolo δ^i_j que significa

$$\delta^i_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j. \end{cases}$$

Un vector $\vec{v}$ también puede considerarse como un tensor de tipo $(1,0)$ que aplica cada $\widetilde{\varphi} \in V^*$ en $\widetilde{\varphi}(\vec{v})$ y sus componentes en una base son simplemente sus coordenadas. De la misma forma un elemento de V^* se puede considerar un tensor de tipo $(0,1)$ cuyas componentes son sus coordenadas en la base dual. Consecuentemente el concepto de tensor engloba a los principales personajes del álgebra lineal del primer curso.

El conjunto de todos los tensores de tipo (r,s) tiene estructura de espacio vectorial, porque podemos multiplicar por números, sumar y restar tensores del mismo tipo.

También se puede definir una especie de multiplicación exterior de dos tensores no necesariamente del mismo tipo, que se reduce a sustituir parte de las variables en uno y la otra parte en el otro, multiplicando los resultados.

Definición: Si T es un tensor de tipo (r, s) y S es un tensor de tipo (u, v), se llama *producto tensorial* de T y S al tensor $T \otimes S$ de tipo $(r + u, s + v)$ cuyo valor en $\Omega = (\widetilde{\varphi}^1, \dots, \widetilde{\varphi}^{r+u}, \vec{v}_1, \dots, \vec{v}_{s+v})$ es

$$(T \otimes S)(\Omega) = T(\widetilde{\varphi}^1, \dots, \widetilde{\varphi}^r, \vec{v}_1, \dots, \vec{v}_s) \cdot S(\widetilde{\varphi}^{r+1}, \dots, \widetilde{\varphi}^{r+u}, \vec{v}_{s+1}, \dots, \vec{v}_{s+v}).$$

<u>Ejemplo</u>: Si $\widetilde{\varphi}$ es el tensor $(0, 1)$ que asigna a cada vector su primera coordenada, $\widetilde{\varphi} \otimes \widetilde{\varphi}$ asigna a cada par de vectores el producto de sus primeras coordenadas.

<u>Ejemplo</u>: Sea T el tensor $(1, 1)$ que corresponde al endomorfismo identidad en $\mathbb{R}^2$, y sea S el que corresponde a intercambiar las dos coordenadas (respecto de la base canónica). Entonces las componentes no nulas de T son $T^1_1 = T^2_2 = 1$, y las de S, $S^2_1 = S^1_2 = 1$. Consecuentemente, las componentes no nulas de $P = T \otimes S$ son $P^{12}_{11} = P^{11}_{12} = P^{22}_{21} = P^{21}_{22} = 1$.

La notación tensorial es en principio un poco aparatosa. Por ejemplo, un tensor $(1, 3)$ muy importante es el llamado tensor de Riemann $R : V^* \times V \times V \times V \longrightarrow \mathbb{R}$, que introduciremos en otro capítulo. En relatividad $\dim V = 4$ y R tiene $4 \cdot 4 \cdot 4 \cdot 4 = 256$ componentes y para aplicarlo a un elemento del dual, digamos con componentes (a_1, a_2, a_3, a_4), y a tres vectores, cuyas coordenadas numeramos con superíndices, (b^1, b^2, b^3, b^4), (c^1, c^2, c^3, c^4), (d^1, d^2, d^3, d^4), debemos escribir

$$\sum_{i=1}^{4} \sum_{j=1}^{4} \sum_{k=1}^{4} \sum_{l=1}^{4} R^i_{jkl} a_i b^j c^k d^l$$

que, ciertamente, contiene muchos sumatorios. Si utilizamos la notación de subíndices y superíndices (correspondientes a vectores y contravectores) introducida aquí, se produce una simplificación sustancial usando el llamado *convenio de sumación de Einstein*[3] que consiste en *sobreenteder un sumatorio cada vez que un subíndice aparece también como superíndice*. Por ejemplo, la expresión anterior se escribe simplemente como

$$R^i_{jkl} a_i b^j b^k b^l.$$

Las relaciones matriciales desde el punto de vista de las coordenadas, se reducen enormemente con este convenio y se vuelven más intuitivas. Así el efecto sobre las coordenadas de una aplicación lineal, digamos $\vec{y} = A\vec{x}$, se escribe

$$y^i = a^i_j x^j.$$

[3]Este convenio fue realmente introducido por Einstein quien bromeó al respecto diciendo: “He hecho un gran descubrimiento en Matemáticas; he suprimido el signo de sumación toda vez que la suma se haga en un índice que aparece dos veces”.

Y la igualdad matricial $D = ABC$ componente a componente, se reduce a

$$d^i_j = a^i_k b^k_l c^l_j.$$

Nótese lo sencillo que es de recordar apelando a una "simplificación" de índices (y recuérdese lo fácil que era equivocarse al programarlo sin este truco en la asignatura correspondiente de cálculo numérico). Lo mismo se aplica para abreviar combinaciones lineales. Por ejemplo, para decir que las coordenadas de $\vec{v}$ en la base $\mathcal{B} = \{\vec{e_1}, \vec{e_2}, \dots, \vec{e_m}\}$ son $a^1, a^2, \dots, a^m$

$$\vec{v} = \sum_{j=1}^{m} a^j \vec{e_j} \qquad \text{se abrevia como} \qquad \vec{v} = a^j \vec{e_j}.$$

En definitiva:

$$\boxed{\text{Un índice duplicado arriba y abajo indica un sumatorio.}}$$

Es importante insistir en que todo funciona como si los índices repetidos se simplificasen. Por ejemplo, R^i_{jkl} es un tensor $(1,3)$ pero como R^i_{jil} sólo depende de dos índices, j y l, es $(0,2)$. También $a^i_k b^l_j$ representa un tensor $(2,2)$ y $a^i_k b^k_j$ representa un tensor $(1,1)$. Este fenómeno de igualar un índice y un subíndice y sumar en ellos, se llama *contracción*. Ahora podemos apreciar la conveniencia de pensar en las constantes como tensores de tipo $(0,0)$. Un tensor de este tipo corresponde por ejemplo a la contracción del producto tensorial de un tensor $(0,1)$ por otro $(1,0)$; lo cual puede entenderse como $\widetilde{\varphi}(\vec{v})$ con $\widetilde{\varphi} \in V^*$, $\vec{v} \in V$, y el resultado de esta operación es constante. La contracción de un tensor está bien definida: no depende de la base en la que se lleva a cabo porque, como veremos con detalle en la tercera sección, las reglas de transformación asociadas a subíndices y superíndices (vectores y contravectores) son inversas.

Ya hemos mencionado que el producto escalar usual de $\mathbb{R}^n$ es un tensor dos veces covariante. También sabemos del curso de primero que servía para medir ya que $d(P,Q) = (\overrightarrow{PQ} \cdot \overrightarrow{PQ})^{1/2}$ y esto lo hacía tan importante que incluso recibían un nombre especial, espacios euclídeos, los espacios vectoriales que tenían algún tipo de producto escalar con tres propiedades básicas: ser bilineal, simétrico y definido positivo. Es lógico por tanto dar un nombre especial a los tensores $(0,2)$ que permiten medir asociados a estos objetos. Aquí surge un problema y es que los físicos prefieren[4] cambiar la condición de ser definido positivo por otra más débil de no degeneración, porque en relatividad

[4]Esto no es más que una licencia estilística para indicar los requerimientos de la teoría de la relatividad, quizá no sea una preferencia real de los físicos. Irónicamente el producto escalar no positivo y el concepto asociado de espacio-tiempo fueron introducidos en la relatividad por H. Minkowski, un matemático que fue profesor de Einstein. El propio Einstein criticó al principio este formalismo por considerarlo superfluo con frases como "Desde que los matemáticos han invadido la teoría de la relatividad ni yo mismo la entiendo" pero al desarrollar la relatividad general se adhirió totalmente a él.

aparece naturalmente y $\overrightarrow{PQ}\cdot\overrightarrow{PQ} < 0$ significa que no se puede alcanzar el punto (evento) Q partiendo de P.

Nuestra aproximación a la relatividad nos lleva a unirnos a los físicos. Un matemático purista podría todavía objetar que la definición habitual se hace en una variedad diferenciable (véase más adelante) pero se puede contraatacar diciendo que $V \times V$ lo es.

Definición: Se dice que G es un *tensor métrico* si es un tensor dos veces covariante y sus componentes g_{ij} conforman una matriz simétrica no singular.

Ejemplo. El producto escalar de toda la vida es un tensor métrico cuyas componentes con la base usual vienen dadas por la delta de Kronecker: $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ en otro caso. En $\mathbb{R}^4$ quizá más famoso incluso que el producto escalar usual está la métrica de Minkowski cuyas únicas componentes no nulas son $g_{11} = 1$, $g_{22} = g_{33} = g_{44} = -1$. Dependiendo de los autores, a veces se consideran estas componentes cambiadas de signo.

Ejercicios de la sección 1

1) Responder brevemente a las siguientes preguntas:

i) Si $T = T(\vec{x}, \vec{y})$ y $S = S(\vec{x}, \vec{y})$ son tensores, ¿lo es $R(\vec{x}, \vec{y}) = T(\vec{x}, \vec{y}) \cdot S(\vec{x}, \vec{y})$?

ii) ¿Es $T(\vec{x}, \vec{y}) = \vec{x} + \vec{y}$ una aplicación bilineal?

iii) ¿Es el producto tensorial conmutativo?

iv) ¿Es un tensor la aplicación que a cada par de vectores en $\mathbb{R}^3$ con la base canónica le asigna la primera coordenada de su producto vectorial?

v) ¿Es un tensor la aplicación que a cada par de vectores en $\mathbb{R}^2$ con la base canónica le asigna el área del paralelogramo que determinan?

vi) ¿Cuántas componentes tiene un tensor de tipo (r, s) con $V = \mathbb{R}^m$?

vii) ¿Por qué si las componentes de dos tensores coinciden en una base deben coincidir en todas?

2) Demostrar que, fijada una base, todo tensor dos veces covariante es de la forma $T(\vec{x}, \vec{y}) = \vec{x}^t A\vec{y}$ con A una matriz.

3) Hallar cuántas componentes nulas y cuántas no nulas tiene el tensor determinante. Estudiar también cuántas son positivas.

4) Si multiplicamos tensorialmente unos cuantos elementos de $\mathcal{B}$ y otros de $\mathcal{B}^*$, hallar cuántas componentes no nulas tiene el tensor resultante. Usar este hecho para probar que todo tensor se puede escribir como combinación lineal de estos productos tensoriales.

5) Para $V = \mathbb{R}^3$ consideremos un tensor de tipo $(0, 3)$, otro de tipo $(1, 2)$ y otro de tipo $(2, 1)$, cuyas componentes, digamos ϵ_{ijk}, ϵ^i_{jk} y ϵ^{ij}_k, en la base canónica son: 0 si i, j, k no es una reordenación de $1, 2, 3$; 1 si i, j, k es una permutación par de $1, 2, 3$ (esto es, se ordena con un número par de intercambios) y -1 si i, j, k es una permutación impar de $1, 2, 3$. Dados $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ y $\vec{F} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, $\vec{F} = (F^1, F^2, F^3)$, explicar qué objetos matemáticos bien conocidos representan las cantidades $\epsilon^{ij}_k \partial F^k/\partial x^j$, $\epsilon^i_{jk} v^j w^k$ y $\epsilon_{ijk} u^i v^j w^k$.

6) Sea un endomorfismo $\vec{x} \mapsto A\vec{x}$, con $A = (a^i_j)$, en $\mathbb{R}^n$.

a) Dar una demostración tensorial de que la traza a^i_i es invariante por cambios de base.

b) Probar que $a^i_i a^j_j - a^i_j a^j_i$ también es invariante e identificar esta cantidad en términos de trazas de matrices.

7) Sea $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ una base de $\mathbb{R}^m$ y sean g_{ij} las componentes del tensor métrico usual, es decir, $g_{ij} = \vec{v}_i \cdot \vec{v}_j$. Demostrar que

$$|\det(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m)| = \sqrt{\det(g_{ij})}.$$

Indicación: Cambiar a una base ortonormal y escribir g_{ij} en términos de la matriz de cambio de base.

1.2. Repaso de Geometría III

El lector después de haber pasado por [Di] o [GoJ] debería haber sacado la conclusión de que el gran cambio del curso de Geometría III con respecto al de Geometría II, extendido en parte en el Cálculo III, es que ahora se consideran los objetos de la geometría diferencial, las variedades, sin referencia a ningún espacio exterior[5]. Ya no falta que exista $\mathbb{R}^3$ para que podamos hablar de la superficie esférica.

La idea de variedad diferenciable n-dimensional es la de un objeto geométrico compuesto por parches que son similares a abiertos de $\mathbb{R}^n$. Partimos de un espacio topológico M al que exigimos que tenga la propiedad de Hausdorff y una base numerable (segundo axioma de numerabilidad [Mu]). La primera propiedad es natural si queremos poder tratar separadamente los puntos, y las segunda va también en este sentido, porque permite asegurar la existencia de particiones de la unidad [Wa], que son totalmente necesarias para hacer el análisis local típico de la geometría diferencial [Ja], [Di], [Sp2].

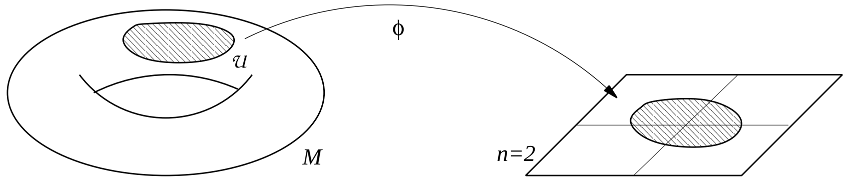
Una carta nos dice la manera de allanar un parche de M en $\mathbb{R}^n$.

Definición: Una *carta* n-dimensional de M es un par $(\mathcal{U}, \phi)$ donde $\mathcal{U}$ es un abierto de M y ϕ es una función $\phi : \mathcal{U} \longrightarrow \mathbb{R}^n$ que es homeomorfismo sobre su imagen.

[5]Quizá al lector le parezca baldío y rebuscado este empeño al saber que existe un teorema bello, simple e interesante (que en una forma más fuerte se debe a H. Whitney) que prueba que las variedades siempre se pueden meter dentro de algún $\mathbb{R}^n$ y por tanto pueden considerarse como subvariedades y ser tratadas con los métodos de Cálculo III (véase [Sp1] Cap. 2, Th. 17). De hecho un profundo teorema debido a J. Nash (uno de los pocos matemáticos que tienen película) afirma que esta inclusión se puede hacer sin cambiar las distancias (comprender este último enunciado requiere esperar un poco hasta que definamos el tensor métrico en variedades).

Que todo esto sea posible no quiere decir que sea adecuado. Incluso desde el punto de vista intuitivo, es mucho más fácil imaginar el espacio proyectivo $\mathbb{P}^2(\mathbb{R})$ (véase [Di] Ej. 2.2.11) como un círculo con los puntos antipodales de la frontera identificados que como una subvariedad de $\mathbb{R}^4$, que es el primer $\mathbb{R}^n$ en el que se puede meter bien.

Notación: A veces, con el abuso obvio, se llama carta a la función ϕ. Con frecuencia se denota con x^i la coordenada i-ésima de ϕ, es decir $\phi = (x^1, x^2, \dots, x^n)$, éstas son las llamadas *funciones coordenadas*. El uso de superíndices esta relacionado con el convenio introducido en la primera sección. Para indicar que un abierto $\mathcal{U}$, típicamente de una carta, contiene al punto p escribiremos $\mathcal{U}(p)$.



Un punto puede estar tapado por varios parches, diferentes abiertos de cartas, debemos asegurarnos de que el análisis no se estropea bajando por una ϕ o por otra.

Definición: Se dice que dos cartas n-dimensionales de M, $(\mathcal{U}, \phi)$ y $(\mathcal{V}, \psi)$, son *compatibles* si $\psi \circ \phi^{-1} : \phi(\mathcal{U} \cap \mathcal{V}) \longrightarrow \psi(\mathcal{U} \cap \mathcal{V})$ es C^∞ con inversa C^∞. Se incluye como caso especial en el que $\mathcal{U}$ y $\mathcal{V}$ son disjuntos.

Para no entretenernos saltaremos al último peldaño. Por razones técnicas es conveniente pensar en todas las posibles cartas n-dimensionales compatibles entre sí y se dice que la colección correspondiente $\{(\mathcal{U}_\alpha, \phi_\alpha)\}_{\alpha \in I}$, $M = \bigcup \mathcal{U}_\alpha$, es una *estructura diferenciable* n-dimensional. Con esto llegamos al objetivo: la definición de variedad[6].

Definición: Se dice que un espacio topológico M con las propiedades anteriores, es una *variedad diferenciable* n-dimensional si está dotado de una estructura diferenciable de dimensión n.

La filosofía subyacente es que como una variedad es un espacio topológico abstracto todas las operaciones de análisis que queramos hacer se llevarán a cabo bajando a $\mathbb{R}^n$ por una carta.

Por ejemplo, se dice que una función $f : M \longrightarrow \mathbb{R}$ es C^∞ si para cada carta $(\mathcal{U}, \phi)$ la función $f \circ \phi^{-1} : \phi(\mathcal{U}) \longrightarrow \mathbb{R}$ lo es y se define para cada $p \in \mathcal{U}$ la *derivada parcial* i-ésima en la variedad como

$$\frac{\partial f}{\partial x^i}\Big|_p = D_i(f \circ \phi^{-1})(\phi(p))$$

[6]Sólo para lectores interesados: Una vez que uno ha puesto la definición en términos suficientemente raros hay una pregunta extraña pero natural. ¿Es posible tener en una esfera o en otro espacio topológico de toda la vida con la topología usual diferentes estructuras diferenciables? En la esfera usual S^2 se sabe que sólo hay una estructura diferenciable pero J. Milnor probó que en la de 7 dimensiones, S^7, la situación es muy distinta, de hecho hay 28 posibilidades. Sólo una nos resulta familiar y por ello se dice que el resto son *esferas exóticas*.

Esto muestra que las variedades C^0 (con cambios de carta continuos) son bien diferentes de las variedades C^∞ aquí definidas. Por otro lado hay un teorema que afirma que las variedades C^1 son en realidad como las C^∞ eliminando "cartas malas" (véase [Hi] §2.10 para el enunciado preciso).

donde el símbolo D_i significa la derivada parcial usual con respecto a la i-esima variable. En general, si M y N son variedades, se puede hablar de funciones C^∞, $f : M \longrightarrow N$ si para cada par de cartas $(\mathcal{U}, \phi)$, $(\mathcal{V}, \psi)$, respectivamente de M y de N se cumple que $\psi \circ f \circ \phi^{-1} : \phi(\mathcal{U}) \longrightarrow \psi(\mathcal{V})$. Las funciones C^∞ entre variedades que tienen inversa C^∞ se llaman *difeomorfismos*.

La propia notación usual para las funciones coordenadas, dándoles nombre de punto, $(x^1, x^2, \ldots, x^n)$, nos recuerda que la versión operativa de los puntos de una variedad es su reflejo en $\mathbb{R}^n$ después de aplicar la función de una carta.

Nota: Éste es un curso de geometría diferencial, no de análisis, por ello daremos por supuesto que la regularidad no constituye ninguna obstrucción en las definiciones. A partir de ahora supondremos, sin indicarlo cada vez, que todas las funciones entre variedades que consideramos son C^∞.

Un problema técnicamente más complejo es la definición del espacio tangente, que en el caso de subvariedades de $\mathbb{R}^n$ es muy fácil (recuérdese el curso de Cálculo III). No es una mera adaptación porque allí los vectores tangentes eras "pelos" orientados que se salían de la subvariedad, mientras que concebimos las variedades como una entidad única, sin referencia a un posible "exterior". Hay varias maneras de superar este obstáculo (véase [Ja]). Aquí mencionaremos las definiciones matemáticas que corresponden a ver los vectores tangentes como velocidades de curvas y como derivadas direccionales. La segunda es más abstracta, introduciendo implícitamente el concepto de *derivación* [ON], pero en Geometría III se mostraba más útil en las demostraciones.

Definición: Se llama *espacio tangente* de M en un punto p al conjunto cociente $T_p(M) = \mathcal{K}_p(M)/\sim$ donde $\mathcal{K}_p(M) = \{\text{Funciones } c : (-\epsilon, \epsilon) \longrightarrow M \text{ con } c(0) = p\}$ y $\sim$ identifica las funciones (curvas) tales que $(\phi \circ c_1)'(0) = (\phi \circ c_2)'(0)$ con $(\mathcal{U}(p), \phi)$ una carta. Se llama *vector tangente* de M en p a cualquiera de sus elementos.

Definición: Se llama *vector tangente* de M en p a cualquier operador $\mathbb{R}$-lineal $v : \mathcal{E}_p(M) \longrightarrow \mathbb{R}$ que satisface $v(fg) = v(f)g(p) + f(p)v(g)$ para todo $f, g \in \mathcal{E}_p(M)$ donde $\mathcal{E}_p(M)$ es el anillo de funciones $M \longrightarrow \mathbb{R}$ definidas en un entorno suficientemente pequeño de p. Se llama *espacio tangente* de M en un punto p al conjunto formado por los vectores tangentes.

El nexo entre ambas definiciones es que a cada $c \in \mathcal{K}_p(M)$ se le puede asignar el operador $v : f \mapsto (f \circ c)'(0)$ (véase [GoJ] §3.2).

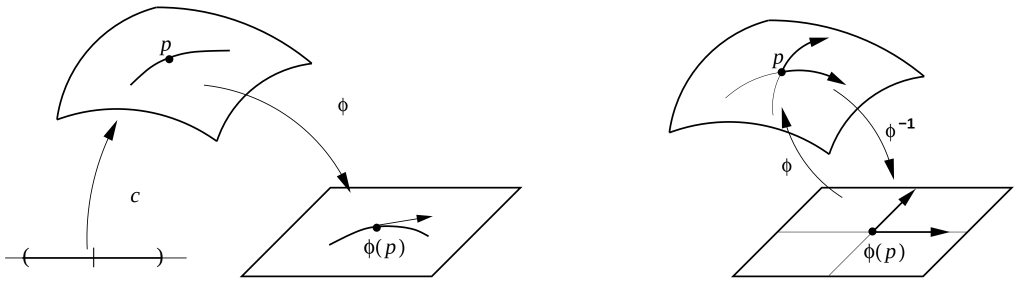
A partir de las curvas que corresponden a los ejes coordenados (una vez que bajamos a $\mathbb{R}^n$) se obtienen unos vectores tangentes que denotaremos con el extraño nombre $\frac{\partial}{\partial x^i}\big|_p$. Para ser rigurosos, si $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es la base canónica, fijada una carta $(\mathcal{U}(p), \phi = (x^1, \ldots, x^n))$, con la primera definición se tiene

$$\frac{\partial}{\partial x^i}\Big|_p = [c_i] \qquad \text{con} \qquad c_i(t) = \phi^{-1}(\phi(p) + t\vec{e}_i), \quad i = 1, 2, \ldots, n.$$

Denominar a estos vectores con el mismo símbolo que el de las derivadas parciales no es casual pues con la segunda definición no son más que la derivadas parcial i-ésimas en la variedad, es decir

$$\frac{\partial}{\partial x^i}\Big|_p \ : \ f \longrightarrow \frac{\partial f}{\partial x^k}\Big|_p. \tag{1.1}$$

Por razones obvias se suelen denotar estos vectores tangentes con la notación abreviada $\partial_i\big|_p$ o incluso ∂_i si no se quiere indicar la carta o el punto.



Como se vio en cursos pasados:

Proposición 1.2.1 *El espacio tangente $T_p(M)$ tiene una estructura natural de espacio vectorial cuya dimensión es la de la variedad diferenciable M.*

Proposición 1.2.2 *Para cada punto p de una variedad diferenciable n-dimensional, M, el conjunto $\{\partial_1\big|_p, \partial_2\big|_p, \dots \partial_n\big|_p\}$ es una base de $T_p(M)$.*

Una vez que tenemos estos resultados y hemos acumulado la miseria debajo de la alfombra de la notación, nos podemos despreocupar de la dificultad y abstracción de los conceptos definidos a la hora de hacer operaciones. Podemos sumar y multiplicar por números coordenada a coordenada como nos enseñaron en primero y uno puede escribir sin remordimientos cosas como: $(2\partial_1|_p + 3\partial_2|_p) + 4(\partial_1|_p - 2\partial_2|_p) = 6\partial_1|_p - 5\partial_2|_p$.

Con $f : M \longrightarrow N$ podemos pasar curvas en curvas lo cual induce una aplicación $T_p(M) \longrightarrow T_{f(p)}(N)$. Aunque ésta es la idea intuitiva es más sintético proceder tomando en cuenta la segunda definición de espacio tangente.

Definición: Sea $f : M \longrightarrow N$. Se llama *aplicación tangente* de f en p y se denota con $df|_p$, a la aplicación lineal $T_p(M) \longrightarrow T_{f(p)}(N)$ que aplica un elemento de $T_p(M)$ (considerado con la segunda definición), digamos $v(\cdot)$ en $v(\cdot \circ f)$.

Ahora todo funciona como con la diferencial de toda la vida, siempre componiendo con las cartas.

Proposición 1.2.3 *Sea* $f : M \longrightarrow N$ *y sean* $(\mathcal{U}(p), \phi)$ *y* $(\mathcal{V}(f(p)), \psi)$ *cartas de* M *y* N *respectivamente en los puntos indicados. La matriz de la aplicación tangente* $df|_p$ *en las bases* $\{\partial/\partial x^1\big|_p, \dots, \partial/\partial x^m\big|_p\}$ *y* $\{\partial/\partial y^1\big|_{f(p)}, \dots, \partial/\partial y^n\big|_{f(p)}\}$ *correspondientes a estas cartas es la matriz jacobiana de* $\psi \circ f \circ \phi^{-1}$ *en* $\phi(p)$.

En otras palabras, quieran lo que quieran decir los símbolos $\partial/\partial x^i$, ya sean derivadas de clases de curvas o derivaciones que actúan sobre funciones, el caso es que formalmente se transforman por medio de una matriz jacobiana, es decir, como en los otros cursos cuando el mismo símbolo tenía otro significado.

Dada una carta $(\mathcal{U}(p), \phi = (x^1, \dots, x^n))$ de M tiene sentido considerar $dx^i|_p$, las aplicaciones tangentes de las funciones coordenadas como funciones de M en $\mathbb{R}$ con la estructura de variedad obvia. Usando las definiciones de vector tangente y aplicación tangente se puede probar que

$$dx^i|_p\left(\frac{\partial}{\partial x^j}\bigg|_p\right) = \delta^i_j.$$

Dicho de otra forma,

$$\{dx^1|_p, dx^2|_p, \dots, dx^n|_p\} \quad \text{es la base dual de} \quad \left\{\frac{\partial}{\partial x^1}\Big|_p, \frac{\partial}{\partial x^2}\Big|_p, \dots \frac{\partial}{\partial x^n}\Big|_p\right\}.$$

Si a uno le gusta poner nombres, ahí van un par de ellos:

Definición: Dada una carta $(\mathcal{U}(p), \phi = (x^1, \dots, x^n))$ de M, al espacio vectorial sobre $\mathbb{R}$ generado por $\{dx^1|_p, dx^2|_p, \dots, dx^n|_p\}$ se le denomina *espacio cotangente* de M en p y se denota con $T^*_p(M)$, por ser el dual de $T_p(M)$. Los elementos de $T^*_p(M)$ se llaman *uno formas* (o *covectores*).

Como cabía esperar, en lo sucesivo descargaremos la notación para las aplicaciones tangentes y las bases introducidas de $T_p(M)$ y $T^*_p(M)$ omitiendo el punto cuando no sea relevante. Por ejemplo, escribiremos por ejemplo dx^1 en lugar de $dx^1|_p$.

Una vez más insistimos en que todos los espacios vectoriales sobre $\mathbb{R}$ son lo mismo, y una vez fijadas las bases las operaciones se realizan coordenada a coordenada como nos enseñaron en primero cuando casi todo era con vectores de $\mathbb{R}^n$. Los elementos del dual no albergan nada nuevo y siguen funcionando como se indicó en la sección anterior (y en el curso de primero) por mucho que pongamos d y ∂ por todos los lados. En un ejemplo:

$$(2dx^1 + 3dx^2)\big(2\frac{\partial}{\partial x^1} - \frac{\partial}{\partial x^2}\big) = 1 \qquad \text{porque} \qquad (2 \quad 3)\begin{pmatrix} 2 \\ -1 \end{pmatrix} = 1.$$

Ejercicios de la sección 2

1) Comprobar que para definir la circunferencia unidad bastan dos cartas. Utilícese un argumento topológico para probar que una no es suficiente.

2) En la superficie esférica unidad en $\mathbb{R}^3$, $S^2 = \{(x, y, z) \in \mathbb{R}^3 \ : \ x^2 + y^2 + z^2 = 1\}$ considérense las cartas $(S^2 - \{\text{N}\}, \phi_{\text{N}})$ y $(S^2 - \{\text{S}\}, \phi_{\text{S}})$ que dan las proyecciones estereográficas en $z = 0$ desde los polos norte N y sur S respectivamente.
a) Hallar una fórmula para ϕ_{N} y ϕ_{S}.
b) Demostrar que son cartas compatibles.

3) Estudiar si con la estructura de variedad correspondiente a las cartas del problema anterior las funciones $f_1, f_2 : S^2 \longrightarrow \mathbb{R}$ dadas por $f_1(x, y, z) = |z - 1|$ y $f_2(x, y, z) = |x|$ son C^∞.

4) Sea $f : S^1 \longrightarrow S^1$ dada por un giro de ángulo α. Describir el efecto de la aplicación tangente sobre ∂_1 en los siguientes casos:
a) En ambas circunferencias se emplea la carta $(S^2 - \{(-1, 0)\}, \phi_1)$ donde ϕ_1 asigna a cada punto el ángulo que determina con OX, normalizado en $(-\pi, \pi)$.
b) En la primera se emplea $(S^2 - \{(-1, 0)\}, \phi_1)$ y en la segunda $(S^2 \cap \{x > 0\}, \phi_2)$ con $\phi_2(x, y) = y$.

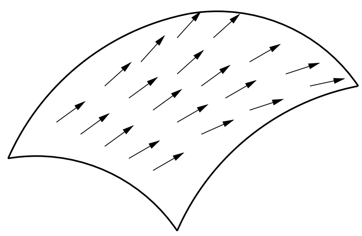
5) Comprobar usando las definiciones dadas en la sección que realmente

$$dx^i|_p \left(\frac{\partial}{\partial x^j}\bigg|_p \right) = \delta^i_j.$$

¿Cómo se deduce de aquí que los $dx^i|_p$ son linealmente independientes? ¿Y que todo elemento de $T^*_p(M)$ es una combinación lineal de los $dx^i|_p$?

1.3. Tensores en el espacio tangente

Nuestra intención es llenar una variedad de tensores, uno en cada plano tangente, conservando cierta suavidad entre ellos, lo que requiere cierta noción de proximidad.



La manera más sintética de concretar este punto pasa por dar una estructura de variedad al conjunto

$$TM = \bigcup_p T_p(M).$$

El objeto resultante es el llamado *fibrado tangente* (véase [Di] §4.1). Una vez que tenemos esta estructura podemos hablar de planos tangentes cercanos y de tensores cercanos. En vez de seguir este camino, sin duda más directo e invariante y que nos introduce a la teoría de fibrados, elegiremos una definición que involucra cartas y componentes. Para ir poco a poco, llenaremos primero de "pelos" tangentes a la variedad.

Definición: Sea M una variedad n-dimensional. Un *campo de vectores* C^∞ en M es una aplicación que asigna a cada punto $p \in M$ un vector de $T_p(M)$, de manera que en cada carta se escribe como $\sum a^i(p)\, \partial_i\big|_p$ con a^i funciones C^∞.

Se podría definir de la misma forma campos de uno formas, de tensores métricos, etc. Veamos el caso general.

Definición: Sea M una variedad n-dimensional. Un *campo tensorial* C^∞ de tipo (r,s) en M, o simplemente un *tensor* de tipo (r,s) en M, es una aplicación que asigna a cada punto $p \in M$ un tensor de tipo (r,s) con $V = T_p(M)$, $V^* = T_p^*(M)$ y que en cada carta tiene componentes C^∞.

Hay otro caso particular de gran interés en este curso.

Definición: Un campo tensorial C^∞ de tensores métricos en una variedad M se dice que es una *métrica*.

Siguiendo el convenio que veníamos manejando en el caso $r = s = 0$, un tensor de tipo $(0,0)$ en M le asigna a cada punto una constante, es decir, es simplemente una función C^∞.

Las componentes de un tensor T de tipo (r,s) en una variedad definen en cada carta $(\mathcal{U}, \phi = (x^1, \dots, x^n))$ funciones C^∞ de $\mathcal{U}$ en $\mathbb{R}$ dadas por

$$p \in \mathcal{U} \mapsto T(p)(dx^{i_1}, dx^{i_2}, \dots, dx^{i_r}, \frac{\partial}{\partial x^{j_1}}\Big|_p, \frac{\partial}{\partial x^{j_2}}\Big|_p, \dots, \frac{\partial}{\partial x^{j_s}}\Big|_p).$$

Habitualmente expresaremos estas componentes en términos de las funciones coordenadas, que a su vez dependen del punto p.

<u>Ejemplo</u>: En S^1 tenemos la carta $(S^1 - \{(-1,0)\}, \theta)$ donde $\theta = \theta(x,y)$ da el argumento (ángulo) de cada punto $(x,y) \in S^1$ en el rango $(-\pi, \pi)$. La fórmula

$$T = (x+y)\frac{\partial}{\partial \theta}$$

define un campo de vectores C^∞ en (la subvariedad) $S^1 - \{(-1,0)\}$ porque $f : S^1 - \{(-1,0)\} \longrightarrow \mathbb{R}$ dada por $f(x,y) = x + y$ es C^∞, ya que $f \circ \theta^{-1}(t) = \cos t + \operatorname{sen} t$ es C^∞ como función de $(-\pi, \pi)$ en $\mathbb{R}$. Como $x = \cos\theta$, $y = \operatorname{sen}\theta$, podemos escribir

$$T = (\cos\theta + \operatorname{sen}\theta)\frac{\partial}{\partial \theta}.$$

No hay un gran abuso en dar la componente en términos de la función coordenada θ pues a fin de cuentas $\theta = \theta(x,y)$. Si uno se pusiera muy pesado y quisisera ver la dependencia completa en el punto (x,y) debería escribir

$$T = (\cos\theta(x,y) + \operatorname{sen}\theta(x,y))\frac{\partial}{\partial\theta}\Big|_{(x,y)}.$$

Dadas dos cartas $\big(\mathcal{U}, \phi = (x^1,\dots,x^m)\big)$, $\big(\mathcal{U}', \phi' = (x'^1,\dots,x'^m)\big)$ que se solapan, $\mathcal{U}\cap\mathcal{U}' \neq \emptyset$, la función $\phi\circ\phi'^{-1}$ pasa de $(x'^1,\dots,x'^m)$ a $(x^1,\dots,x^m)$ y por razones obvias la matriz de su diferencial se suele escribir $\partial x^i/\partial x'^j$ y su inversa $\partial x'^i/\partial x^j$. En cada carta se tendrán campos $\partial/\partial x^1,\dots,\partial/\partial x^m$, $dx^1,\dots,dx^m$ (usando ϕ) y $\partial/\partial x'^1,\dots,\partial/\partial x'^m$, $dx'^1,\dots,dx'^m$ (usando ϕ') que dan las bases del espacio tangente y cotangente.

Lema 1.3.1 *Con la notación anterior*

$$1)\ \frac{\partial}{\partial x'^j} = \frac{\partial x^i}{\partial x'^j}\frac{\partial}{\partial x^i}, \qquad\qquad 2)\ dx'^i = \frac{\partial x'^i}{\partial x^j}dx^j.$$

Demostración: Si consideramos la aplicación tangente de la función $\mathrm{Id} : M \longrightarrow M$, 1) es consecuencia inmediata de la Proposición 1.2.3. Para dar una prueba independiente, nótese que la regla de la cadena asegura que para cada función $f : M \longrightarrow \mathbb{R}$ se cumple, $D(f\circ\phi'^{-1}) = D(f\circ\phi^{-1})\cdot D(\phi\circ\phi'^{-1})$ y empleando (1.1) se tiene el resultado, ya que las componentes de las matrices fila $D(f\circ\phi'^{-1})$ y $D(f\circ\phi^{-1})$ representan la acción de $\partial/\partial x'^j$ y $\partial/\partial x^i$ sobre f.

Para comprobar 2) basta ver que ambos miembros aplicados a cualquier $\partial/\partial x'^l$ dan el mismo resultado. Para el primer miembro éste es, por definición, δ^i_l y para el segundo

$$\frac{\partial x'^i}{\partial x^j}dx^j\Big(\frac{\partial}{\partial x'^l}\Big) = \frac{\partial x'^i}{\partial x^j}dx^j\Big(\frac{\partial x^k}{\partial x'^l}\frac{\partial}{\partial x^k}\Big) = \frac{\partial x'^i}{\partial x^j}\frac{\partial x^k}{\partial x'^l}\delta^j_k = \frac{\partial x'^i}{\partial x^k}\frac{\partial x^k}{\partial x'^l} = \delta^i_l$$

donde en el primer paso se ha usado 1) y en el último que la primera matriz es inversa de la segunda. □

Estas relaciones prueban que para cualquier tensor

$$T\Big(dx'^{i_1},\dots,dx'^{i_r},\frac{\partial}{\partial x'^{j_1}},\dots,\frac{\partial}{\partial x'^{j_s}}\Big)$$

coincide con

$$T\Big(\frac{\partial x'^{i_1}}{\partial x^k}dx^k,\dots,\frac{\partial x'^{i_r}}{\partial x^k}dx^k,\frac{\partial x^l}{\partial x'^{j_1}}\frac{\partial}{\partial x^l},\dots,\frac{\partial x^l}{\partial x'^{j_s}}\frac{\partial}{\partial x^l}\Big)$$

Por tanto, cuando cambiamos de carta (o parametrización) las componentes de un tensor de tipo (r,s) en una variedad cambian por la fórmula

$$\boxed{T'^{i_1 i_2\dots i_r}_{j_1 j_2\dots j_s} = \Big(\frac{\partial x'^{i_1}}{\partial x^{k_1}}\cdot\frac{\partial x'^{i_2}}{\partial x^{k_2}}\cdot\dots\cdot\frac{\partial x'^{i_r}}{\partial x^{k_r}}\Big)\cdot\Big(\frac{\partial x^{l_1}}{\partial x'^{j_1}}\cdot\frac{\partial x^{l_2}}{\partial x'^{j_2}}\cdot\dots\cdot\frac{\partial x^{l_s}}{\partial x'^{j_s}}\Big)T^{k_1 k_2\dots k_r}_{l_1 l_2\dots l_s}} \tag{1.2}$$

Esta fórmula es tan característica de los tensores que en muchos libros, sobre todo en los más orientados a la Física, se definen los tensores y campos de tensores como conjuntos de números o funciones sujetos a esta regla de transformación, que a veces se llama *tensorialidad* por antonomasia. No hay que asustarse con una expresión tan compleja. En primer lugar, es fácil de recordar notando que los índices repetidos se deben "simplificar". Y por otra parte, no tiene un significado profundo, simplemente representa lo que ocurre cuando cambiamos de base las variables de un tensor; lo que hay de singular es que los cambios de carta corresponden a cambios de base en el espacio tangente y cotangente cuya matriz es un poco fea: la jacobiana (o su inversa). Ahora podemos apreciar por qué la contracción está bien definida, basta aplicar la regla de la cadena para darse cuenta de que la contracción de un tensor se transforma como un tensor (ejercicio).

Ejemplo: En cada punto de $\mathbb{R}^2$ tenemos un tensor métrico en el plano tangente dado por $dx \otimes dx + dy \otimes dy$ con las coordenadas usuales (omitimos por brevedad el punto), esto es un campo de tensores métricos (nótese que no es más que el producto escalar usual en cada punto), es decir, una métrica. Si ahora cambiamos a coordenadas polares $x = r\cos\theta$, $y = r\,\mathrm{sen}\,\theta$ entonces podemos calcular los nuevos coeficientes del tensor métrico usando la fórmula anterior, para $(x^1, x^2) = (x, y)$ y $(x'^1, x'^2) = (r, \theta)$, o simplemente sustituir, según el lema anterior,

$$dx = \cos\theta\, dr - r\,\mathrm{sen}\,\theta\, d\theta, \qquad dy = \mathrm{sen}\,\theta\, dr + r\cos\theta\, d\theta$$

para obtener

$$(\cos\theta\, dr - r\,\mathrm{sen}\,\theta\, d\theta) \otimes (\cos\theta\, dr - r\,\mathrm{sen}\,\theta\, d\theta) + \\ (\mathrm{sen}\,\theta\, dr + r\cos\theta\, d\theta) \otimes (\mathrm{sen}\,\theta\, dr + r\cos\theta\, d\theta) = dr \otimes dr + r^2 d\theta \otimes d\theta.$$

En general los tensores no se comportan bien al derivarlos componente a componente porque en (1.2) aparecerían derivadas segundas que estropean la tensorialidad. Más adelante introduciremos una derivada especial que tiene carácter tensorial. Veamos un ejemplo trivial en el que sí se puede derivar y nos debería hacer dudar del nómbre "vector" gradiente.

Ejemplo: Una función $f : M \longrightarrow \mathbb{R}$ es por definición un tensor de tipo $(0, 0)$, su única componente es la propia función. Sus derivadas parciales definen un tensor porque

$$\frac{\partial f}{\partial x'^j} = \frac{\partial x^l}{\partial x'^j}\frac{\partial f}{\partial x^l}.$$

Comparando con (1.2) vemos que las componentes del gradiente en variedades (que obviamente generaliza al habitual) corresponden a un tensor de tipo $(0, 1)$, no un tensor $(1, 0)$ que representaría un vector. Esto es natural porque $df = \frac{\partial f}{\partial x^1}\, dx^1 + \frac{\partial f}{\partial x^2}\, dx^2 + \cdots + \frac{\partial f}{\partial x^n}\, dx^n$ es una uno forma.

Si todavía queda algún escéptico, tómese $f(x, y, z) = x + y + z$ definida en $\mathbb{R}^3$ con la carta trivial. La transformación (cambio de carta) $x' = 2x$, $y' = 2y$, $z' = 2z$ pasa el vector de $T_{\mathbf{0}}(\mathbb{R}^3)$ de coordenadas $(1, 1, 1)$ al de coordenadas $(2, 2, 2)$. El gradiente de $x+y+z$ es $(1, 1, 1)$ pero el de $x'/2+y'/2+z'/2(= x+y+z$ en las nuevas coordenadas) es $(1/2, 1/2, 1/2)$. Por mucho que nos empeñemos el vector gradiente no es un vector[7] en sentido estricto.

Una vez entendido todo esto, los elementos de la relatividad especial deberían ser fáciles.

Después de los trabajos de Lorentz y Einstein, Minkowski dio un interesante giro en la manera de entender la relatividad especial introduciendo el concepto de *espacio-tiempo* que geométricamente no es más que $\mathbb{R}^4$ con la métrica de Minkowski mencionada en la primera sección. Para acercarnos rápidamente al significado físico, nos desprendemos de dos coordenadas y consideramos $\mathbb{R}^2$ con una carta $(\mathbb{R}^2, \phi = (t, x))$ y la métrica

$$G = -dt \otimes dt + c^2 dx \otimes dx$$

donde físicamente t indica el tiempo, x el espacio y c es una constante que representa la velocidad de la luz en el vacío (en el sistema interancional es alrededor de $2{,}99 \cdot 10^8 m/s$ pero hay otras unidades, llamadas relativistas, con las que vale 1 para que la fórmula de G quede bonita). La reformulación de Minkowski de la relatividad se basa en el postulado de que G debe permanecer invariante para todos los *observadores inerciales* (intuitivamente los que no están sometidos a fuerzas). Al igual que los movimientos del plano eran las transformaciones que dejaban invariante el producto escalar usual $dx \otimes dx + dy \otimes dy$, para la relatividad simplemente[8] había que pensar en un producto escalar extraño del que se deducía matemáticamente a partir de unos postulados básicos lo que Einstein había obtenido con experimentos imaginarios con varillas y espejos. Minkowski afirma orgulloso en 1907 (véase [Ei-Lo-Mi-We]) "A partir de ahora el espacio por sí mismo y el tiempo por sí mismo, están condenados a desvanecerse en meras sombras y sólo una especie de unión de ambos conservará una realidad independiente".

En términos geométricos, si un observador inercial utiliza para medir espacios y

---

[7]R. P. Feynman, premio Nobel de Física, dedica toda la sección 2-5 de su magnífico libro [Fe-Le-Sa] a demostrar al lector que el vector gradiente es un vector. ¿Dónde ha quedado el argumento de autoridad? El truco está en que Feynman sólo considera transformaciones dadas por matrices ortogonales (realmente sólo giros) y recuérdese que estas matrices cumplen $A = (A^{-1})^t$, por tanto intercambiar índices y numeradores por denominadores no tiene efecto sobre (1.2). Geométricamente el gradiente es un vector normal, y sigue siéndolo cuando sólo hacemos movimientos en $\mathbb{R}^n$ pero como hemos visto, el gradiente no se comporta como un vector por cambios de carta generales.

[8]Como se ha indicado en una nota anterior este "simplemente" fue al principio discutido incluso por el propio Einstein, pero el desarrollo de la relatividad general mostró que el planteamiento de Minkowski no sólo era matemáticamente elegante, sino que era el camino más natural y sencillo para comprender los avances posteriores.

tiempo una carta $(\mathbb{R}^2, \phi' = (t', x'))$ con $t' = t'(t, x)$, $x' = x'(t, x)$ se debe cumplir

$$-dt \otimes dt + c^2 dx \otimes dx = -dt' \otimes dt' + c^2 dx' \otimes dx'.$$

Efectuando el cambio de carta e igualando coeficientes,

$$-1 = -\big(\frac{\partial t'}{\partial t}\big)^2 + c^2\big(\frac{\partial x'}{\partial t}\big)^2, \qquad \frac{\partial t'}{\partial t}\frac{\partial t'}{\partial x} = c^2 \frac{\partial x'}{\partial t}\frac{\partial x'}{\partial x}, \qquad c^2 = -\big(\frac{\partial t'}{\partial x}\big)^2 + c^2\big(\frac{\partial x'}{\partial x}\big)^2.$$

Cuya solución general es

$$\begin{pmatrix} \partial t'/\partial t & \partial t'/\partial x \\ \partial x'/\partial t & \partial x'/\partial x \end{pmatrix} = \begin{pmatrix} \sqrt{1 + c^2\lambda^2} & c^2\lambda \\ \lambda & \sqrt{1 + c^2\lambda^2} \end{pmatrix} \quad \text{con } \lambda = \lambda(t, x),$$

salvo la indeterminación en el signo de la raíz que no consideramos por corresponder a una sencilla simetría. También componiendo con una traslación podemos suponer $\phi' \circ \phi(0, 0) = (0, 0)$. Para seguir, pidamos un poco de ayuda a la interpretación física: los sistemas inerciales tienen velocidades relativas constantes unos con respecto de otros porque si hubiera aceleraciones estarían sometidos a fuerzas. Si v es la velocidad relativa, al segundo observador le parecerá que las partículas con $(t, x) = (t, vt)$ están quietas. En términos matemáticos estamos pidiendo $x'(t, vt) = 0$, que derivando implica $\partial x'/\partial t + v\partial x'/\partial x = 0$ y combinado con las ecuaciones anteriores permite deducir $\lambda = -v/\sqrt{c^4 - c^2v^2}$. En definitiva, con las simplificaciones supuestas, los cambios de carta entre observadores inerciales son las llamadas *transformaciones de Lorentz*

$$\begin{pmatrix} t' \\ x' \end{pmatrix} = \frac{1}{\sqrt{1 - v^2/c^2}} \begin{pmatrix} 1 & -v/c^2 \\ -v & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix}.$$

Esto es contraintuitivo porque todos diríamos que si un observador O' se mueve a velocidad v delante de O, debería cumplirse $x' = x - vt$ y por supuesto $t' = t$. La relación anterior contradice estas ideas preconcebidas. También se sigue que no hay sistemas inerciales con $v \geq c$. En los fenómenos que podemos observar, normalmente v^2/c^2 es prácticamente cero y por ello las transformaciones de Lorentz escapan al alcance inmediato de nuestros sentidos.

Ejercicios de la sección 3

1) Considerando $\mathbb{R}^3$ como variedad, escribir la uno forma $x\,dx + y\,dy + z\,dz$ en coordenadas esféricas.

2) En un ejemplo de la sección se definió un campo de vectores sobre $S^1 - \{(-1, 0)\}$. Extenderlo a un campo de vectores C^∞ en S^1 comprobando que lo es empleando alguna carta compatible.

3) Se dice que un tensor de tipo $(0,2)$ es *simétrico* si $T_{ij} = T_{ji}$ donde T_{ij} son sus componentes.

a) Demostrar que este concepto de simetría está bien definido, es decir, que no depende de la carta empleada para calcular las componentes.

b) Comprobar que sin embargo no se puede extender a tensores de tipo $(1,1)$, concretamente, construir un ejemplo para el que $T^i_j = T^j_i$ se cumpla usando una carta pero no otra.

$_*$c) ¿Qué matrices simétricas lo siguen siendo en cualquier otra base?

4) Dar un ejemplo concreto en $\mathbb{R}^2$ con la carta trivial que muestre que las derivadas parciales de las componentes de un campo de vectores no tienen carácter tensorial: no se transforman como las componentes de un tensor de tipo $(1,1)$.

5) Según habíamos visto, el cambio a polares lleva la métrica usual de $\mathbb{R}^2$, $dx\otimes dx + dy\otimes dy$, a $dr \otimes dr + r^2 d\theta \otimes d\theta$. Hallar ahora un cambio de coordenadas (de carta) en $\mathbb{R}^2$ que pase la *métrica de Minkowski* en $\mathbb{R}^2$, $dx \otimes dx - dy \otimes dy$, a $dr \otimes dr - r^2 d\theta \otimes d\theta$. *Indicación:* Los dos problemas son similares salvo el "cambio" $y \mapsto y\sqrt{-1}$, $\theta \mapsto \theta\sqrt{-1}$.

6) Sea M una variedad bidimensional. Un campo de uno formas en M se expresa en cada carta $(\mathcal{U}, \phi = (x^1, x^2))$ como $T = T_1 dx^1 + T_2 dx^2$. Sea $\mathcal{D}$ el operador que asigna a T el tensor de tipo $(0,2)$ en M dado por

$$\Big(\frac{\partial T_2}{\partial x^1} - \frac{\partial T_1}{\partial x^2}\Big) dx^1 \otimes dx^2 + \Big(\frac{\partial T_1}{\partial x^2} - \frac{\partial T_2}{\partial x^1}\Big) dx^2 \otimes dx^1.$$

Demostrar que $\mathcal{D}$ está bien definido, es decir, que no depende de la carta escogida.

1.4. Formas diferenciales

Los contenidos de esta sección están ligados al nombre de É. Cartan (no confundir con su hijo H. Cartan, también matemático renombrado), quien introdujo el concepto de forma diferencial tal como ahora lo conocemos y además definió una nueva operación, la derivada exterior, que resulta fundamental para escribir y describir algunos resultados de geometría diferencial, incluido el teorema de Stokes que veremos en el próximo capítulo. La estructura algebraica subyacente en la que se basó se llama álgebra exterior y fue introducida por H. Grassman con anterioridad.

Aquí no nos pararemos en las estructuras e iremos directamente a la definición.

Definición: Se llama *k-forma alternada* a un tensor k veces covariante, T, que es antisimétrico en cualquier par de argumentos, es decir

$$T(\dots, \vec{v_i}, \dots, \vec{v_j}, \dots) = -T(\dots, \vec{v_j}, \dots, \vec{v_i}, \dots).$$

El conjunto de k-formas alternadas sobre un espacio vectorial V se denota con $\mathrm{Alt}^k(V)$.

Las formas diferenciales corresponden al caso en que V es el espacio tangente de una variedad.

Definición: Una *k-forma diferencial* es un campo de k-formas alternadas en una variedad. El conjunto de k-formas diferenciales sobre una variedad M se denota con $\Omega^k(M)$.

En analogía con lo que se hacía en la teoría general de tensores, se conviene que $\mathrm{Alt}^0(V)$ son las constantes, esto es, $\mathbb{R}$ y por tanto $\Omega^0(M) = C^\infty(M)$. El caso $k = 1$ es también un poco singular porque las definiciones no ponen ninguna restricción y decir 1-forma alternada es lo mismo que decir tensor una vez covariante. Nótese que esto es coherente con nuestra denominación de "uno formas" en la sección de repaso. Habitualmente se suelen representar las formas diferenciales (y también a veces las formas alternadas) con letras griegas minúsculas, especialmente ω y η. Evidentemente si $m > \dim V$ toda m-forma alternada es nula, en particular en una variedad n-dimensional $\Omega^m(M) = \{0\}$ para $m > n$.

El conjunto $\mathrm{Alt}^k(V)$ tiene una estructura de espacio vectorial sobre $\mathbb{R}$ con las operaciones habituales de suma y multiplicación por números reales. En el caso de $\Omega^k(M)$ esos números reales dependerán del punto sobre el que estemos considerando el espacio tangente y por tanto son funciones $f : M \longrightarrow \mathbb{R}$ que, como siempre, supondremos C^∞. La inconveniencia de que estas funciones no formen un cuerpo estropea la estructura de espacio vectorial.

Los siguientes ejemplos se podrían deducir de resultados posteriores. Ahora nos servirán para entender un poco mejor los conceptos.

Ejemplo: Hallar todos los elementos de $\mathrm{Alt}^2(\mathbb{R}^2)$.

Si $T \in \mathrm{Alt}^2(\mathbb{R}^2)$ y $\vec{v}_1 = a_1^1\vec{e}_1 + a_1^2\vec{e}_2$, $\vec{v}_2 = a_2^1\vec{e}_1 + a_2^2\vec{e}_2$, entonces usando la linealidad y $T(\vec{x},\vec{x}) = 0$

$$\begin{aligned} T(\vec{v}_1,\vec{v}_2) &= a_1^1a_2^1T(\vec{e}_1,\vec{e}_1) + a_1^1a_2^2T(\vec{e}_1,\vec{e}_2) + a_1^2a_2^1T(\vec{e}_2,\vec{e}_1) + a_1^2a_2^2T(\vec{e}_2,\vec{e}_2) \\ &= (a_1^1a_2^2 - a_1^2a_2^1)T(\vec{e}_1,\vec{e}_2) = T(\vec{e}_1,\vec{e}_2)\det(\vec{v}_1,\vec{v}_2). \end{aligned}$$

Entonces el único tensor que hay en $\mathrm{Alt}^2(\mathbb{R}^2)$ es el determinante salvo multiplicar por una constante, $T(\vec{e}_1,\vec{e}_2)$, es decir, $\mathrm{Alt}^2(\mathbb{R}^2) = \{\lambda \det\}$. Nótese que $T(\vec{e}_1,\vec{e}_2)$ es una constante una vez fijada una base y también el determinante depende de la base elegida. Si se quiere una formulación independiente de las coordenadas de $\vec{v}_1$ y $\vec{v}_2$ en una base de $\mathbb{R}^2$, podemos elegir una base del dual $\{\widetilde{\varphi}^1,\widetilde{\varphi}^2\}$ y decir que todo elemento de $\mathrm{Alt}^2(\mathbb{R}^2)$ es de la forma

$$T = \lambda\,\widetilde{\varphi}^1 \otimes \widetilde{\varphi}^2 - \lambda\,\widetilde{\varphi}^2 \otimes \widetilde{\varphi}^1.$$

Esto es exactamente lo mismo que antes porque si $\{\vec{e}_1,\vec{e}_2\}$ es la base cuya base dual es $\{\widetilde{\varphi}^1,\widetilde{\varphi}^2\}$, $\widetilde{\varphi}^i(\vec{v}_j) = a_j^i$.

De la misma forma, si M es una variedad de dimensión 2, entonces en cada carta $(\mathcal{U},\phi = (x^1,\ldots,x^n))$ los elementos de $\Omega^2(M)$ son de la forma

$$\omega = f\,dx^1 \otimes dx^2 - f\,dx^2 \otimes dx^1$$

con $f : M \longrightarrow \mathbb{R}$.

Podemos leer el resultado del ejemplo anterior diciendo que si $\{\widetilde{\varphi}^1, \widetilde{\varphi}^2\}$ es una base de V^* y $\dim V = 2$, entonces $\{\widetilde{\varphi}^1 \otimes \widetilde{\varphi}^2 - \widetilde{\varphi}^2 \otimes \widetilde{\varphi}^1\}$ es una base (con un solo elemento) de $\mathrm{Alt}^2(V)$. Intentemos hacer un ejemplo en dimensión mayor.

Ejemplo: Hallar una base de de $\mathrm{Alt}^2(\mathbb{R}^3)$.

Escribamos como antes $\vec{v}_j = a_j^i \vec{e}_i$. El desarrollo de $T(\vec{v}_1, \vec{v}_2)$ da lugar a nueve términos. Olvidándonos de los tres con coeficientes nulos $T(\vec{e}_i, \vec{e}_i)$ y agrupando $T(\vec{e}_i, \vec{e}_j)$ con su negativo $T(\vec{e}_j, \vec{e}_i)$, se llega a

$$T(\vec{v}_1, \vec{v}_2) = \begin{vmatrix} a_1^2 & a_1^3 \\ a_2^2 & a_2^3 \end{vmatrix} T(\vec{e}_2, \vec{e}_3) + \begin{vmatrix} a_1^3 & a_1^1 \\ a_2^3 & a_2^1 \end{vmatrix} T(\vec{e}_3, \vec{e}_1) + \begin{vmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{vmatrix} T(\vec{e}_1, \vec{e}_2).$$

De nuevo determinantes, que son menores de orden dos de la matriz de coordenadas de $\vec{v}_1$ y $\vec{v}_2$. Empleando la base dual $\{\widetilde{\varphi}^1, \widetilde{\varphi}^2, \widetilde{\varphi}^3\}$ tenemos que estos menores son $M_1(\vec{v}_1, \vec{v}_2)$, $M_2(\vec{v}_1, \vec{v}_2)$ y $M_3(\vec{v}_1, \vec{v}_2)$ donde

$$M_1 = \widetilde{\varphi}^2 \otimes \widetilde{\varphi}^3 - \widetilde{\varphi}^3 \otimes \widetilde{\varphi}^2, \qquad M_2 = \widetilde{\varphi}^3 \otimes \widetilde{\varphi}^1 - \widetilde{\varphi}^1 \otimes \widetilde{\varphi}^3, \qquad M_3 = \widetilde{\varphi}^1 \otimes \widetilde{\varphi}^2 - \widetilde{\varphi}^2 \otimes \widetilde{\varphi}^1.$$

Y podemos escribir $T = \lambda^1 M_1 + \lambda^2 M_2 + \lambda^3 M_3$. Por otro lado M_1, M_2 y M_3 son formas alternadas (empléese el lema) y es fácil ver que son linealmente independientes haciéndolas actuar sobre $\{\vec{e}_1, \vec{e}_2\}$. En resumen, hemos demostrado que $\{M_1, M_2, M_3\}$ es una base de $\mathrm{Alt}^2(\mathbb{R}^3)$. Como curiosidad, nótese que M_1, M_2 y M_3 dan las coordenadas del producto vectorial.

Para desarrollos posteriores va a ser conveniente cambiar de nombre a los determinantes.

Definición: Dados $\widetilde{\varphi}^1, \widetilde{\varphi}^2, \dots, \widetilde{\varphi}^k \in V^*$ su *producto exterior*, $\widetilde{\varphi}^1 \wedge \widetilde{\varphi}^2 \wedge \cdots \wedge \widetilde{\varphi}^k$, se define como la forma de $\mathrm{Alt}^k(V)$ dada por

$$(\widetilde{\varphi}^1 \wedge \widetilde{\varphi}^2 \wedge \cdots \wedge \widetilde{\varphi}^k)(\vec{v}_1, \vec{v}_2, \dots \vec{v}_k) = \det\big(\widetilde{\varphi}^i(\vec{v}_j)\big)_{1 \le i,j \le n}.$$

La definición se extiende de la manera obvia a uno formas.

Con esta notación tenemos en el primero de los ejemplos $T = \lambda \widetilde{\varphi}^1 \wedge \widetilde{\varphi}^2$ y $\omega = f dx^1 \wedge dx^2$, mientras que en el segundo $T = \lambda^1 \widetilde{\varphi}^2 \wedge \widetilde{\varphi}^3 + \lambda^2 \widetilde{\varphi}^3 \wedge \widetilde{\varphi}^1 + \lambda^3 \widetilde{\varphi}^1 \wedge \widetilde{\varphi}^2$. Todavía más, resulta que no hay en el mundo de las formas alternadas nada muy diferente de los determinantes.

Proposición 1.4.1 *Sea* $\{\widetilde{\varphi}^1, \widetilde{\varphi}^2, \dots, \widetilde{\varphi}^n\}$ *una base de* V^**, entonces*

$$\{\widetilde{\varphi}^{i_1} \wedge \widetilde{\varphi}^{i_2} \wedge \cdots \wedge \widetilde{\varphi}^{i_k} \ : \ 1 \le i_1 < i_2 < \ldots < i_k \le n\}$$

es una base de $\mathrm{Alt}^k(V)$*.*

Demostración: Sea $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es una base V cuya base dual es $\{\widetilde{\varphi}^1, \widetilde{\varphi}^2, \dots, \widetilde{\varphi}^n\}$. Consideremos el tensor

$$T = \sum_{i_1<i_2<\dots<i_k} a_{i_1 i_2 \dots i_k} \widetilde{\varphi}^{i_1} \wedge \widetilde{\varphi}^{i_2} \wedge \dots \wedge \widetilde{\varphi}^{i_k}.$$

Si T fuera el tensor nulo para ciertos coeficientes, calculando $T(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$ se tendría $a_{i_1 i_2 \dots i_k} = 0$, por tanto los elementos del conjunto son linealmente independientes.

Por otro lado, si $\omega \in \mathrm{Alt}^k(V)$ y sus componentes en la base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ son $\omega_{i_1 i_2 \dots i_k}$ entonces eligiendo $a_{i_1 i_2 \dots i_k} = \omega_{i_1 i_2 \dots i_k}$ se tiene que T y ω tienen las mismas componentes $i_1 i_2 \dots i_k$ siempre que $i_1 < i_2 < \dots < i_k$. En el resto de los casos también deben coincidir por la antismetría de las formas alternadas al intercambiar dos argumentos. $\square$

Evidentemente algo similar ocurre en $\Omega^k(M)$.

Corolario 1.4.2 *Cualquier elemento de $\Omega^k(M)$ se puede escribir como*

$$\omega = \sum_{i_1<i_2<\dots<i_k} f_{i_1 i_2 \dots i_k}\, dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$$

Nota: Como explicamos en la sección anterior, las componentes $f_{i_1 i_2 \dots i_k}$ son en principio funciones de $p \in M$ pero habitualmente esa dependencia se expresa a través de las funciones coordenadas y escribiremos $f_{i_1 i_2 \dots i_k} = f_{i_1 i_2 \dots i_k}(x^1, x^2, \dots, x^n)$.

Es natural definir

$$\left(\widetilde{\varphi}^{i_1} \wedge \widetilde{\varphi}^{i_2} \wedge \dots \wedge \widetilde{\varphi}^{i_k}\right) \wedge \left(\widetilde{\varphi}^{j_1} \wedge \widetilde{\varphi}^{j_2} \wedge \dots \wedge \widetilde{\varphi}^{j_l}\right) = \widetilde{\varphi}^{i_1} \wedge \widetilde{\varphi}^{i_2} \wedge \dots \wedge \widetilde{\varphi}^{i_k} \wedge \widetilde{\varphi}^{j_1} \wedge \widetilde{\varphi}^{j_2} \wedge \dots \wedge \widetilde{\varphi}^{j_l}.$$

Con ello y la Proposición 1.4.1 o el Corolario 1.4.2 habremos extendido por la distributiva la definición del *producto exterior* a una operación

$$\wedge : \mathrm{Alt}^k(V) \times \mathrm{Alt}^l(V) \longrightarrow \mathrm{Alt}^{k+l}(V) \qquad \text{y} \qquad \wedge : \Omega^k(M) \times \Omega^l(M) \longrightarrow \Omega^{k+l}(M).$$

Para respetar los convenios se debe interpretar que el producto exterior por números, que son 0-formas, es el producto usual (por ejemplo $2 \wedge \omega = \omega \wedge 2 = 2\omega$).

En los textos se suele dar una definición más invariante del producto exterior que muestra su relación con el producto tensorial [Sp2]. Nuestra definición es cuestionable pero permite deducir sin dificultad dos propiedades básicas: la asociativa

$$\omega \wedge (\eta \wedge \lambda) = (\omega \wedge \eta) \wedge \lambda$$

y la anticonmutativa (o *superconmutativa*, si uno es físico)

$$\omega \wedge \eta = (-1)^{rs} \eta \wedge \omega, \tag{1.3}$$

donde $\omega \in \mathrm{Alt}^k(V)$, $\eta \in \mathrm{Alt}^l(V)$ o $\omega \in \Omega^k(M)$, $\eta \in \Omega^l(M)$.

Ejemplo: Sean $\omega = dx + y\ dz \in \Omega^1(\mathbb{R}^3)$ y $\eta = z\ dx \wedge dy \in \Omega^2(\mathbb{R}^3)$ definidas en la carta trivial $(\mathbb{R}^3, \phi = (x, y, z))$, entonces

$$\omega \wedge \eta = z\ dx \wedge dx \wedge dy + yz\ dz \wedge dx \wedge dy, \qquad \eta \wedge \omega = z\ dx \wedge dy \wedge dx + zy\ dx \wedge dy \wedge dz.$$

Se cumple $dx \wedge dx \wedge dy = dx \wedge dy \wedge dx = 0$ porque corresponden a determinantes con dos filas iguales. De la misma forma $dz \wedge dx \wedge dy = -dx \wedge dz \wedge dy = dx \wedge dy \wedge dz$ porque los determinantes cambian de signo al intercambiar dos de sus filas. Con ello hemos comprobado $\omega \wedge \eta = (-1)^{1\cdot 2}\ \eta \wedge \omega$ en consonancia con (1.3).

Si las formas alternadas y las formas diferenciales no son más que combinaciones de determinantes, ¿por qué no escribimos simplemente esos determinantes y nos olvidamos de estas definiciones tan raras? La respuesta es que los determinantes aparecen en algunos teoremas, por ejemplo en el de Stokes, de una manera complicada y más vale inventar una notación para poder proceder simbólicamente. No hay nada nuevo en esta forma de actuar y los propios determinantes son un buen ejemplo: la relación $\det(A) \cdot \det(B) = \det(AB)$ es a la vez bonita, simple y no trivial (¿alguien recuerda la prueba?) pero si no existiera una notación especial para el determinante y escribiéramos todo el desarrollo, digamos por ejemplo en el caso 3×3, ¿nos diría algo esa relación? ¿merecería los adjetivos anteriores?[9]

En la siguiente definición no usaremos a propósito el convenio de sumación para mayor claridad.

Definición: Dado un elemento de $\Omega^k(M)$ que en una carta es de la forma

$$\omega = \sum_{i_1, i_2, \dots, i_k} f_{i_1 i_2 \dots i_k}\ dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$$

se llama *derivada exterior* de ω al elemento $d\omega \in \Omega^{k+1}(M)$ dado por

$$d\omega = \sum_j \sum_{i_1, i_2, \dots, i_k} \frac{\partial f_{i_1 i_2 \dots i_k}}{\partial x^j}\ dx^j \wedge dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}$$

(en el caso especial $k = 0$, $df = \sum \frac{\partial f}{\partial x^j}\ dx^j$).

[9] La utilidad o conveniencia de una notación o un modo de cálculo no son en absoluto evidentes a priori ni siquiera para los expertos pues a veces dependen de desarrollos ulteriores de las Matemáticas. Por ejemplo, cuando Grassman creó el álgebra exterior que después fue retomada por Cartan para desarrollar la teoría de formas diferenciales, los matemáticos de su tiempo no le prestaron mucha atención, tanto es así que en los últimos años de su vida prácticamente dejó las Matemáticas y se dedicó a investigar en Lingüística.

En principio no está claro con esta definición que d se aplique a todo elemento de $\Omega^k(M)$ y, todavía peor, que sea coherente con los cambios de carta. Es posible evitar este último problema comenzando la casa por el tejado con un tratamiento axiomático: se imponen las propiedades de la Proposición 1.4.3 en las que no aparecen cartas y se prueba que sólo hay un operador decente d con esas propiedades. Al calcular su expresión en coordenadas se obtiene la fórmula anterior [Bi-Go]. Daremos una breve demostración directa sin entrar mucho en detalles porque en el siguiente capítulo tendremos una visión más clara de este punto (véase una prueba más sencilla y natural en [GoG] §1.3).

Por la antisimetría podemos suponer que en la definición anterior la sumación es sobre $i_1 < i_2 < \ldots < i_k$. Las cantidades $f_{i_1 i_2 \ldots i_k}$ son componentes de un tensor k veces covariante y por tanto responden a los cambios de carta por medio de la fórmula:

$$f_{i_1 i_2 \ldots i_k} = \frac{\partial x'^{j_1}}{\partial x^{i_1}} \frac{\partial x'^{j_2}}{\partial x^{i_2}} \cdots \frac{\partial x'^{j_k}}{\partial x^{i_k}} f'_{j_1 j_2 \ldots j_k}$$

donde hay que entender que si no se cumple $j_1 < j_2 < \ldots < j_k$ entonces $f'_{j_1 j_2 \ldots j_k}$ es $\operatorname{sgn}(\sigma) f'_{\sigma(j_1)\sigma(j_2)\ldots\sigma(j_k)}$ con σ la permutación que ordena $j_1, j_2, \ldots, j_k$. Al derivar con respecto de x^j no tenemos una relación similar entre las derivadas de $f_{i_1 i_2 \ldots i_k}$ y de $f'_{j_1 j_2 \ldots j_k}$ porque aparecen derivadas parciales segundas de los cambios de carta. Sin embargo estos términos no influyen en la expresión de $d\omega$ porque la antisimetría del producto de uno formas asegura $\sum_{i,j} \frac{\partial^2 x^k}{\partial x^i \partial x^j}\, dx^i \wedge dx^j = 0$.

Por último veamos dos propiedades importantes que de hecho determinan la derivada exterior. La primera muestra la relación entre los definiciones de producto exterior y derivada exterior y la segunda será crucial para definir la cohomología de de Rham.

Proposición 1.4.3 *Sean $\omega \in \Omega^k(M)$ y $\eta \in \Omega^l(M)$, entonces*

$$1) \quad d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta, \qquad\qquad 2) \quad d(d\omega) = 0.$$

Demostración: Por el Corolario 1.4.2, empleando la linealidad de d y la distributiva, podemos limitarnos al caso $\omega = f\, dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k}$, $\eta = g\, dx^{j_1} \wedge dx^{j_2} \wedge \cdots \wedge dx^{j_l}$. Se tiene

$$\begin{aligned} d(\omega \wedge \eta) &= \frac{\partial (fg)}{\partial x^m}\, dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \\ &= \frac{\partial f}{\partial x^m} g\, dx^m \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \\ &\quad + (-1)^k \frac{\partial g}{\partial x^m} f\, dx^{i_1} \wedge \cdots \wedge dx^{i_k} \wedge dx^m \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_l} \end{aligned}$$

donde se ha usado (1.3). El primer sumando es $d\omega \wedge \eta$ y el segundo $(-1)^k \omega \wedge d\eta$.

Es fácil ver que $d(d\omega)=0$ se cumple para $k=0$:

$$d(df) = d\big(\frac{\partial f}{\partial x^i}\, dx^i\big) = \frac{\partial^2 f}{\partial x^i \partial x^j}\, dx^i \wedge dx^j = 0$$

porque $dx^i \wedge dx^j = -dx^j \wedge dx^i$ y las derivadas parciales cruzadas coinciden. Por otro lado, de 1) se deduce con un pequeño cálculo que $d\big(d(\omega\wedge\eta)\big) = d(d\omega)\wedge\eta - \omega\wedge d(d\eta)$. La prueba se sigue por inducción ya que el corolario anterior permite escribir toda forma como productos exteriores de funciones y diferenciales de funciones[10]. $\Box$

<u>Ejemplo</u>: Sean $\omega, \eta \in \Omega^1(\mathbb{R}^3)$ definidas en la carta trivial $(\mathbb{R}^3, \phi = (x,y,z))$ por

$$\omega = x\, dx + yz\, dy + x^2 y\, dz, \qquad \eta = xy\, dz.$$

Comprobar las relaciones de la proposición anterior.

Comencemos por la segunda haciendo el cálculo muy despacio y con paréntesis innecesarios:

$$\begin{aligned} d\omega &= (dx\wedge dx) + (z\, dy\wedge dy + y\, dz\wedge dy) + (2xy\, dx\wedge dz + x^2\, dy\wedge dz) \\ &= 2xy\, dx\wedge dz + (x^2 - y)\, dy\wedge dz. \end{aligned}$$

En la segunda derivada exterior ya procedemos más rápido:

$$d(d\omega) = 2x\, dy\wedge dx\wedge dz + 2x\, dx\wedge dy\wedge dz = 0.$$

Calculemos ahora los tres términos que participan en la primera relación. La derivada del producto:

$$d(\omega\wedge\eta) = x^2\, dy\wedge dx\wedge dz + y^2 z\, dx\wedge dy\wedge dz = (y^2 z - x^2)\, dx\wedge dy\wedge dz.$$

La derivada del primero por el segundo, para lo cual aprovechamos el cálculo de $d\omega$:

$$d\omega\wedge\eta = (2xy\, dx\wedge dz + (x^2-y)\, dy\wedge dz)\wedge(xy\, dz) = 0.$$

Y el primero por la derivada del segundo:

$$\begin{aligned} \omega\wedge d\eta &= (x\, dx + yz\, dy + x^2 y\, dz)\wedge(y\, dx\wedge dz + x\, dy\wedge dz) \\ &= y^2 z\, dy\wedge dx\wedge dz + x^2\, dx\wedge dy\wedge dz = (x^2 - y^2 z)\, dx\wedge dy\wedge dz. \end{aligned}$$

Entonces se tiene $d(\omega\wedge\eta) = d\omega\wedge\eta + (-1)^k\omega\wedge\eta$, como esperábamos.

[10]Por ejemplo, para $k=2$ una forma diferencial es suma de cosas del tipo $f\, dx^{i_1}\wedge dx^{i_2}$ y si suponemos el resultado probado para $k=0$ y $k=1$, se tiene $d\big(d(f\, dx^{i_1}\wedge dx^{i_2})\big) = d\big(d(f\, dx^{i_1})\big)\wedge dx^{i_2} - f\, dx^{i_1}\wedge d\big(d(dx^{i_2})\big) = 0$. Con nuestros convenios el caso especial $k=1$, el primer paso de la inducción, no es conflicitivo porque $f\wedge dx^{i_1} = f\, dx^{i_1}$.

Ejercicios de la sección 4

1) Responder brevemente a las siguientes preguntas:
i) Si $\omega(\vec{e}_1 + 2\vec{e}_2, \vec{e}_1 + \vec{e}_2) = 2$, ¿cuánto vale $\omega(\vec{e}_2, \vec{e}_1)$?
ii) Si M es una variedad n-dimensional y $\omega \in \Omega^{n-3}(M)$ y $\eta \in \Omega^1(M)$, ¿cuánto vale $d\omega \wedge \eta \wedge d\eta$?
iii) ¿Cuál el la dimensión de $\text{Alt}^k(V)$?
iv) Si $\omega = \widetilde{\varphi}^1 \wedge \widetilde{\varphi}^2 + \widetilde{\varphi}^3 \wedge \widetilde{\varphi}^4$ con $\widetilde{\varphi}^1, \widetilde{\varphi}^2, \widetilde{\varphi}^3, \widetilde{\varphi}^4 \in V^*$, ¿cómo se pueden simplificar las expresiones $\omega \wedge \omega$ y $\omega \wedge \omega \wedge \omega$?

2) Para $n > 1$ definamos en $\mathbb{R}^n$ con la base usual el *producto vectorial generalizado* de $n-1$ vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}$ como cero si son linealmente dependientes, y como el vector $\vec{w}$ que cumple

$$\vec{x} \cdot \vec{w} = \det(\vec{x}, \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1}) \qquad \text{para todo} \quad \vec{x} \in \mathbb{R}^n$$

si son linealmente independientes. Nótese que para $n = 3$ el producto vectorial usual tiene esta propiedad y está caracterizada por ella.

a) Demostrar que este producto vectorial está bien definido. Es decir, que no pueden existir dos $\vec{w}$ con la propiedad anterior. Probar también que el producto vectorial generalizado es siempre ortogonal a cada uno de los vectores de partida.

b) Demostrar que la función M_i que asigna a $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{n-1})$ la i-esima coordenada de su producto vectorial generalizado cumple $M_i \in \text{Alt}^{n-1}(\mathbb{R}^n)$.

c) Expresar M_i en términos de productos exteriores de elementos de la base dual de la usual. *Indicación:* Tómese como $\vec{x}$ el i-ésimo vector de la base canónica.

3) Si $\omega \in \Omega^0(\mathbb{R}^3)$, es decir, si ω es una función, entonces los coeficientes de $d\omega$ vienen dados por el gradiente (usamos la carta trivial). Encontrar relaciones similares con el rotacional y la divergencia en $\Omega^1(\mathbb{R}^3)$ y $\Omega^2(\mathbb{R}^3)$.

4) Comprobar en $\mathbb{R}^3$ (con la carta trivial) la relación $d(\omega \wedge \eta) = d\omega \wedge \eta - \omega \wedge d\eta$ para las uno formas $\omega = x^2y^2z^2dy + x^3dz$ y $\eta = zdy + dz$.

5) Sea $\omega = dx \wedge dy \in \Omega(\mathbb{R}^2)$. Para $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ calcular $\omega(df(\frac{\partial}{\partial x}), df(\frac{\partial}{\partial y}))$ donde df es la aplicación tangente.

6) Sea $\omega_1, \omega_2, \ldots, \omega_n$ uno formas alternadas. Probar que son linealmente dependientes sobre $\mathbb{R}$ si y sólo si $\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_n = 0$. *Indicación:* Si fueran independientes, serían base de un subespacio de V^* y tendrían una base (bi-)dual en V.

7) Sea $\omega = \sum f_{i_1 i_2 \ldots i_k}\, dx^{i_1} \wedge dx^{i_2} \wedge \cdots \wedge dx^{i_k} \in \Omega^k(M)$ donde $i_1 < i_2 < \ldots < i_k$. Demostrar que la componente $j_1, j_2, \ldots, j_{k+1}$, con $j_1 < j_2 < \ldots < j_k$, del tensor $d\omega$ es

$$\sum_{s=1}^{k+1} (-1)^{s-1} \frac{\partial f_{j_1 \ldots \widehat{j_s} \ldots j_{k+1}}}{\partial x^{j_s}}$$

donde el circunflejo indica que se omite ese índice.

8) Probar que un tensor k veces covariante es una k-forma alternada si y sólo si se anula siempre que se aplique a k vectores linealmente dependientes.