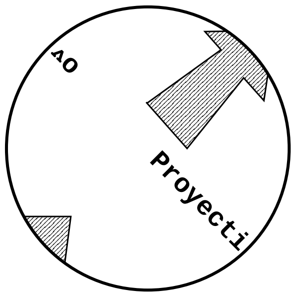
Capítulo 1

Bases de geometría diferencial

H. Whitney comenzó un famoso artículo [Whi36] de 1936 escribiendo: *una variedad diferenciable se define en general de dos formas: como un conjunto de puntos con entornos homeomorfos a un espacio euclídeo* $\mathbb{R}^n$*, estando las coordenadas en entornos no disjuntos relacionadas por transformaciones diferenciables; o como un subconjunto de* $\mathbb{R}^n$ *definido en cada punto expresando algunas coordenadas en términos de otras por medio de funciones diferenciables.* Dede entonces ha pasado mucho tiempo pero la herencia de esta dualidad tiene su reflejo en los currículos académicos actuales. Casi siempre hemos preferido ecuaciones pero la geometría de variedades, originada por Riemann y desarrollada a lo largo del siglo XX, pretende estudiar la geometría desde dentro, sin referencia a ningún espacio exterior y por ello ha hecho suya la primera de las definiciones.



Whitney obtuvo un resultado (véase el Teorema 1.1.5 más adelante) que asegura que las variedades siempre se pueden meter dentro de algún $\mathbb{R}^n$. En sus palabras: *la primera definición no es más general que la segunda.* Quizá al lector le parezca rebuscado entonces el empeño de estudiar siempre los objetos geométricos desde dentro. Si hay un espacio ambiente donde escribir ecuaciones ¿por qué no usarlo?

Pensemos por ejemplo en un universo dado por el disco unidad en el que los habitantes que salen por la frontera vuelven a aparecer en el punto diametralmente opuesto. Localmente ninguno de ellos notaría nada distinto a $\mathbb{R}^2$, tendrían que recorrer todo un diámetro para darse cuenta por ejemplo de que su universo es cerrado. La descripción geométrica de este objeto, que no es más que el *espacio proyectivo* $\mathbb{P}^2(\mathbb{R})$, es sencilla y permite desarrollar cierta intuición que perderíamos en gran medida trabajando con las ecuaciones de una subvariedad de $\mathbb{R}^4$ que es el primer $\mathbb{R}^n$ en el que se puede meter bien [dC92, p.32 Ex.5].

En resumidas cuentas, un espacio ambiente puede ser demasiado artificial para ciertos objetos. Por ejemplo, todos representamos un punto sobre la Tierra por su latitud y su longitud en vez de escribir las tres coordenadas en algún sistema de referencia del $\mathbb{R}^3$

circundante. En términos cosmológicos también esto es natural. Si el universo es por definición todo, ¿para qué imaginar algo que lo contiene?

Una dificultad en este enfoque es que estamos acostumbrados a ver las velocidades como vectores, flechas, que salen fuera de la variedad. Por ello las definiciones de espacio tangente y de las aplicaciones lineales y multilineales que se viven en él, requerirán una gran dosis de abstracción.

1.1. El concepto de variedad

La idea de variedad diferenciable n-dimensional es la de un objeto geométrico compuesto por parches que son similares a abiertos de $\mathbb{R}^n$.

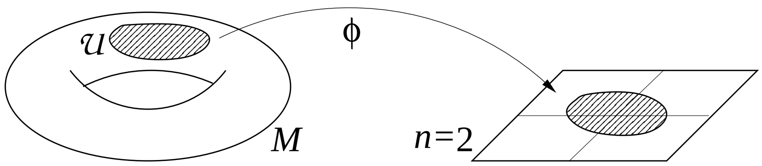
Para que este objeto capture parte de las ideas que tenemos en mente sobre continuidad y comportamiento local, postularemos que sea un espacio topológico con ciertas propiedades básicas.

Hipótesis topológica: Diremos que un espacio topológico M satisface la hipótesis topológica si tiene la propiedad de Hausdorff y admite una base numerable.

Sólo recordaremos de estas propiedades fundamentalmente técnicas que la *propiedad de Hausdorff* dice que dados dos puntos siempre hay entornos disjuntos de ellos. La existencia de una base numerable o *segundo axioma de numerabilidad* [Mun75] asegura que existe una familia numerable de abiertos tal que cualquier entorno de un punto contiene a un entorno que está en dicha familia. En el caso de $\mathbb{R}$ se podrían tomar por ejemplo todos los intervalos con extremos racionales.

En algunos textos esta hipótesis topológica no se incluye en la definición de variedad. Su propósito es poder llevar a cabo el análisis local típico de la geometría diferencial y se manifiesta en la existencia de las particiones de la unidad (Proposición 1.1.4), las cuales permiten modificar ligeramente una función en un entorno de un punto sin que los puntos fuera de él se enteren. Esto se opone a lo que ocurre en la geometría algebraica, donde las funciones son polinómicas o racionales y una modificación en las cercanías de un punto conlleva un cambio global de la función.

Una carta nos dice la manera de allanar un parche de M en $\mathbb{R}^n$. Los parches allanados son como los mapas de carreteras o los de un callejero, y querremos que cubran todo el territorio aunque cada uno de ellos sólo muestre una pequeña porción.



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.

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* y se dice que $\big(x^1(p), \dots, x^n(p)\big)$ son las coordenadas de p en la carta $(\mathcal{U}, \phi)$. El lector encontrará extraño el uso de superíndices pero cierto convenio relacionado con los tensores que veremos más adelante sugiere que es una notación apropiada. 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.

Por razones técnicas es conveniente pensar en todas las cartas n-dimensionales posibles compatibles entre sí (se habla de *atlas maximales*) 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.

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.

Nota avanzada: 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*. Una interpretación es 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 [Hir76] §2.10 para el enunciado preciso).

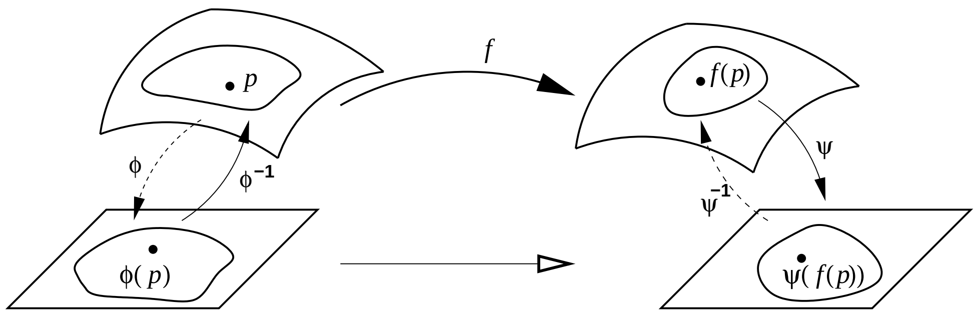
Las cartas permiten definir propiedades locales en la variedad bajando a $\mathbb{R}^n$. Por ejemplo, se dice que una función $f : M \longrightarrow \mathbb{R}$ es C^∞ si para cada carta $(\mathcal{U}, \phi)$ la función $f \circ \phi^{-1} : \phi(\mathcal{U}) \longrightarrow \mathbb{R}$ lo es y se define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

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

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.



Observación: 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^∞ o al menos tantas veces diferenciables como sea necesario para que no haya problemas de regularidad.

Un concepto técnicamente más complejo de generalizar es el espacio tangente, que en el caso de subvariedades de $\mathbb{R}^n$ tenía un significado sencillo (recuérdese el curso de Cálculo III). No es posible una mera adaptación directa 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 [Jän01]). 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* [O'N83], 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 la relación $\sim$ 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 conjunto de funciones $f : \mathcal{U}(p) \longrightarrow \mathbb{R}$ con $\mathcal{U}(p)$ un entorno 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 [Gon05] §3.2).

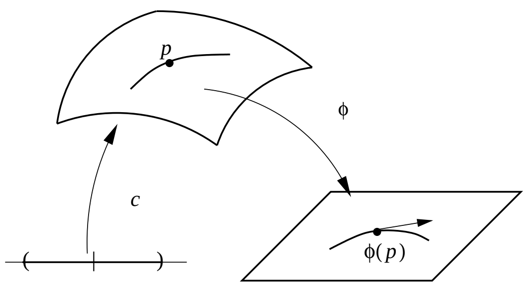
A partir de las curvas que corresponden a los ejes coordenados (una vez que bajamos a $\mathbb{R}^n$) se obtienen unos vectores tangentes que denotaremos con el extraño nombre $\frac{\partial}{\partial x^i}\big|_p$. Para ser rigurosos, si $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es la base canónica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

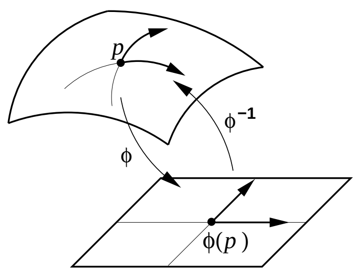
$$\frac{\partial}{\partial x^i}\Big|_p = [c_i] \qquad \text{donde} \qquad c_i(t) = \phi^{-1}(\phi(p) + t\vec{e}_i), \quad i = 1, 2, \dots, n.$$

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

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

A menudo la notación para estos vectores tangentes se abrevia como $\partial_i\big|_p$ o incluso como ∂_i si no se quiere indicar la carta o el punto.





Como se vio en cursos pasados:

Proposición 1.1.1 *El espacio tangente $T_p(M)$ tiene una estructura natural de espacio vectorial cuya dimensión es la de la variedad diferenciable M.*

Proposición 1.1.2 *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 vectores y multiplicarlos por números coordenada a coordenada como nos enseñaron en primero, escribiendo 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 el vector $v(\cdot)$ (considerado con la segunda definición) en $v(\cdot \circ f)$.

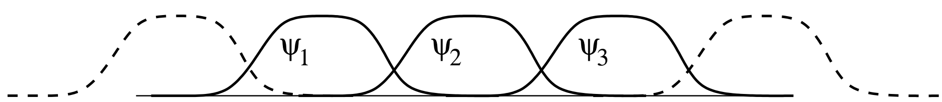
Ahora todo funciona como con la diferencial de toda la vida, siempre componiendo con las cartas.

Proposición 1.1.3 *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 un significado más restringido.

A pesar de nuestra insistencia en el análisis local, en geometría diferencial también hay operaciones globales que requieren pegar adecuadamente las cartas. Supongamos por ejemplo que deseamos definir la integral sobre una variedad. Sea cual sea su definición, si bajamos a $\mathbb{R}^n$ con cartas cuyos abiertos recubren la variedad, está claro que no basta con integrar allí y sumar los resultados. Un punto puede pertenecer a varios abiertos y en ese caso hay que ponderar su contribución para contarlo sólo una vez. Analíticamente, hay que partir la función idénticamente uno en funciones que vivan en los diferentes abiertos.

Definición: Una *partición de la unidad* subordinada a un recubrimiento abierto $\bigcup \mathcal{U}_\alpha = M$ es una colección de funciones $\{\psi_\alpha\}$, C^∞ y $0 \le \psi_\alpha \le 1$, tales que: 1) En cada punto $p \in M$ sólo hay un número finito de ellas con $\psi_\alpha(p) \neq 0$ y su suma es uno. 2) Cada ψ_α tiene soporte incluido en alguno de los $\mathcal{U}_\beta$.



La buena noticia acerca de las particiones de la unidad es su existencia. Este hecho está intrínsicamente ligado a la hipótesis topológica. Incluso hay una especie de equivalencia entre ambas [BC70, Prop.3.4.3], [Mun75].

Proposición 1.1.4 *Cualquier recubrimiento abierto de una variedad diferenciable M admite una partición de la unidad subordinada a él.*

Demostración: Digamos que el recubrimiento es $\{\mathcal{U}_\alpha\}$. La hipótesis topológica permite suponer, quizá considerando abiertos más pequeños incluidos en los $\mathcal{U}_\alpha$, que $\overline{\mathcal{U}}_\alpha$ (el cierre) es compacto, que para todo $p \in M$ el conjunto $\{\alpha \ : \ p \in \mathcal{U}_\alpha\}$ es finito y que existe un recubrimiento $\{\mathcal{V}_\alpha\}$ con $\overline{\mathcal{V}}_\alpha \subset \mathcal{U}_\alpha$ gozando de las mismas propiedades (en nuestro caso también se puede suponer que ambas colecciones son numerables). Todo esto es un hecho topológico general que responde a las palabras mágicas de que Hausdorff, localmente compacto y base numerable implican paracompacto y existencia de refinamientos[1].

Si $(\mathcal{U}_\alpha, \phi_\alpha)$ son cartas (siempre podemos trocear los $\mathcal{U}_\alpha$ para que estén incluidos en los dominios de las funciones de las cartas), la compacidad asegura que $\phi_\alpha(\mathcal{V}_\alpha) \subset \mathbb{R}^n$ se puede recubrir por un número finito de bolas, $B_{\alpha,1}, B_{\alpha,2}, \ldots, B_{\alpha,k(\alpha)} \subset \mathcal{U}_\alpha$. Sean las funciones $\eta_{\alpha,i} : \mathbb{R}^n \longrightarrow [0,1]$ que son idénticamente 1 en $B_{\alpha,k}$ y cero fuera de alguna bola concéntrica contenida en $\mathcal{U}_\alpha$. La prueba termina considerando

$$\psi_\alpha(p) = \frac{\mu_\alpha(p)}{\displaystyle\sum_{\beta \,:\, p\in\mathcal{U}_\beta} \mu_\beta(p)} \qquad \text{con} \qquad \mu_\alpha(p) = \begin{cases} \displaystyle\sum_{i=1}^{k(\alpha)} \eta_{\alpha,i}\big(\phi_\alpha(p)\big) & \text{si } p \in \mathcal{U}_\alpha \\ 0 & \text{si } p \in M - \mathcal{U}_\alpha \end{cases}$$

que está bien definida porque todo p pertence a algún $\mathcal{V}_\alpha$ lo que asegura que algún $\eta_{\alpha,i}\big(\phi_\alpha(p)\big)$ es no nulo. $\square$

Las particiones de la unidad constituyen el ingrediente principal para probar que las variedades compactas se pueden considerar subvariedades de algún $\mathbb{R}^n$. Enunciamos aquí el resultado general de Whitney aunque sólo daremos un esquema de la prueba en el caso compacto, siguiendo [Gua08], que es bastante más sencillo que el caso general [Hir76].

Teorema 1.1.5 (de Whitney) *Una variedad diferenciable de dimensión n admite una inmersión en $\mathbb{R}^{2n+1}$.*

Observación: La palabra *inmersión* traduce tanto a *imbedding* como a *embedding* que son distintos en la terminología anglosajona americana. El segundo concepto es más

[1] Esbozamos la prueba en el caso de variedades: Elegimos una colección numerable de cartas $\{(\mathcal{O}_i, \phi_i)\}_{i=1}^{\infty}$ con $\bigcup \mathcal{O}_i = M$ y $\overline{\mathcal{O}}_i$ es compacto (tómense por ejemplo preimágenes de bolas en $\mathbb{R}^n$ y empléese que hay una base numerable). De la misma forma escogemos un recubrimiento $\{\mathcal{U}'_i\}_{i=1}^{\infty}$ tal que $\overline{\mathcal{U}'_i}$ es compacto y está contenido en algún $\mathcal{U}_{\alpha_i}$. Para $k \ge 1$ sean el abierto $\mathcal{W}_k = \bigcup_{i=1}^{k+2} \mathcal{O}_i - \bigcup_{i=1}^{k-1} \overline{\mathcal{O}}_i$, el compacto $C_k = \bigcup_{i=1}^{k+1} \overline{\mathcal{O}}_i - \bigcup_{i=1}^{k} \mathcal{O}_i$ y otro abierto $\mathcal{W}'_k \supset C_k$ con $\overline{\mathcal{W}'_k} \subset \mathcal{W}_k$. Consideremos un subrecubrimiento finito de C_k por abiertos $\mathcal{U}'_j \cap \mathcal{W}'_k$. La colección $\{\mathcal{V}_j\}$ formada por todos estos abiertos (cuando varía k) cumple que $\{j \ : \ p \in \mathcal{V}_j\}$ es finito porque cada $p \in M$ está en algún $C_k \subset \mathcal{W}_k$ y $\mathcal{W}_i \cap \mathcal{W}_j = \emptyset$ si $i \ge j+3$. Para obtener el recubrimiento grande basta considerar $\mathcal{U}_{\alpha_j} \cap \mathcal{W}_k$ al recubrir C_k.

fuerte pues requiere que la aplicación sea homeomorfismo sobre su imagen. Es en ese sentido fuerte en el que se emplea aquí. Es decir, nuestra inmersión es una función cuya aplicación tangente es inyectiva (no pega direcciones distintas) y que tiene una inversa bien definida sobre su imagen (no pega puntos, ni los acerca infinitamente).

Demostración (Esbozo para el caso compacto): Por la compacidad, existe una colección finita de cartas $\{(\mathcal{O}_i, \phi_i)\}$ con $M = \bigcup \mathcal{O}_i$. Revisando el argumento topológico de la Proposición 1.1.4, vemos que da lugar a recubrimientos finitos $\{\mathcal{V}_i\}_{i=1}^N$ y $\{\mathcal{U}_i\}_{i=1}^N$. Entonces se puede suponer que la partición de la unidad es $\{\psi_i\}_{i=1}^N$ con $\operatorname{supp} \psi_i \subset \mathcal{U}_i$ y $(\mathcal{U}_i, \phi_i)$ cartas. Sea $\widetilde{\phi}_i(p) = \psi_i(p)\phi_i(p)$ si $p \in \mathcal{U}_i$ y cero en otro caso. Claramente $\widetilde{\phi}_i : M \longrightarrow \mathbb{R}^n$ es C^∞. Consideremos $f : M \longrightarrow \mathbb{R}^D$ con $D = (n+1)N$ definida por

$$f(p) = \big(\widetilde{\phi}_1(p), \widetilde{\phi}_2(p), \dots, \widetilde{\phi}_N(p), \psi_1(p), \psi_2(p), \dots, \psi_N(p)\big).$$

Dado $p \in M$ existe i tal que $\psi_i(p) \neq 0$ y si $f(p) = f(q)$ se tendría $\psi_i(p) = \psi_i(q) \neq 0$ y $\widetilde{\phi}_i(p) = \widetilde{\phi}_i(q)$ y la inyectividad de ϕ_i implica $p = q$. Por topología básica, una función inyectiva definida en un conjunto compacto es homeomorfismo sobre su imagen. Por otro lado, la matriz jacobiana $D\big(f \circ \phi_i^{-1}\big)$ tiene rango (máximo) n para todo $x \in \phi_i(\mathcal{U}_i)$.

En conclusión, tenemos una inmersión $f : M \longrightarrow \mathbb{R}^D$ pero D está relacionado con el número de cartas y parece muy difícil de controlar. Las líneas que restan esbozan cómo demostrar que cualquier $D > 2n+1$ se puede reducir componiendo f con la proyección P sobre el hiperplano $\{\vec{x} \in \mathbb{R}^n \ : \ \vec{x} \cdot \vec{v} = 0\}$ para algún vector unitario $\vec{v} \in S^{D-1}$.

Si $P \circ f$ dejara de ser inyectiva entonces $f(p)$ y $f(q)$ deberían determinar un vector proporcional a $\vec{v}$ para ciertos $p, q \in M$. Dicho de otra forma, $\vec{v}$ estaría en la imagen de la función

$$a : M \times M - \{(p,p) \ : \ p \in M\} \longrightarrow S^{D-1} \qquad \text{con } \ a(p,q) = \frac{f(q) - f(p)}{\|f(q) - f(p)\|}.$$

Por otro lado, si la jacobiana $D\big(P \circ f \circ \phi_i^{-1}\big)$ dejara de tener rango máximo entonces $\vec{v}$ estaría en la imagen de $D\big(f \circ \phi_i^{-1}\big)$ ya que $\vec{v}$ genera el núcleo de $D(P) = P$. En otras plabras, se tendría que $\vec{v}$ está en la imagen de

$$b_i : \phi_i(\mathcal{U}_i) \times S^{n-1} \longrightarrow S^{D-1} \qquad \text{con } \ b_i(x, \vec{w}) = \frac{\big(D\big(f \circ \phi_i^{-1}\big)(x)\big)(\vec{w})}{\|\big(D\big(f \circ \phi_i^{-1}\big)(x)\big)(\vec{w})\|}.$$

Ahora bien, la variedad origen de a tiene dimensión $2n$ y las de las b_i tienen dimensión $2n-1$. La unión de sus imágenes en S^{D-1} será, por tanto, a lo más $2n$ dimensional y si $D - 1 > 2n$ siempre hay posibilidad de escoger $\vec{v}$ fuera de este conjunto, garantizando que la proyección reduce D.

Este argumento final con dimensiones necesita alguna justificación. Basta una versión sencilla del llamado *Lema de Sard* [Gua08, Cor.2.10], [Spi82]. □

1.2. Vectores, covectores y tensores

En primer lugar vamos a ver algunas definiciones que no tiene que ver directamente con la geometría de variedad sino que constituyen una extensión natural del álgebra lineal estudiada en cursos anteriores.

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) $f(\vec{v_1}, \dots, \lambda\vec{v_i}, \dots, \vec{v_n}) = \lambda f(\vec{v_1}, \dots, \vec{v_i}, \dots, \vec{v_n}) \quad \text{con } \lambda \in \mathbb{R}$

b) $f(\vec{v_1}, \dots, \vec{v_i} + \vec{v_i}', \dots, \vec{v_n}) = f(\vec{v_1}, \dots, \vec{v_i}, \dots, \vec{v_n}) + f(\vec{v_1}, \dots, \vec{v_i}', \dots, \vec{v_n}).$

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 $T(\widetilde{\varphi}, \vec{v_1}, \vec{v_2}, \dots, \vec{v_n}) = \widetilde{\varphi}\big(f(\vec{v_1}, \vec{v_2}, \dots, \vec{v_n})\big)$ 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 aceptación 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 $T: V^* \times \overset{r \text{ veces}}{\cdots} \times V^* \times V \times \overset{s \text{ veces}}{\cdots} \times V \longrightarrow \mathbb{R}$.

[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.

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.

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

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

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

<u>Ejemplo</u>: 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.

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.

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

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

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. 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 $a^1, a^2, \dots, a^m$

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

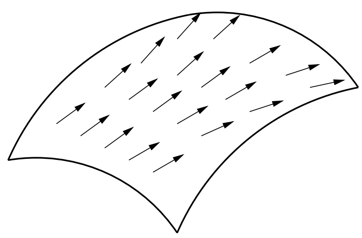
[3] Este convenio fue realmente introducido por Einstein quien bromeó al respecto diciendo: “He hecho un gran descubrimiento en Matemáticas; he suprimido el signo de sumación toda vez que la suma se haga en un índice que aparece dos veces”.

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.

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*. 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^∞.

Para definir los tensores en variedades hay que introducir los duales de los vectores. No hace falta proceder en abstracto, nos podemos apoyar en conceptos anteriores:

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, $\{dx^1|_p, dx^2|_p, \dots, dx^n|_p\}$ es la base dual de $\big\{\frac{\partial}{\partial x^1}\big|_p, \frac{\partial}{\partial x^2}\big|_p, \dots \frac{\partial}{\partial x^n}\big|_p\big\}$.

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*.

Si por ejemplo en $dx^1|_p$ no fijamos p, en cuyo caso escribiremos habitualmente dx^1, ya tenemos un campo de uno formas definido en un abierto de la variedad. Como las uno formas son operadores lineales que actúan sobre los vectores, lo que tenemos es un campo de tensores $(0,1)$. No hay razón para posponer más la definición 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^∞.

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

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

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 1.2.1 *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 1.1.3. 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. $\square$

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

$$\boxed{T'^{i_1i_2\dots i_r}_{j_1j_2\dots j_s}=\big(\frac{\partial x'^{i_1}}{\partial x^{k_1}}\cdot\frac{\partial x'^{i_2}}{\partial x^{k_2}}\cdot\ldots\cdot\frac{\partial x'^{i_r}}{\partial x^{k_r}}\big)\cdot\big(\frac{\partial x^{l_1}}{\partial x'^{j_1}}\cdot\frac{\partial x^{l_2}}{\partial x'^{j_2}}\cdot\ldots\cdot\frac{\partial x^{l_s}}{\partial x'^{j_s}}\big)T^{k_1k_2\dots k_r}_{l_1l_2\dots l_s}}\tag{1.2}$$

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).

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 nombre "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[4] en sentido estricto.

1.3. 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

[4]R. P. Feynman, premio Nobel de Física, dedica toda la sección 2-5 de su magnífico libro [FLS64] 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.

de forma diferencial tal como ahora lo conocemos y además definió una nueva operación, la derivada exterior, que resulta fundamental para escribir sintéticamente algunos resultados de geometría diferencial, en especial los relacionados con la integración, como el teorema de Stokes. La estructura algebraica subyacente en la que se basó se llama álgebra exterior y fue introducida por H. Grassmann con anterioridad.

Las definiciones de partida son muy simples.

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 imponen 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 anterior para referirnos a los covectores. 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.

<u>Ejemplo</u>: Hallar todos los tensores $T \in \mathrm{Alt}^2(\mathbb{R}^3)$.

Consideremos dos vectores genéricos $\vec{v}_j = a^i_j \vec{e}_i$, $j=1,2$. El desarrollo de $T(\vec{v}_1,\vec{v}_2)$ da lugar a nueve términos. Olvidándonos de los tres con coeficientes nulos $T_{ii} = T(\vec{e}_i,\vec{e}_i)$ y agrupando $T_{ij} = T(\vec{e}_i,\vec{e}_j)$ con su negativo $T_{ji} = T(\vec{e}_j,\vec{e}_i)$, se llega a

$$T(\vec{v}_1,\vec{v}_2) = \begin{vmatrix} a^2_1 & a^3_1 \\ a^2_2 & a^3_2 \end{vmatrix} T_{23} + \begin{vmatrix} a^3_1 & a^1_1 \\ a^3_2 & a^1_2 \end{vmatrix} T_{31} + \begin{vmatrix} a^1_1 & a^2_1 \\ a^1_2 & a^2_2 \end{vmatrix} T_{12}.$$

Cada elección de las constantes T_{23}, T_{31} y T_{12} da lugar a un elemento de $\mathrm{Alt}^2(\mathbb{R}^3)$ distinto.

Acabamos de ver que las 2-formas alternadas en $\mathbb{R}^3$ no son más que combinaciones lineales de determinantes formados a partir de las coordenadas de los vectores sobre los que se aplican. Esto no es un caso aislado. En cierto modo, las k-formas alternadas y diferenciales no son más que determinantes escritos con una notación conveniente.

Veamos primero esa notación y comprobaremos después que los determinantes cubren todas nuestras necesidades.

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

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

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

Proposición 1.3.1 *Sea $\{\widetilde{\varphi}^1, \widetilde{\varphi}^2, \dots, \widetilde{\varphi}^n\}$ una base de V^*, entonces*

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

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

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

Si T fuera el tensor nulo para ciertos coeficientes, calculando $T(\vec{e}_{i_1}, \vec{e}_{i_2}, \dots, \vec{e}_{i_n})$ se tendría $a_{i_1 i_2 \ldots i_k} = 0$, 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 $\omega_{i_1 i_2 \ldots i_k}$ entonces eligiendo $a_{i_1 i_2 \ldots i_k} = \omega_{i_1 i_2 \ldots i_k}$ se tiene que T y ω tienen las mismas componentes $i_1 i_2 \ldots i_k$ siempre que $i_1 < i_2 < \ldots < i_k$. En el resto de los casos también deben coincidir por la antisimetría de las formas alternadas al intercambiar dos argumentos. $\Box$

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

Corolario 1.3.2 *Cualquier elemento de $\Omega^k(M)$ se puede escribir como*

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

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

$$\big(\widetilde{\varphi}^{i_1} \wedge \widetilde{\varphi}^{i_2} \wedge \cdots \wedge \widetilde{\varphi}^{i_k}\big) \wedge \big(\widetilde{\varphi}^{j_1} \wedge \widetilde{\varphi}^{j_2} \wedge \cdots \wedge \widetilde{\varphi}^{j_l}\big) = \widetilde{\varphi}^{i_1} \wedge \widetilde{\varphi}^{i_2} \wedge \cdots \wedge \widetilde{\varphi}^{i_k} \wedge \widetilde{\varphi}^{j_1} \wedge \widetilde{\varphi}^{j_2} \wedge \cdots \wedge \widetilde{\varphi}^{j_l}.$$

Con ello y la Proposición 1.3.1 o el Corolario 1.3.2 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 [Spi82]. Nuestra definición no está dentro de la estética libre de coordenadas que prefiere la geometría actual 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)

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

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

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?[5]

[5] 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 Grassmann 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 abandonó las Matemáticas y se dedicó a la Lingüística.

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

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

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

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

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

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 1.3.3 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 [BG80]. Se deja al lector interesado que indague esta línea o que trate por sí mismo de probar que la definición de $d\omega$ no depende de la carta empleada. El punto crucial es que las derivadas parciales segundas cruzadas desaparecen por la anticonmutatividad del producto exterior.

Veamos dos propiedades importantes que, como hemos apuntado, determinan la derivada exterior. La primera muestra la relación entre los definiciones de producto exterior y derivada exterior y la segunda es crucial en topología diferencial [Cha08, Cap.2].

Proposición 1.3.3 *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 1.3.2, 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[6]. □

Para terminar introduciremos, sin profundizar, el concepto de integral en variedades. Incluso si nos restringimos a $\mathbb{R}^n$, una definición invariante requiere objetos que se multipliquen por el determinante jacobiano al cambiar de coordenadas. Las formas diferenciales satisfacen este requerimiento y su razón de ser original es que constituyen los objetos matemáticos que se pueden integrar. Nótese que si

$$\omega = g\, dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n \in \Omega(\mathbb{R}^n) \qquad \text{con } g = g(x^1, x^2, \dots, x^n)$$

y hacemos el cambio $x^j = f^j(y^1, y^2, \dots, y^n)$, $1 \le j \le n$, entonces $dx^j = \frac{\partial f^j}{\partial y^i} dy^i$ y sustituyendo y utilizando las propiedades de los determinantes

$$\omega = \det(Jf)\; g \circ f\; dy^1 \wedge dy^2 \wedge \cdots \wedge dy^n.$$

Entonces tiene sentido definir la integral de $\omega \in \Omega(\mathbb{R}^n)$, digamos de soporte compacto[7] si queremos evitar problemas con la existencia, como

$$\int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} g. \tag{1.4}$$

Realmente hay un pequeño problema con el signo, sobre el que volveremos más adelante, para que esta definición sea invariante del todo por cambios de coordenadas.

Antes de seguir, generalicemos los cambios de variable incluso cuando f no es un difeomorfismo.

[6] 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 conflictivo porque $f \wedge dx^{i_1} = f\, dx^{i_1}$.

[7] Cuando hablamos del soporte de $\omega = g\, dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n$ o de su no anulación en un punto nos referimos a los conceptos análogos para g. Esto es coherente porque los cambios de carta multiplican g por una función no nula.

Definición: Sean M y N variedades y $f : M \longrightarrow N$. Se llama *imagen recíproca* (o *pullback*) a la aplicación lineal $f^* : \Omega^k(N) \longrightarrow \Omega^k(M)$ dada por

$$f^*\omega(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) = \omega\big(df(\vec{v}_1), df(\vec{v}_2), \dots, df(\vec{v}_k)\big) \qquad \text{para } \vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in T_p(M).$$

Nota: Esta definición se extiende sin cambios a cualquier tensor de tipo $(0, s)$.

Si $\vec{v}_j \in T_p(M)$ entonces $df(\vec{v}_j) \in T_{f(p)}(N)$, por tanto si ω está soportada en un entorno de un punto, $f^*\omega$ lo está en entornos de sus preimágenes for f. En coordenadas todo está más claro,

$$\omega = g_{i_1 i_2 \dots i_n} dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_n} \;\Rightarrow\; f^*\omega = g_{i_1 i_2 \dots i_n} \circ f \; d(x^{i_1} \circ f) \wedge d(x^{i_2} \circ f) \wedge \dots \wedge d(x^{i_n} \circ f).$$

Si f es un difeomorfismo de $\mathbb{R}^n$ con $\det(Jf) > 0$, la fórmula de cambio de variable es

$$\int_{\mathbb{R}^n} \omega = \int_{\mathbb{R}^n} f^*\omega. \tag{1.5}$$

Ahora podemos utilizar las cartas para definir la integral de una forma en un parche de una variedad. Concretamente, si $(\mathcal{U}, \phi)$ es una carta de una variedad m-dimensional, M, y $\omega \in \Omega^n(M)$, digamos de soporte compacto incluido en $\mathcal{U}$ para que no haya problemas al integrar, entonces se define

$$\int_{\mathcal{U}} \omega = \int_{\phi(\mathcal{U})} (\phi^{-1})^*\omega. \tag{1.6}$$

El segundo miembro, es una integral como la de (1.4). Por (1.5) con $f = \phi \circ \widetilde{\phi}^{-1}$, al cambiar la carta $(\mathcal{U}, \phi)$ por otra $(\mathcal{V}, \widetilde{\phi})$ el resultado es el mismo si $\det(Jf) > 0$ (suponiendo que el soporte de ω está en $\mathcal{U} \cap \mathcal{V}$). Excluir la posibilidad $\det(Jf) < 0$, lleva a considerar sólo la "mitad" de las cartas. Una elección de un subconjunto (digamos maximal) de cartas tales que los cambios $\phi \circ \widetilde{\phi}^{-1}$ tengan siempre jacobiano positivo, se dice que define una *orientación*. En $\mathbb{R}^2$, por ejemplo, las cartas $\big(\mathbb{R}^2, (x, y)\big)$ y $\big(\mathbb{R}^2, (y, x)\big)$ no pertenecen a la misma orientación y darían signos distintos al definir la integral. En variedades conexas hay dos posibles orientaciones.

La extensión de la definición de integral de un abierto (1.6) a toda una variedad se lleva a cabo de la manera natural empleando las particiones de la unidad [Spi82].

Un resultado central sobre integración en variedades es el teorema de Stokes que establece una bella dualidad entre la derivada exterior y la frontera topológica. Requiere introducir el concepto de variedad con borde. Por ejemplo, el disco cerrado unidad D en $\mathbb{R}^2$ no es una variedad porque un punto en la frontera $\partial D = S^1$ no tiene entornos homeomorfos a un abierto de $\mathbb{R}^2$, sino al semiespacio $\mathbb{R} \times [0, \infty)$. En general, se dice que M es una *variedad con borde* de dimensión n, si tiene puntos con entornos homeomorfos a un semiespacio $\mathbb{R}^{n-1} \times [0, \infty)$ que conforman una variedad ∂M de dimensión $n - 1$, mientras que el resto de los puntos, $M-\partial M$, satisfacen la definición habitual de variedad de dimensión n. Geométricamente ∂M es el borde de M porque el semiespacio establece un plano $x^n = 0$ que no se puede rebasar.

Al estar ∂M en el "borde" de $M - \partial M$, una orientación en $M - \partial M$ condiciona una orientación en ∂M llamada *orientación inducida*. Las definiciones rigurosas de estos conceptos se pueden consultar por ejemplo en [Cha08].

Teorema 1.3.4 (Teorema de Stokes) *Sea M una variedad n-dimensional orientable con borde y $\omega \in \Omega^{n-1}(M)$ con soporte compacto, entonces*

$$\int_{\partial M} i^*\omega = \int_M d\omega$$

donde $i : \partial M \longrightarrow M$ es la inclusión y ∂M tiene la orientación inducida por la de M.

Con el consiguiente abuso de notación, a veces se escribe $\int_{\partial M} \omega$ en el primer miembro en lugar de $\int_{\partial M} i^*\omega$, porque la única manera de integrar algo n-dimensional en un conjunto $n-1$ dimensional es usar la relación entre las variables y eso es justamente lo que indica $i^*\omega$.

Los caso $n = 1, 2$ y 3 del teorema de Stokes dan todas las versiones del teorema fundamental del cálculo estudiadas en cursos anteriores. Por ejemplo, si $n = 2$ y $\omega = P\,dx + Q\,dy$ se tiene (suprimiendo la inclusión, como en esos cursos)

$$\int_{\partial M} P\,dx + Q\,dy = \int_M \Big(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\Big)\,dx \wedge dy$$

que es el *teorema de Green*.