

## 1.4 Noción de espacio tangente

*Javier Finat*

### Índice

|        |                                                               |    |
|--------|---------------------------------------------------------------|----|
| 0.1.   | Notación e Introducción . . . . .                             | 3  |
| 1.     | <b>Vectores tangentes</b> . . . . .                           | 4  |
| 1.1.   | Curvas tangentes . . . . .                                    | 4  |
| 1.1.1. | Vector velocidad de una curva . . . . .                       | 4  |
| 1.1.2. | Definición conjuntista de Espacio Tangente . . . . .          | 6  |
| 1.1.3. | Pegado de curvas con tangentes compatibles . . . . .          | 6  |
| 1.1.4. | Reinterpretación usando operadores . . . . .                  | 7  |
| 1.2.   | Gérmenes de aplicaciones . . . . .                            | 8  |
| 1.2.1. | Definición . . . . .                                          | 8  |
| 1.2.2. | Proposición . . . . .                                         | 8  |
| 1.2.3. | Comportamiento funtorial . . . . .                            | 8  |
| 1.2.4. | Ejemplo . . . . .                                             | 8  |
| 1.2.5. | Una nota sobre la aproximación analítica local . . . . .      | 9  |
| 1.3.   | El módulo de las derivaciones . . . . .                       | 9  |
| 1.3.1. | Equivalencia entre descripciones . . . . .                    | 10 |
| 1.3.2. | Trivialidad local para el espacio tangente . . . . .          | 11 |
| 1.3.3. | El espacio tangente a la circunferencia . . . . .             | 11 |
| 1.3.4. | El espacio tangente al toro . . . . .                         | 12 |
| 2.     | <b>Vector Tangente a una Aplicación en un Punto</b> . . . . . | 15 |
| 2.1.   | Aplicación Tangente como derivación . . . . .                 | 15 |
| 2.1.1. | La derivación como funcional . . . . .                        | 15 |
| 2.1.2. | Expresión Local de la Aplicación Tangente . . . . .           | 16 |
| 2.1.3. | Campo vectorial gradiente . . . . .                           | 17 |
| 2.1.4. | Versión coordenada local . . . . .                            | 17 |
| 2.1.5. | Espacio cotangente. Definición . . . . .                      | 18 |
| 2.2.   | Dualidad entre derivaciones y diferenciales . . . . .         | 18 |
| 2.2.1. | El caso cartesiano . . . . .                                  | 18 |
| 2.2.2. | Formas no-degeneradas . . . . .                               | 19 |
| 2.3.   | Calculando espacios tangentes . . . . .                       | 19 |
| 2.3.1. | Representaciones implícitas . . . . .                         | 20 |
| 2.3.2. | Nota sobre la extensión al caso singular . . . . .            | 20 |
| 2.3.3. | Representaciones paramétricas . . . . .                       | 21 |
| 2.3.4. | Representaciones simbólicas . . . . .                         | 23 |
| 2.4.   | Práctica: Gradiente complejo . . . . .                        | 26 |
| 2.4.1. | Caso de una variable compleja . . . . .                       | 26 |

---

|        |                                                                        |    |
|--------|------------------------------------------------------------------------|----|
| 2.4.2. | Operadores diferenciales estándar . . . . .                            | 26 |
| 2.4.3. | Funciones algebraicas y superficies de Riemann . . . . .               | 27 |
| 3.     | <b>Algunas aplicaciones del campo gradiente</b> . . . . .              | 28 |
| 3.1.   | Elementos de Optimización Multiobjetivo ó Vectorial . . . . .          | 29 |
| 3.1.1. | Conjuntos convexos . . . . .                                           | 29 |
| 3.1.2. | Generalidades sobre Optimización . . . . .                             | 30 |
| 3.1.3. | Optimización de Pareto . . . . .                                       | 30 |
| 3.2.   | Aproximaciones a la Optimización Multiobjetivo . . . . .               | 31 |
| 3.2.1. | Método de los pesos objetivos . . . . .                                | 31 |
| 3.2.2. | Elementos para la resolución . . . . .                                 | 32 |
| 3.3.   | Aprendizaje en sistemas expertos . . . . .                             | 33 |
| 3.3.1. | Aprendizaje supervisado . . . . .                                      | 33 |
| 3.3.2. | Modelos de back-propagation . . . . .                                  | 34 |
| 3.3.3. | Modelos asociativos (Hopfield) . . . . .                               | 35 |
| 3.3.4. | Aprendizaje no supervisado: Modelos auto-organizativos . . . . .       | 36 |
| 3.4.   | Campos de Potencial en Ingeniería . . . . .                            | 36 |
| 3.4.1. | Guiado de Robots mediante campos de potencial . . . . .                | 37 |
| 3.4.2. | Elementos de Procesamiento de Imagen . . . . .                         | 37 |
| 4.     | <b>Ecuaciones diferenciales ordinarias de orden superior</b> . . . . . | 38 |
| 4.1.   | Notación . . . . .                                                     | 38 |
| 4.2.   | Construcción local de vectores tangentes de orden superior . . . . .   | 38 |
| 4.3.   | Un ejemplo elemental: El Péndulo Simple . . . . .                      | 39 |
| 4.3.1. | Oscilaciones forzadas . . . . .                                        | 40 |
| 4.3.2. | Acoplamiento de dos sectores . . . . .                                 | 40 |
| 4.3.3. | Acoplamiento entre tres sectores . . . . .                             | 40 |

## 0.1. Notación e Introducción

En todo lo que sigue las variedades y aplicaciones son siempre de clase  $C^r$  para  $r \geq 2$ , salvo que expresamente se indique lo contrario. Al principio de este apartado, se denota mediante  $X$  (en lugar de  $M$ ) a la variedad con la que se trabaja, para insistir en que la mayor parte de los resultados presentados son extrapolables al caso de variedades algebraicas y también al de variedades analíticas ó algebraicas.

El *objetivo fundamental* de este apartado es construir una aproximación lineal (de primer orden) a variedades  $X$  de clase  $C^r$  en cada punto  $x \in X$  dada por el espacio tangente  $T_x X$ ; asimismo se muestra cómo extender esta noción a espacios de funciones. En términos conjuntistas el espacio tangente  $T_x X$  se describe como el conjunto de vectores tangentes a curvas que presentan un contacto de orden  $\geq 2$  en el punto. Hay que desarrollar una noción intrínseca de vector tangente, es decir, que no dependa del sistema coordenado elegido.

Para ello, es necesario incorporar un formalismo algebraico local basado en la noción de *germen* para evitar problemas de ambigüedad. De una forma intuitiva, un germen de una función en un punto  $p$  es una clase de equivalencia que se obtiene para funciones que tienen “el mismo” comportamiento en un punto; cuando esta idea se extiende a derivadas de orden  $k \geq 1$  se obtiene la noción de (germen de)  $k$ -jet asociada a los polinomios de Taylor. Así, p.e. el 1-jet corresponde a (los gérmenes de) funciones que tienen la “misma componente lineal” en el punto  $p$ ; desde el punto de vista geométrico corresponde a las curvas que pasan por el punto  $p$  y tienen la misma tangente en el punto (contacto de orden 2). Esta noción se extiende a polinomios de Taylor de orden  $k$  (truncaciones del desarrollo de Taylor) con la interpretación correspondiente al contacto de orden  $k$  en el punto  $p$  de las curvas que pasan por  $p$ <sup>1</sup>

La primera sección del capítulo trata sobre las diferentes nociones de vector tangente. La segunda extiende esta noción a vector tangente a una Aplicación en un Punto que es la extensión natural de la derivada direccional, pero aplicada ahora a espacios de funciones. Las secciones siguientes son posibles tópicos a desarrollar como trabajo individual y tratan sobre algunas aplicaciones del campo gradiente, EDO de orden superior o elementos de la dinámica clásica newtoniana. La elección del trabajo a realizar depende de las preferencias sobre Topología, Análisis o Mecánica Clásica de las personas matriculadas.

---

<sup>1</sup> En el caso geométrico, nótese que no es necesario que las curvas estén contenidas en la variedad; basta con que tengan el contacto requerido.

## 1. Vectores tangentes

En esta sección se presentan las versiones coordenada e intrínseca de vector tangente. El carácter intrínseco está vinculado a la independencia con respecto al sistema coordenado elegido o, si se prefiere, al abierto del espacio topológico (que es un abierto de trivialización en el caso de variedades). para ello se introduce la siguiente relación de equivalencia:

Supongamos que  $p \in X$  es un punto de un espacio topológico  $X$  y denotemos mediante  $C^r(X, \mathbb{K})$  al conjunto de funciones de clase  $C^r$  de  $X$  a valores en un cuerpo  $\mathbb{K}$ . Deseamos comparar  $f, g \in C^r(X, \mathbb{K})$  de manera independiente con respecto al sistema coordenado en  $p \in X$  que están definidas localmente en los abiertos  $U, V$  con  $p \in U \cap V$ . Para ello, se introduce la relación de equivalencia siguiente:

$$f \sim_p g \Rightarrow \exists W \subset U \cap V \mid f|_W = g|_W$$

*Definición.*- A la clase de equivalencia por dicha relación se le llama el *germen* de  $f$  en  $p \in X$  y se le representa mediante  $[f]$ .

A menudo y para no complicar la notación se denota al germen  $[f]$  de la misma forma que a su representante, pasando de funciones a gérmenes y recíprocamente, dependiendo del contexto. Esta noción se extiende de forma natural a las derivadas de cualquier orden para dichas funciones.

Sea  $(U, \phi)$  una carta con coordenadas locales  $x^1, \dots, x^n$ , correspondientes a las componentes de la  $C^r$ -equivalencia  $\phi : U \rightarrow \mathbb{R}^n$ . Sea  $\gamma : I \rightarrow U$  una curva de clase  $C^r$  (con  $r \geq 1$ ) que pasa por  $\underline{p}_0 = \gamma(0) \in U$ . Para un pequeño entorno de  $\phi(\underline{p}_0)$  se tiene un desarrollo de la forma

$$(x^i \circ \gamma)(t) = x^i(\gamma(0)) + \frac{d}{dt} \Big|_{t=0} (x^i \circ \gamma)t + o(t^2) .$$

### 1.1. Curvas tangentes

Empezamos motivando las definiciones formales con una presentación geométrica de la noción de vector tangente en términos cinemáticos. A continuación se desarrolla una versión intrínseca

#### 1.1.1. Vector velocidad de una curva

Llamamos *vector velocidad de la curva*  $\gamma$  en el punto  $\underline{p}_0$  al vector dado por

$$v^i := \frac{d}{dt} \Big|_{t=0} (x^i \circ \gamma) =: \dot{x}^i \Big|_{t=0} , \quad 1 \leq i \leq n .$$

*Definición.*- Diremos que dos curvas  $\gamma_1$  y  $\gamma_2$  que pasan por  $\underline{p}_0$  son *tangentes* en  $\underline{p}_0$  si la distancia  $d(\gamma_1(t), \gamma_2(t)) = o(t)$  cuando  $t \rightarrow 0$ .

Esta noción de tangencia equivale a que las dos curvas tengan el mismo vector velocidad en el punto  $\underline{p}_0$  (comprobarlo como ejercicio) y es obviamente una relación de equivalencia.

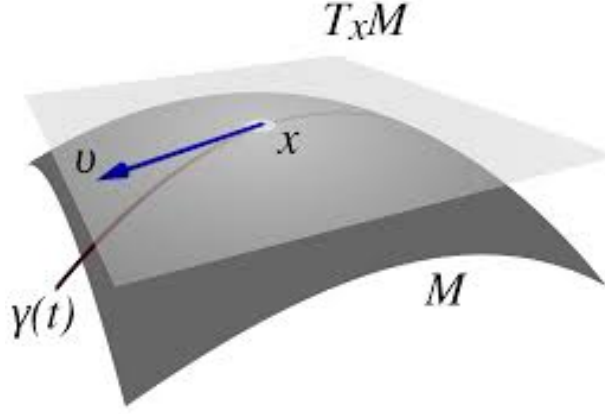


Fig. 1: Vector tangente a una curva

Por ello, *la condición de tangencia no depende de la carta elegida* (usar la regla que da la diferenciación de una función compuesta y razonar de manera similar a como se ha mostrado en el §1,2,3). Esto justifica que, en ocasiones, no representemos la carta coordenada sobre la que estamos trabajando, cuando hablamos de "vectores tangentes".

Diremos que un *vector*  $\mathbf{v}$  es *tangente* a  $M$  en  $p$  si es el vector velocidad de una curva  $\gamma : I \rightarrow M$  que pasa por  $p$ . En este sentido, el vector tangente es una clase de equivalencia de órbitas cuyo representante es el germen del vector velocidad asociado a la función en un punto.

La descripción intrínseca de las condiciones de tangencia en términos de vectores tangentes a curvas contenidas en la variedad no es gratuita, pues puede no resultar práctico utilizar coordenadas locales ó bien podemos necesitar la noción de tangencia sobre espacios de dimensión infinita no numerable (para localizar extremales asociados a funcionales de tipo integral en problemas variacionales ó de optimización, por ejemplo).

No obstante, la expresión explícita del vector tangente depende del sistema coordenado elegido. Si  $\underline{p}_0 \in U \cap V$ , entonces la diferencial  $d(\psi \circ \phi^{-1})$  del cambio de carta expresa la transformación del vector tangente en  $\underline{p}_0$  en términos de la base correspondiente al nuevo sistema de coordenadas en  $(\psi, V)$ . De una manera más explícita, si denotamos mediante  $x^1, \dots, x^n$  y mediante  $y^1, \dots, y^n$  a coordenadas locales en  $\phi(U)$  y  $\psi(V)$ , respectivamente, entonces dicha diferencial está representada por la matriz jacobiana del cambio de base que escribiremos abreviadamente como

$$\left( \frac{\partial y^i}{\partial x^j} \right)_{1 \leq i, j \leq n}.$$

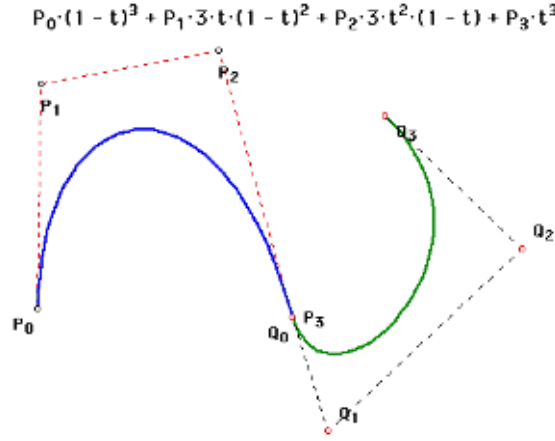


Fig. 2: Pegado de curvas con la misma tangente (Bézier)

### 1.1.2. Definición conjuntista de Espacio Tangente

Dada una variedad  $X$ , llamamos *espacio tangente en*  $\mathbf{p} \in X$  y lo denotamos por  $T_{\mathbf{p}}X$  al conjunto de vectores que son tangentes a  $X$  en  $\mathbf{p}$ . Dicho conjunto de vectores se puede expresar de forma local en términos de *derivadas parciales* como

$$\xi := \sum_{i=1}^{i=n} \xi^i \frac{\partial}{\partial x^i}$$

donde  $(x^1, \dots, x^n)$  es un sistema coordenado centrado en  $\mathbf{p}$ , ó más geométricamente como el conjunto de vectores tangentes en  $\mathbf{p}$  a las curvas  $\gamma : I \rightarrow X$  que pasan por  $\mathbf{p} \in X$ , es decir, tales que  $\gamma(0) = \mathbf{p}$ .

### 1.1.3. Pegado de curvas con tangentes compatibles

En aplicaciones de diseño CAD/CAM es frecuente utilizar curvas de Bézier que imponen restricciones en puntos de control para curvas de grado bajo que tengan la misma tangente. Este tópico tiene interés para modelado de objetos planares y se extiende al caso de superficies en 3D usando B-splines que pueden ser descritas en términos de “producto tensorial” de curvas de Bézier o de snakes. En la fig.2 se muestra un ejemplo particularmente simple basado en curvas de grado  $\leq 3$

La generación automática de snakes es apropiada para reconocimiento de siluetas o cuestiones de animación basadas en la variación de puntos de control.

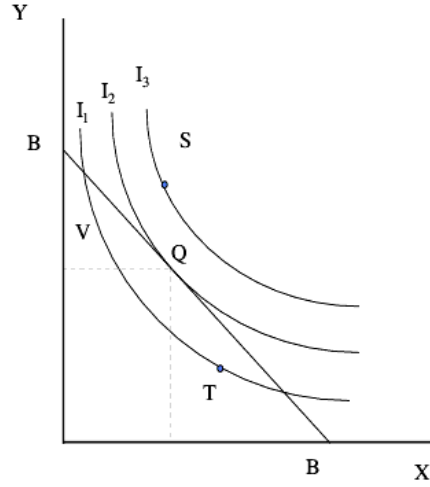


Fig. 3: Preferencia Óptima del consumidor sobre curvas de indiferencia

#### 1.1.4. Reinterpretación usando operadores

La descripción conjuntista de vectores tangentes se puede reinterpretar de una manera funcional en términos de operadores. Esta versión es la que se muestra en las aplicaciones, donde la diferencia para el valor de una variable entre dos estados se representa mediante un operador vectorial que asociamos a la "tasa de variación" de la variable. Esta tasa de variación se representa matemáticamente como la *derivada direccional* correspondiente a la variable seleccionada cuyo comportamiento espacio-temporal deseamos evaluar.

Una aplicación básica de esta idea aparece en relación con la formulación en términos de *estática comparativa* de las preferencias de un consumidor asociadas a una *restricción presupuestaria*  $r$  en Microeconomía. La elección entre dos bienes en función de los precios relativos se puede representar mediante una colección de funciones (visualizables como hipérbolas). La restricción presupuestaria de cada consumidor se representa mediante una recta  $r$ .

El punto en el que esta recta es tangente a las "curvas de indiferencia"  $I_j$  determina la asignación óptima de recursos por parte del consumidor. El desplazamiento de la recta asociada a la restricción presupuestaria o la variación en los precios relativos de los bienes a consumir da lugar a desplazamientos en la localización del equilibrio (ver Fig.3). En este caso, los operadores pueden actuar sobre las rectas o bien sobre las curvas.

Este modelo se extiende a un número mayor de bienes, a preferencias no-convexas (no unicidad de equilibrios), a modelos macroeconómicos o de finanzas o también a la propagación de shocks y la localización de equilibrios dinámicos en términos de hiperplanos o, con más generalidad, hipersuperficies tangentes.

## 1.2. Gérmenes de aplicaciones

Dado el conjunto de las aplicaciones locales  $f : U \rightarrow Y$  (donde  $U$  es un abierto de  $X$  que contiene a  $p_0$ ), introducimos la relación de equivalencia siguiente:

$$(f, U) \sim (g, V) \Leftrightarrow \exists W \subset U \cap V \text{ con } p_0 \in W \mid f|_W = g|_W$$

### 1.2.1. Definición

A la clase de equivalencia por esta relación se le llama el *germen de la aplicación  $f$  en el punto  $x \in X$* .

Al conjunto de gérmenes de aplicaciones en un punto  $x \in X$  se le denota mediante  $\mathcal{O}_{X,x}$  (para el caso analítico) ó mediante  $\mathcal{E}_{M,p}$  para el caso diferenciable. Ambos conjuntos tiene estructura de álgebra real con respecto a las operaciones suma y multiplicación definidas mediante la suma y la multiplicación de sus representantes (comprobarlo como ejercicio).

### 1.2.2. Proposición

El conjunto de gérmenes de funciones  $f : X \rightarrow \mathbb{R}$  en  $x \in X$  tiene estructura de anillo local, con ideal maximal  $\mathcal{M}$  igual al conjunto de funciones que se anulan en  $x \in X$  (ver *Ejercicio 1.1.7.6*). Los elementos inversibles ó unidades son los no pertenecientes al ideal maximal.

### 1.2.3. Comportamiento funtorial

Una aplicación  $F : X \rightarrow Y$  entre  $C^r$ -variedades induce un homomorfismo entre álgebras (análogamente para los anillos) locales en  $x$  e  $y = f(x)$  dado por la composición:

$$F^* : \mathcal{O}_{Y,y} \rightarrow \mathcal{O}_{X,x} \mid F^*(\varphi) := \varphi \circ F$$

verificándose que  $Id^* = Id$  y  $(G \circ F)^* = F^* \circ G^*$  (carácter contravariante para el funtor). En particular, si  $F$  representa un germen inversible con respecto a la composición (localmente biyectivo), entonces

$$F \circ F^{-1} = Id \Rightarrow (F \circ F^{-1})^* = Id \Rightarrow (F^*)^{-1} = (F^{-1})^*$$

### 1.2.4. Ejemplo

El caso más simple corresponde a gérmenes de funciones definidas sobre  $\mathbb{R}^n$ , es decir  $X = \mathbb{R}^n$  e  $Y = \mathbb{R}$ ; el anillo local  $\mathcal{O}_{X,x}$  está dado por el conjunto de las funciones definidas sobre  $\mathbb{R}^n$  cuyo ideal maximal corresponde a las funciones que se anulan en un punto  $x$  (que podemos tomar como el origen). Denotamos mediante  $\mathcal{O}_n$  a este anillo local.



Usando una carta local  $(\varphi, U)$  la derivación bajo el signo integral permite reducirnos al caso anterior. En efecto, para un pequeño entorno de  $x \in X$  tenemos

$$f(\underline{x}) - f(\underline{0}) = \int_0^1 \frac{d}{dt} f(tx^1, \dots, tx^n) dt = \sum_{i=1}^n x^i \int_0^1 D_i f(tx^1, \dots, tx^n) dt$$

por lo que en la categoría  $C^r$  (para  $r \geq 2$ ) podemos escribir

$$f(\underline{x}) = f(\underline{0}) + \sum_{i=1}^n x^i f_i(\underline{x}) \quad \text{donde} \quad f_i(\underline{x}) = \int_0^1 D_i f(t\underline{x}) dt$$

La descripción del germen de una aplicación  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$  en términos de sus componentes  $(f_1, \dots, f_p)$  con  $f_j = f_j(x^1, \dots, x^n)$  para  $j = 1, \dots, p$ , dota al conjunto de gérmenes de aplicaciones  $F : (\mathbb{R}^n, x) \rightarrow (\mathbb{R}^p, f(x))$  de la estructura de  $\mathcal{O}_n$ -módulo.

### 1.2.5. Una nota sobre la aproximación analítica local

La reformulación en términos de anillos ó de álgebras locales asociadas a  $C^r$ -variedades permite estudiarlas mediante las propiedades de espacios de funciones definidas sobre variedades, incluso en el caso singular. La extensión del estudio al caso singular no es posible si utilizamos la aproximación tradicional de la Geometría Diferencial de Variedades.

El procedimiento de “pegado” de datos locales para obtener  $C^r$ -variedades se extiende de forma natural al “pegado” de estructuras definidas por los anillos, módulos ó álgebras locales asociadas a los anillos locales  $A$  de funciones regulares  $\mathcal{O}_{X,x}$  ó los módulos  $M$  construidos sobre dichos anillos. El caso más simple corresponde a la representación de los gérmenes de aplicaciones  $f \in C^r(n, p)$  en un punto como el módulo dado por la suma de  $p$  copias del anillo  $\mathcal{O}_{X,x}$ .

Obviamente, además de los anillos (o módulos) de funciones regulares sobre un punto, se pueden considerar otros objetos obtenidos a partir de las operaciones básicas (producto tensorial, p.e.) definidos sobre dichos anillos que es necesario “pegar” usando condiciones de compatibilidad similares a las del caso regular. Los objetos iniciales que verifican las propiedades naturales de compatibilidad reciben el nombre de “prehaces”. Una reformulación de la Geometría Diferencial en estos términos se puede ver en Warner <sup>2</sup>.

## 1.3. El módulo de las derivaciones

Una *derivación*  $\xi$  es una aplicación lineal  $\xi : \mathcal{O}_{X,x} \rightarrow \mathbb{R}$  (análogamente para el caso diferenciable  $\mathcal{E}_{M,p}$ ) que satisface la *regla de Leibnitz*, es decir,

<sup>2</sup> Una reformulación de la Geometría Algebraica y Analítica se puede ver en Grothendieck y Dieudonné: *Éléments de Géométrie Algébrique, I*, Springer-Verlag, 1974

$$\xi(f + g) = (\xi f)g + f(\xi g) \quad \forall f, g \in \mathcal{O}_{X,x}$$

La expresión anterior debe ser entendida de forma simbólica, es decir, cada una de las funciones se evalúa en un punto  $x \in X$  (en  $p \in M$  para el caso diferenciable). Si trabajamos en la categoría  $C^r$  para  $r < \infty$ , hay que tener presente que una derivación llevará funciones de clase  $C^r$  en funciones de clase  $C^{r-1}$ .

Con esta notación, el espacio tangente se expresa algebraicamente como

$$T_x X := \text{Der}_{\mathbf{k}}(\mathcal{O}_{X,x}) = \text{Hom}_{\mathbf{k}}(\mathcal{O}_{X,x}, \mathbf{k}) \simeq \mathcal{O}_{X,x} \left\langle \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right\rangle ,$$

donde el último isomorfismo es como  $\mathbf{k}$ -álgebras. En particular, si  $X = \mathbb{R}^n$  y  $\mathbf{k} = \mathbb{R}$ , entonces el espacio tangente en cualquier punto es isomorfo al  $\mathcal{O}_{X,x}$ -módulo de las derivaciones generado sobre el anillo local  $\mathcal{O}_{X,x}$  por las derivadas parciales

$$\left\langle \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^n} \right\rangle ,$$

donde

$$\frac{\partial}{\partial x^i} : \mathcal{O}_{X,x} \rightarrow \mathbb{R} \quad | \quad f \mapsto \frac{\partial}{\partial x^i} f(0)$$

es decir, cualquier derivación sobre  $\mathbb{R}^n$  se escribe en coordenadas locales como un operador diferencial en derivadas parciales de primer orden (y recíprocamente):

$$\sum_{i=1}^n f^i \frac{\partial}{\partial x^i} \quad \text{donde} \quad f^i \in C^r(\mathbb{R}^n, \mathbb{R}) , \quad \forall i = 1, \dots, n .$$

Es bien conocido del Análisis que el vector tangente a cualquier curva  $\gamma : I \rightarrow \mathbb{R}^n$  se puede expresar en términos de las coordenadas locales de  $\mathbb{R}^n$  usando la fórmula precedente y la regla de la cadena. Recíprocamente, la integración de un operador diferencial de la forma anterior (con condiciones iniciales prefijadas) permite obtener una curva cuyo vector tangente satisface al operador. Por tratarse de una propiedad *local*, esta situación se extiende obviamente al caso más general de variedades:

### 1.3.1. Equivalencia entre descripciones

*Proposición.*- Las descripciones precedentes de espacio tangente para una variedad  $X$  son equivalentes

*Ejercicio (avanzado).*- Demuestra la proposición precedente (Indicación: Ver Brocker y Janich: *Introducción a la Topología Diferencial*, Ed. AC, 1977, para detalles).

*Nota.*- La mayor parte del curso está dedicada a técnicas para el caso liso, por lo que marginaremos las propiedades algebraicas que son sin embargo cruciales para la Geometría Analítica y la Geometría Algebraica. Las condiciones

de lisitud y la dimensión finita de las variedades con que trabajamos permiten una aproximación más intuitiva basada en coordenadas locales.

### 1.3.2. Trivialidad local para el espacio tangente

Como la descripción del espacio tangente es *local* (en virtud del Teorema de la Función Implícita), si  $x^1, \dots, x^m$  son coordenadas locales para un abierto coordinado  $U$  de  $p \in M$  suficientemente pequeño, existe una  $C^r$ -equivalencia local que permite describir localmente el espacio tangente  $TU$  al abierto coordinado  $U$  como un producto cartesiano  $U \times T_p M$ .

En particular, si  $\pi : TU \rightarrow U$  denota la proyección sobre la primera componente, entonces la  $C^r$ -equivalencia precedente se expresa mediante  $TU \simeq U \times T_p M$  y la restricción de esta  $C^r$ -equivalencia a la fibra  $\pi^{-1}(p')$  sobre cualquier punto  $p' \in U \subset M$ , induce un isomorfismo

$$\pi^{-1}(p') \simeq \pi^{-1}(p) \simeq \mathbf{K} \left\langle \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m} \right\rangle \simeq \mathbf{K}^m.$$

Por ello, identificaremos usualmente al espacio tangente de una variedad  $m$ -dimensional  $M$  con el espacio cartesiano  $\mathbb{K}^m$ , aunque esta identificación no es canónica (depende de la carta elegida). La identificación es canónica cuando la variedad es el propio espacio producto  $\mathbb{K}^m$ . Para fijar ideas, restringimos la atención al caso  $\mathbb{K} = \mathbb{R}$ , aunque los argumentos algebraicos se extienden a cualquier otro cuerpo.

Si se parte de un espacio de Banach  $B$  en lugar de un espacio cartesiano, el razonamiento es similar, con una identificación canónica de la fibra con el propio espacio  $B$ ; en el §2, se muestra que el fibrado tangente de un espacio de Banach (en el caso geométrico, un espacio cartesiano con la métrica euclídea ó la hermitica ordinaria) es “trivial”, es decir, es una variedad producto de la base con la fibra genérica. Nótese que una variedad de Banach  $\mathcal{B}$  construida sobre un espacio de Banach  $B$  (ver capítulo 3) es, en general, no trivial, lo cual extiende al caso infinito-dimensional los argumentos presentados para el caso geométrico. Esta extensión no es gratuita, pues responde al estudio de soluciones de ecuaciones funcionales definidas sobre espacios de funciones.

### 1.3.3. El espacio tangente a la circunferencia

Denotemos mediante  $\mathbb{S}^1 := \{p = (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$  a la circunferencia de radio unidad centrada en el origen  $O$ .

El espacio tangente  $T_p \mathbb{S}^1$  está dado por la recta normal a  $\mathbb{S}^1$  al radio vector  $\overline{Op}$ . Dicho subespacio en  $p = (x, y)$  está generado por el vector  $(-y, x)$  que varia de forma  $C^r$  con respecto al punto base y es siempre no-nulo; nótese que  $(x, y) \cdot (-y, x)^T = 0$  lo cual permite reinterpretar al vector libre  $(x, y)$  como el vector normal  $\mathbf{n}_p \mathbb{S}^1$  a  $\mathbb{S}^1$  en  $p$ .

Por ello, el “conjunto” de los vectores tangentes  $\mathbf{v} = t_P \mathbb{S}^1$  a la circunferencia se puede expresar globalmente como el producto

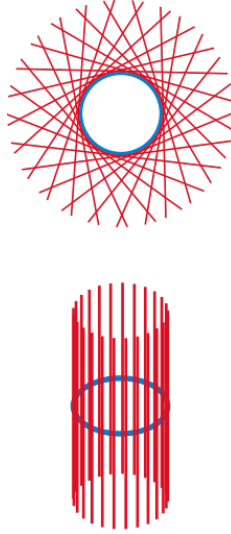


Fig. 4: Rectas tangentes a una circunferencia y representación espacial global

$$\mathbb{S}^1 \times \mathbb{R}^1 = \{(p, \mathbf{v}) \in \mathbb{S}^1 \times \mathbb{R} \mid \mathbf{v} = t_p \mathbb{S}^1\}$$

La representación polar  $e^{i\theta}$  proporciona una parametrización de la circunferencia  $\mathbb{S}^1$  por el ángulo polar  $\theta = \arctan(y/x)$ . En el lenguaje de los campos vectoriales el vector tangente está representado por la derivación  $\partial/\partial\theta$ .

En este caso, como la fibra es idéntica para todos los puntos (aunque la orientación cambie), el “pegado” de los datos locales correspondientes a las cartas que recubren  $\mathbb{S}^1$  se lleva a cabo utilizando las mismas cartas que las correspondientes a la estructura de  $\mathbb{S}^1$  como variedad (en este caso bastan dos cartas). Al resultado de “pegar” las rectas tangentes se le etiqueta como “fibrado tangente” (más adelante se da una definición más formal)

De una forma sintética y global se expresa esta condición diciendo que el fibrado tangente a la circunferencia  $\mathbb{S}^1$ , obtenido como producto cartesiano global  $\mathbb{S}^1 \times \mathbb{R}^1$  es el fibrado trivial sobre  $\mathbb{S}^1$  al que se denota mediante

$$\tau_{\mathbb{S}^1} = \mathbb{S}^1 \times \mathbb{R}^1 = \varepsilon_{\mathbb{S}^1}^1$$

donde  $\tau_{\mathbb{S}^1}$  denota el fibrado tangente a  $\mathbb{S}^1$ .

#### 1.3.4. El espacio tangente al toro

Denotemos mediante  $\mathbb{T}^m := \mathbb{S}^1 \times \dots \times \mathbb{S}^1$  al toro  $m$ -dimensional producto de  $m$  copias de  $\mathbb{S}^1$ . Cada punto  $p \in \mathbb{T}^m$  está parametrizado por  $(\theta_1, \dots, \theta_m)$ . Como el toro es un producto cartesiano, no existe ninguna relación algebraica

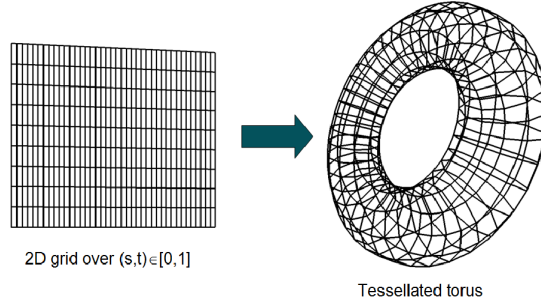


Fig. 5: Trivialización del Fibrado tangente al toro 2D

entre las coordenadas. Por ello, el vector tangente en cada punto es una suma formal de vectores l.i. que expresamos de forma más compacta como

$$\mathbf{v} = (\mathbf{v}_1, \dots, \mathbf{v}_m) \in T_p \mathbb{T}^m \quad \text{donde} \quad \mathbf{v}_i \in T_{\theta_i} \mathbb{S}^1 \quad 1 \leq i \leq m$$

Como cada vector tangente  $\mathbf{v}_i \in T_{\theta_i} \mathbb{S}^1$  se puede representar mediante una derivación, el conjunto de vectores tangentes es un espacio vectorial generado por las derivaciones  $\partial/\partial\theta_i$  para  $1 \leq i \leq m$ . El argumento mostrado en el apartado anterior se extiende de forma inmediata a

$$\tau_{\mathbb{T}^m} = \mathbb{T}^m \times \mathbb{R}^m = \prod_{i=1}^m (\mathbb{S}^1 \times \mathbb{R}^1) \simeq \oplus^m \varepsilon_{\mathbb{S}^1}^1$$

donde  $\tau_{\mathbb{T}^m}$  denota el fibrado tangente a  $\mathbb{T}^m$ . En este caso, las derivaciones definidas sobre el toro son combinaciones lineales formales de  $\partial/\partial\theta_i$  que expresamos en términos de coordenadas polares como

$$\sum_{i=1}^m a_i \frac{\partial}{\partial\theta_i} \quad \text{donde} \quad a_i \in C^r(\mathbb{T}^m, \mathbb{R})$$

con la ventaja en este caso de que las derivaciones son globales, es decir,  $\partial/\partial\theta_i$  se extiende a la copia  $i$ -ésima de  $\mathbb{S}^1$ .

*Ejercicio.*- Cualquier junta esférica de un robot descompone en producto de juntas rotacionales (parametrizadas por la circunferencia  $\mathbb{S}^1$ ). Por ello, el espacio de configuraciones  $\mathcal{C}$  de un robot que no tenga juntas traslacionales se puede reducir al acoplamiento de un número finito de  $m$  juntas rotacionales; el acoplamiento se lleva a cabo en términos de sistemas de barras que suponemos de longitud fija. Describe la cinemática del robot en términos de la parametrización de un toro  $m$ -dimensional.

*Observaciones.*- 1) La reducción de una junta esférica a juntas rotacionales es una simplificación matemática que puede resultar abusiva para algunas aplicaciones; en particular, no incluye la presencia de juntas traslacionales de las que disponen un gran número de robots. En ocasiones es necesario trabajar

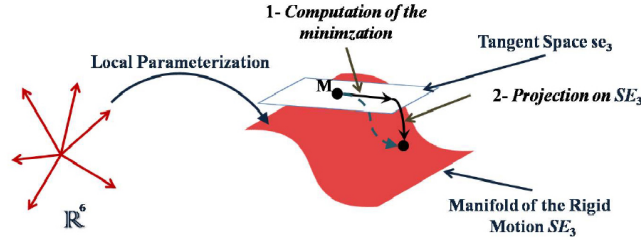


Fig. 6: Geodésicas en el grupo Euclídeo

directamente con transformaciones euclídeas en el espacio ordinario. El grupo de transformaciones euclídeas  $SE(3)$  es un grupo 6-dimensional definido como el producto semidirecto del grupo de rotaciones por el grupo de traslaciones que es una variedad no-lineal.

En este caso, las trayectorias óptimas (geodésicas) a realizar por el efector final del robots deben describirse directamente sobre la variedad, aunque pueden ser aproximadas por su discretización dada sobre copias “suficientemente próximas” del espacio tangente. El control para la ejecución de tareas por parte de los mecanismos del robot se lleva a cabo sobre el espacio tangente en puntos próximos (ver Fig.5).

Algunos robots tienen componentes que no son reducibles a sistemas de barras y juntas rotacionales; un ejemplo típico está dado por plataformas de Stewart que cuentan con dos “platos” conectados entre sí mediante un sistema de barras extensibles (juntas prismáticas) que no están contenidas en ninguno de los planos correspondientes a dichas plataformas con juntas esféricas en los extremos. Estos robots se utilizan p.e. para simuladores de vuelo.

La cinemática ocupa un lugar intermedio entre la Geometría y la Dinámica del Robot. En la cinemática no se considera el origen de las fuerzas ó momentos asociados a los mecanismos. En particular, la cinemática prescinde de efectos fundamentales correspondientes a fenómenos inerciales y los mecanismos de anticipación o compensación que permiten mejorar el rendimiento de los dispositivos robóticos. Estos aspectos requieren herramientas más avanzadas de dinámica que son una extensión natural de la mecánica newtoniana (ver final de este capítulo) y su reformulación lagrangiana (que se desarrolla a partir del capítulo 5 del módulo 2).

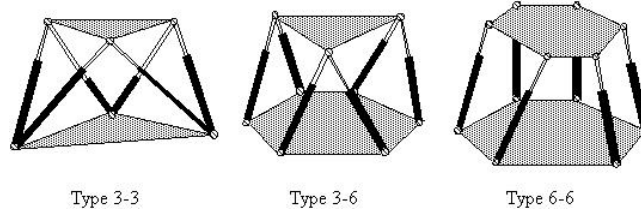


Fig. 7: Plataforma de Stewart simétricas

## 2. Vector Tangente a una Aplicación en un Punto

*Notación.* - En toda esta sección  $M$  denotará la  $C^r$ -variedad correspondiente al espacio de partida  $N$  ó de llegada  $P$  para una  $C^r$ -aplicación  $f : N \rightarrow P$ .

### 2.1. Aplicación Tangente como derivación

La diferencial  $d_{\underline{p}}f$  de la aplicación  $f$  en  $\underline{p} \in N$  está dada por una asignación que a cada vector tangente  $\xi_{\underline{p}} \in T_{\underline{p}}N$  le lleva en un vector  $\eta_{f(\underline{p})} \in T_{f(\underline{p})}P$ . En esta subsección se interpreta y se formaliza esta idea.

#### 2.1.1. La derivación como funcional

Interpretamos ahora dichos vectores como la evaluación de las derivaciones que actúan como operadores sobre el espacio de funciones definido en cada uno de los puntos. Cada uno de los vectores está determinado por el valor que toma el operador dado por la derivación en cada punto del espacio de funciones. Por ello, para definir la aplicación tangente, debemos mostrar cuál es la imagen de cada función  $g \in C^{r-1}(P, \mathbb{R})$ . De una manera más formal, escribiremos esta asignación mediante

$$(d_{\underline{p}}f)(\xi)(g) := \xi(g \circ f) =: \xi(f^*(g)) ,$$

donde  $f^*(g) := g \circ f$  representa la *imagen recíproca de  $g$  via  $f$* , obtenida por composición. De una forma más simbólica, si hacemos  $f_* := d_{\underline{p}}f$ , la expresión anterior se escribe también en forma algebraica como

$$(f_*\xi)(g) = \xi(f^*g) .$$

Esta formulación es especialmente apropiada para cuestiones relacionadas con el comportamiento funtorial de las derivaciones y las diferenciales. La definición que acabamos de mostrar es intrínseca, por lo que es fácilmente generalizable a espacios de funciones.

La equivalencia formulada en la Proposición precedente permite traducir esta formulación a los términos locales que son más familiares por el uso que de ellos se hace en Análisis Matemático.

### 2.1.2. Expresión Local de la Aplicación Tangente

Dada una aplicación  $f : N \rightarrow P$  entre  $C^r$ -variedades, los atlas  $\mathcal{A}_N = \{(U_i, \phi_{U_i})\}_{i \in I}$  y  $\mathcal{A}_P = \{(V_j, \phi_{V_j})\}_{j \in J}$ , permiten compatibilizar los datos relativos a la construcción de los espacios tangentes para cada punto. Para simplificar la notación, identifiquemos al punto  $\underline{p}$  con sus coordenadas locales  $\underline{x}$  y hagamos  $\underline{y} = f(\underline{x}) \in V \subset P$  la imagen de  $\underline{x} \in U \subset N$  via  $f$ .

Como  $\phi_U$  y  $\psi_V$  son  $C^r$ -equivalencias sobre sus imágenes respectivas, se tiene que la restricción de sus diferenciales a las fibras inducen isomorfismos entre los espacios tangentes. Por ello, tiene sentido construir la diferencial de  $f$  relativa a las cartas  $(U, \phi_U)$  y  $(V, \psi_V)$  en las  $C^r$ -variedades de partida  $N$  y llegada  $P$  como la restricción a  $(d_x \phi_U)(T_x U)$  de la composición  $d\psi_V \circ df \circ d\phi_U^{-1}$ .

Si  $x^1, \dots, x^n$  e  $y^1, \dots, y^p$  denotan coordenadas locales en  $\phi(U)$  y  $\psi(V)$ , respectivamente, entonces la regla de la cadena usual en Análisis de Varias Variables, permite expresar localmente el valor que toma en  $\xi$  la diferencial  $d_{\underline{p}}f$  de  $f$  en  $\underline{p}$  como

$$(d_{\underline{p}}f)(\xi) = \sum_{j=1}^p \xi(y^j \circ f) |_{\underline{p}} \frac{\partial}{\partial y^j}.$$

En particular, las imágenes via  $d_{\underline{p}}f$  de los elementos  $\xi^i = \partial/\partial x^i$  de la base de  $T_{\underline{p}}N$  están dadas como

$$f_*\left(\frac{\partial}{\partial x^i}\right) |_{\underline{p}} = \sum_{j=1}^p \frac{\partial}{\partial x^i}(y^j \circ f) |_{\underline{p}} \frac{\partial}{\partial y^j} = \sum_{j=1}^p \frac{\partial f^j}{\partial x^i}(\underline{p}) \frac{\partial}{\partial y^j} \quad 1 \leq i \leq n$$

Usando nuevamente la regla de la cadena para la intersección  $U_\alpha \cap U_\beta$  de los dominios de dos cartas coordenadas  $(U_\alpha, \phi_\alpha)$  y  $(U_\beta, \phi_\beta)$ , obtenemos que

$$(d_{\underline{p}}f |_{U_\alpha})(\xi_\alpha) = (d_{\underline{p}}f |_{U_\beta})(\xi_\beta)$$

(comprobarlo), por lo que *la diferencial de  $f$  en  $\underline{p} \in N$  no depende de la carta elegida*. Una vez fijados sistemas coordenados locales en los espacios de partida y de llegada, en el lenguaje del Análisis Clásico la diferencial está representada localmente por la *matriz jacobiana*. En este sentido, es frecuente encontrar en Física ó Ingeniería descripciones antiguas de vectores tangentes como aquellos que se transforman de acuerdo con reglas de transformación dadas por matrices jacobianas. El origen de este enfoque se justifica por el carácter intrínseco de los tensores que representan los vectores tangentes (es decir las derivaciones) ó sus duales (vectores cotangentes ó diferenciales); más adelante, volveremos sobre ello.



Esta propiedad permite trasladar las propiedades usuales del Análisis de Varias Variables al caso de Variedades (en particular, su caracterización algebraica como una aplicación lineal que satisface la regla de Leibnitz).

### 2.1.3. Campo vectorial gradiente

Si  $f : M \rightarrow \mathbb{R}$  es una función de clase  $C^r$  (para  $r \geq 1$ ) definida sobre una  $C^r$ -variedad  $M$ , entonces  $(d_{\underline{p}}f)(\xi) = \xi(f)(\underline{p})$  para cualquier  $\xi \in T_{\underline{p}}M$  (comprobarlo como ejercicio). Por ello,

*Lema.*- Cualquier campo vectorial en un punto tiene asociada una aplicación lineal, es decir, se puede representar como un *operador lineal* dado sobre el conjunto  $C_{\underline{p}}^r(N, \mathbb{R})$  de las aplicaciones definidas en un punto  $\underline{p} \in N$ . El operador lineal dado por un campo vectorial no depende del sistema coordenado elegido.

*Ejercicio.*- Comprueba que

- La expresión habitual de  $df$  es local, pero es una noción intrínseca.
- Si dos funciones  $f \in C^r(U, \mathbb{R})$ ,  $g \in C^r(V, \mathbb{R})$  coinciden sobre un abierto  $W \subseteq U \cap V$ , entonces el valor de cualquier campo sobre ellas es el mismo.
- El conjunto de campos (considerados como operadores lineales) con las operaciones habituales (suma y producto por un escalar) tiene una estructura de espacio vectorial al que se llama *espacio tangente en un punto* (Indicación: utiliza las propiedades algebraicas descritas en §1.4.3)

### 2.1.4. Versión coordenada local

Denotemos mediante  $(\xi^1, \dots, \xi^n)$  a las componentes de  $\xi$  en  $T_{\underline{p}}N$ , y tomamos coordenadas locales  $(x^1, \dots, x^n)$  en una carta  $(U, \phi)$  con  $\underline{p} \in U \subset N$ . Entonces,

$$(d_{\underline{p}}f)(\xi) = \sum_{i=1}^n \frac{\partial f}{\partial x^i} \xi^i = \text{grad}_{\underline{p}}(f) \cdot \underline{\xi},$$

donde

$$\text{grad}_{\underline{p}}(f) := \left( \frac{\partial f}{\partial x^1}, \dots, \frac{\partial f}{\partial x^n} \right)$$

denota el *gradiente de  $f$* . El gradiente es un operador vectorial que actúa sobre cada vector  $\underline{\xi} = (\xi^1, \dots, \xi^n)$  mediante  $f \mapsto \left( \frac{\partial}{\partial x^i} \right)_{1 \leq i \leq n}(f)$ . Es inmediato verificar que  $\text{grad}(f)$  es una aplicación lineal sobre  $T_{\underline{p}}N$ . Por ello, el gradiente así construido define un elemento del *espacio vectorial dual*  $T_{\underline{p}}^*N$  del espacio tangente en cada punto.

### 2.1.5. Espacio cotangente. Definición

Dada una  $C^r$ -variedad  $M$ , al espacio  $T_p^*M := \text{Hom}(T_pM, \mathbb{R})$  dual de  $T_pM$  se le llama el *espacio cotangente* ó también *espacio de las fases* (en Física, Ingeniería, Economía, etc). En particular, si  $f = x_j$  tenemos que

$$(d_p f)\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial f}{\partial x^i}(p) \Rightarrow (d_p x^j)\left(\frac{\partial}{\partial x^i}\right) = \delta_i^j \quad 1 \leq i, j \leq n$$

lo cual proporciona una versión coordenada local para la dualidad entre los espacios tangente y cotangente en  $p \in M$

*Nota.-* En las aplicaciones practicas, casi nunca se conoce cuáles son las leyes del movimiento o las características de las fuerzas que actúan sobre los sistemas. En este caso, sólo podemos llegar a conocer los campos de forma indirecta, es decir, midiendo el efecto que producen sobre el sistema; esa medida numérica en cada punto es precisamente el valor que proporciona la diferencial. Por consiguiente, el conjunto de dichas medidas sobre una colección de tasas de variación linealmente independientes genera el espacio cotangente que, por dualidad, nos proporciona el espacio tangente.

## 2.2. Dualidad entre derivaciones y diferenciales

El operador delta de Kronecker  $dx_i(\partial/\partial x_j) = \delta_{ij}$  da lugar a una dualidad formal entre los módulos de derivaciones y diferenciales sobre cualquier anillo base  $A$ ; estamos especialmente interesados en

- anillos de  $C^r$ -funciones regulares que denotamos mediante  $\mathcal{E}_M$  (caso diferenciable) ó  $\mathcal{O}_X$  (caso algebraico  $k[\underline{x}]$  ó analítico  $k[[\underline{x}]]$ );
- anillos coordenados de variedades del tipo que denotamos mediante  $\mathcal{E}_M/\mathcal{I}$  (caso diferenciable) ó  $\mathcal{O}_X/\mathcal{I}$  (caso algebraico  $k[\underline{x}]/I$  ó analítico  $k[[\underline{x}]]/I$  donde  $I$  es un ideal);
- anillos locales que denotamos mediante  $\mathcal{E}_M/\mathcal{M}$  (caso diferenciable) ó  $\mathcal{O}_X/\mathcal{M}$  (caso algebraico ó analítico) siendo  $\mathcal{M}$  el ideal maximal  $(\underline{x} - \underline{x}_0)$  en el punto  $\underline{p}_0 \in M$  ó  $\underline{x}_0 \in X$

### 2.2.1. El caso cartesiano

En virtud de la igualdad  $d_p x^j(\frac{\partial}{\partial x^i}) = \delta_i^j$ , las diferenciales  $dx^j$  para  $j = 1, \dots, m$ , desempeñan un papel dual al que desempeñan las derivaciones ordinarias  $\frac{\partial}{\partial x^i}$ . Esta igualdad se expresa globalmente como una dualidad

$$\mathbb{R}^m \simeq T_p M \simeq \mathbb{R} \left\langle \frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m} \right\rangle \rightarrow \mathbb{R} \langle dx_1 \dots dx_m \rangle \simeq T_p^* M \simeq \mathbb{R}^m$$

En ocasiones (abuso de notación) se denota mediante el mismo símbolo a un vector de un espacio y a su dual. Así, por ejemplo el gradiente está definido

por la expresión  $(d_{\underline{p}}f)(\xi) = \xi(f)(\underline{p})$  por lo que es un elemento del espacio cotangente  $T_{\underline{p}}^*M$ .

Sin embargo, lo más frecuente es referirse a él en términos coordenados locales adoptando la expresión

$$(d_{\underline{p}}f)(\xi) = \text{grad}_{\underline{p}}(f) \cdot \underline{\xi}$$

donde  $\text{grad}_{\underline{p}}(f)$  es la evaluación del campo vectorial  $(\frac{\partial}{\partial x^1}, \dots, \frac{\partial}{\partial x^m})$  sobre  $f$ , es decir, como un elemento de  $T_{\underline{p}}M$ . En este último caso, se sobrentiende que estamos tomando el vector dual y consideramos el gradiente como un operador diferencial  $C^r(M, \mathbb{R}) \rightarrow C^{r-1}(M, \mathbb{R})$ . Esta diversidad puede dar lugar a alguna confusión que se salva diciendo que el contexto es el que determina una acepción u otra.

### 2.2.2. Formas no-degeneradas

El isomorfismo  $T_{\underline{p}}M \simeq T_{\underline{p}}^*M := \text{Hom}_{\mathbb{R}}(T_{\underline{p}}M, \mathbb{R})$  equivale a la existencia de una aplicación bilineal no-degenerada

$$T_{\underline{p}}M \times T_{\underline{p}}M \rightarrow \mathbb{R}$$

Una métrica  $g$  sobre una variedad  $M$  está dada por una aplicación bilineal simétrica no-degenerada  $g_{ij}(\underline{p})$  que depende de forma  $C^\infty$  del punto  $\underline{p} \in M$ . La métrica  $g$  así definida induce un isomorfismo  $T_{\underline{p}}M \simeq T_{\underline{p}}^*M$  entre los espacios tangente y cotangente que depende del punto base elegido  $\underline{p} \in M$ , es decir, la métrica  $g$  no es constante. En particular, la dualidad entre derivaciones y diferenciales asociada a

$$d_{\underline{p}}x^j(\frac{\partial}{\partial x^i}) = \delta_i^j$$

sólo es “natural” para la métrica euclídea. Así, para una variedad  $M$  con una métrica  $g = ds^2$  sobre  $M$  se tiene un isomorfismo no-canónico  $T_{\underline{p}}M \simeq T_{\underline{p}}^*M$  inducido por la métrica (este punto de vista se desarrolla a partir del §4).

## 2.3. Calculando espacios tangentes

El operador diferencial gradiente proporciona la primera aproximación al cálculo del espacio tangente de una variedad regular, es decir, no singular; al final del primera apartado de esta subsección se muestra una primera aproximación al cálculo de “elementos tangentes” correspondiente al caso singular, en términos del “cono tangente”.

El Teorema de la Función Implícita permite reparametrizar localmente una variedad en términos de funciones coordenadas locales correspondientes a las variables que tienen un menor de tamaño máximo con determinante no-nulo; la interpretación geométrica de este enfoque motiva el estudio de la Geometría

del Discriminante que ilustramos con un ejemplo básico asociado a familias de variedades dependientes de parámetros.

El último apartado de esta subsección está dedicado a mostrar una estrategia sintética para el caso en el que no se dispone (ó resulta muy engorroso de utilizar) una descripción implícita ó paramétrica para una variedad.

### 2.3.1. Representaciones implícitas

La forma más sencilla para expresar una superficie en  $\mathbb{R}^3$  está dada en forma implícita por  $z = f(x, y)$  (donde  $z$  se puede interpretar como una función "altura" ó "profundidad", p.e.). En este caso, el campo gradiente para  $g(x, y, z) = f(x, y) - z$  está dado por  $(f_x, f_y, -1)$  que representa el vector normal  $\mathbf{N}$  (dirigido hacia el "exterior") de la variedad  $V(g) := \{(x, y, z) \in \mathbb{R}^3 \mid g(x, y, z) = 0\}$ . En este caso, el plano tangente está dado por la condición de ortogonalidad

$$\langle \mathbf{N}_{\mathbf{x}_0}, \mathbf{x} - \mathbf{x}_0 \rangle = 0$$

Esta construcción y su interpretación geométrica se extienden al caso  $n$ -dimensional para

- Una *hipersuperficie*  $f(x_1, \dots, x_n) = 0$  con hiperplano tangente  $(n-1)$ -dimensional dado por  $\langle \mathbf{N}_{\mathbf{x}_0}, \mathbf{x} - \mathbf{x}_0 \rangle = 0$  donde ahora  $\mathbf{N}_{\mathbf{x}_0} = \text{grad}(f)(\mathbf{x}_0) = (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n})$  que genéricamente es  $\neq 0$  (Teorema de Sard)
- La *intersección de  $p$  hipersuperficies*  $f_1, \dots, f_k$  (sistema no-redundante ó "funcionalmente independiente") con  $f_i = f_i(x_1, \dots, x_n)$ . El rango de la matriz jacobiana  $\text{Jac}(F) = \text{Jac}(f_1, \dots, f_k)$  es genéricamente  $m = \min(n, k)$  (más adelante diremos que las hipersuperficies se cortan "transversalmente"). Por ello, el espacio vectorial tangente  $T_{\underline{x}_0}(\cap_{i=1}^k V(f_i))$  a la intersección  $\cap_{i=1}^k V(f_i)$  de las variedades  $V(f_i)$  es la intersección de los espacios tangentes  $\cap_{i=1}^k T_{\underline{x}_0}(V(f_i))$  que es un subespacio de codimensión  $k$ , es decir, un elemento de  $\text{Grass}(n-k, n) \simeq \text{Grass}(k, n)$ . En otras palabras, el subespacio  $L^{n-k}$  tangente a  $\cap_{i=1}^k V(f_i)$  en cada punto regular  $\underline{x}_0$  está representado en cada punto regular  $\underline{x}_0 \in \mathbb{R}^n$  (matriz jacobiana de rango máximo) por el producto exterior  $(\mathbf{N}_1 \wedge \dots \wedge \mathbf{N}_k)(\underline{x}_0)$  donde  $\mathbf{N}_j = \text{grad}(f_j) = (\frac{\partial f_j}{\partial x_1}, \dots, \frac{\partial f_j}{\partial x_n})$ . Por la condición de regularidad de  $F$  en  $\underline{x}_0$  podemos reordenar las líneas de la matriz jacobiana para que el primer  $(m \times m)$ -menor tange determinante no-nulo donde  $m = \min(n, k)$

### 2.3.2. Nota sobre la extensión al caso singular

Los argumentos anteriores se aplican al caso de puntos regulares de  $f$  que son "genéricos" (forman un denso) en virtud del Teorema de Sard. Las Geometrías Algebraica y Analítica se ocupan del análisis correspondiente al caso singular. Una primera aproximación al caso singular está dada por los ejemplos siguientes básicos:

- *Curva nodal*: El campo gradiente para la curva  $y^2 = x^2 + x^3$  (dibujarla) está dado por  $(2x + 3x^2, -2y)$  que es genéricamente no-nulo salvo en el origen. La noción de espacio tangente presentada al comienzo del capítulo no es aplicable pues el término lineal es idénticamente nulo. Por ello, se recurre a la *forma inicial* dada por los términos de menor grado  $x^2 - y^2 = (x - y)(x + y)$  (diagonales de los cuadrantes) que son las rectas tangentes a las dos ramas que presenta la curva nodal en el origen. A esta forma inicial se le llama el *cono tangente de Zariski* y desempeña el mismo papel que el espacio tangente para cada una de las ramas. *Ejercicio.*- Obtener una parametrización regular  $(x(t), y(t))$  de la curva (fuera del origen) a partir de la familia de rectas  $y = tx$  para  $t \in \mathbb{R}$  y verificar que esta parametrización está dada por cocientes de polinomios en  $t$ ; se dice entonces que la curva es *racional* (el cuerpo de funciones racionales es el mismo que el de la recta).
- *Curva cuspidal*: El campo gradiente para la curva  $y^2 = x^3$  (parábola semicuspidal de Newton en la terminología del s.XVIII) está dado por  $(3x^2, -2y)$  que es genéricamente no-nulo salvo en el origen. De forma análoga al caso anterior se tiene un cono tangente dado por  $y^2 = 0$  (una única rama correspondiente al eje  $Ox$  contado con "multiplicidad" 2), por lo que no es aplicable el argumento anterior. Sin embargo, la parametrización es bastante más sencilla, pues basta tomar  $(x(t), y(t)) = (t^2, t^3)$ . En este caso, cada una de las dos ramas de la curva cuspidal admite una parametrización analítica (serie de potencias convergentes) correspondientes a cada una de las raíces cuadradas de  $x^3$ .

El análisis de las singularidades que pueden presentar hipersuperficies más generales se aborda de forma sucesiva en términos de singularidades aisladas, singularidades estables por perturbación (incluyendo puntos múltiples ordinarios, puntos pinzados y las intersecciones entre ambos tipos de lugares) y singularidades arbitrarias (caso que aún presenta muchos problemas abiertos). Las singularidades correspondientes a la intersección de  $k$  hipersuperficies requieren una estratificación por el rango de la diferencial (representada por la matriz jacobiana), presentando una mayor diversidad de tipos que los considerados para el caso de una única hipersuperficie.

Una herramienta general para abordar este tipo de problemas consiste en identificar parametrizaciones locales, resultado que es consecuencia del Teorema Preparatorio de Weierstrass (ver R.C.Gunning and H.Rossi: "Analytic Functions of Several Complex Variables", Prentice Hall, 197?). En el apartado siguiente se aborda el estudio de representaciones paramétricas para el caso regular:

### 2.3.3. Representaciones paramétricas

Una representación paramétrica real está dada por una  $C^r$ -aplicación  $F : \mathbb{R}^n \times \Lambda^\ell \rightarrow \mathbb{R}^p$  que se puede interpretar como una "familia" de aplicaciones  $f_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}^p$  dependientes de un "espacio" de parámetros  $\Lambda$ . Denotemos

mediante  $\pi_1 : \mathbb{R}^n \times \Lambda^\ell \rightarrow \mathbb{R}^n$  la proyección sobre la primera componente y mediante  $\pi_\Lambda : \mathbb{R}^n \times \Lambda^\ell \rightarrow \Lambda$  la proyección sobre la segunda componente. Ambas proyecciones son genéricamente regulares. Un objeto de gran interés es el estudio de la *Geometría del Lugar Discriminante* asociado a  $\pi_1$  dado por

$$\{(\underline{x}, \underline{\lambda}) \in \mathbb{R}^n \times \Lambda^\ell \mid F(\underline{x}, \underline{\lambda}) = 0, \frac{\partial F(\underline{x}, \underline{\lambda})}{\partial \underline{x}} = 0\}$$

que consiste en “eliminar” los parámetros  $\underline{\lambda}$  ó bien en la *Geometría del Lugar discriminante asociado a  $\pi_\Lambda$*  dado por

$$\{(\underline{x}, \underline{\lambda}) \in \mathbb{R}^n \times \Lambda^\ell \mid F(\underline{x}, \underline{\lambda}) = 0, \frac{\partial F(\underline{x}, \underline{\lambda})}{\partial \underline{\lambda}} = 0\}$$

que consiste en “eliminar” las variables  $\underline{x}$ . Para entender este proceso, consideremos un ejemplo muy sencillo que permite “conectar” los dos ejemplos considerados en el apartado anterior como elementos de una familia de cúbicas dada por la deformación “universal” de  $y = f(x^3)$  (singularidad de tipo  $A_2$  dentro de la clasificación diferencial de gérmenes de aplicaciones):

*Ejemplo.-* Se considera la familia  $f(x; a, b) := x^3 - 3ax + b$  (representarlo gráficamente como un “pliegue” geológico con su correspondiente sinclinal y anticlinal). El lugar discriminante asociado a la proyección sobre el plano  $(a, b)$  de parámetros corresponde a “eliminar” la  $x$ , resolviendo el sistema

$$f(x; a, b) := x^3 - 3ax + b = 0, \quad \frac{\partial f(x; a, b)}{\partial x} := 3x^2 - 3a = 0$$

por lo que  $a = x^2 \Rightarrow x = \pm a^{1/2}$ ; sustituyendo obtenemos  $f(a, b) = \pm a^{3/2} - 3(\pm a^{3/2}) + b = 0$ , por lo que

$$b = \pm 2a^{3/2} \quad \Rightarrow \quad b^2 = 4a^3$$

es decir, se obtiene una curva cuspidal  $K$  en el plano de parámetros  $(a, b) \in \Lambda$  con la siguiente interpretación:

- $\forall (a, b) \in \Lambda - K$  la imagen recíproca es un polinomio que tiene 3 raíces simples;
- la imagen recíproca de  $(a, b) \in K - \underline{0}$  es un polinomio que tiene 2 raíces múltiples y una simple
- la imagen recíproca de  $\underline{0}$  es un polinomio con una única raíz triple

Por ello, la curva cuspidal del plano de parámetros “separa” los diferentes casos que se pueden presentar para los polinomios en una sola variable de grado 3 en términos del lugar discriminante, es decir, de las raíces múltiples de dichos polinomios. La elevación del lugar discriminante a la superficie cúbica  $V(f)$  en  $\mathbb{R}^3$  dada por  $f(x; a, b) = 0$  da una curva  $Ram(f) \subset V(f)$  a la que se llama el *lugar de ramificación* que se visualiza como el *contorno aparente* asociado a la proyección sobre el plano  $(a, b)$

La generalización de este ejemplo al caso de singularidades de tipo  $A_k$  (en la terminología de V.I. Arnol'd) con deformación “universal” dada por  $f(x; a_1, \dots, a_k) := x^{k+1} + a_1 x^{k-1} + \dots + a_k$  que depende de  $k$  parámetros (verificar que siempre existe una transformación que permite “eliminar” el coeficiente correspondiente al término de grado  $k$ ). La extensión a funciones que dependen de dos o más variables  $(x_1, x_2, \dots)$  ya no es tan elemental y se aborda en la Clasificación de Singularidades de Aplicaciones <sup>3</sup>

### 2.3.4. Representaciones simbólicas

Las descripciones correspondientes a los dos apartados anteriores utilizan la forma implícita ó paramétrica para la representación global ó local de una variedad. En este apartado se incluye un “ejemplo” que muestra cómo calcular el espacio tangente a una variedad  $X$  en un punto  $\underline{x}_0$  de una forma “intrínseca”, es decir, sin utilizar ningún tipo de representación en términos coordenados. La estrategia consiste en construir el espacio tangente a partir de la definición, es decir, como el conjunto de los vectores tangentes a las curvas  $\gamma : I \rightarrow X$  que tienen un contacto de orden  $\geq 2$  con la variedad  $X$  en el punto  $\gamma(0) = \underline{x}_0$ .

*Ejemplo: Espacio tangente a la grassmanniana en un punto* Denotemos por  $L_0$  a un subespacio vectorial  $(k+1)$ -dimensional de  $V^{n+1}$  representado por un punto  $\mathbf{P}_{L_0} \in \text{Grass}(k+1, n+1)$  que denotamos abreviadamente como  $G^{n-k}$  (grassmanniana de subespacios de codimensión  $n-k$ ). La grassmanniana  $G^{n-1}$  es intersección de una colección de hipersuperficies cuádricas, pero esta representación carece de utilidad para calcular el espacio tangente y da lugar a una casuística considerable. Por ello, es conveniente adoptar una estrategia alternativa basada en la descomposición  $V^{n+1} = L_0^{k+1} \oplus Q^{n-1}$  inducida por la sucesión exacta

$$0 \rightarrow L_0^{k+1} \rightarrow V^{n+1} \rightarrow Q^{n-1} = V^{n+1}/L_0^{k+1} \rightarrow 0$$

Denotemos mediante  $\{e_0, \dots, e_k\}$  a una base  $B(L_0^{k+1})$  de  $L_0^{k+1}$  que completamos con  $\{e_{k+1}, \dots, e_n\}$  de  $Q^{n-k}$  que forman una base  $B(Q^{n-k})$ , hasta obtener una base  $\{e_0, \dots, e_n\}$  de  $V^{n+1}$ . Denotemos mediante  $L_{ij}^{k+1}$  al subespacio  $(k+1)$ -dimensional que resulta de reemplazar el generador  $i$ -ésimo  $e_i$  de  $B(L_0^{k+1})$  por el vector  $j$ -ésimo  $e_j$  de  $B(Q^{n-k})$  para  $0 \leq i \leq k$  y para  $k+1 \leq j \leq n$ ; esta construcción da lugar a  $(k+1)(n-k)$  elementos de la grassmanniana verificando que  $\dim(L_0^{k+1} \cap L_{ij}^{k+1}) = k$  por lo que el haz  $\lambda_0 L_0^{k+1} + \lambda_1 L_{ij}^{k+1}$  determina una recta tangente a la grassmanniana  $G^{n-k}$ ; por la elección de los generadores, todos los subespacios  $(k+1)$ -dimensionales están asociados a vectores l.i.  $e_0 \wedge \dots \wedge \hat{e}_j^i \wedge \dots \wedge e_k$  en el álgebra exterior  $\wedge^{k+1} V$  donde  $\hat{e}_j^i$  denota el resultado de reemplazar el elemento  $i$ -ésimo  $e_i$  de  $B(L_0^{k+1})$  por el elemento  $j$ -ésimo  $e_j$  de  $B(Q^{n-k})$ . Como la dimensión de  $G^{n-k}$  es  $(k+1)(n-k)$  los multivectores  $e_0 \wedge \dots \wedge \hat{e}_j^i \wedge \dots \wedge e_k$  son l.i., son necesariamente un sistema de generadores de  $G^{n-k}$ . De una manera más sintética, la construcción anterior muestra que

<sup>3</sup> Estos tópicos se abordan en el módulo local de Topología Diferencial

$$T_{L_0^{k+1}} \text{Grass}(k+1, n+1) = \text{Hom}_{\mathbb{K}}(L_0^{k+1}, Q^{m-k})$$

y están representados por la  $(k+1) \times (n-k)$ -matriz correspondiente a las  $(n-k)$  últimas columnas de la representación paramétrica usual de una grassmanniana como variedad proyectiva en el abierto coordenado  $D_+(p_{01\dots k}) = \{(p_I) \in \mathbb{P}^N \mid p_{01\dots k} \neq 0\}$  (una vez realizada la división por dicha coordenada  $p_{01\dots k}$  que permite reemplazar la primera caja por la matriz identidad  $I_{k+1}$ ). Para  $k=0$  se obtiene que  $\text{Grass}(1, n+1) \simeq \mathbb{P}^n$ , por lo que el espacio tangente al espacio proyectivo en un punto  $L^1$  está dado por (el proyectivizado de)  $\text{Hom}_{\mathbb{K}}(L^1, Q^n)$

*Ejercicio.*- Describe de forma explícita el espacio tangente a la grassmanniana  $\text{Grass}(2, 4)$  (rectas en  $\mathbb{P}^3$ ) en un espacio 2-dimensional  $L^{(2)}$  (cuyo proyectivizado es una recta proyectiva  $\ell$ ).

*Nota (para matemáticos con conocimientos de Macroeconomía).*- Esta descripción se puede aplicar al modelo keynesiano IS-LM en Macroeconomía donde  $I$  es la inversión,  $S$  el ahorro,  $L$  la oferta monetaria (liquidity) y  $M$  la demanda monetaria para los 2-espacio  $IS$  (mercado de bienes) y  $LM$  (mercado monetario), respectivamente. Las políticas fiscales modifican el equilibrio en el plano  $IS$ , mientras que las políticas monetarias modifican el equilibrio en el plano  $LM$ . Si se toman las funciones ISLM como las coordenadas locales del modelo se obtiene una representación “natural”. Una política inteligente consiste en modificar de forma alternativa unas y otras, algo muy alejado de la política del Banco Central Europeo y de sus acólitos que se empeña en actuar sólo sobre la componente  $LM$ , evitando cualquier política fiscal, lo cual conduce de forma inevitable al estancamiento o al retroceso <sup>4</sup>.

*Ejercicio.*- Denotemos mediante  $\mathcal{B}(\ell_1, \dots, \ell_r)$  denota la variedad de banderas de “nacionalidad”  $(\ell_1, \dots, \ell_r)$ , es decir a la colección de subespacios encajados

$$(9) \subset L^{\ell_1} \subset L^{\ell_1} \oplus L^{\ell_2} \subset \dots \subset L^{\ell_1} \oplus L^{\ell_2} \oplus \dots \oplus L^{\ell_r} = V^n$$

para cualquier partición  $(\ell_1, \dots, \ell_r)$  de  $n+1$ . Si  $B \in \mathcal{B}(\ell_1, \dots, \ell_r)$ , demostrar que que

$$T_B \mathcal{B}(\ell_1, \dots, \ell_r) \simeq \prod_{1 \leq i < j \leq r} \text{Hom}(L^{\ell_i}, L^{\ell_j})$$

*indicación:* El estabilizador de  $B$  es el producto  $GL(\ell_1) \times \dots \times GL(\ell_r)$  (matrices diagonales por cajas de tamaño  $\ell_1, \dots, \ell_r$  y la variedad  $\mathcal{B}(\ell_1, \dots, \ell_r)$  está parametrizada por las matrices triangulares por cajas de tamaño  $(\ell_i \times \ell_j)$  que están situadas por encima de las cajas de la diagonal principal. Razonando por inducción sobre  $r$  (el caso  $r=1$  es trivial y el caso  $r=2$  es la grassmanniana) se concluye.

*Nota avanzada.*- La descripción anterior es válida para cualquier tipo de jerarquía geométrica que podemos introducir sobre variedades en el sentido más

<sup>4</sup> Para más detalles ver G.Mankiw: *Macroeconomía* (2994) o mis notas sobre Modelos Geométricos para Macroeconomía; estos tópicos se desarrollan más adelante en relación con aplicaciones de Foliaciones



amplio de la palabra, incluyendo el caso estratificado o variedades de Hilbert  $\mathcal{H}$  ó de Banach  $\mathcal{B}$  modeladas sobre espacios de Hilbert  $H$  ó de Banach  $B$ . En particular,

- Para una variedad estratificada, la descripción anterior permite comparar los límites de espacios tangentes de diferentes dimensiones asociados a diferentes estratos.
- Si se tiene una jerarquía entre soluciones asociadas a diferentes operadores cuyas soluciones dan lugar a conjuntos parcialmente ordenados (posets), el esquema precedente permite aproximar las soluciones de un sistema por las de otro sistema de ecuaciones

Este tipo de observaciones plantea el problema de identificar soluciones óptimas (que minimicen algún tipo de funcional definido sobre una variedad ó sobre un espacio de funciones) atendiendo a las jerarquías naturales de una variedad eventualmente estratificada o bien de espacios de soluciones asociados a diferentes tipos de operadores. La descripción de la variedad de banderas como cociente de un producto de grupos lineales es poco apropiada, pues  $GL(r; \mathbb{R})$  es un abierto, lo cual da lugar a una convergencia "muy lenta" para las soluciones de sistemas. Para remediar este problema, resulta más apropiado ortogonalizar o bien ortonormalizar las bases:

- *Ortogonalización* (Gram-Schmid): está justificada porque cualquier matriz regular con coeficientes reales descompone en producto de una simétrica y una ortogonal. El conjunto de las matrices simétricas es contractible por ser un espacio vectorial. Por ello, la ortogonalización permite pasar a un subgrupo cerrado.
- *Ortonormalización* en términos de grupos consiste en pasar a  $SO(n) = SL(n) \cap O(n)$  que es un grupo compacto, lo cual acelera la convergencia, si bien a costa de "reescalar" modelos y procesos (algo que puede resultar inapropiado dependiendo del problema a resolver).

Esta "reducción" del grupo estructural (paso del grupo lineal general al ortogonal o al especial ortogonal) implica que los cálculos de Optimización deben ser llevados a cabo en el grupo ortogonal  $O(n)$  (cerrado) ó el especial ortogonal  $SO(n)$  (compacto); en otras palabras, es necesario resolver el problema de optimización en el espacio tangente (álgebra de Lie) del grupo. Esta cuestión se aborda en el módulo 2. Una pequeña introducción a cuestiones de optimización se presenta en la sección siguiente.

Toda la construcción llevada a cabo en este apartado se traslada casi de forma inmediata al caso complejo, si bien la interpretación intuitiva es algo más complicada, debido al carácter que tienen las derivaciones holomorfas y antiholomorfas. La subsección siguiente lleva a cabo una primera introducción al problema que se irá abordando en diferentes módulos de esta asignatura.

## 2.4. Práctica: Gradiente complejo

El resultado clave para obtener una parametrización local real en un pequeño entorno de un punto es el Teorema de las Funciones Implícitas. La regularidad de la diferencial en un punto está representada por la no-anulación del determinante de un menor maximal de la matriz jacobiana correspondiente a un  $C^r$ -cambio de coordenadas. Realizando las adaptaciones oportunas para  $\mathbb{C}^n \simeq \mathbb{R}^{2n}$  se obtiene una descripción análoga. Para fijar ideas, nos restringimos a  $n \leq 2$ . Esta práctica debe ser completada con resultados básicos de Geometría Diferencial aplicados a Geometría Algebraica que se presentan en el módulo 4 (Formas Diferenciales), en el marco de variedades casi-complejas y su correspondiente extensión a variedades Kähler. En esta subsección nos limitamos a presentar el formalismo básico de los campos holomorfos y anti-holomorfos.

### 2.4.1. Caso de una variable compleja

Denotemos mediante  $\bar{z} = x - iy$  al conjugado de  $z \in \mathbb{C}$  con  $x = \frac{1}{2}[z + \bar{z}]$ ,  $y = \frac{1}{2i}[z - \bar{z}]$ . El espacio vectorial  $\mathbb{C}$  está generado por  $x, y$  ó bien por  $z, \bar{z}$  como  $\mathbb{R}$ -espacio vectorial. Análogamente,

$$\frac{\partial}{\partial x} = \frac{\partial z}{\partial x} \frac{\partial}{\partial z} + \frac{\partial \bar{z}}{\partial x} \frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial z} + \frac{\partial}{\partial \bar{z}}$$

y

$$\frac{\partial}{\partial y} = \frac{\partial z}{\partial y} \frac{\partial}{\partial z} + \frac{\partial \bar{z}}{\partial y} \frac{\partial}{\partial \bar{z}} = i \left[ \frac{\partial}{\partial z} - \frac{\partial}{\partial \bar{z}} \right]$$

cuyas inversas nos dan

$$\frac{\partial}{\partial z} = \frac{1}{2} \left[ \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right], \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left[ \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right]$$

que expresan el cambio de base para las  $\mathbb{R}$ -derivaciones complejas.

### 2.4.2. Operadores diferenciales estándar

*Lema* El polinomio  $f(z, \bar{z})$  no depende de  $\bar{z}$  si y sólo si  $\frac{\partial f}{\partial \bar{z}}$  es idénticamente nulo (comprobadlo como ejercicio).

Una función  $f(z) = u + iv$  es *analítica compleja* si y sólo si

$$0 = \frac{\partial f}{\partial \bar{z}} = \frac{1}{2} \left[ \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) + i \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \right]$$

que equivale a las *ecuaciones de Cauchy-Riemann*:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}$$

a partir de las que obtenemos  $\Delta u = \Delta v = 0$  donde

$$\Delta := \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial z \partial \bar{z}}$$

es el *operador de Laplace*. Las soluciones del operador de Laplace se llaman *funciones armónicas*. Este análisis se extiende de forma natural al caso de  $n$  variables complejas. Veamos el caso  $n = 2$ .

### 2.4.3. Funciones algebraicas y superficies de Riemann

*Proposición* Sea  $f \in \mathbb{C}[z, w] \simeq \mathbb{C}[z][w]$ . Denotemos mediante  $V(f) = \{(z, w) \in \mathbb{C}^2 \mid f(z, w) = 0\}$  a la variedad asociada a  $f \in \mathbb{C}[z][w]$  y consideremos la ecuación  $f(z, w) = 0$  en  $w$ . Si  $f_w = \frac{\partial f}{\partial w} \neq 0$  en  $\mathbf{p}_0 \in V(f)$ , entonces existe un pequeño entorno  $U$  de  $\mathbf{p}_0$  en  $\mathbb{C}^2$  tal que

1.  $g(z)$  es analítica compleja en  $U$
2. La función  $w = g(z)$  es una solución de la ecuación  $f(z, w) = 0$  en  $U$ , es decir,  $f(z, g(z)) = 0$  y esta solución es única en  $U$ .

Se dice entonces que  $w = g(z)$  es una *función algebraica* y al lugar de ceros  $V(f) = \{(z, w) \in \mathbb{C}^2 \mid f(z, w) = 0\}$  de  $f$  se le llama la *superficie de Riemann* para la función algebraica  $w = g(z)$  en los puntos en los que está definida.

Obviamente, el papel jugado por las variables  $z, w$  es intercambiable. Por ello, la condición de solubilidad local para  $f(z, w) = 0$  depende sólo de que el gradiente  $\text{grad}(f) = (\frac{\partial f}{\partial z}, \frac{\partial f}{\partial w})$  sea no nulo. La extensión de esta construcción al caso de  $n$  variables representa un importante tópico de la Teoría de Funciones de Varias Variables Complejas.

### 3. Algunas aplicaciones del campo gradiente

*Advertencia previa:* Toda esta sección está orientada a proporcionar elementos para realizar trabajos de forma individual (modalidad no-presencial)

A pesar de su simplicidad los sistemas procedentes de campos gradientes son de gran interés para las aplicaciones de las Matemáticas a otras Ciencias. A este tipo de sistemas se les llama *sistemas potenciales ó conservativos* (proceden de un potencial ó conservan una cantidad invariante, como veremos más adelante). Están dados vectorialmente en el entorno de un punto  $\underline{x}$  por un par de la forma  $(\underline{p}, \text{grad}(V\underline{x}))$ . A la función  $V : N \rightarrow \mathbb{R}$  se le llama el *potencial*; en un sistema conservativo se dice que la dinámica "deriva" de un potencial  $V$ . El lugar geométrico de los puntos  $\underline{p} \in N$  tales que  $V(\underline{x}) = r$  representa obviamente una solución particular de la ecuación del campo gradiente y se puede representar como una "hipersuperficie de nivel" correspondiente al valor regular  $r$ ; si  $c \in \mathbb{R}$  es un valor crítico, entonces  $f^{-1}(c)$  no es una variedad diferenciable (la matriz jacobiana no es inversible), aunque sea una variedad algebraica ó analítica.

Admitamos que para una variedad compacta  $M$ , las funciones de Morse son densas en  $C^r(M; \mathbb{R})$  para la topología compacta-abierto <sup>5</sup>. Como consecuencia, mediante una perturbación genérica se puede conseguir que todos los puntos críticos de la función  $f$  sean no-degenerados, es decir, tales que la forma cuadrática  $Hess(f)(\underline{x})$  es regular, a las que se llama *funciones de Morse*. Por ello, podemos restringirnos al caso de funciones con puntos críticos no-degenerados que presentan la misma topología (modulo homeomorfismos). La topología de las hipersuperficies de nivel cambia de forma cualitativa cada vez que se alcanza un punto crítico.

Una extensión inmediata de este enfoque concierne al estudio de extremos relativos simultáneos para una colección de funciones. En presencia de varias restricciones simultáneas  $f_1, \dots, f_p$ , el campo gradiente  $\text{grad}(f_i)(\underline{x})$  en cada punto define el vector normal a una hipersuperficie del espacio ambiente  $\mathbb{R}^n$  en cada punto que es el borde de la restricción  $\{\underline{x} \in \mathbb{R}^n \mid f(\underline{x}) \geq 0\}$ . La optimización (minimización ó maximización) simultánea de las restricciones recibe el nombre de *Optimización Multiobjetivo*. Se expresa como una linealización simultánea que afecta a una "distribución" de campos vectoriales conservativos (proceden de una función  $f_i$  para  $i = 1, \dots, p$ ). Este problema se aborda en primer lugar, con algunas ilustraciones a aspectos elementales de la Teoría Económica.

Además de los métodos lineales que se suponen conocidos de Programación Lineal (asociados al cálculo de la envolvente convexa en términos geométricos), la Optimización Multiobjetivo utiliza frecuentemente el método de *descenso más rápido* que se supone conocido también de Análisis Numérico. Se ilustra este principio mediante aplicaciones básicas relacionadas con Sistemas Expertos (aprendizaje supervisado) y algunas extensiones del enfoque basado en el "descenso más rápido" a cuestiones básicas de Robótica y de Visión Computacional, a las que se dedican los dos últimos apartados de esta sección.

<sup>5</sup> Para una demostración ver M.Hirsch: *Differential Topology*, Chap.5, Springer-Verlag, 1976

### 3.1. Elementos de Optimización Multiobjetivo ó Vectorial

*Nota previa.- Toda esta subsección se incluye como posible práctica para aquellas personas matriculadas que tengan algún interés en aplicaciones a Economía ó en Ingeniería.*

Para justificar el enfoque desarrollado en la subsección, observemos que

- Las situaciones teóricas más significativas para el análisis de variedades con borde están relacionadas con las condiciones iniciales para los sistemas de EDO ó bien para las condiciones frontera en los sistemas de EDP que introducen restricciones para sistemas en términos de desigualdades.
- En cualquier problema de optimización se buscan soluciones óptimas para un sistema mixto de ecuaciones e inecuaciones (algebraicas ó diferenciales). Las soluciones son extremales de un funcional y no tienen por qué ser únicas (ver más abajo)
- Los espacios de configuraciones  $\mathcal{C}$  y de trabajo  $\mathcal{W}$  para un robot son variedades con borde.
- Las regiones que aparecen como resultado de una segmentación (extracción de contornos) de cualquier imagen en Visión Computacional son variedades con borde.

#### 3.1.1. Conjuntos convexos

Un conjunto  $C$  es convexo si el segmento  $\overline{PQ}$  que une dos puntos cualesquiera  $P, Q$  está contenido en el conjunto. Esta definición no es constructiva (debido al cuantificador universal). La *envolvente convexa*  $\mathcal{H}(R)$  de una región  $R$  es el mínimo conjunto convexo que contiene a  $R$

La intersección de un número finito de semiespacios es un conjunto convexo. Cualquier convexo se puede construir como una intersección de los semiespacios que le contienen. Este resultado elemental proporciona una primera aproximación constructiva al cálculo de la envolvente convexa  $\mathcal{H}(R)$  de una región (para una aproximación algorítmica más eficiente ver O'Rourke: *Computational Geometry in C*, Camb. Univ Press ó bien Berg et al: *Computational Geometry*, Springer-Verlag)

Un conjunto convexo es una  $C^0$ -variedad con borde que facilita la gestión de regiones acotadas. Un conjunto convexo no es una  $C^r$ -variedad para  $r \geq 1$  (razonarlo). La optimización con respecto a una colección de restricciones lineales se realiza sobre conjuntos convexos. Si las restricciones son no-lineales, se tiene una optimización no-convexa. En las fases mas tempranas de la Visión Computacional ó en el diseño y planificación más tosco de movimientos en robótica, se utiliza la envolvente convexa de regiones para una acotación inicial del problema. La existencia de diferentes tipos de restricciones en diferentes ramas de

la Teoría Económica motiva la introducción de métodos de Optimización en Macro, Micro, Comercio Internacional y Finanzas.

### 3.1.2. Generalidades sobre Optimización

El *objetivo general* es minimizar ó maximizar un funcional (no necesariamente lineal) con respecto a un sistema de restricciones dadas por igualdades y desigualdades. El funcional puede estar definido sobre el soporte de la variedad ó sobre alguna de sus linealizaciones (espacio tangente ó su dual, el espacio, cotangente)

Las igualdades están soportadas sobre variedades ordinarias, mientras que las desigualdades están soportadas sobre variedades con borde. Suponemos que las restricciones se satisfacen de forma simultánea; esta hipótesis debe ser corregida en los problemas reales (como los de control, ejecución de tareas, seguimiento de trayectorias), donde existe un proceso secuencial para las restricciones con retrasos que deben ser tomados en consideración.

*Inconveniente analítico:* Falta de unicidad para las soluciones. Según Hadamard, se trata de un problema "mal planteado". Esta etiqueta ha funcionado con demasiada frecuencia como un freno para las conexiones naturales entre diferentes ramas de las matemáticas ó de éstas con otras ciencias aplicadas. La existencia de equilibrios (soluciones óptimas) no-aislados a lo largo de trayectorias es bien conocida desde mediados del s.XVIII. Asimismo, incluso cuando las soluciones óptimas son aisladas, no tienen por qué ser únicas. La utilización de funciones multivaloradas es bien conocida desde mediados del s.XIX (Riemann). Las restricciones no-lineales pueden dar lugar a la existencia de múltiples estados de equilibrio simultáneos, no sólo para cualquier tipo de interacción, sino también para diferentes estados de macrosistemas (como ocurre en la economía real, p.e.).

En el resto de esta sección se comentan otras aproximaciones generales al problema motivadas originalmente por desarrollos de la Teoría Económica y con un gran número de aplicaciones variadas a Biología, Química e Ingeniería. Actualmente, las aplicaciones más avanzadas conciernen al desarrollo de estrategias evolutivas de Optimización Multiojetivo para el aprendizaje en entornos colaborativos como los correspondientes a videojuegos; en este caso, el aprendizaje del sistema se lleva a cabo de forma implícita por parte de jugadores que pueden jugar on-line, lo cual da lugar a una modificación de "pesos relativos" para diferentes combinaciones de funciones objetivo. Debido a su carácter más avanzado, este tópico no se aborda en el Curso de Geometría Diferencial.

### 3.1.3. Optimización de Pareto

Frecuentemente, la función objetivo para la que deseamos calcular los extremos no es única. En este caso, la solución óptima no existe y se dice que tenemos "subóptimos de Pareto", Intuitivamente, corresponden a soluciones para las cuales cualquier modificación parcial *local* de los parámetros da lugar a

una situación peor que la correspondiente al subóptimo encontrado. De una manera más formal, el problema consiste en calcular

$$\min_{\mathbf{x} \in \mathbf{X}} p(\mathbf{f}(\mathbf{x})) \text{ con } \mathbf{x} \in \mathbb{R}^n, \mathbf{f} \in \mathbb{R}^m \text{ bajo las restricciones}$$

$$\mathbf{h}(\mathbf{x}) = 0, \mathbf{h} \in \mathbb{R}^q, \quad \mathbf{g}(\mathbf{x}) \leq 0, \mathbf{g} \in \mathbb{R}^p$$

### 3.2. Aproximaciones a la Optimización Multiobjetivo

*Nota previa.- Toda esta subsección se incluye como posible práctica para aquellas personas matriculadas que tengan algún interés en aplicaciones a Economía ó en Ingeniería.*

Un problema de *Optimización Multiobjetivo* consiste en la minimización simultánea de una colección de funciones  $(f_1(\underline{x}), f_2(\underline{x}), \dots, f_k(\underline{x}))^T$  sometidas a restricciones  $\underline{x} \in X$  donde  $X$  es una  $C^r$ -variedad cuyos "estados" se etiquetan como "vectores decisión"; esta terminología justifica la denominación de *optimización vectorial* para la optimización multiobjetivo. Cada decisión óptima corresponde a una solución del problema: en general, la solución no tiene por qué ser única.

En el caso clásico, se supone que  $X$  es un subconjunto convexo de  $\mathbb{R}^n$ , en cuyo caso se tiene solución única. En general,  $X$  es una variedad no necesariamente convexa. En las aplicaciones a Robótica,  $X$  es un producto de grupos clásicos. En problema de optimización para restricciones lineales cambiantes bajo información incompleta (parametrizadas por subespacios  $k$ -dimensionales de un espacio vectorial  $N$ -dimensional), el espacio ambiente es una grassmanniana  $Grass(k, N)$  que es una variedad no-lineal (ver más abajo la sección dedicada a variedades cociente). La interpolación ordinaria entre soluciones particulares da resultados (asociados a subespacios lineales) que no tienen sentido como puntos de  $Grass(k, N)$ . Por ello, es necesario resolver el problema en la "linealización" de la variedad (calculando su fibrado tangente)), cuestión que se aborda en el módulo 2.

#### 3.2.1. Método de los pesos objetivos

Consiste en seleccionar un "vector de pesos  $\mathbf{w}$  que minimice la expresión

$$p(\mathbf{f}(\mathbf{x})) := \mathbf{w} \cdot \mathbf{f}(\mathbf{x}),$$

donde  $0 \leq w_i \leq 1$  y  $\sum w_i = 1$  siendo  $w_i$  los *pesos relativos*.

Los pesos relativos pueden estar fijados por las condiciones del problema (optimización convencional) ó ser ajustados mediante algún proceso de aprendizaje; esta última es la opción más interesante en relación con sistemas evolutivos, incluyendo el aprendizaje (supervisado ó no) en Redes Neuronales Artificiales (ANNs). La Optimización sobre ANN se realiza en términos de una versión aleatorizada (los vectores son variables aleatorias) de la acción izquierda-derecha presentada en el capítulo anterior que incorpora un suavizado para facilitar el

pegado de datos; no existe un estándar para la elección de las funciones para el suavizado (la tangente hiperbólica proporciona un modelo con resultados razonables).

El mayor problema es la lentitud en la convergencia para los métodos de aprendizaje. Para resolver este problema se superponen diferentes tipos de algoritmos sobre las ANN basados en Algoritmos Genéticos (para acelerar el proceso de optimización), Mapas Auto-organizativos (para acelerar el proceso de reconocimiento sobre patrones) ó Programación Evolutiva (para incorporar aprendizaje sobre modelos cambiantes). Todas estas técnicas utilizan variantes más ó menos sofisticadas de procesos de optimización soportados sobre  $C^r$ -variedades dependientes ó no de parámetros.

Las condiciones sobre el vector  $\mathbf{w}$  de pesos se introducen inicialmente para garantizar la convexidad de las regiones donde se va a realizar el proceso de optimización; la convexidad acelera la convergencia. La interpretación geométrica es elemental en términos de geometría afín (coordenadas baricéntricas) y el método clásico a implementar es el del simplex. La convexidad puede no ser "natural" desde el punto de vista del espacio ambiente ó de las restricciones a resolver. En ese caso, hay que recurrir a técnicas de optimización no-convexa; un cuello de botella frecuente está relacionado con la eficiencia de la implementación computacional en presencia de incertidumbre o de condiciones ambientales muy cambiantes.

Frecuentemente, la formulación multiplicativa del problema de optimización permite acelerar la convergencia. En lugar de utilizar un enfoque aditivo, se puede utilizar un enfoque multiplicativo para la función multiobjetivo a minimizar; en este caso, los pesos son los exponentes de cada una de las componentes. Tomando logaritmos se pasa a la formulación aditiva; la exponencial proporciona la transformación inversa que permite pasar de la formulación aditiva a la multiplicativa. La aproximación diferencial se lleva a cabo en términos de las soluciones de EDO cuya solución general es de tipo exponencial. La integración para cada una de las componentes por separado da lugar a un grupo de homeomorfismos locales a lo largo de la trayectoria solución de cada EDO. Esta simplicidad es engañosa pues, en general, los sistemas diferenciales que describen los comportamientos variables no son "completamente integrables", es decir, no es posible realizar un desacoplamiento completo en diferentes funciones, por lo que es necesario refinar los métodos lineales que se acaban de describir. La integrabilidad de sistemas de aborda en el módulo 4 de esta asignatura.

### 3.2.2. Elementos para la resolución

Ilustramos estas ideas con dos estrategias sencillas para abordar el problema de optimización no-lineal. El enfoque está adaptado a aplicaciones en ANN (Redes Neuronales Artificiales) con aprendizaje supervisado desde el punto de vista de la aproximación lineal ó cuadrática (acorde con el enfoque presentado al final del capítulo 2):

- *Formulación Min-Max:* Elegir umbrales  $f_{i_0}$  y construir



$$p(\mathbf{f}(\mathbf{x})) := \max_{1 \leq i \leq m} \left( \frac{f_i(\mathbf{x})}{f_{i_0}(\mathbf{x})} \right) \quad f_{i_0} > 0 \quad \forall i = 1, \dots, m$$

Permite un comportamiento adaptativo a condiciones iniciales ó en el borde de cambiantes, donde los cambios pueden estar controlados por valores críticos de ciertos parámetros.

- *Función cuadrática de pesado:* Elegir

$$p(\mathbf{f}(\mathbf{x})) := \sqrt{\sum_{i=1}^m w_i f_i^2(\mathbf{x})}$$

donde los pesos  $0 \leq w_i \leq 1$  con  $\sum w_i = 1$  (con RPR reward-penalty rules) por lo que pueden ser reinterpretadas como las coordenadas afines de las funciones cuadráticas. Una extensión obvia de este enfoque se basa en elegir la norma  $L^p$  en lugar de la cuadrática (las regiones equipotenciales adquieren diferentes representaciones según el valor de  $p$ )

Este enfoque se utiliza para facilitar la navegación de dispositivos móviles (robots, vehículos inteligentes) con ruedas en entornos parcialmente estructurados.

### 3.3. Aprendizaje en sistemas expertos

*Nota previa.- Toda esta subsección se incluye como posible práctica a realizar para aquellas personas matriculadas que tengan algún interés en aplicaciones de Sistemas Expertos.*

Básicamente hay dos tipos de aprendizaje que se etiquetan como supervisado y no-supervisado. En el primero se dispone de una función objetivo a la que se ajustan los valores actuales mediante un mecanismo de realimentación (back-propagation estándar o bien temporal); un esquemas más avanzado corresponde a redes de tipo auto-asociativo en presencia de ruido (Hopfield. En el caso no-supervisado, no se dispone de una función objetivo previamente definida y es necesario desarrollar

#### 3.3.1. Aprendizaje supervisado

El *aprendizaje supervisado* consiste en el ajuste automático de los pesos de una red neuronal que minimice la distancia con respecto a una función objetivo prefijada.

El criterio habitual para la minimización de las distancias es de tipo cuadrático para las diferencias de primer orden entre los valores observados y los esperados como salida de la red (output). Por ello, la función objetivo recibe el nombre de *energía* (a veces también se le llama desviación de la función). El *descenso del gradiente* proporciona el método más simple para ajustar los pesos ó ponderaciones locales  $w_i$

El vector de pesos  $\mathbf{w}(k+1)$  en el paso  $k+1$  se calcula a partir del vector de pesos  $\mathbf{w}(k)$ . La evaluación de la función objetivo  $E$  (energía) sobre el vector de pesos  $\mathbf{w}(k+1)$  y su diferencia con respecto al vector  $\mathbf{w}(k)$ , proporciona una estimación tosca que etiquetamos como  $\nabla E(\mathbf{w}(k))$ . Los datos a estimar para el campo vectorial así descrito son

- La *dirección del ascenso más rápido* dada por  $\nabla E(\mathbf{w}(k))$ : Los pesos se modifican mediante "pequeños incrementos" en la dirección opuesta.
- La *magnitud* del incremento dada por la tasa de aprendizaje  $\varepsilon > 0$  entre los pasos  $k$  y  $k+1$  expresada como

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \varepsilon \nabla E(\mathbf{w}(k))$$

Según este esquema la *modificación del peso*  $\delta w_{ij}(k) = w_{ij}(k+1) - w_{ij}(k)$  se expresa como

$$\delta w_{ij} = -\varepsilon \frac{\partial E}{\partial w_{ij}}$$

donde  $\frac{\partial E}{\partial w_{ij}}$  representa la componente correspondiente del gradiente. De una forma ideal, el proceso converge hacia un mínimo absoluto (supuesto que exista), evitando los mínimos locales relativos mediante procedimientos de "perturbación". Esta aproximación presente múltiples deficiencias relacionadas con la lentitud en la convergencia, las dificultades para identificar si un mínimo es absoluto (trampas en falsos mínimos) ó la dependencia con respecto a los valores iniciales (una vez realizado el aprendizaje y fijados los pesos, si hay cambios en los datos del objeto a reconocer, es necesario entrenar nuevamente la red).

### 3.3.2. Modelos de back-propagation

Este método es uno de los más frecuentes para calcular el gradiente de una función objetivo para ANN que son independientes del tiempo. Si  $y_i^{des}$  es el valor deseado en la capa de salida para la neurona  $i$ -ésima, entonces la función de energía (error) a minimizar es

$$E = \frac{1}{2} [y_i - y_i^{des}]^2$$

Los errores que se calculan en los nodos de salida mediante

$$\Delta_i = \sigma'(x_i) [y_i - y_i^{des}]$$

En ocasiones, estos errores se han estimado previamente y ya se dispone de una distribución que suponemos conocida de una fase experimental previa. Con estos datos los cambios de pesos se calculan utilizando los errores mediante

$$\delta w_{ij} = -\varepsilon \frac{\partial E}{\partial w_{ij}} = -\varepsilon \Delta_i y_j$$

Para los nodos ocultos se tiene una distribución de errores dada por

$$\Delta_i = \sigma'(x_i) \sum_k w_{ki} \Delta_k$$

Este modelo es una primera aproximación estática al problema que debe ser completada con modelos de propagación temporal con *pesos variables*. Esta incorporación se lleva a cabo fácilmente en términos de una *función de error*

$$F = \frac{1}{2\tau} \int_0^\tau \sum_{i \in \Omega} [y_i - y_i^{des}] dt$$

asociada al conjunto  $\Omega$  de nodos de salida y donde  $\tau$  es la longitud del intervalo de tiempo. En este contexto, es necesario reemplazar pesos constantes por *pesos variables* que ahora están dados por

$$\delta w_{ij} = -\varepsilon \frac{\partial F}{\partial w_{ij}} = -\varepsilon \frac{1}{\tau} \int_0^\tau \Delta_i \sigma'_i(x_i) y_j dt$$

En este caso, las señales de error en un sistema dependiente del tiempo se calculan resolviendo el sistema de ecuaciones en derivadas parciales

$$\frac{d\Delta_i}{dt} = -\frac{\partial F}{\partial y_i} + \Delta_i - \sum_k w_{ki} \sigma'_k(x_k) \Delta_k$$

con las condiciones frontera  $[\Delta_i]_{t=\tau} = 0$ . Nótese que para los nodos de salida se tiene  $\frac{\partial F}{\partial y_i} = y_i - y : i^{des}$ , mientras que para los demás nodos  $\frac{\partial F}{\partial y_i} = 0$ .

*Práctica:* Desarrollar según S.Haykin: *Neural Networks. A comprehensive Foundation*, IEEE Computer Society, 1994

*comentario:* Esta presentación no tiene en cuenta las distribuciones de error que pueden aparecer asociadas a datos cambiantes ó eventualmente móviles. Para incorporarlas, adaptar los argumentos relacionados con la distancia de Mahalonobis que se presenta al final del capítulo 5 en el módulo 3 (movimiento) del CEViC en relación con problemas de seguimiento. *Práctica avanzada:* Llevar a cabo dicha formulación

### 3.3.3. Modelos asociativos (Hopfield)

Una *red de Hopfield* asocia a un conjunto inicial (incompleto ó corrompido por el ruido) un patrón binario del conjunto de aprendizaje. Al tipo de memoria resultante se le llama *memoria autoasociativa* y se ha utilizado en algunos procesos de aprendizaje (incluida la simulación de la pérdida temporal de memoria asociativa) para reforzar modelos de aprendizaje (desarrollar según mis notas de colaboración con el Lab Créare de Paris VI-VII).

*Práctica:* Desarrollar modelos no-lineales adaptativos para el aprendizaje por reforzamiento en relación con memorias de tipo asociativo y corrección de trastornos en la memoria. Una referencia para el marco general desde el punto

de vista de sistemas expertos podría ser R.Peretto: *An introduction to the Modeling of Neural Networks (reprt)*, Cambridge Univ. Press, 1994; otros aspectos más específicos desde el punto de vista geométrico es S. Watanabe: *Algebraic Geometry and Statistical Learning Theory*, Camb. Univ. Press, 2009)

### 3.3.4. Aprendizaje no supervisado: Modelos auto-organizativos

El modelo más sofisticado y con mayor potencial para el aprendizaje no-supervisado de máquinas corresponde a los mapas auto-organizativos. En este caso, no hay datos sobre el comportamiento esperado (capa de salida), ni tampoco sobre la evaluación de los outputs durante el proceso de aprendizaje. La cuestión principal es el desarrollo de modelos de aprendizaje que se basan en transformaciones de los datos que presenten diferentes tipos de estabilidad. El nivel más bajo para evaluar dicha estabilidad aparece asociado a la “conservación de la forma”; esta cuestión se aborda con diferentes niveles de complejidad que afectan a la invariancia topológica. Desde finales de los noventa se han desarrollado diferentes aproximaciones al problema que incluyen la homología persistente en el marco de la Topología Algebraica (ver último capítulo del módulo 2 del Curso de Topología Algebraica).

El modelo más conocido de mapas auto-organizativos se debe a Kohonen. En este modelo, cada componente de un vector de entrada actúa mediante conexiones con un número finito (pero cambiante) de nodos de la capa siguiente (oculta) según configuraciones geométricas previamente asignadas. El crecimiento en los “estímulos” debe generar una información redundante que facilite el aprendizaje. Por consiguiente, a partir de una inicialización de pesos interesa reforzar el aprendizaje para alcanzar una “estabilidad” en el plazo más breve posible. Las formas de activación inicialmente propuestas siguen esquemas de tipo Poisson (para el caso discreto) ó de tipo Gaussiano (debido a sus buenas propiedades como  $\sigma$ -álgebra). Otras elecciones utilizan representaciones de tipo hiperbólico (reemplazando las exponenciales por las tangentes hiperbólicas de las distribuciones centradas en torno a valores medios); estas soluciones facilitan un suavizado y una estabilización más rápidas que las soluciones basadas en el solapamiento de distribuciones Gaussianas.

*Práctica avanzada.-* Desarrollar mejoras para las estrategias de prueba-y-error que se sigue en estos modelos (ref: T.Kohonen: *Self-Organizing Maps (2nd ed)*, Information Sciences, Springer-Verlag, 1997)

## 3.4. Campos de Potencial en Ingeniería

*Nota previa.-* Toda esta subsección se incluye como posible práctica para aquellas personas matriculadas que tengan algún interés en aplicaciones a Ingeniería .

Se presentan dos ejemplos básicos que muestran el alcance de los campos vectoriales en Ingeniería.

### 3.4.1. Guiado de Robots mediante campos de potencial

En este apartado describimos la utilización del campo vectorial gradiente para el guiado de plataformas en escenarios planares con puntos que deben ser recorridos de forma sucesiva por una plataforma móvil. Cada punto de referencia es un atractor hasta que es alcanzado. En ese momento, el campo vectorial gradiente se invierte, pasando a ser repulsivo en lugar de atractivo. La plataforma móvil se desplaza siguiendo líneas del campo gradiente que son ortogonales a las hipersuperficies de nivel.

Para ello se lleva a cabo una aproximación jerarquizada que consiste en alcanzar ciertos hitos u objetivos parciales que actúan como “atractores” y que se modelan como mínimos relativos para un potencial. Por ello, los objetivos parciales alcanzados corresponden a “pozos de potencial” que invierten su signo una vez que el objetivo ha sido alcanzado. La inversión del signo corresponde a la conversión de un atractor en un repulsor.

Para generar el guiado es necesario construir una representación simplificada del espacio topológico

$$\mathcal{F} := \mathbb{R}^3 - \cup_{i=1}^k B_i$$

al que se llama *espacio libre de obstáculos*  $B_i$ . Cada obstáculo  $B_i$  se modela en términos de primitivas geométricas muy simples en el plano (rectángulos, elipses) o en el espacio (paralelepípedos, elipsoides) para facilitar la actualización de información

*desarrollar según Latombe, p.e.*

### 3.4.2. Elementos de Procesamiento de Imagen

En este apartado describimos algunos operadores diferenciales basados en el campo vectorial gradiente que proporcionan una información tosca sobre la estructura de una escena ó que facilitan el seguimiento de objetos en movimiento. La primera aplicación se basa en un modelo de diferencias finitas con diferentes tipos de máscaras (pesos en píxeles adyacentes), mientras que el segundo sigue un esquema similar pero adaptado a un modelo discreto del flujo óptico donde suponemos que disponemos de información adicional sobre el modelo ó la escena para inicializar correctamente el seguimiento.

La clave de los procedimientos para la extracción de aristas consiste en identificar las discontinuidades de la función de intensidad (aristas de objetos contenidos en una imagen). Dichas discontinuidades se reinterpretan en términos de “saltos” en el gradiente que es necesario agrupar y suavizar. La introducción de una orientación para las aristas obtenidas permite introducir un Histograma de Gradientes Orientados (HOG) que es clave para agrupar aristas y facilitar el seguimiento de objetos eventualmente deformables o móviles.

## 4. Ecuaciones diferenciales ordinarias de orden superior

*Nota previa.- Toda esta subsección se incluye como posible práctica para aquellas personas matriculadas que tengan algún interés en las EDO y sus aplicaciones a Física Matemática.*

### 4.1. Notación

Una ecuación diferencial ordinaria (en adelante, EDO) de orden  $n$  dada por

$$\frac{d^n x}{dt^n} = F\left(t, x, \frac{dx}{dt}, \dots, \frac{d^{n-1}x}{dt^{n-1}}\right),$$

es equivalente a un sistema de EDO de primer orden dado por

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = x_3, \quad \dots, \quad \dot{x}_{n+1} = F(t, x_1, \dots, x_n).$$

Por ello, dar una solución  $\phi : I \rightarrow \mathbb{R}$  de la ecuación original verificando condiciones iniciales prefijadas, equivale a dar una solución del sistema de  $n$  ecuaciones de primer orden, verificando las condiciones iniciales

$$\phi(u_0) = u_1, \quad \frac{d\phi}{dt} \Big|_{t=t_0} = u_2, \quad \dots, \quad \frac{d^{n-1}\phi}{dt^{n-1}} \Big|_{t=t_0} = u_n.$$

La situación más favorable se presenta cuando estudiamos EDO de orden  $n$  que son lineales en  $x_2, \dots, x_n$  (según la notación del sistema precedente). En este caso existen diferentes procedimientos de reducción ó transformación que nos permiten "separar" variables (ver por ejemplo, Ince y Sneddon, §5 para una introducción). Aunque éste no es un problema elemental, supongamos que hemos realizado cambios en las variables que nos permiten resolver el sistema por iteración.

### 4.2. Construcción local de vectores tangentes de orden superior

El procedimiento descrito equivale geométricamente a construir elevaciones sucesivas de caminos  $\phi_i$  con  $i = 1, \dots, n$  que son soluciones para cada uno de los sistemas parciales de EDO de primer orden que aparecen en las  $i$  primeras filas del sistema original, de modo que dichas elevaciones sean "compatibles" sobre la intersección de sus dominios de definición.

De una manera más explícita  $\phi_1 : I \rightarrow I \times R$  asocia a cada  $u_0 \in U \subseteq I$  un par  $(u_0, u_1) \in \text{grafo}(u_1) \subset I \times \mathbb{R}$  tal que  $\phi(u_0) = u_1$ . Por ello,  $\phi_1$  sólo fija el valor de la solución en el punto de acuerdo con las condiciones iniciales.

A continuación, como  $\dot{\phi} \Big|_{t=t_0}$  fija el valor  $u_2$  del vector tangente en el punto, tenemos que la aplicación  $\phi_2 : I \rightarrow I \times T\mathbb{R}$  está dada por el par cuyas componentes son  $(u_0, u_1)$  y  $(u_1, u_2)$ , donde la segunda componente  $(u_1, u_2)$  fija de forma única la solución que pasa por el punto. Este par puede ser interpretado

como una terna ordenada dada por  $(u_0, u_1, u_2)$  con la misma significación que la expuesta.

La aplicación  $\phi_3 : I \rightarrow T^2\mathbb{R}$  asigna a cada punto  $u_0$  una terna de pares cuyas componentes  $(u_0, u_1)$ ,  $(u_1, u_2)$  y  $(u_2, u_3)$  corresponden a la condición para que una curva  $u_1$  pase por un punto  $u_0$ , tenga un vector tangente  $u_2$  en  $u_0$  y la velocidad de dicho vector tangente (ó aceleración, según la terminología física tradicional) sea  $u_3$  en  $u_0$ . Del mismo modo que antes, dicha terna puede ser rescrita como una 4-upla.

La iteración de esta construcción es obvia. Interesa resaltar que esta construcción se puede llevar a cabo en situaciones más generales relativas no sólo a una ecuación diferencial de orden superior, sino a un sistema de EDO de diferentes órdenes. Como veremos en el §2, el lenguaje de fibrados tangentes sucesivos proporciona un lenguaje geométrico apropiado para abordar el estudio de las propiedades de este tipo de sistemas dados sobre variedades más generales que  $\mathbb{R}$ . En Topología Diferencial se lleva a cabo un estudio más sistemático en términos de jets.

### 4.3. Un ejemplo elemental: El Péndulo Simple

Consideremos la ecuación de segundo orden

$$\ddot{x}(t) := \frac{d^2x}{dt^2} = -x ,$$

Esta ecuación es equivalente al sistema de EDO de primer orden

$$\dot{x}_1 = x_2 \quad , \quad \dot{x}_2 = -x_1 ,$$

que también podemos representar matricialmente como

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Para pequeñas oscilaciones de un péndulo, las funciones  $\phi(t) = \sin(t)$  y  $\psi(t) = \cos(t)$  son soluciones independientes de la ecuación original que corresponden a las condiciones iniciales  $(0, 0, 1)$  y  $(0, 1, 0)$  para  $(t, x, \dot{x})$ , respectivamente. Describir la solución general como una función trigonométrica, representarla gráficamente y mostrar la proyección de la dinámica sobre el espacio base. ¿A qué tipo topológico de punto corresponde el origen, si consideramos las soluciones correspondiente a la familia uniparamétrica de EDO dada por  $\ddot{x}(t) = \lambda^2 x$ ?

La primera aplicación de este modelo simplificado de péndulo corresponde al *movimiento de un cuerpo sólido* pendiente de una cuerda perfectamente elástica (se supone que no hay disipación). Muestra que el campo dado por  $V(x) := (x, -kx)$  (para  $k > 0$ ) es conservativo y describir el movimiento.

La integración de la ecuación original del péndulo (dada como  $\ddot{x}(t) + \sin(x) = 0$ ) es más difícil, pues requiere el uso de funciones elípticas. Sin embargo, el comportamiento cualitativo de las soluciones que hemos mostrado es el mismo que el de la ecuación original para pequeñas oscilaciones. Esta es una primera

muestra de un procedimiento de linealización, que desarrollaremos a partir del §2.

#### **4.3.1. Oscilaciones forzadas**

En este apartado se presenta el modelo de oscilador de Van der Pol y se discuten las soluciones.

#### **4.3.2. Acoplamiento de dos sectores**

Desarrolla una adaptación del modelo presentado al acoplamiento de dos sectores en Macroeconomía.

#### **4.3.3. Acoplamiento entre tres sectores**

El propósito de este apartado es ilustrar los fenómenos de inestabilidad que pueden surgir para la dinámica correspondiente al acoplamiento de tres péndulos y la necesidad de identificar regiones de estabilidad (como complementarias de las regiones de caos).