



Álgebra y cálculo en varias variables para Químicas: Notas de un curso

Primer curso del grado en Químicas de la UAM

Beatriz Molina-Samper

UNIVERSIDAD DE VALLADOLID. FACULTAD DE CIENCIAS

Beatriz Molina-Samper
e-mail: beatriz.molina@uva.es

Álgebra y cálculo en varias variables para Químicas ©2024 by Beatriz Molina-Samper is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Última versión: 15 de mayo de 2022



Índice general

I Álgebra lineal

1	Sistemas de ecuaciones lineales	9
2	Espacios vectoriales	13
2.1	Definiciones y primeros ejemplos	13
2.2	Bases, dimensión y coordenadas de un espacio vectorial	15
2.3	Cambio de base	18
3	Transformaciones lineales	21
3.1	Definiciones	21
3.2	Subespacios invariantes y cambios de base	24
3.3	Diagonalización y direcciones propias	25
4	Espacios euclídeos	29
4.1	Definición y primeras propiedades	29
4.2	Bases ortonormales	32
4.3	Proyección ortogonal	34
4.4	Método de ortogonalización de Gramm-Schmidt	35
4.5	Transformaciones ortogonales	37
4.6	Transformaciones ortogonales en \mathbb{R}^2	37

4.7	Transformaciones ortogonales en \mathbb{R}^3	38
4.8	Transformaciones ortogonales generales	39

5 Sistemas de EDO's lineales con coeficientes constantes

41

5.1	El caso lineal y con coeficientes constantes	42
5.1.1	El caso homogéneo.	43
5.1.2	Resolución del caso no homogéneo:	47
5.2	Sistemas de ecuaciones diferenciales y EDO's de orden superior	48
5.3	Sistemas dinámicos discretos y continuos	49

II Cálculo de varias variables

6 Funciones de varias variables 53

6.1	Funciones de varias variables. Definición	53
6.2	Funciones de dos variables: Curvas de nivel	55
6.3	Funciones en tres variables: Superficies de nivel	57
6.4	Representación gráfica: Superficies de nivel	60
6.5	Coordenadas: polares, cilíndricas, esféricas.	61
6.5.1	Coordenadas polares	61
6.5.2	Coordenadas cilíndricas	63
6.5.3	Coordenadas esféricas	64
6.6	Parametrización de curvas	66

7 Cálculo diferencial 67

7.1	Derivadas parciales	67
7.2	Derivadas direccionales	70
7.3	Diferenciabilidad. Plano tangente	71
7.4	Propiedades del gradiente. Puntos críticos	74
7.4.1	Direcciones de máximo crecimiento y decrecimiento	74
7.4.2	El gradiente mediante curvas de nivel	76
7.4.3	Puntos críticos	78
7.5	Derivadas parciales de orden superior	80
7.6	Matriz Hessiana. Cuádrica aproximante	81
7.7	Análisis de puntos críticos	82
7.8	Aplicaciones vectoriales. Matriz Jacobiana	86
7.9	Cambio de variable. Regla de la cadena	88
7.9.1	Regla de la cadena (sobre una curva)	88
7.9.2	Regla de la cadena (a lo largo de una superficie)	89
7.9.3	Derivación implícita	90

8	Cálculo integral	91
8.1	Integrales iteradas	92
8.2	Parametrización de regiones	94
8.3	Cambio de coordenadas	97
8.3.1	Coordenadas polares	97
8.3.2	Otros cambios de variable	98
8.4	Integrales triples. Volumen y masa	99
8.4.1	Coordenadas cilíndricas	101
8.4.2	Coordenadas esféricas	102

III

Cálculo vectorial

9	Aplicaciones vectoriales	107
9.1	Rotacional y divergencia. Potencial	108
9.2	Integrales de línea. Trabajo	110
9.2.1	El caso de un campo conservativo	112



Álgebra lineal

1	Sistemas de ecuaciones lineales ...	9
2	Espacios vectoriales	13
2.1	Definiciones y primeros ejemplos	
2.2	Bases, dimensión y coordenadas de un espacio vectorial	
2.3	Cambio de base	
3	Transformaciones lineales	21
3.1	Definiciones	
3.2	Subespacios invariantes y cambios de base	
3.3	Diagonalización y direcciones propias	
4	Espacios euclídeos	29
4.1	Definición y primeras propiedades	
4.2	Bases ortonormales	
4.3	Proyección ortogonal	
4.4	Método de ortogonalización de Gramm-Schmidt	
4.5	Transformaciones ortogonales	
4.6	Transformaciones ortogonales en \mathbb{R}^2	
4.7	Transformaciones ortogonales en \mathbb{R}^3	
4.8	Transformaciones ortogonales generales	
5	Sistemas de EDO's lineales con coeficientes constantes	41
5.1	El caso lineal y con coeficientes constantes	
5.2	Sistemas de ecuaciones diferenciales y EDO's de orden superior	
5.3	Sistemas dinámicos discretos y continuos	

1. Sistemas de ecuaciones lineales

Este capítulo se dedica a la resolución de sistemas de ecuaciones lineales. En dimensión dos este problema responde a encontrar los puntos de corte de varias rectas, en dimensión tres de varios planos y en general a encontrar los puntos comunes a varios “hiperplanos”.

Trabajaremos en el espacio \mathbb{R}^n . Una ecuación lineal con n incógnitas representa un hiperplano de \mathbb{R}^n : una recta en el plano ($n = 2$), un plano en el espacio tridimensional ($n = 3$), etc. Esto es, la expresión

$$E := a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \quad a_i, b \in \mathbb{R}$$

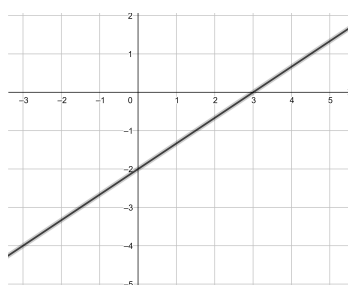
representa el conjunto de puntos

$$\{(u_1, u_2, \dots, u_n) \in \mathbb{R}^n : a_1u_1 + a_2u_2 + \cdots + a_nu_n = b\}.$$

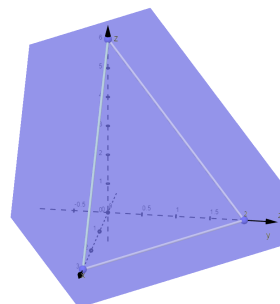
Cuando n es pequeño escribimos letras distintas para representar a las variables (en lugar de numerar con subíndices); por ejemplo (x, y) en \mathbb{R}^2 o (x, y, z) en \mathbb{R}^3 .

■ **Ejemplo 1.1** Consideremos las ecuaciones lineales

$$2x - 3y = 6 \text{ en } \mathbb{R}^2$$



$$2x + 3y + z = 6 \text{ en } \mathbb{R}^3$$

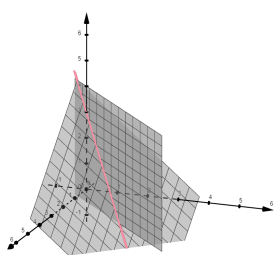


Las soluciones de un *sistema de m ecuaciones lineales con n incógnitas* representa el conjunto de puntos de \mathbb{R}^n donde se cortan m hiperplanos de \mathbb{R}^n . Este viene dado por:

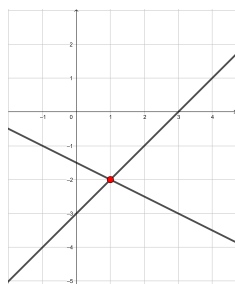
$$S := \begin{cases} E_1 := a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ E_2 := a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ E_m := a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

Éste puede tener solución o no: cuando tiene solución se dice *compatible* y cuando no la tiene *incompatible*. Además, en el caso de tener solución, puede que ésta sea única o no: cuando es única decimos que es *compatible determinado* y cuando no es única (de hecho tiene infinitas) decimos que es *compatible indeterminado*.

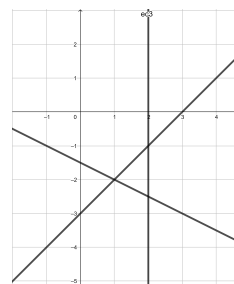
■ **Ejemplo 1.2** Las siguientes figuras se corresponden con:



Sistema compatible
indeterminado



Sistema compatible
determinado



Sistema
incompatible

Decimos que dos sistemas de ecuaciones S_1 y S_2 son *equivalentes* cuando su conjunto de soluciones coincide, es decir, ambos están representando el mismo subconjunto de \mathbb{R}^n .

Cuando los sistemas de ecuaciones que consideramos tienen pocas variables, los métodos conocidos de “sustitución”, “igualación” o “reducción” funcionan relativamente bien para encontrar sistemas equivalentes y así resolverlos. Sin embargo, cuando tenemos muchas variables, necesitamos sistematizar mejor estos métodos y una manera de hacerlo es utilizando la *eliminación de Gauss*.

Cuando tratamos con un número grande de variables, resulta muy cómodo representar el sistema S en *forma matricial* $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \ddots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Llamamos a A la *matriz de coeficientes* y a \mathbf{b} el *vector de términos independientes*. Si queremos representar el sistema con una sola matriz, lo hacemos mediante $A^* = (A | \mathbf{b})$ de tamaño $m \times (n + 1)$, que llamamos *matriz ampliada del sistema*.

Decimos que un sistema S es un *sistema escalonado* (*row echelon form*) cuando su matriz ampliada está en forma escalonada. Esto es cuando cumple las siguientes propiedades:

1. Las filas compuestas únicamente por ceros se sitúan abajo del todo.
2. El primer elemento no nulo en una fila, llamado el *pivote de la fila*, vale 1.
3. Una columna que contiene un elemento pivote, tiene el resto de entradas nulas; en particular, una columna contiene como máximo un pivote.

■ **Ejemplo 1.3** Las siguientes matrices son escalonadas:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & b_2 \\ 0 & 0 & 1 & b_3 \end{array} \right), \quad \left(\begin{array}{ccc|c} 1 & a_1 & 0 & b_1 \\ 0 & 0 & 1 & b_2 \\ 0 & 0 & 0 & 1 \end{array} \right), \quad \left(\begin{array}{ccc|c} 1 & 0 & a_1 & b_1 \\ 0 & 1 & a_2 & b_2 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

■ **Observación 1.1** Estas matrices se llaman escalonadas porque se puede trazar una “escalera” cambiando de peldaño cada vez que encontramos un punto pivote.

Cuando un sistema S es escalonado, es sencillo determinar cómo es en cuanto a sus soluciones se refiere (y también resolverlo).

Teorema 1.0.1 — Rouché-Frobenius. Sea S un sistema escalonado.

1. El sistema S es compatible si y solo si no existe un punto pivote en la última columna de la matriz ampliada.
2. Supongamos que S es compatible. Es compatible determinado si y solo si todas las columnas de la matriz de coeficientes contienen puntos pivote.

Observemos que, el hecho de que exista un punto pivote en la última columna de la matriz ampliada significa que una de las ecuaciones de nuestro sistema es

$$0x_1 + 0x_2 + \cdots + 0x_n = 1,$$

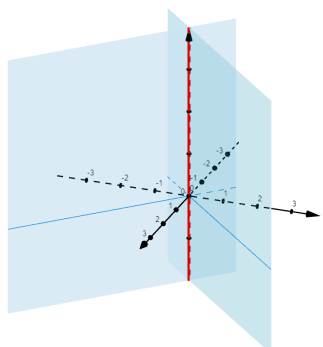
que claramente no tiene sentido, es decir, no hay puntos de \mathbb{R}^n que la satisfagan. Por otro lado, ya en el caso compatible, exigir que todas las columnas de la matriz de coeficientes contengan puntos pivote equivale a pedir que nuestro sistema sea

$$\begin{cases} x_1 = b_1 \\ x_2 = b_2 \\ \vdots \\ x_n = b_n \end{cases}$$

cuya solución es única y viene dada por $(b_1, b_2, \dots, b_n) \in \mathbb{R}^n$.

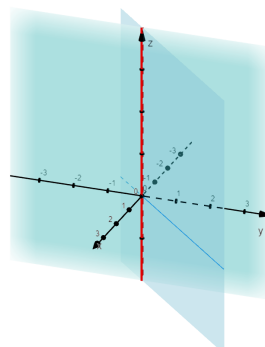
Dado un sistema de ecuaciones S , el objetivo del *método de eliminación de Gauss-Jordan* es encontrar un sistema escalonado equivalente a S . Esto se consigue utilizando *operaciones elementales*:

1. Intercambiar dos ecuaciones. Es claro que al intercambiar dos ecuaciones, el problema que pretendemos resolver continúa siendo el mismo.
2. Multiplicar una ecuación por un número real $\lambda \neq 0$. Las soluciones de la ecuación λE determina el mismo subconjunto de \mathbb{R}^n que las de E .
3. Sumar dos ecuaciones. Un punto $(p_1, p_2, \dots, p_n) \in \mathbb{R}^n$ satisface las ecuaciones E y F si y solamente si satisface E y $E + F$. Gráficamente lo vemos con el siguiente ejemplo en \mathbb{R}^3 .



$$\begin{cases} E & := & x - y = 0 \\ F & := & x + y = 0 \end{cases}$$

$$S = \{(0, 0, t) \in \mathbb{R}^3, t \in \mathbb{R}\}$$



$$\begin{cases} E & := & x - y = 0 \\ E + F & := & 2x = 0 \end{cases}$$

$$S = \{(0, 0, t) \in \mathbb{R}^3, t \in \mathbb{R}\}$$

En la práctica muchas veces combinamos las operaciones 2 y 3 para “ahorrarnos pasos”, así podemos cambiar directamente la ecuación E por $E + rF$, donde r es un número real no nulo y F es otra ecuación del sistema. También podemos combinar varias a la vez cambiando E por $E + r_1F_1 + r_2F_2 + \dots + r_kF_k$, donde los r_i son números reales y las F_i son otras ecuaciones del sistema.

■ **Ejemplo 1.4** Consideremos los siguientes sistemas de ecuaciones:

$$a) \begin{cases} x - y = 2 \\ x + y = 4 \end{cases} \quad b) \begin{cases} x + 3y + z = -3 \\ 3x + 9y + 4z = -7 \\ 2x - y + z = 6 \end{cases} \quad c) \begin{cases} x + 3y - z = 2 \\ -x + y + 3z = 0 \\ 2y + z = 2 \end{cases}$$

Tenemos que:

- a) Es compatible determinado y tiene por solución $\{(3, 1)\}$.
- b) Es compatible determinado y tiene por solución $\{(1, -2, 2)\}$.
- c) Es incompatible.

Para concluir este tema, diremos que un sistema $A\mathbf{x} = \mathbf{b}$ es *homogéneo* si el vector de términos independientes es nulo. Estos sistemas siempre son compatibles, dado que el origen de \mathbb{R}^n es una solución para ellos (geoméricamente estamos considerando hiperplanos que pasan todos ellos por el origen). Este tipo de sistemas serán muy útiles para definir subespacios vectoriales de \mathbb{R}^n .

2. Espacios vectoriales

En este capítulo introducimos el concepto de espacio vectorial, independencia lineal y base. Hablaremos de transformaciones lineales entre espacios vectoriales y nos centraremos en el estudio de éstas cuando no cambiamos de espacio (endomorfismos), en concreto estudiaremos cuando son diagonalizables y por tanto tienen una base formada por vectores propios.

2.1 Definiciones y primeros ejemplos

Los espacios vectoriales se definen relativos a un “cuerpo”, por ejemplo \mathbb{R}^n es un espacio vectorial sobre los números reales (el cuerpo \mathbb{R}) y \mathbb{C}^n es un espacio vectorial sobre los números complejos (el cuerpo \mathbb{C}). Así, comenzamos esta sección introduciendo el concepto de cuerpo. Para ello nos centramos en propiedades conocidas de los números reales.

Sobre el conjunto de los números reales \mathbb{R} , tenemos las operaciones conocidas de suma y producto (multiplicación):

$$\begin{array}{lcl} + : \mathbb{R} \times \mathbb{R} & \rightarrow & \mathbb{R} \\ (a, b) & \mapsto & a + b \end{array} \quad \begin{array}{lcl} \cdot : \mathbb{R} \times \mathbb{R} & \rightarrow & \mathbb{R} \\ (a, b) & \mapsto & a \cdot b = ab \end{array}$$

y para ellas se tienen las siguientes propiedades.

S) Para la suma:

S1) (Asociativa) Para todos $a, b, c \in \mathbb{R}$ se tiene $(a + b) + c = a + (b + c)$.

S2) (Conmutativa) Para todos $a, b \in \mathbb{R}$ se tiene $a + b = b + a$.

S3) Existe un elemento $0 \in \mathbb{R}$, llamado *elemento neutro para la suma o cero*, tal que para todo $a \in \mathbb{R}$ se tiene $a + 0 = 0 + a = a$.

S4) Para todo $a \in \mathbb{R}$ existe un elemento $b \in \mathbb{R}$, llamado *elemento opuesto de a* de modo que $a + b = b + a = 0$. Lo denotaremos por $-a$.

P) Para el producto:

P1) (Asociativa) Para todos $a, b, c \in \mathbb{R}$ se tiene que $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.

P2) (Conmutativa) Para todo $a, b \in \mathbb{R}$ se tiene $a \cdot b = b \cdot a$.

P3) Existe un elemento $1 \in \mathbb{R}$, llamado *elemento neutro para el producto o uno*, tal que para todo $a \in K$ se tiene $a \cdot 1 = 1 \cdot a = a$.

P4) Para todo $a \in \mathbb{R}$ con $a \neq 0$ existe un elemento $b \in K$, llamado *elemento inverso* de modo que $a \cdot b = b \cdot a = 1$. Lo denotaremos por a^{-1} .

Finalmente, la suma y el producto cumplen la propiedad distributiva:

D) Dados $a, b, c \in \mathbb{R}$ se tiene que $(a + b) \cdot c = a \cdot c + b \cdot c$.

Un *cuerpo* \mathbb{K} es una terna $\mathbb{K} = (K, +, \cdot)$ dada por un conjunto no vacío K y dos aplicaciones u operaciones binarias

$$+ : K \times K \rightarrow K; \quad \cdot : K \times K \rightarrow K,$$

llamadas *suma* y *producto* cumpliendo las propiedades S1)-S4), P1)-P4) y D.

Ejemplos de cuerpos son los números racionales \mathbb{Q} , los números reales \mathbb{R} o los complejos \mathbb{C} , pero hay muchos más: por ejemplo, existen cuerpos con un número finito de elementos. Nosotros no vamos a entrar en esto y nos centraremos especialmente en el caso $\mathbb{K} = \mathbb{R}$ y en alguna ocasión $\mathbb{K} = \mathbb{C}$.

Sobre el conjunto \mathbb{R}^n tenemos las operaciones de *suma vectorial* y *producto por escalares* (elementos del cuerpo \mathbb{R}) definidas por:

$$\begin{aligned} + : \quad & \mathbb{R}^n \times \mathbb{R}^n & \rightarrow & \mathbb{R}^n \\ & ((a_1, a_2, \dots, a_n), (b_1, b_2, \dots, b_n)) & \mapsto & (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n) \\ \cdot : \quad & \mathbb{R} \times \mathbb{R}^n & \rightarrow & \mathbb{R}^n \\ & (\lambda, (a_1, a_2, \dots, a_n)) & \mapsto & (\lambda a_1, \lambda a_2, \dots, \lambda a_n). \end{aligned}$$

Por comodidad en la notación escribiremos $\vec{a} = (a_1, a_2, \dots, a_n)$.

Se cumplen las siguientes propiedades.

S) Para la suma vectorial se cumple S1), S2), S3) y S4).

M) Para el producto por escalares se tiene:

M1) $1\vec{a} = \vec{a}$, para todo $\vec{a} \in \mathbb{R}^n$.

M2) $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$, para todo $\vec{a} \in \mathbb{R}^n$ y $\lambda, \mu \in \mathbb{R}$.

M3) $(\lambda + \mu)\vec{a} = \lambda\vec{a} + \mu\vec{a}$, para todo $\vec{a} \in \mathbb{R}^n$ y $\lambda, \mu \in \mathbb{R}$.

M4) $\lambda(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b}$, para todo $\vec{a}, \vec{b} \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$.

Decimos que $(\mathbb{R}^n, +, \cdot)$ es un espacio vectorial y a sus elementos los llamamos *vectores*.

De manera completamente general, un *espacio vectorial sobre un cuerpo* \mathbb{K} es una terna $\mathcal{V} = (V, +, \cdot)$ dada por un conjunto no vacío V y dos aplicaciones

$$+ : V \times V \rightarrow V \quad \cdot : K \times V \rightarrow V,$$

llamadas *suma vectorial* y *producto por escalares* cumpliendo las propiedades propiedades [S1-S4] y [M1-M4]. Al conjunto V lo llamamos *conjunto subyacente a \mathcal{V}* y frecuentemente abusamos de la notación y escribimos directamente V para referirnos al espacio vectorial cuando las operaciones se sobreentienden. Cuando $\mathbb{K} = \mathbb{R}$, decimos que \mathcal{V} es un *espacio vectorial real*. Cuando $\mathbb{K} = \mathbb{C}$, decimos que \mathcal{V} es un *espacio vectorial complejo*.

■ **Ejemplo 2.1** Algunos ejemplos de espacios vectoriales son los siguientes:

1. El conjunto $M_{n,m}(\mathbb{R})$ de las matrices con n filas y m columnas y entradas en \mathbb{R} es un \mathbb{R} -espacio vectorial.
2. El conjunto $\mathbb{R}[X]$ de los polinomios en la variable X con coeficientes en \mathbb{R} .
3. Las funciones reales de variable real $F(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{R}\}$ con las operaciones:

$$\begin{aligned} f + g : \mathbb{R} &\rightarrow \mathbb{R} & \lambda f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto f(x) + g(x) & x &\mapsto \lambda f(x) \end{aligned}$$

Sea $\mathcal{V} = (V, +, \cdot)$ un espacio vectorial. Si $W \subset V$ es tal que $\mathcal{W} = (W, +, \cdot)$ es un espacio vectorial, decimos que \mathcal{W} es un *subespacio vectorial* de \mathcal{V} . Notemos que para comprobar si un subconjunto $W \subset V$ es un subespacio vectorial basta comprobar que las operaciones son cerradas dentro de W , esto es que

$$\vec{w}_1 + \vec{w}_2 \in W, \quad \lambda \vec{w}_1 \in W$$

para todos $\vec{w}_1, \vec{w}_2 \in W$.

■ **Ejemplo 2.2** El conjunto de soluciones de un sistema de ecuaciones homogéneo $Ax = 0$ es un subespacio vectorial de \mathbb{R}^n . De hecho, todo subespacio vectorial de \mathbb{R}^n puede describirse de esta manera, es decir, como intersección de hiperplanos que pasan por el origen.

2.2 Bases, dimensión y coordenadas de un espacio vectorial

Sea V un \mathbb{R} -espacio vectorial y sea $C = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subset V$ un subconjunto finito de vectores en V . Una *combinación lineal de C* es cualquier vector $\vec{v} \in V$ que se pueda escribir como

$$\vec{v} = \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_k \vec{v}_k, \quad \lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{R}.$$

El conjunto de combinaciones lineales de C es un subespacio vectorial de V que llamamos *subespacio generado por C* y que denotamos por $L(C) = L(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$.

■ **Ejemplo 2.3** $(4, 5, 2)$ es una combinación lineal de $(1, 2, 1)$ y $(2, 1, 0)$, pues

$$(4, 5, 2) = 2(1, 2, 1) + (2, 1, 0).$$

Por otro lado $C = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\} \subset V$ es un conjunto de vectores *linealmente independientes* si ninguno de sus elementos es combinación lineal del resto. Esto se expresa diciendo que $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$ es la única solución al sistema

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_k \vec{v}_k = \vec{0}.$$

Cuando C no es un conjunto de vectores linealmente independientes, decimos que es un conjunto de vectores *linealmente dependientes*.

■ **Ejemplo 2.4** Los vectores $(1, 2, 1)$, $(2, 1, 0)$ y $(4, 5, 2)$ de \mathbb{R}^3 son linealmente dependientes pues ya hemos visto que

$$(4, 5, 2) = 2(1, 2, 1) + (2, 1, 0).$$

Los vectores $(1, 2, 1)$, $(2, 1, 0)$ y $(3, 4, 1)$ sí son linealmente independientes pues el sistema de ecuaciones lineales

$$\begin{cases} 1\lambda_1 + 2\lambda_2 + 3\lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 + 4\lambda_3 = 0 \\ \lambda_1 + 0\lambda_2 + \lambda_3 = 0 \end{cases}$$

tiene por única solución $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

Un subconjunto finito $C = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \subset V$ es un *sistema de generadores del espacio vectorial* V si $V = L(C)$. Una *base* β de V es un sistema de generadores cuyos elementos son linealmente independientes.

■ **Ejemplo 2.5** El conjunto $C = \{(0, 1, 2), (0, 2, 1), (2, 0, 0)\}$ es una base de \mathbb{R}^3 .

La *dimensión* $\dim V$ de un espacio vectorial V es el número de elementos en una base de V . Cuando el espacio vectorial consta solamente de un elemento, necesariamente el $\vec{0}$, es decir $V = \{\vec{0}\}$, decimos que tiene dimensión cero. Cuando la dimensión del espacio vectorial V es uno habitualmente decimos que V es una *recta* y cuando la dimensión de V es dos decimos que V es un *plano*.

Todo conjunto de n -vectores linealmente independientes en un espacio vectorial de dimensión n determina una base. Del mismo modo, si tenemos n vectores que son generadores de un espacio vectorial de dimensión n , entonces son necesariamente linealmente independientes.

■ **Ejemplo 2.6** Consideremos el subespacio vectorial $P \subset \mathbb{R}^3$ determinado por

$$P = \{(x, y, z) \in \mathbb{R}^3; x + 2y + 3z = 0\}$$

Tomando $y = \lambda$ y $z = \mu$ como parámetros, observamos que todos los elementos de P son de la forma

$$\vec{v} = (-2\lambda - 3\mu, \lambda, \mu).$$

Así tenemos que

$$\vec{v} = \lambda(-2, 1, 0) + \mu(-3, 0, 1).$$

De esta manera $\vec{v}_1 = (-2, 1, 0)$ y $\vec{v}_2 = (-3, 0, 1)$ son dos vectores linealmente independientes que generan todo P , es decir, determinan una base de P . Por tanto P es un subespacio vectorial de \mathbb{R}^3 de dimensión 2, un plano.

Teorema 2.2.1 Todas las bases de un espacio vectorial tienen el mismo número de elementos.

Teorema 2.2.2 Todos los espacios vectoriales tienen una base (no necesariamente finita).

El espacio vectorial $V = \mathbb{R}^n$ tiene una base especial, que llamamos *base estándar* o *base canónica*, la dada por $\beta_{st} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, donde

$$\vec{e}_i = (0, 0, \dots, 1, \dots, 0)$$

con el 1 apareciendo en la posición i -ésima.

Sea V un \mathbb{R} -espacio vectorial de dimensión n y sea $\beta = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base de V . Tenemos una aplicación

$$\begin{aligned} \varphi_\beta : \quad V &\rightarrow \mathbb{R}^n \\ \lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_n \vec{v}_n &\mapsto (\lambda_1, \lambda_2, \dots, \lambda_n) \end{aligned}$$

A φ_β la llamamos aplicación de coordenadas en la base β y $\varphi_\beta(\vec{v}) = (\lambda_1, \lambda_2, \dots, \lambda_n)$ son las *coordenadas de \vec{v} en la base β* . Al número $\lambda_i \in \mathbb{K}$ lo llamamos *coordenada i -ésima de \vec{v} en la base β* , para todo $i = 1, 2, \dots, n$.

■ **Ejemplo 2.7** Consideremos el espacio vectorial \mathbb{R}^3 y la base $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ dada por

$$\vec{v}_1 = (2, 5, 0), \vec{v}_2 = (-1, 0, 2), \vec{v}_3 = (1, 2, 1).$$

Las coordenadas de $\vec{v} = (0, 1, 1)$ en la base β son $\varphi_\beta(\vec{v}) = (0, 1, -1)$. Esto es

$$\vec{v} = 0\vec{v}_1 + 1\vec{v}_2 + (-1)\vec{v}_3.$$

■ **Ejemplo 2.8** Consideramos el subespacio vectorial $P \subset \mathbb{R}^3$ determinado por

$$P = \{(x, y, z) \in \mathbb{R}^3; x + 2y + 3z = 0\}$$

del Ejemplo 2.6 y la base $\beta = \{\vec{v}_1, \vec{v}_2\}$ allí obtenida, esto es

$$\vec{v}_1 = (-2, 1, 0), \vec{v}_2 = (-3, 0, 1),$$

tenemos que $\varphi_\beta(\vec{v}_1) = (1, 0)$, $\varphi_\beta(\vec{v}_2) = (0, 1)$ y $\varphi_\beta((1, 1, 1)) = \varphi_\beta(\vec{v}_1 + \vec{v}_2) = (1, 1)$.

Sea $C = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ un subconjunto de vectores de un espacio vectorial V de dimensión n y sea β una base de V . Podemos determinar una base de $L(C)$ como sigue:

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_k \vec{v}_k = 0$$

nos proporciona un sistema de n ecuaciones con k incógnitas en \mathbb{R}^n al tomar

$$\lambda_1 \varphi_\beta(\vec{v}_1) + \lambda_2 \varphi_\beta(\vec{v}_2) + \dots + \lambda_k \varphi_\beta(\vec{v}_k) = 0.$$

La dimensión de $L(C)$ será igual a n menos la dimensión del espacio de soluciones de este sistema (el número de parámetros). Así, por ejemplo, si el sistema tiene una única solución (necesariamente el $\vec{0}$) significa que los vectores son linealmente independientes (ocurre si $k = n$ y $V = L(C)$).

■ **Observación 2.1** Al colocar la matriz del sistema, los vectores que proporcionan una base son aquellos correspondientes a columnas en las que aparecen “pivotes” tras aplicar el método de Gauss-Jordan.

■ **Ejemplo 2.9** Determiné una base del subespacio $L(C)$ con

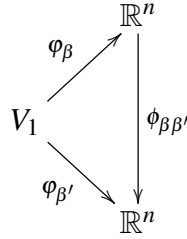
$$C = \{(1, 2, 1), (2, 1, 0), (4, 5, 2)\}.$$

2.3 Cambio de base

Consideremos ahora un \mathbb{R} -espacio vectorial V y supongamos fijas dos bases

$$\beta = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}, \quad \beta' = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n\}$$

A β la llamaremos *base antigua* y β' será la *base nueva*. Estamos interesados en describir la *aplicación de cambio de base* $\phi_{\beta\beta'} = \phi_{\beta'} \circ \phi_{\beta}^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, que visualizamos en el siguiente diagrama



y viene dada por

$$(\lambda_1, \lambda_2, \dots, \lambda_n) \mapsto \phi_{\beta'}(\vec{v}), \quad \phi_{\beta}(\vec{v}) = (\lambda_1, \lambda_2, \dots, \lambda_n).$$

Esto nos interesa, dado que cuando estudiemos transformaciones (movimientos) en un espacio vectorial, veremos que en determinadas bases éstas tendrán una expresión más sencilla, de hecho las podremos descomponer en varias transformaciones entre subespacios vectoriales de dimensiones pequeñas.

■ **Ejemplo 2.10** Sea $\beta = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ una base de un espacio vectorial real V y consideremos una nueva base $\beta' = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ dada por las relaciones

$$\vec{u}_1 = \vec{v}_1 + \vec{v}_3, \quad \vec{u}_2 = \vec{v}_2, \quad \vec{u}_3 = \vec{v}_2 + \vec{v}_3$$

Tomamos el vector $\vec{x} \in V$ dado por $\vec{x} = 3\vec{u}_1 + 2\vec{u}_2$, esto es tal que $\phi_{\beta}(\vec{x}) = (3, 2, 0)$. Para ver las coordenadas de \vec{x} en la base β' sustituimos

$$\vec{x} = 3\vec{u}_1 + 2\vec{u}_2 = 3(\vec{v}_1 + \vec{v}_3) + 2\vec{v}_2 = 3\vec{v}_1 + 2\vec{v}_2 + 3\vec{v}_3.$$

Así $\phi_{\beta, \beta'}((3, 2, 0)) = (3, 2, 3)$.

Más generalmente, dado $(\lambda_1, \lambda_2, \lambda_3)$ en \mathbb{R}^3 consideremos el vector $\vec{x} \in V$ tal que $\phi_{\beta}(\vec{x}) = (\lambda_1, \lambda_2, \lambda_3)$. Tenemos que

$$\begin{aligned} \vec{x} &= \lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 + \lambda_3 \vec{u}_3 = \lambda_1(\vec{v}_1 + \vec{v}_3) + \lambda_2 \vec{v}_2 + \lambda_3(\vec{v}_2 + \vec{v}_3) \\ &= (1\lambda_1 + 0\lambda_2 + 0\lambda_3)\vec{v}_1 + (0\lambda_1 + 1\lambda_2 + 1\lambda_3)\vec{v}_2 + (1\lambda_1 + 0\lambda_2 + 1\lambda_3)\vec{v}_3 \end{aligned}$$

Esto es

$$\phi_{\beta, \beta'}((\lambda_1, \lambda_2, \lambda_3)) = \phi_{\beta'}(\vec{x}) = (\lambda_1, \lambda_2 + \lambda_3, \lambda_1 + \lambda_3)$$

Que se corresponde con

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 + \lambda_3 \\ \lambda_1 + \lambda_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$$

Vamos a escribir las ideas del ejemplo anterior para el caso general. Escribiendo la base antigua en función de la nueva, sabemos que tenemos las relaciones

$$\begin{aligned}\vec{v}_1 &= \mu_{11}\vec{w}_1 + \mu_{21}\vec{w}_2 + \cdots + \mu_{n1}\vec{w}_n &\iff \varphi_{\beta'}(\vec{v}_1) &= (\mu_{11}, \mu_{21}, \dots, \mu_{n1}) \\ \vec{v}_2 &= \mu_{12}\vec{w}_1 + \mu_{22}\vec{w}_2 + \cdots + \mu_{n2}\vec{w}_n &\iff \varphi_{\beta'}(\vec{v}_2) &= (\mu_{12}, \mu_{22}, \dots, \mu_{n2}) \\ &\vdots \\ \vec{v}_n &= \mu_{1n}\vec{w}_1 + \mu_{2n}\vec{w}_2 + \cdots + \mu_{nn}\vec{w}_n &\iff \varphi_{\beta'}(\vec{v}_n) &= (\mu_{1n}, \mu_{2n}, \dots, \mu_{nn})\end{aligned}$$

Dado $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{K}^n$, sea \vec{v} el vector tal que $\varphi_{\beta}(\vec{v}) = (\lambda_1, \lambda_2, \dots, \lambda_n)$, esto es, tal que

$$\vec{v} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \cdots + \lambda_n\vec{v}_n$$

Sustituyendo las expresiones de cada \vec{v}_j llegamos a que

$$\vec{v} = \lambda_1 \sum_{i=1}^n \mu_{i1}\vec{w}_i + \cdots + \lambda_n \sum_{i=1}^n \mu_{in}\vec{w}_i = \sum_{j=1}^n \lambda_j \mu_{1j}\vec{w}_1 + \cdots + \sum_{j=1}^n \lambda_j \mu_{nj}\vec{w}_n$$

Esto es, si $\varphi_{\beta'}(\vec{v}) = (\mu_1, \mu_2, \dots, \mu_n)$, tenemos que

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} = M_{\beta\beta'} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}, \quad M_{\beta\beta'} = \begin{pmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{n1} & \mu_{n2} & \cdots & \mu_{nn} \end{pmatrix}$$

La matriz $M_{\beta,\beta'}$ describe la aplicación $\varphi_{\beta\beta'}$ de cambio de base y la llamamos *matriz de cambio de base de β a β'* . Obsérvese que la columna i -ésima de la matriz $M_{\beta,\beta'}$ está definida por las coordenadas del vector \vec{v}_i en la base β' .

3. Transformaciones lineales

Una transformación lineal entre dos espacios vectoriales es una aplicación que transforma subespacios vectoriales en subespacios vectoriales. En particular lleva rectas en rectas y deja fijo el origen. Podemos decir que una transformación lineal tiene la propiedad de “no curvar los objetos” y de “no trasladarlos”. De hecho, decimos que los movimientos asociados a estas transformaciones son rígidos.

3.1 Definiciones

Comenzamos considerando el plano real \mathbb{R}^2 , utilizando el conocimiento que ya tenemos sobre él. Vamos a considerar transformaciones del plano \mathbb{R}^2 en sí mismo, es decir, transformaciones $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. Podríamos pensar en

$$f(x, y) = (x, y^2),$$

pero si nos fijamos, esta manda la recta $\ell = \{(x, x) : x \in \mathbb{R}\}$ en la parábola $f(\ell) = \{(x, x^2) : x \in \mathbb{R}\}$, por tanto no va a ser un caso de transformación lineal, dado que no manda rectas en rectas. Esto nos “indica” que la expresión de f debe ser lineal (con exponentes igual a 1) en x e y . Podríamos pensar ahora en

$$f(x, y) = (x, y + 1),$$

que manda la recta $\ell = \{(x, x) : x \in \mathbb{R}\}$ en la misma recta trasladada una unidad “hacia arriba”, es decir, en $f(\ell) = \{(x, x + 1) : x \in \mathbb{R}\}$. Esta tampoco es una transformación lineal, porque no deja fijo el origen (hablaríamos de transformación afín = transformación lineal + traslación). De manera más general, una transformación lineal $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tiene una expresión

$$f(x, y) = (f_1(x, y), f_2(x, y)) = (a_{11}x + a_{12}y, a_{21}x + a_{22}y),$$

y por tanto tiene asociada a una expresión matricial

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

■ **Ejemplo 3.1** Los siguientes son ejemplos de transformaciones lineales de \mathbb{R}^2 en sí mismo:

- **Simetría:** La aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $f(x,y) = (y,x)$. Es una simetría respecto a la recta $y = x$.
- **Homotecia:** La aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $f(x,y) = (x,2y)$. Es una homotecia en las direcciones $(1,0)$ y $(0,1)$ de razones 1 y 2, respectivamente.
- **Rotación:** La aplicación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $f(x,y) = (-y,x)$. Es una rotación de $\pi/2$ radianes, es decir 90° , en sentido positivo (antihorario).

De manera más general, una *transformación lineal* entre dos \mathbb{R} -espacios vectoriales V y W es una aplicación $f : V \rightarrow W$ tal que se cumple:

$$f(\lambda \vec{v}_1 + \mu \vec{v}_2) = \lambda f(\vec{v}_1) + \mu f(\vec{v}_2),$$

para todo $\vec{v}_1, \vec{v}_2 \in V$ y $\lambda, \mu \in \mathbb{R}$. De esta definición se deduce que f envía rectas en rectas y que se cumple $f(\vec{0}) = \vec{0}$.

Proposición 3.1.1 Una transformación lineal $f : V \rightarrow W$ queda determinada por la imagen de los elementos de una base de V .

■ **Ejemplo 3.2** La transformación lineal $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por

$$f((1,0,0)) = (1,0), \quad f((0,1,0)) = (0,1), \quad f((0,0,1)) = (0,0)$$

envía a un punto $(x,y,z) \in \mathbb{R}^3$ en $(x,y) \in \mathbb{R}^2$, pues tenemos que:

$$\begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

■ **Ejemplo 3.3** Dado un número real $k \in \mathbb{R}$, la transformación $f_k : \mathbb{R}[X] \rightarrow \mathbb{R}$ que a cada polinomio $p(X)$ le hace corresponder su evaluación en $X = k$.

Sea $f : V \rightarrow W$ una transformación lineal entre espacios vectoriales:

- Llamamos *núcleo* de f y lo denotamos por $\ker f$ al subespacio vectorial de V dado por

$$\ker f = \{\vec{v} \in V : f(\vec{v}) = \vec{0}\}$$

- Llamamos *imagen* de f y lo denotamos por $\text{Im} f$ al subespacio vectorial de W dado por

$$\text{Im} f = \{\vec{w} \in W : \exists \vec{v} \in V, f(\vec{v}) = \vec{w}\}$$

■ **Ejemplo 3.4** Sea $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ la aplicación lineal dada por la proyección sobre las dos primeras componentes, esto es $f(x,y,z) = (x,y)$. Tenemos que

$$\ker f = L((0,0,1)) \quad \text{Im} f = L((1,0), (0,1)) = \mathbb{R}^2$$

■ **Ejemplo 3.5** Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ la aplicación lineal dada por $f(x, y) = (x, x, y)$. Se tiene que

$$\ker f = \{\vec{0}\} \quad \text{Im} f = L((1, 1, 0), (0, 0, 1))$$

Si $f : V \rightarrow W$ cumple que $\text{Im} f = W$ decimos que f es *sobreyectiva*. En este caso tenemos que $\dim W \leq \dim V$. Si $f : V \rightarrow W$ cumple que $\ker f = \vec{0}$, decimos que f es *inyectiva*. Se puede probar que esto equivale a decir que vectores distintos tienen imágenes distintas. En este caso tenemos que $\dim V \leq \dim W$. Cuando f es sobreyectiva e inyectiva, decimos que es *biyectiva*, y en este caso tenemos necesariamente que $\dim V = \dim W$.

■ **Ejemplo 3.6** La aplicación $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $f(x, y, z) = (-y, x, 2z)$ es biyectiva. Es una rotación del plano $z = 0$ (plano xy) junto con una homotecia en la dirección del eje z .

En todos los ejemplos que hemos introducido hasta ahora nos hemos estado apoyando en la base estándar de \mathbb{R}^n . Veamos ahora en general cómo podemos “codificar” f , utilizando coordenadas, una vez se han fijado bases

$$\beta_1 = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}, \quad \beta_2 = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_m\}$$

en V y W , respectivamente. Es decir, queremos dar la *lectura de f en las bases β_1 y β_2* . Veamos un ejemplo en primer lugar para ilustrar el concepto.

■ **Ejemplo 3.7** Consideramos el espacio vectorial \mathbb{R}^3 y los vectores

$$\vec{v}_1 = (1, 1, 0), \quad \vec{v}_2 = (0, 0, 1); \quad \vec{w}_1 = (0, 1, 0), \quad \vec{w}_2 = (0, 0, 1).$$

Llamamos $\beta_1 = \{\vec{v}_1, \vec{v}_2\}$ y $\beta_2 = \{\vec{w}_1, \vec{w}_2\}$ y consideramos los subespacios vectoriales de dimensión dos $P_1 = L(\beta_1)$ y $P_2 = L(\beta_2)$. Sea $f : P_1 \rightarrow P_2$ la aplicación lineal descrita por

$$f(\vec{v}_1) = \vec{w}_1, \quad f(\vec{v}_2) = \vec{w}_2$$

Así, la matriz de f en las bases β_1 y β_2 es

$$A_{\beta_1\beta_2}^f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Volviendo al caso general, sabemos que

$$\begin{aligned} f(\vec{v}_1) &= a_{11}\vec{w}_1 + a_{21}\vec{w}_2 + \dots + a_{m1}\vec{w}_m \\ f(\vec{v}_2) &= a_{12}\vec{w}_1 + a_{22}\vec{w}_2 + \dots + a_{m2}\vec{w}_m \\ &\vdots \\ f(\vec{v}_n) &= a_{1n}\vec{w}_1 + a_{2n}\vec{w}_2 + \dots + a_{mn}\vec{w}_m \end{aligned}$$

Dado $(\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}^n$, sea \vec{v} el vector tal que $\vec{v} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + \dots + \lambda_n\vec{v}_n$. Por ser f una aplicación lineal, tenemos que

$$f(\vec{v}) = \lambda_1 f(\vec{v}_1) + \lambda_2 f(\vec{v}_2) + \dots + \lambda_n f(\vec{v}_n)$$

Sustituyendo las expresiones de cada $f(\vec{v}_j)$ llegamos a que

$$f(\vec{v}) = \lambda_1 \sum_{i=1}^m a_{i1}\vec{w}_i + \dots + \lambda_n \sum_{i=1}^m a_{in}\vec{w}_i = \sum_{j=1}^n \lambda_j a_{1j}\vec{w}_1 + \dots + \sum_{j=1}^n \lambda_j a_{mj}\vec{w}_m$$

Esto es, si $f(\vec{v}) = \mu_1 \vec{w}_1 + \mu_2 \vec{w}_2 + \cdots + \mu_m \vec{w}_m$, tenemos que

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix} = A_{\beta_1 \beta_2}^f \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}, \quad A_{\beta_1 \beta_2}^f = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

La matriz $A_{\beta_1 \beta_2}^f$ describe la lectura de f en las bases β_1 y β_2 y la llamamos *matriz de f en las bases β_1 y β_2* . Cuando $V = W$ y $\beta_1 = \beta_2 = \beta$ escribimos $A_{\beta_1 \beta_2}^f = A_{\beta}^f$.

Notemos que la columna i -ésima de la matriz A_{β_1, β_2}^f está dada por las coordenadas del vector $f(\vec{v}_i)$ en la base β_2 .

3.2 Subespacios invariantes y cambios de base

Nosotros nos vamos a centrar a partir de ahora en transformaciones lineales que actúan sobre el mismo espacio vectorial V . El objetivo es estudiar estas transformaciones y determinar bases de V para las cuales la expresión de f sea sencilla. Cuando iteramos f obtenemos un *sistema dinámico discreto* sobre V .

Sea $f: V \rightarrow V$ una transformación lineal de V . Un subespacio vectorial $W \subset V$ se dice *invariante por f* si se tiene que $f(W) \subset W$.

■ **Ejemplo 3.8** Consideremos $V = \mathbb{R}^3$ y la base $\beta = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, donde

$$\vec{u}_1 = (1, 0, 0), \quad \vec{u}_2 = (0, 1, 1), \quad \vec{u}_3 = (0, 1, 0).$$

Tomemos la transformación lineal $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ descrita por

$$f(\vec{u}_1) = -\vec{u}_2, \quad f(\vec{u}_2) = \vec{u}_1, \quad f(\vec{u}_3) = \vec{u}_3$$

Tenemos que $W_1 = L(\vec{u}_1, \vec{u}_2)$ y $W_2 = L(\vec{u}_3)$ son subespacios vectoriales de \mathbb{R}^3 invariantes por f . De hecho f describe una rotación dentro de W_1 y es la identidad en restricción a W_2 . Tenemos que la matriz de f en la base β viene dada por

$$A_{\beta}^f = \left(\begin{array}{cc|c} 0 & 1 & 0 \\ -1 & 0 & 0 \\ \hline 0 & 0 & 1 \end{array} \right)$$

La separación en “dos bloques” que observamos es la que nos indica que podemos partir la base en dos para separar dos subespacios invariantes. En este caso por un lado \vec{u}_1 y \vec{u}_2 y por el otro \vec{u}_3 .

Notemos en el ejemplo anterior que si consideramos la base estándar de \mathbb{R}^3 dada por $\vec{e}_1 = (1, 0, 0)$, $\vec{e}_2 = (0, 1, 0)$ y $\vec{e}_3 = (0, 0, 1)$, tenemos que

$$f(\vec{e}_1) = -\vec{e}_2 - \vec{e}_3, \quad f(\vec{e}_2) = \vec{e}_2, \quad f(\vec{e}_3) = f(\vec{u}_2 - \vec{u}_3) = \vec{u}_1 - \vec{u}_3 = \vec{e}_1 - \vec{e}_2$$

Así, la lectura de f en la base estándar viene dada por

$$A_{\beta_{st}}^f = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

De esta manera, pese a que la base estándar nos resulta más familiar y nos sentimos más cómodos trabajando con ella, no nos permite distinguir la “separación de la dinámica” en esos dos bloques que sí detectamos con la otra base. Por eso es de especial importancia saber realizar los cambios de base o referencia, para poder así visualizar mejor el “comportamiento a largo plazo de f ”, es decir, nuestro sistema dinámico.

Lo que acabamos de observar para el ejemplo anterior nos sucederá frecuentemente. Nos darán la transformación lineal $f : V \rightarrow V$ mediante su lectura en una base β , es decir, nos proporcionarán la matriz A_β^f y estaremos interesados en expresar f en otra base β' , es decir, nos interesará la matriz $A_{\beta'}^f$. La relación entre ambas viene dada por

$$A_{\beta'}^f = M_{\beta\beta'} A_\beta^f M_{\beta\beta'}^{-1}.$$

donde $M_{\beta\beta'}$ es la matriz de cambio de base de β a β' , que recordemos que se obtiene al colocar las coordenadas en β' de los vectores de β por columnas.

■ **Ejemplo 3.9** Sea V un espacio vectorial de dimensión 2 y $\beta = \{\vec{u}_1, \vec{u}_2\}$ una base de V . Consideremos la transformación lineal $f : V \rightarrow V$ cuya expresión en la base β viene dada por

$$A_\beta^f = \begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix}$$

Esto es, tenemos que

$$f(\vec{u}_1) = 3/5\vec{u}_1 + 4/5\vec{u}_2, \quad f(\vec{u}_2) = 4/5\vec{u}_1 - 3/5\vec{u}_2$$

Consideremos la base $\beta' = \{\vec{w}_1, \vec{w}_2\}$, donde $\vec{w}_1 = 2\vec{u}_1 + \vec{u}_2$ y $\vec{w}_2 = \vec{u}_1 - 2\vec{u}_2$. La expresión de f en la base β' viene dada por

$$A_{\beta'}^f = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2/5 & 1/5 \\ 1/5 & -2/5 \end{pmatrix} \begin{pmatrix} 3/5 & 4/5 \\ 4/5 & -3/5 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$$

La expresión de f en la base β no deja claro cual es el movimiento sobre el plano que describe f . En cambio, cuando leemos f en la base β' de inmediato vemos que deja fijos todos los elementos de $L(\vec{w}_1)$ y por otro lado que preserva el subespacio $L(\vec{w}_2)$ enviando a cada elemento en su opuesto. De hecho f es una simetría respecto a la recta $L(\vec{w}_1)$.

3.3 Diagonalización y direcciones propias

Un caso especial de transformaciones lineales $f : V \rightarrow V$ son aquellas cuya dinámica se puede “descomponer” en n subespacios invariantes de dimensión uno, donde n es la dimensión de V . En el caso de $V = \mathbb{R}^2$ esto es lo que ocurre con homotecias en dos direcciones independientes, pero no con los giros.

Sea $f : V \rightarrow V$ una transformación lineal. Los subespacios invariantes de f que tienen dimensión 1 se denominan *direcciones propias* y sus elementos se llaman *vectores propios* o *autovectores*. Nótese que $L(\vec{v})$ es una dirección propia si y solamente si

$$f(\vec{v}) \in L(\vec{v}) \iff f(\vec{v}) = \lambda \vec{v}, \lambda \in \mathbb{R}.$$

El valor λ se denomina *autovalor* o *valor propio* de $L(\vec{v})$. El valor propio explica cómo se “contrae o expande” la dirección propia mediante la transformación f . Además, si es negativo, explica la existencia de una “simetría o reflexión”.

■ **Ejemplo 3.10** Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ determinada por $f(x, y) = (x + y, y)$. El vector $\vec{v} = (0, 1)$ es propio con autovalor asociado $\lambda = 1$, esto es $f(\vec{v}) = \vec{v}$.

Sea V un espacio vectorial de dimensión n . Sea $f : V \rightarrow V$ una transformación lineal de modo que hay una descomposición en n direcciones propias

$$L(\vec{v}_1), L(\vec{v}_2), \dots, L(\vec{v}_n)$$

tales que $f(\vec{v}_i) = \lambda_i \vec{v}_i$, para todo $i = 1, 2, \dots, n$. La expresión de f en la base $\beta = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ viene determinada por la matriz diagonal

$$A_{\beta}^f = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Diremos que una transformación lineal $f : V \rightarrow V$ es *diagonalizable* si tiene una base formada por vectores propios. Equivalentemente, si existe una base para la cual la matriz que describe la expresión de f en dicha base es diagonal.

■ **Ejemplo 3.11** La transformación lineal dada por $f(x, y) = (x + y, y)$ no es diagonalizable, pues tiene una única dirección propia generada por $\vec{v} = (0, 1)$.

Veamos ahora cómo determinar si una transformación lineal es diagonalizable y como hallar una base de vectores propios. Un vector es propio si y solo si pertenece al núcleo de la aplicación

$$f_{\lambda} = \lambda \text{id}_V - f : V \rightarrow V,$$

para algún $\lambda \in \mathbb{R}$; esto es, si $f_{\lambda}(\vec{v}) = f(\vec{v}) - \lambda \vec{v} = \vec{0}$. Llamamos *polinomio característico* $P_f(\lambda)$ de f al determinante de la matriz asociada a f_{λ} en alguna base de V . Esta matriz la podemos obtener mediante

$$A_{\beta}^{f_{\lambda}} = \lambda \text{Id}_{n \times n} - A_{\beta}^f,$$

donde n es la dimensión de V . Se llama polinomio característico porque es un polinomio en λ (al dejar λ variar en \mathbb{R}) y además porque es independiente de la base en la que se haya tomado la lectura de f . Nos interesan los valores λ para los cuales f_{λ} no es inyectiva (la dimensión de su núcleo es mayor que uno), y estos son exactamente las raíces del polinomio característico.

■ **Observación 3.1** En muchos libros de texto se define $f_{\lambda} = f - \lambda \text{id}_V$. El polinomio así obtenido coincide con el nuestro multiplicado salvo quizá por el signo, pero esto no afecta al procedimiento posterior, siempre y cuando se sea consistente a la hora de aplicarlo.

Proposición 3.3.1 Dado $\lambda \in \mathbb{R}$ se tiene $P_f(\lambda) = 0$ si y solo si existe un vector $\vec{v} \in V$ con $\vec{v} \neq \vec{0}$ tal que

$$f_{\lambda}(\vec{v}) = \lambda \vec{v} - f(\vec{v}) = \vec{0}.$$

■ **Ejemplo 3.12** Considérese la transformación lineal $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$f(x, y) = (x + 5y, 2x + 4y)$$

Tenemos que

$$A_{\beta_{\text{st}}}^f = \begin{pmatrix} 1 & 4 \\ 2 & 5 \end{pmatrix}, \quad A_{\beta_{\text{st}}}^{f_{\lambda}} = \begin{pmatrix} \lambda - 1 & -4 \\ -2 & \lambda - 5 \end{pmatrix},$$

y por tanto $P_f(\lambda) = \lambda^2 - 5\lambda - 6 = (\lambda + 1)(\lambda - 6)$. De esta manera los valores propios de f son $\lambda = -1$ y $\lambda = 6$.

Las siguientes observaciones caracterizan cuándo f es diagonalizable.

1. Si $P_f(\lambda)$ no puede descomponerse como producto de factores lineales, entonces f no es diagonalizable. Obsérvese que esto puede suceder en \mathbb{R} , pero jamás en \mathbb{C} (teorema fundamental del álgebra).
2. Suponamos que $P_f(\lambda) = (\lambda - \lambda_1)^{k_1}(\lambda - \lambda_2)^{k_2} \cdots (\lambda - \lambda_\ell)^{k_\ell}$. Tenemos que f es diagonalizable si y solamente si

$$k_i = \dim(\ker f_{\lambda_i}) = \dim V - \text{rango}(A_{\beta}^{f_{\lambda_i}}),$$

para todo $i = 1, 2, \dots, n$. En particular, si $P_f(\lambda)$ tiene n raíces distintas, es diagonalizable.

■ **Ejemplo 3.13** Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la transformación lineal dada en su expresión en la base estándar por la matriz

$$A_{\beta_{\text{st}}}^f = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$$

El polinomio característico es $P_f(\lambda) = (\lambda + 1)(\lambda - 3)$, por tanto f es una transformación lineal diagonalizable, ya que $P_f(\lambda)$ se descompone como producto de factores lineales y tiene dos raíces distintas. Tenemos que

$$\ker f_{-1} = L((-2, 1)), \quad \ker f_3 = L((2, 1)),$$

Así $\beta = \{(-2, 1), (2, 1)\}$ es una base de vectores propios de f y

$$A_{\beta}^f = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} -1/4 & 2/4 \\ 1/4 & 2/4 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}$$

■ **Ejemplo 3.14** Consideramos la transformación lineal $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por $f(x, y) = (x, x + y)$. Tenemos que

$$P_f(\lambda) = (\lambda - 1)^2, \quad \ker f_1 = L((0, 1)).$$

Como $\dim(\ker f_1) = 1 < 2$, concluimos que f no es diagonalizable.

■ **Ejemplo 3.15** Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la transformación lineal que se obtiene al realizar un giro de $\pi/2$ en sentido antihorario. Esto es $f(x, y) = (-y, x)$. El polinomio característico de f viene dado por

$$P_f(\lambda) = \lambda^2 + 1$$

y por tanto f no tiene direcciones propias reales. En cambio, si consideramos la inclusión $\mathbb{R}^2 \subset \mathbb{C}^2$ observamos dos direcciones propias complejas $L((i, 1))$ y $L((-i, 1))$ asociadas a los valores propios complejos i y $-i$, respectivamente.

■ **Ejemplo 3.16** Sea $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ la transformación lineal dada en su expresión en la base estándar por la matriz

$$A_{\beta_{\text{st}}}^f = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & -4 \\ 0 & 0 & 3 \end{pmatrix}$$

Tenemos que

$$A_{\beta_{\text{st}}}^{f_\lambda} = \begin{pmatrix} \lambda - 1 & -2 & -4 \\ -2 & \lambda - 1 & 4 \\ 0 & 0 & \lambda - 3 \end{pmatrix}$$

y por tanto el polinomio característico es $P_f(\lambda) = (\lambda + 1)(\lambda - 3)^2$. Dado que $P_f(\lambda)$ se descompone como producto de factores lineales, para ver que f es una transformación diagonalizable debemos comprobar que

$$2 = 3 - \text{rango}(A_{\beta_{\text{st}}}^{f_3}),$$

pero esto es cierto. Obtenemos que

$$\ker f_{-1} = L((1, -1, 0)), \quad \ker f_3 = L((1, 1, 0), (2, 0, 1)),$$

Así $\beta = \{(1, -1, 0), (1, 1, 0), (2, 0, 1)\}$ es una base de vectores propios de f y

$$A_{\beta}^f = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Finalmente, escribamos cómo se obtiene dicha expresión diagonal utilizando la matriz de cambio de base entre la base β_{st} y la base β . Más precisamente, vamos a escribir

$$A_{\beta}^f = M_{\beta_{\text{st}}\beta} A_{\beta_{\text{st}}}^f M_{\beta_{\text{st}}\beta}^{-1}.$$

Esto es

$$A_{\beta}^f = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 & -1 \\ 1/2 & 1/2 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 \\ 2 & 1 & -4 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Finalmente, tenemos que si $f : V \rightarrow V$ es una transformación lineal sobre un espacio vectorial real V y β es una base de V tal que la matriz de f en β es simétrica, entonces el polinomio característico $P_f(\lambda)$ tiene n raíces reales (es decir, se puede descomponer dentro de \mathbb{R} como producto de factores lineales) y además f es una transformación lineal diagonalizable.

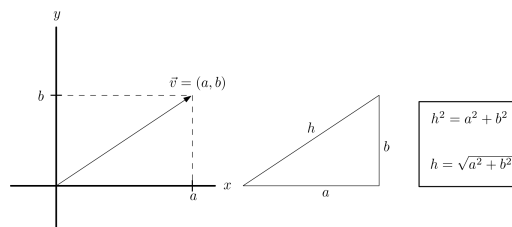
4. Espacios euclídeos

En diversas ocasiones hemos utilizado el conocimiento que ya hemos adquirido a lo largo de los años sobre el plano \mathbb{R}^2 y el espacio \mathbb{R}^3 y hemos hablado de ángulos y distancias, especialmente para construir transformaciones lineales. Al hacer esto, hemos utilizado, sin decirlo explícitamente, una “estructura” adicional sobre \mathbb{R}^2 y \mathbb{R}^3 : la estructura de espacio euclídeo.

De hecho, observemos que en la definición general de espacio vectorial no se incluye ningún concepto de medida. Por lo tanto, para hablar de la “longitud” de un vector y del “ángulo” entre dos vectores, debemos añadir a nuestro espacio vectorial una operación adicional: el producto escalar. Hablamos así de los espacios euclídeos.

4.1 Definición y primeras propiedades

En \mathbb{R}^2 tenemos la idea de longitud que nos proporciona el teorema de Pitágoras como ilustramos a continuación:



Podemos decir que la *norma* (o longitud) del vector $\vec{v} = (a, b)$ de \mathbb{R}^2 es la hipotenusa del triángulo rectángulo de catetos con longitudes a y b , esto es $\sqrt{a^2 + b^2}$. De manera general definimos la *norma euclídea* $\|\vec{v}\|_u$ de un vector $\vec{v} = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ mediante la expresión

$$\|\vec{v}\|_u = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}.$$

La norma de un vector nos proporciona una idea de su “longitud”.

En el espacio vectorial \mathbb{R}^n definimos el *producto escalar usual*

$$\langle \cdot, \cdot \rangle_u : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

$$(\vec{u}, \vec{v}) \mapsto \langle \vec{u}, \vec{v} \rangle_u$$

que a cada par de vectores $\vec{u} = (u_1, u_2, \dots, u_n)$ y $\vec{v} = (v_1, v_2, \dots, v_n)$ los envía en el número real $\langle \vec{u}, \vec{v} \rangle_u = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$. Obsérvese que se tienen las siguientes propiedades para el producto escalar usual:

- E1) (Conmutativa) $\langle \vec{u}, \vec{v} \rangle_u = \langle \vec{v}, \vec{u} \rangle_u$, para todo $\vec{u}, \vec{v} \in \mathbb{R}^n$.
- E2) (Distributiva) $\langle \vec{u}, \vec{v} + \vec{w} \rangle_u = \langle \vec{u}, \vec{v} \rangle_u + \langle \vec{u}, \vec{w} \rangle_u$, para todo $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$.
- E3) $\lambda \langle \vec{u}, \vec{v} \rangle_u = \langle \lambda \vec{u}, \vec{v} \rangle_u$, para todo $\vec{u}, \vec{v} \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$.
- E4) (Definida positiva) $\langle \vec{v}, \vec{v} \rangle_u > 0$, para todo $\vec{v} \neq 0$.

El resultado de realizar el producto escalar de un vector \vec{v} sobre sí mismo nos permite recuperar la norma euclídea, de hecho, podríamos definir la norma euclídea a partir del producto escalar mediante la relación

$$\|\vec{v}\|_u = \sqrt{\langle \vec{v}, \vec{v} \rangle_u}$$

Proposición 4.1.1 — Desigualdad de Cauchy-Schwarz. Para cada par de vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$, se tiene que

$$|\langle \vec{u}, \vec{v} \rangle_u| \leq \|\vec{u}\|_u \cdot \|\vec{v}\|_u,$$

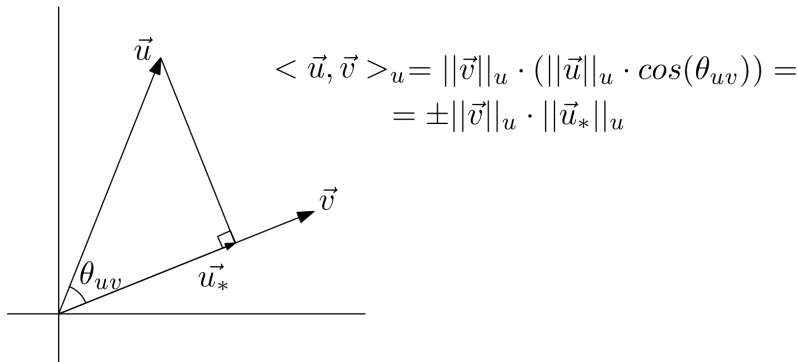
o equivalentemente

$$-1 \leq \frac{\langle \vec{u}, \vec{v} \rangle_u}{\|\vec{u}\|_u \cdot \|\vec{v}\|_u} \leq 1.$$

La desigualdad de Cauchy-Schwarz nos lleva a definir el *ángulo* θ_{uv} entre dos vectores \vec{u} y \vec{v} como aquel en el intervalo $[0, \pi]$ (o entre 0° y 180°) tal que

$$\cos \theta_{uv} = \frac{\langle \vec{u}, \vec{v} \rangle_u}{\|\vec{u}\|_u \cdot \|\vec{v}\|_u},$$

que coincide con la idea de ángulo que usualmente conocemos. De hecho, podemos describir el producto escalar usual de dos vectores como el producto de la longitudes de uno de esos vectores con la proyección del otro sobre el primero. La siguiente imagen ilustra esta relación:



Proposición 4.1.2 — Teorema de Pitágoras. Para cualquier par de vectores \vec{u}, \vec{v} de \mathbb{R}^n tales que $\theta_{uv} = \pi/2 = 90^\circ$ se tiene que

$$\|\vec{u} + \vec{v}\|_u^2 = \|\vec{u}\|_u^2 + \|\vec{v}\|_u^2.$$

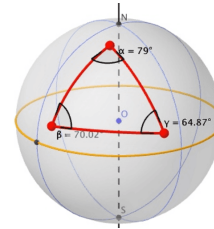
Demostración. Utilizando las propiedades [E1-E4] tenemos que

$$\begin{aligned} \|\vec{u} + \vec{v}\|_u^2 &= \langle \vec{u} + \vec{v}, \vec{u} + \vec{v} \rangle_u = \langle \vec{u}, \vec{u} \rangle_u + \langle \vec{u}, \vec{v} \rangle_u + \langle \vec{v}, \vec{u} \rangle_u + \langle \vec{v}, \vec{v} \rangle_u = \\ &= \|\vec{u}\|_u^2 + \|\vec{v}\|_u^2 + 2\langle \vec{u}, \vec{v} \rangle_u. \end{aligned}$$

Por otro lado, dado que $\langle \vec{u}, \vec{v} \rangle_u = \|\vec{u}\|_u \cdot \|\vec{v}\|_u \cos \theta_{uv}$ y $\cos \pi/2 = 0$, obtenemos que $\langle \vec{u}, \vec{v} \rangle_u = 0$ y llegamos al resultado. ■

Ahora, y de forma completamente general, un *espacio euclídeo* es un \mathbb{K} -espacio vectorial V junto con una operación de producto escalar $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ cumpliendo las propiedades [E1-E4]. Dado cualquier producto escalar se puede definir su norma asociada $\|\cdot\|$, se cumple para él la desigualdad de Cauchy-Schwarz y se puede hablar del ángulo θ_{uv} entre dos vectores $\vec{u}, \vec{v} \in V$. Asimismo, en cualquier espacio euclídeo se satisface el teorema de Pitágoras y se satisface el V postulado de Euclides (de ahí el nombre a estos espacios), que garantiza que la suma de los ángulos de un triángulo es 180° .

Curiosidad: La suma de los ángulos de un triángulo esférico (un triángulo construido sobre una esfera) ¡es estrictamente mayor a 180° ! La esfera es un ejemplo de espacio no euclídeo.¹



Por esto también ocurre que los “mapas planos” de la tierra no reflejan bien la realidad (podemos usar google maps para comprobarlo).

■ **Ejemplo 4.1** Incluso dentro del propio plano \mathbb{R}^2 podemos considerar otros productos escalares distintos al usual, como por ejemplo la aplicación que envía a cada par de vectores $\vec{u} = (u_1, u_2)$ y $\vec{v} = (v_1, v_2)$ de \mathbb{R}^2 en el número

$$(u_1, u_2) \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

que es un producto escalar sobre \mathbb{R}^2 distinto del usual. Para éste tenemos que:

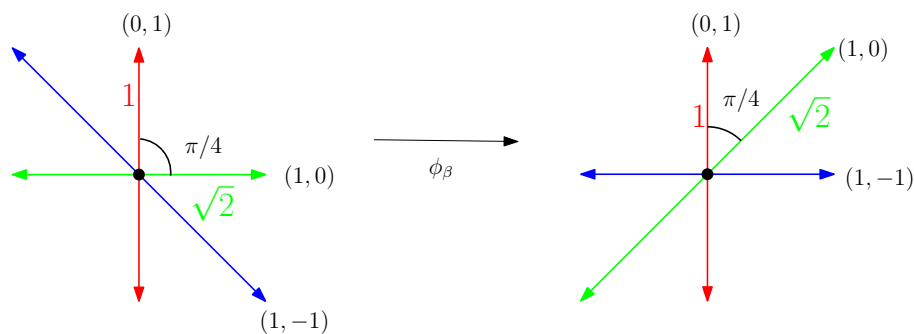
$$\|(1, 0)\| = \sqrt{2}, \quad \|(0, 1)\| = 1, \quad \theta_{(1,0),(0,1)} = \pi/4 = 45^\circ$$

\mathbb{R}^2 con el producto escalar

$$(u_1, u_2) \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

\mathbb{R}^2 con el producto escalar usual

(como espacio modelo)



¹La imagen se ha extraído de <https://www.geogebra.org/m/XcxqGfb3>

Nos puede parecer que este producto escalar es muy raro, que mide las distancias y los ángulos de una manera muy extraña, pero si consideramos la base de \mathbb{R}^2 dada por $\beta = \{(1, -1), (0, 1)\}$, rápidamente observamos que

$$(u_1, u_2) \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \langle \phi_\beta(\vec{u}), \phi_\beta(\vec{v}) \rangle_u,$$

donde $\vec{u} = (u_1, u_2)$ y $\vec{v} = (v_1, v_2)$, pues

$$\vec{x} = (x_1, x_2) = x_1(1, -1) + (x_1 + x_2)(0, 1),$$

y por tanto $\phi_\beta(\vec{x}) = (x_1, x_1 + x_2)$. Esto lo que no está diciendo es que se ha deformado el sistema de referencia (ya no es la base estándar de \mathbb{R}^2 habitual, sino la base β). Veremos más adelante formas de cambiar de referencia (de base) sin deformar distancias ni ángulos, es decir, respetando el producto escalar fijado.

La existencia de espacios euclídeos raros permite hablar de longitudes y ángulos en lugares que a priori poco tienen que ver con la geometría.

4.2 Bases ortonormales

Consideremos un \mathbb{R} -espacio vectorial V dotado de una operación de producto escalar $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$, esto es, un espacio euclídeo. A modo de guía, podemos tener en mente \mathbb{R}^n con el producto escalar usual $\langle \cdot, \cdot \rangle_u$.

Decimos que un vector $\vec{v} \in V$ es un *vector unitario* si tiene longitud uno, es decir si $\|\vec{v}\| = 1$. Llamamos *normalización* de un vector $\vec{v} \in V$ al único vector unitario $\vec{v}_{\text{un}} \in V$ que forma 0° con él (¡ojo! 0° , no 180°). Este viene dado por

$$\vec{v}_{\text{un}} = \frac{\vec{v}}{\|\vec{v}\|}.$$

Dados dos vectores $\vec{u}, \vec{v} \in V$, decimos que son *ortogonales* (o *perpendiculares*) si el ángulo θ_{uv} entre ellos es de 90° , es decir, si $\langle \vec{u}, \vec{v} \rangle = 0$ y decimos que son *ortonormales* si además de ser ortogonales ambos son unitarios.

Notemos que si $S = \{v_1, v_2, \dots, v_k\} \subset V$ es un conjunto de vectores ortogonales dos a dos, entonces es un conjunto de vectores linealmente independientes. En efecto, si escribimos

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0},$$

utilizando las propiedades [E1-E4] del producto escalar, tenemos que

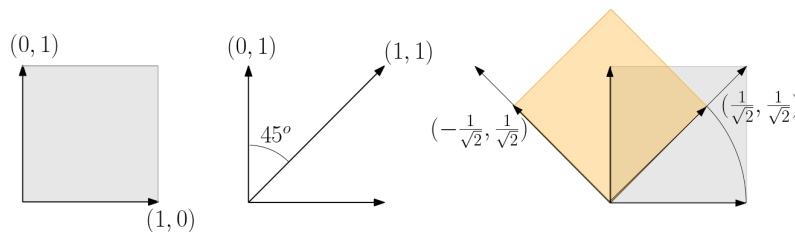
$$\left\langle \sum_{i=1}^k \alpha_i \vec{v}_i, \sum_{i=1}^k \alpha_i \vec{v}_i \right\rangle = \sum_{i=1}^k \alpha_i^2 \|\vec{v}_i\|^2 + 2 \sum_{i < j} \alpha_i \alpha_j \langle \vec{v}_i, \vec{v}_j \rangle = \sum_{i=1}^k \alpha_i^2 \|\vec{v}_i\|^2 = 0.$$

y dado que $\alpha_i^2 \geq 0$ y $\|\vec{v}_i\| > 0$ para todo $i = 1, 2, \dots, k$ tenemos que necesariamente $\alpha_i = 0$ para todo $i = 1, 2, \dots, k$. Por tanto $L(S) = L(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$ es un subespacio vectorial de V de dimensión k .

Definición 4.2.1 Sea V un \mathbb{R} -espacio vectorial. Una base $\beta = \{v_1, v_2, \dots, v_n\}$ de V es *ortogonal* si sus vectores son ortogonales dos a dos. Si además cada uno de ellos es un vector unitario, decimos que la base es *ortonormal*.

■ **Ejemplo 4.2** Consideremos \mathbb{R}^2 con el producto escalar usual:

1. El ejemplo más sencillo de base ortogonal (y ortonormal, de hecho) es la base estándar $\beta_{\text{st}} = \{(1, 0), (0, 1)\}$.
2. La base $\beta_1 = \{(1, 1), (0, 1)\}$ no es una base ortogonal, pues el ángulo entre esos dos vectores es de 45° .
3. La base $\beta_2 = \{(1, 1), (-1, 1)\}$ es ortogonal, pero no ortonormal, puesto que $\|(1, 1)\| = \|(-1, 1)\| = \sqrt{2}$. Si normalizamos estos dos vectores obtenemos la base ortonormal $\beta_3 = \left\{ \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) \right\}$.



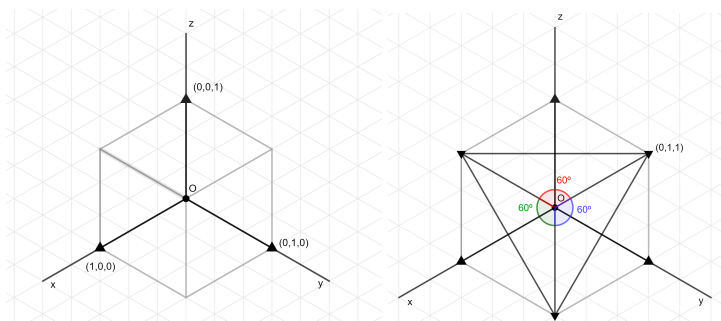
Podemos observar que hacer un cambio de base entre bases ortonormales no modifica formas, ángulos, longitudes y áreas. La figura refleja una imagen bastante fiel de lo que sucede (el cuadrado de área 1 va a parar a un cuadrado de área 1) al cambiar de la base β_{st} a la base β_3 . La matriz de cambio de base de β_{st} a la base β_3 viene dada por

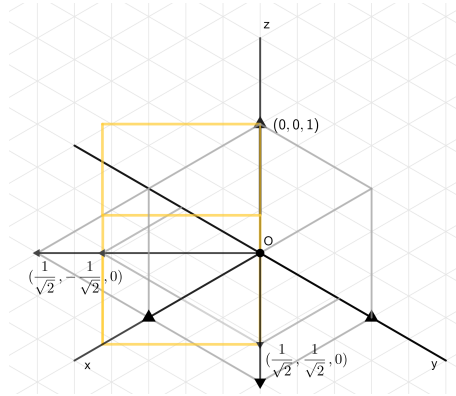
$$M_{\beta_{\text{st}}\beta_3} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix},$$

y observemos que $M_{\beta_{\text{st}}\beta_3}^{-1} = M_{\beta_3\beta_{\text{st}}} = M_{\beta_{\text{st}}\beta_3}^t$

■ **Ejemplo 4.3** Consideremos \mathbb{R}^3 con el producto escalar usual:

1. El ejemplo más sencillo de base ortogonal (y ortonormal, de hecho) es la base estándar $\beta_{\text{st}} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.
2. La base $\beta_1 = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$ no es una base ortogonal, pues el ángulo entre cada dos de estos vectores es de 60° .
3. La base $\beta_2 = \{(1, 1, 0), (1, -1, 0), (0, 0, 1)\}$ es ortogonal, pero no ortonormal, pues $\|(1, 1, 0)\| = \|(1, -1, 0)\| = \sqrt{2}$. Si normalizamos estos dos vectores obtenemos la base ortonormal $\beta_3 = \left\{ \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0 \right), \left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0 \right), (0, 0, 1) \right\}$.





De nuevo observamos que hacer un cambio de base entre bases ortonormales no modifica formas, ángulos, longitudes, áreas y volúmenes. La figura refleja este hecho si tomamos en cuenta la perspectiva, de nuevo, el cubo de volumen 1 va a parar a un cubo de volumen 1 al cambiar de la base β_{st} a la base β_3 . La matriz de cambio de base de β_{st} a la base β_3 viene dada por

$$M_{\beta_{st}\beta_3} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

y observemos que $M_{\beta_{st}\beta_3}^{-1} = M_{\beta_3\beta_{st}} = M_{\beta_{st}\beta_3}^t$

Lo que hemos observado en los ejemplos ocurre en general y la matriz de cambio de base $M_{\beta\beta'}$ entre dos bases ortonormales β y β' cumple que

$$M_{\beta\beta'}^{-1} = M_{\beta\beta'}^t$$

Decimos que es una *matriz ortogonal* (¡OJO! Que no nos confunda el nombre, las bases son de vectores unitarios).

4.3 Proyección ortogonal

Las propiedades de perpendicularidad nos resultan útiles para hablar de proyecciones y así quizá poder trabajar en dimensiones más pequeñas.

Sea W un subespacio vectorial de V . Diremos que un vector $\vec{v} \in V$ es *ortogonal* a W si lo es a cada uno de sus vectores. Usaremos la notación $\vec{v} \perp W$. Para ver si un vector es ortogonal a un subespacio, basta tomar una base de W y comprobar que \vec{v} es ortogonal a cada uno de los vectores de la base.

■ **Ejemplo 4.4** Consideremos \mathbb{R}^3 con el producto escalar usual. El vector $(-1, 1, 0)$ es ortogonal al subespacio

$$W = \{(x, y, z) \in \mathbb{R}^3; x = y\}$$

Para ello basta ver que es ortogonal a los vectores $(1, 1, 0)$ y $(0, 0, 1)$, dado que estos determinan una base de W . Tenemos que

$$\langle (1, 1, 0), (-1, 1, 0) \rangle_u = 0, \quad \langle (0, 0, 1), (-1, 1, 0) \rangle_u = 0.$$

Proposición 4.3.1 Dado un vector $\vec{v} \in V$ y un subespacio vectorial $W \subset V$, existe un único $\vec{w} \in W$ tal que $\vec{v} - \vec{w}$ sea ortogonal a W .

Demostración. Supongamos que \vec{w}_1, \vec{w}_2 son dos vectores en W para los que se tiene $\vec{v} - \vec{w}_1 \perp W$ y $\vec{v} - \vec{w}_2 \perp W$. Observemos que

$$\langle \vec{v} - \vec{w}_i, \vec{w}_i \rangle = \langle \vec{v}, \vec{w}_i \rangle - \langle \vec{w}_i, \vec{w}_i \rangle = 0, \quad i = 1, 2.$$

Por tanto $\langle \vec{v}, \vec{w}_i \rangle = \|\vec{w}_i\|^2$, para $i = 1, 2$. Por otro lado

$$\langle \vec{v} - \vec{w}_1, \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_2 \rangle - \langle \vec{w}_1, \vec{w}_2 \rangle, \quad \langle \vec{v} - \vec{w}_2, \vec{w}_1 \rangle = \langle \vec{v}, \vec{w}_1 \rangle - \langle \vec{w}_2, \vec{w}_1 \rangle.$$

De aquí se deduce que $\|\vec{w}_1\|^2 = \langle \vec{v}, \vec{w}_1 \rangle = \langle \vec{w}_1, \vec{w}_2 \rangle = \langle \vec{v}, \vec{w}_2 \rangle = \|\vec{w}_2\|^2$. Por tanto $\|\vec{w}_1\| = \|\vec{w}_2\|$ y $\cos \theta_{\vec{w}_1, \vec{w}_2} = 1$, es decir \vec{w}_1 y \vec{w}_2 son vectores paralelos que forman un ángulo de 0° (van en el mismo sentido) y tienen la misma longitud. Necesariamente $\vec{w}_1 = \vec{w}_2$. ■

La proposición anterior nos motiva a introducir la siguiente definición. Dado un vector $\vec{v} \in V$ y un subespacio vectorial $W \subset V$, definimos la *proyección ortogonal de \vec{v} sobre W* como el único vector $\text{pr}_W(\vec{v}) \in W$ tal que

$$\vec{v} - \text{pr}_W(\vec{v}) \perp W$$

Obsérvese que si \vec{v} ya es ortogonal a W , entonces $\text{pr}_W(\vec{v}) = \vec{0}$.

■ **Ejemplo 4.5** Consideremos el espacio \mathbb{R}^3 con el producto escalar usual:

1. La proyección ortogonal de $(1, 1, 1)$ sobre el plano $P = \{(x, y, z) : z = 0\}$ es el vector $(1, 1, 0)$. En efecto, tenemos que $(1, 1, 1) - (1, 1, 0) = (0, 0, 1)$, que es ortogonal a P .
2. La proyección ortogonal de $(1, 0, 0)$ sobre el plano $P = \{(x, y, z) : x = y\}$ es el vector $(1/2, 1/2, 0)$. En efecto, tenemos que

$$(1, 0, 0) - (1/2, 1/2, 0) = (1/2, -1/2, 0),$$

que es ortogonal a P .

Nos gustaría tener un método más efectivo para calcular las proyecciones. Se puede probar que si $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k\}$ es una **base ortonormal** de W , entonces

$$\text{pr}_W(\vec{v}) = \langle \vec{v}, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{v}, \vec{u}_2 \rangle \vec{u}_2 + \dots + \langle \vec{v}, \vec{u}_k \rangle \vec{u}_k.$$

Así, surge de manera natural la pregunta de si siempre existe una base ortonormal para un espacio vectorial y cómo obtenerla. Esto es lo que asegura el método de ortogonalización de Gramm-Schmidt.

4.4 Método de ortogonalización de Gramm-Schmidt

Proposición 4.4.1 Todo espacio euclídeo admite una base ortonormal.

La prueba de este resultado es algorítmica y se basa en el *método de ortogonalización de Gramm-Schmidt* que describimos a continuación:

Sea $\beta = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ una base cualquiera de V y sea $L_k = L(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$ para todo k . Vamos a ir construyendo nuestra base ortonormal

$$\beta' = \{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n\}$$

1. Normalizamos el primer vector $\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|}$. Observemos que $L_1 = L(\vec{w}_1)$.
2. Fabricamos un vector ortonormal a \vec{w}_1 tomando

$$\vec{z}_2 = \vec{v}_2 - \text{pr}_{L_1}(\vec{v}_2) = \vec{v}_2 - \langle \vec{w}_1, \vec{v}_2 \rangle \vec{w}_1, \quad \vec{w}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|}.$$

...

- k+1. Supongamos que tenemos construidos (iterando el proceso anterior) vectores ortonormales $\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k$ tales que $L_k = L(\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k)$. Construimos \vec{w}_{k+1} mediante

$$\vec{z}_{k+1} = \vec{v}_{k+1} - \text{pr}_{L_k}(\vec{v}_{k+1}) = \vec{v}_{k+1} - \sum_{i=1}^k \langle \vec{w}_i, \vec{v}_{k+1} \rangle \vec{w}_i, \quad \vec{w}_{k+1} = \frac{\vec{z}_{k+1}}{\|\vec{z}_{k+1}\|}.$$

Veamos el método de Gramm-Schmidt mediante un ejemplo.

■ **Ejemplo 4.6** Consideremos \mathbb{R}^3 con el producto escalar usual y el subespacio vectorial dado por $P = \{(x, y, z) \in \mathbb{R}^3 : x - y - z = 0\} \subset \mathbb{R}^3$. Buscamos una base ortonormal de P . De entrada sabemos que

$$P = \{\lambda(1, 1, 0) + \mu(1, 0, 1), \quad \lambda, \mu \in \mathbb{R}\},$$

y por tanto $\beta = \{\vec{v}_1, \vec{v}_2\}$ con $\vec{v}_1 = (1, 1, 0)$ y $\vec{v}_2 = (1, 0, 1)$ es una base de P , pero que no es ortonormal y tan siquiera ortogonal. Vamos a usar el método de Gramm-Schmidt para hallar una base β' que sí lo sea. Tenemos que $\beta' = \{\vec{w}_1, \vec{w}_2\}$. Buscamos \vec{w}_1 :

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{(1, 1, 0)}{\sqrt{2}} = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)$$

Buscamos \vec{w}_2 :

$$\begin{aligned} \vec{z}_2 &= \vec{v}_2 - \text{pr}_{L(\vec{w}_1)}(\vec{v}_2) = (1, 0, 1) - \langle (1, 0, 1), \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right) \rangle \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right) = \\ &= (1, 0, 1) - \frac{\sqrt{2}}{2} \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right) = \left(\frac{1}{2}, -\frac{1}{2}, 1\right). \end{aligned}$$

Ahora

$$\vec{w}_2 = \frac{\vec{z}_2}{\|\vec{z}_2\|} = \frac{\left(\frac{1}{2}, -\frac{1}{2}, 1\right)}{\frac{\sqrt{6}}{2}} = \left(\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right)$$

Notemos que

$$\begin{aligned} \vec{v}_1 &= (1, 1, 0) = \sqrt{2} \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right) = \sqrt{2} \vec{w}_1 \\ \vec{v}_2 &= (1, 0, 1) = \frac{\sqrt{2}}{2} \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right) + \frac{\sqrt{6}}{2} \left(\frac{\sqrt{6}}{6}, -\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}\right) = \frac{\sqrt{2}}{2} \vec{w}_1 + \frac{\sqrt{6}}{2} \vec{w}_2 \end{aligned}$$

Así, la aplicación de cambio de base de β a β' viene dada por la matriz

$$M_{\beta\beta'} = \begin{pmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{6}}{2} \end{pmatrix}.$$

Por cómo viene dado el proceso de ortogonalización de Gramm-Schmidt tenemos que $M_{\beta\beta'}$ es una matriz triangular superior siempre.

4.5 Transformaciones ortogonales

Dedicamos una última sección al estudio de transformaciones ortogonales. Estas transformaciones conservan el producto escalar y como consecuencia ángulos, distancias, áreas, volúmenes, etc.

Sea V un espacio vectorial con una operación de producto escalar $\langle \cdot, \cdot \rangle$. Una transformación lineal $f : V \rightarrow V$ decimos que es una *transformación ortogonal* si

$$\langle f(\vec{u}), f(\vec{v}) \rangle = \langle \vec{u}, \vec{v} \rangle, \quad \vec{u}, \vec{v} \in V.$$

Fijada una base ortonormal $\beta = \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ de V , tenemos que

$$(A_\beta^f)^{-1} = (A_\beta^f)^t,$$

es decir, la matriz de f en la base β es una matriz ortogonal. En particular, la imagen de los vectores de β dan de nuevo una base ortonormal de V . Esto es

$$\beta' = \{f(\vec{v}_1), f(\vec{v}_2), \dots, f(\vec{v}_n)\}$$

es una base ortonormal de V . Recíprocamente, sea $f : V \rightarrow V$ una transformación lineal y β una base ortonormal de V . Si la matriz de f en la base V es ortogonal, entonces f es una transformación ortogonal.

■ **Ejemplo 4.7** La transformación $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por $f(x, y, z) = (x, -z, y)$ es ortogonal. Envía la base estándar $\beta_{\text{st}} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ en

$$f(\vec{e}_1) = \vec{e}_1, \quad f(\vec{e}_2) = \vec{e}_3, \quad f(\vec{e}_3) = -\vec{e}_2$$

que determina de nuevo una base ortogonal. De otra manera, vemos que la matriz

$$A = A_{\beta_{\text{st}}}^f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

es una matriz ortogonal. Esto es, se tiene que $A^{-1} = A^t$.

4.6 Transformaciones ortogonales en \mathbb{R}^2

Toda transformación ortogonal de \mathbb{R}^2 es o bien una *rotación de ángulo θ* alrededor del origen o bien una *simetría axial*.

1. La rotación de ángulo θ (en sentido antihorario) tiene en cualquier base ortonormal (en particular en la base estándar) la expresión

$$A = A_{\beta_{\text{st}}}^f = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Observemos que la rotación de ángulo θ no es diagonalizable.

2. La simetría axial (de eje) respecto a una recta $\ell = L(\vec{v})$ se determina como sigue. Se considera una base ortonormal $\beta = \{\vec{v}, \vec{v}^\perp\}$, donde \vec{v}^\perp es uno de los dos vectores ortogonales a \vec{v} . Su matriz en la base β viene dada por

$$A_\beta^f = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Una simetría axial sí es diagonalizable (con valores propios ± 1).

En cualquier otra base ortonormal β' encontramos que la matriz de $A_{\beta'}^f$ es ortogonal y simétrica. Recíprocamente, si una matriz ortogonal y simétrica representa una transformación f en una base ortonormal (por ejemplo en la base estándar), entonces f es una transformación ortogonal y diagonalizable: es necesariamente una simetría axial.

■ **Observación 4.1** Una matriz A es ortogonal y simétrica si y solo si cumple que $A^2 = \text{Id}$. Esto significa que al aplicar dos veces la transformación lineal a la que representa, no nos hemos movido, recuperamos la transformación identidad.

■ **Ejemplo 4.8** La transformación $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada en su matriz en la base estándar mediante

$$A_{\beta_{\text{st}}}^f = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

es una simetría axial respecto a la recta $x = y$.

Se tiene que:

- 1) La composición de rotaciones es una rotación
- 2) La composición de dos simetrías axiales es una rotación
- 3) La composición de una rotación con una simetría axial es una simetría axial

4.7 Transformaciones ortogonales en \mathbb{R}^3

Toda transformación ortogonal de \mathbb{R}^3 en \mathbb{R}^3 es composición de una de los siguientes tipos:

1. *Rotación de ángulo θ* alrededor de un eje ℓ (eje de giro). En una base ortonormal adecuada $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ tenemos que

$$A_\beta^f = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}.$$

El eje de giro es $\ell = L(\vec{v}_1)$ y el plano que queda invariante mediante el giro es $P = L(\vec{v}_2, \vec{v}_3)$.

2. *Simetría respecto a un plano P* . En una base ortonormal $\beta = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ adecuada, tenemos que

$$A_\beta^f = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

y el plano de simetría es $P = L(\vec{v}_2, \vec{v}_3)$.

Una simetría sí es una transformación diagonalizable (con valores propios ± 1). En cualquier otra base ortonormal β' encontramos que la matriz de $A_{\beta'}^f$ es ortogonal y simétrica. Recíprocamente, si una matriz ortogonal y simétrica representa una transformación f en una base ortonormal (por ejemplo en la base estándar), entonces f es ortogonal y diagonalizable: es necesariamente una simetría (que puede ser composición de varias simetrías respecto a un plano).

■ **Ejemplo 4.9** La transformación $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada en la base estándar por

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{-\sqrt{2}}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{-\sqrt{2}}{2} & 0 \end{pmatrix}$$

es un giro con eje de giro dado por la recta de ecuaciones $z = 0$ y $x = y$ y el plano que queda invariante es el dado por $x = -y$. Detectamos el giro como sigue. Primero vemos que esta transformación es ortogonal (la matriz lo es) y que hay un único valor propio real igual a $+1$. Buscamos un vector propio unitario \vec{v} . En nuestro caso podemos tomar $\vec{v} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0)$; éste nos da el eje de giro. Ahora el plano que queda invariante es el que es ortogonal a \vec{v} .

■ **Ejemplo 4.10** La transformación $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada en la base estándar por

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

es ortogonal. Además, como la matriz es simétrica, es una simetría. Podemos ver que tiene el valor propio 1 doble, por tanto es una simetría respecto a un plano. Los vectores $\vec{u} = (1, 0, 0)$ y $\vec{v} = (0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ son vectores propios unitarios y ortogonales asociados al valor propio 1. Son los que determinan el plano de la simetría, en este caso el dado por $y = z$.

4.8 Transformaciones ortogonales generales

Sea V un espacio vectorial de dimensión n con un producto escalar $\langle \cdot, \cdot \rangle$, es decir, un espacio euclídeo, y sea $f : V \rightarrow V$ una transformación lineal. Fijemos una base ortonormal β de V y denotemos por A_{β}^f la matriz de f en V . Tenemos que f es una *transformación ortogonal diagonalizable* si y solo si A_{β}^f es ortogonal y simétrica, es decir, si se cumple

$$A_{\beta}^f A_{\beta}^f = \text{Id}_{n \times n}.$$

5. Sistemas de EDO's lineales

Las ecuaciones diferenciales están presentes en muchos ámbitos de las ciencias y las ciencias sociales, como la biología, la economía, la física y por supuesto, la química. Esto es natural, dado que las ecuaciones diferenciales representan el cambio de ciertas cantidades físicas respecto a parámetros como el tiempo, la temperatura o la presión.

Una ecuación diferencial ordinaria es una expresión que relaciona una variable independiente x , una variable dependiente $y = y(x)$ y sus derivadas sucesivas que habitualmente denotamos mediante $y^{(i)}(x)$ (o también $y'(x)$, $y''(x)$, etc.). Por ejemplo, la expresión

$$y''(x) + y'(x) + y(x) = 0$$

es una ecuación diferencial ordinaria. El *orden* de una ecuación diferencial es el término de la derivada de grado más alto en su expresión. Dada una ecuación diferencial ordinaria de orden n , decimos que una función n veces derivable $g(x)$ es una solución si $y = g(x)$ verifica la ecuación diferencial.

■ **Ejemplo 5.1** Resolvamos las siguientes ecuaciones diferenciales ordinarias:

$$a) y' = 2y, \quad b) y'' - 3y' + 2y = 0, \quad c) y'' - y = 2e^x$$

a) Es de primer orden y sus soluciones vienen dadas por $f_c(x) = ce^{2x}$, con $c \in \mathbb{R}$. En efecto, tenemos que

$$f'_c(x) = 2ce^{2x} = 2f_c(x)$$

b) Es de segundo orden, tiene coeficientes constantes y es homogénea. Su ecuación característica (pronto veremos la relación con el polinomio característico de una transformación lineal) es

$$t^2 - 3t + 2 = (t - 1)(t - 2)$$

y sus soluciones son $f_{ab}(x) = ae^x + be^{2x}$, con $a, b \in \mathbb{R}$. Tenemos que

$$f'_{ab}(x) = ae^x + 2be^{2x}, \quad f''_{ab}(x) = ae^x + 4be^{2x},$$

y, por tanto, $f''_{ab}(x) - 3f'_{ab}(x) + 2f_{ab}(x) = (a - 3a + 2a)e^x + (4b - 6b + 2b)e^{2x} = 0$.

c) Es de segundo orden, tiene coeficientes constantes y es no homogénea. La ecuación característica es $t^2 - 1$ y la solución homogénea viene dada por

$$y_{ab}^H = ae^x + be^{-x}$$

Buscamos una solución particular. No podemos usar $g(x) = \alpha e^x$, dado que verifica la ecuación homogénea. Tomamos $g(x) = \alpha x e^x$ y obtenemos

$$g'(x) = \alpha e^x + \alpha x e^x, \quad g''(x) = \alpha e^x + \alpha e^x + \alpha x e^x$$

Sustituyendo en la expresión $y'' - y = 2e^x$ llegamos a que

$$2\alpha e^x + \alpha x e^x - \alpha x e^x = 2e^x$$

y por tanto $\alpha = 1$. Así, una solución particular es $y^P = x e^x$ y las soluciones generales vienen dadas por

$$f_{ab}(x) = y_{ab}^H + y^P = ae^x + be^{-x} + x e^x.$$

Tras este breve repaso por algunos aspectos conocidos de las ecuaciones diferenciales, estamos en condiciones de introducir los sistemas de ecuaciones diferenciales. A partir de ahora tomaremos $n = 3$ por comodidad en la escritura, pero todo lo que hagamos se extiende a cualquier número de variables. Cuando cambiamos la variable dependiente $y : \mathbb{R} \rightarrow \mathbb{R}$ por una función vectorial

$$\mathbf{y} = (y_1, y_2, y_3) : \mathbb{R} \rightarrow \mathbb{R}^3,$$

es decir, pasamos de tener una sola variable dependiente a tener 3 variables dependientes, la expresión $y' = f(x, y)$ deja de ser una ecuación diferencial para convertirse en un *sistema de ecuaciones diferenciales ordinarias (de orden uno)*

$$\mathbf{y}' = \mathbf{f}(x, \mathbf{y}).$$

Más precisamente, tenemos relaciones

$$\begin{cases} y'_1 = f_1(x, y_1, y_2, y_3) \\ y'_2 = f_2(x, y_1, y_2, y_3) \\ y'_3 = f_3(x, y_1, y_2, y_3) \end{cases} \quad (5.1)$$

5.1 El caso lineal y con coeficientes constantes

Hay muy pocas ecuaciones diferenciales y sistemas de ecuaciones diferenciales que se puedan resolver de manera explícita. En este curso nos vamos a centrar en el caso de sistemas de ecuaciones diferenciales lineales ordinarias con coeficientes constantes.

Que las ecuaciones consideradas sean *lineales* (en las variables dependientes) significa que las f_i de la ecuación (5.1) son de la forma

$$f_i(x, y_1, y_2, y_3) = a_{i1}(x)y_1(x) + a_{i2}(x)y_2(x) + a_{i3}(x)y_3(x) + b_i(x),$$

para $i = 1, 2, 3$. Por otro lado, *con coeficientes constantes* indica que las $a_{ij}(x)$ son funciones constantes, es decir, $a_{ij}(x) = a_{ij} \in \mathbb{R}$. Así, los sistemas de ecuaciones diferenciales que nos interesan se expresan en forma matricial como

$$\begin{pmatrix} y_1'(x) \\ y_2'(x) \\ y_3'(x) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} y_1(x) \\ y_2(x) \\ y_3(x) \end{pmatrix} + \begin{pmatrix} b_1(x) \\ b_2(x) \\ b_3(x) \end{pmatrix}$$

Escribiremos $\mathbf{y}'(x) = \mathbf{A}\mathbf{y}(x) + \mathbf{b}(x)$ y llamaremos a \mathbf{A} la *matriz de coeficientes*. Cuando $\mathbf{b}(x) = \mathbf{0}$, esto es, cuando $b_i(x) = 0$ para $i = 1, 2, 3$, se dice que el sistema es *homogéneo*.

■ **Ejemplo 5.2** Considera los siguientes sistemas de EDO's lineales

$$a) \begin{cases} y_1' = xy_1 + e^x y_2 \\ y_2' = y_1 + 2xy_2 + x \end{cases}, \quad b) \begin{cases} y_1' = 2y_1 + 9y_2 \\ y_2' = y_1 + 2y_2 \end{cases}, \quad c) \begin{cases} y_1' = 2y_1 + 9y_2 + 10x + 1 \\ y_2' = y_1 + 2y_2 + 4 \end{cases}$$

Tenemos que:

- a) Es un sistema lineal pero no es de coeficientes constantes.
- b) Es un sistema lineal de coeficientes constantes y homogéneo.
- c) Es un sistema lineal de coeficientes constantes pero no homogéneo.

■ **Ejemplo 5.3** La función vectorial $\mathbf{f}_1(x) = (y_1, y_2) = (-3e^{-x}, e^{-x})$ es solución del sistema de ecuaciones homogéneo

$$\begin{cases} y_1' = 2y_1 + 9y_2 \\ y_2' = y_1 + 2y_2 \end{cases}$$

En efecto, vemos que

$$\begin{aligned} y_1' &= 3e^{-x} = 2(-3e^{-x}) + 9e^{-x} \\ y_2' &= -e^{-x} = -3e^{-x} + 2e^{-x} \end{aligned}$$

También es solución del sistema la función vectorial $\mathbf{f}_2(x) = (3e^{5x}, e^{5x})$.

Veamos cómo determinar las soluciones de algunos sistemas de ecuaciones lineales con coeficientes constantes. Daremos la solución completa para el caso homogéneo y algunas pinceladas rápidas respecto al caso no homogéneo.

5.1.1 El caso homogéneo.

El conjunto de soluciones de un sistema de ecuaciones diferenciales lineales con coeficientes constantes homogéneo tiene estructura de espacio vectorial real. Tenemos que si

$$\mathbf{f} = (f_1, f_2, f_3) : \mathbb{R} \rightarrow \mathbb{R}^3, \quad \mathbf{g} = (g_1, g_2, g_3) : \mathbb{R} \rightarrow \mathbb{R}^3$$

son dos soluciones de un sistema de EDO's homogéneo, entonces también lo son

$$\mathbf{f} + \mathbf{g} = (f_1 + g_1, f_2 + g_2, f_3 + g_3), \quad \lambda \mathbf{f} = (\lambda f_1, \lambda f_2, \lambda f_3),$$

donde λ es un número real.

■ **Ejemplo 5.4** Consideremos de nuevo el sistema de EDO's lineales homogéneo

$$\begin{cases} y_1' = 2y_1 + 9y_2 \\ y_2' = y_1 + 2y_2 \end{cases}$$

y dos de sus soluciones $\mathbf{f}_1(x) = (-3e^{-x}, e^{-x})$ y $\mathbf{f}_2(x) = (3e^{5x}, e^{5x})$. Tenemos que

$$\mathbf{f}_{ab} = a\mathbf{f}_1 + b\mathbf{f}_2 = (y_1, y_2) = (-3ae^{-x} + 3be^{5x}, ae^{-x} + be^{5x})$$

son también soluciones del sistema. En efecto

$$\begin{aligned} y_1' &= 3ae^{-x} + 15be^{5x} = 2(-3ae^{-x} + 3be^{5x}) + 9(ae^{-x} + be^{5x}) \\ y_2' &= -ae^{-x} + 5be^{5x} = (-3ae^{-x} + 3be^{5x}) + 2(ae^{-x} + be^{5x}) \end{aligned}$$

Definimos la *exponencial* de una matriz A utilizando el desarrollo de Taylor de la función exponencial, pero aplicado a la matriz. Es decir

$$e^A = \exp A = \text{Id} + A + \frac{1}{2!}A^2 + \dots + \frac{1}{n!}A^n + \dots$$

■ **Ejemplo 5.5** Calculemos la exponencial de las siguientes matrices:

$$a) D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad b) A = PDP^{-1}, \quad c) A = \begin{pmatrix} x & x \\ 0 & x \end{pmatrix}$$

a) El producto de matrices diagonales es diagonal. Más precisamente

$$D^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}, \quad D^3 = \begin{pmatrix} \lambda_1^3 & 0 \\ 0 & \lambda_2^3 \end{pmatrix}, \dots, \quad D^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$$

De esta manera, llegamos a que

$$\exp D = \begin{pmatrix} 1 + \lambda_1 + \dots + \frac{1}{n!}\lambda_1^n + \dots & 0 \\ 0 & 1 + \lambda_2 + \dots + \frac{1}{n!}\lambda_2^n + \dots \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}.$$

b) Tenemos que $A^n = (PDP^{-1})^n = PDP^{-1}PDP^{-1} \dots PDP^{-1} = PD^nP^{-1}$, para todo $n \in \mathbb{N}$ y por tanto

$$\exp A = I + PDP^{-1} + P\frac{D^2}{2!}P^{-1} + \dots + P\frac{D^n}{n!}P^{-1} = P(\exp D)P^{-1}.$$

c) Observemos que

$$A^2 = \begin{pmatrix} x^2 & 2x^2 \\ 0 & x^2 \end{pmatrix}, \quad A^3 = \begin{pmatrix} x^3 & 3x^3 \\ 0 & x^3 \end{pmatrix}, \quad \dots A^n = \begin{pmatrix} x^n & nx^n \\ 0 & x^n \end{pmatrix}$$

Por tanto

$$\exp A = \begin{pmatrix} 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots & x + x^2 + \frac{x^3}{2!} + \dots + \frac{x^{n+1}}{n!} + \dots \\ 0 & 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots \end{pmatrix} = \begin{pmatrix} e^x & xe^x \\ 0 & e^x \end{pmatrix}.$$

Estudiar el sistema de ecuaciones diferenciales $\mathbf{y}'(x) = A\mathbf{y}(x)$ se corresponde a estudiar la transformación lineal f de \mathbb{R}^3 definida por e^A . Más precisamente, el espacio vectorial de soluciones del sistema $\mathbf{y}'(x) = A\mathbf{y}(x)$ viene descrito por las combinaciones lineales de las funciones vectoriales

$$\mathbf{g}_1(x) = e^{Ax}(1, 0, 0)^T \quad \mathbf{g}_2(x) = e^{Ax}(0, 1, 0)^T \quad \mathbf{g}_3(x) = e^{Ax}(0, 0, 1)^T$$

Esto lo vemos fácilmente en el caso en que A es una matriz diagonal, esto es

$$A = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

Estamos pidiendo que $y_i'(x) = \lambda_i y_i(x)$ para $i = 1, 2, 3$, y las soluciones de estas ecuaciones son exactamente las de la forma

$$y_i(x) = ke^{\lambda_i x}, \quad k \in \mathbb{R}$$

Si tenemos fijada una *condición inicial* $\mathbf{y}(x_0) = (\mu_1, \mu_2, \mu_3)$, entonces existe un único $\vec{v} = (a, b, c) \in \mathbb{R}^3$ tal que

$$\mathbf{g}_{\vec{v}}(x) = a\mathbf{g}_1(x) + b\mathbf{g}_2(x) + c\mathbf{g}_3(x)$$

cumple que $\mathbf{g}_{\vec{v}}(x_0) = \mathbf{y}(x_0)$. Si $x_0 = 0$, de hecho tenemos que $(a, b, c) = (\mu_1, \mu_2, \mu_3)$.

■ **Ejemplo 5.6** Vamos a considerar el sistema de ecuaciones

$$\mathbf{y}'(x) = A\mathbf{y}(x), \quad A = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$$

Se nos pide que determinemos todas las soluciones de este sistema y demos la única solución que cumple la condición inicial $\mathbf{y}(0) = (1, 1)$.

Las soluciones vienen dadas por las combinaciones lineales de

$$\mathbf{g}_1(x) = \begin{pmatrix} e^{-x} & 0 \\ 0 & e^{5x} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{-x} \\ 0 \end{pmatrix}$$

$$\mathbf{g}_2(x) = \begin{pmatrix} e^{-x} & 0 \\ 0 & e^{5x} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ e^{5x} \end{pmatrix}$$

Es decir, son de la forma $\mathbf{g}_{ab} = (ae^{-x}, be^{5x})$ con $(a, b) \in \mathbb{R}^2$. La solución particular que buscamos es $\mathbf{g}_{11} = (e^{-x}, e^{5x})$

■ **Ejemplo 5.7** Una vez más, vamos a considerar el sistema de ecuaciones

$$\mathbf{y}'(x) = A\mathbf{y}(x), \quad A = \begin{pmatrix} 2 & 9 \\ 1 & 2 \end{pmatrix}$$

Se nos pide que determinemos todas las soluciones de este sistema y que demos la única solución que cumple la condición inicial $\mathbf{y}(0) = (0, 2)$.

Las transformación lineal $f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ descrita por A es diagonalizable. Su polinomio característico viene dado por

$$P_{f_A}(\lambda) = (\lambda + 1)(\lambda - 5)$$

y una base de vectores propios es $\beta = \{(-3, 1), (3, 1)\}$ con valores propios asociados -1 y 5 , respectivamente. La matriz de f_A en la base β es

$$D = A_{\beta}^{f_A} = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$$

y la matriz de cambio de base de β a β_{st} es

$$P = M_{\beta\beta_{st}} = \begin{pmatrix} -3 & 3 \\ 1 & 1 \end{pmatrix},$$

Por tanto tenemos la relación $A = PDP^{-1}$, esto es

$$\begin{pmatrix} -3 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} -3 & 3 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 9 \\ 1 & 2 \end{pmatrix}$$

y así, en vista del ejercicio anterior,

$$\begin{aligned} e^A &= Pe^DP^{-1} = \begin{pmatrix} -3 & 3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} e^{-x} & 0 \\ 0 & e^{5x} \end{pmatrix} \begin{pmatrix} -1/6 & 1/2 \\ 1/6 & 1/2 \end{pmatrix} = \\ &= \begin{pmatrix} \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} & -\frac{3}{2}e^{-x} + \frac{3}{2}e^{5x} \\ -\frac{1}{6}e^{-x} + \frac{1}{6}e^{5x} & \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} \end{pmatrix} \end{aligned}$$

De esta manera, las soluciones de $\mathbf{y}'(x) = A\mathbf{y}(x)$ vienen dadas por las combinaciones lineales de

$$\begin{aligned} \mathbf{g}_1(x) &= \begin{pmatrix} \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} & -\frac{3}{2}e^{-x} + \frac{3}{2}e^{5x} \\ -\frac{1}{6}e^{-x} + \frac{1}{6}e^{5x} & \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} \\ -\frac{1}{6}e^{-x} + \frac{1}{6}e^{5x} \end{pmatrix} \\ \mathbf{g}_2(x) &= \begin{pmatrix} \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} & -\frac{3}{2}e^{-x} + \frac{3}{2}e^{5x} \\ -\frac{1}{6}e^{-x} + \frac{1}{6}e^{5x} & \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2}e^{-x} + \frac{3}{2}e^{5x} \\ \frac{1}{2}e^{-x} + \frac{1}{2}e^{5x} \end{pmatrix} \end{aligned}$$

La solución particular que buscamos es

$$\mathbf{g}_{11}(x) = (-3e^{-x} + 3e^{5x}, e^{-x} + e^{5x}).$$

■ **Ejemplo 5.8** Consideremos el sistema de ecuaciones homogéneo

$$\mathbf{y}'(x) = A\mathbf{y}(x), \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Se nos pide que determinemos todas las soluciones de este sistema y que demos la única solución que cumple la condición inicial $\mathbf{y}(0) = (0, 2)$.

Notemos que este caso ya no es diagonalizable. Hemos calculado en un ejercicio previo la matriz exponencial de Ax . Así, una base del espacio vectorial de soluciones del sistema viene dada por

$$\mathbf{f}_1(x) = e^{Ax}(1,0)^T = \begin{pmatrix} e^x & xe^x \\ 0 & e^x \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^x \\ 0 \end{pmatrix}$$

$$\mathbf{f}_2(x) = e^{Ax}(0,1)^T = \begin{pmatrix} e^x & xe^x \\ 0 & e^x \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} xe^x \\ e^x \end{pmatrix}$$

La solución particular que buscamos viene determinada por

$$\mathbf{g}_{02}(x) = (2xe^x, 2e^x).$$

5.1.2 Resolución del caso no homogéneo:

Consideremos el sistema

$$\mathbf{y}'(x) = A\mathbf{y}(x) + \mathbf{b}(x)$$

en el caso en que $\mathbf{b}(x) \neq \mathbf{0}$. Las soluciones de este sistema se determinan sumando una solución particular \mathbf{y}^P a las soluciones del sistema homogéneo determinado por $\mathbf{y}'(x) = A\mathbf{y}(x)$ que denotamos mediante $\mathbf{y}_{\vec{v}}^H$ con $\vec{v} \in \mathbb{R}^3$. Tienen por tanto una estructura de espacio afín: son una traslación mediante \mathbf{y}^P del espacio vectorial formado por las $\mathbf{y}_{\vec{v}}^H$. Es decir, son de la forma

$$\mathbf{g}_{\vec{v}} = \mathbf{y}_{\vec{v}}^H + \mathbf{y}^P$$

Una vez más, si tenemos fijada una *condición inicial* $\mathbf{y}(x_0) = (\mu_1, \mu_2, \mu_3)$, entonces existe un único $\vec{v} = (a, b, c) \in \mathbb{R}^3$ tal que $\mathbf{g}_{\vec{v}}(x_0) = \mathbf{y}(x_0)$. Cuidado, porque aquí cuando $x_0 = 0$ no tenemos en general que $\vec{v} = (\mu_1, \mu_2, \mu_3)$.

La dificultad del caso no homogéneo reside en determinar la solución particular, ya que depende de qué tan complicada sea $\mathbf{b}(x)$. Cuando ésta es una función sencilla (senos, exponenciales, polinomios,..) es posible determinar la solución particular planteando la “función de prueba” como una del mismo tipo que $\mathbf{b}(x)$, de forma similar a como sabemos hacer para resolver EDO's de orden dos.

■ **Ejemplo 5.9** Determina la solución del sistema de ecuaciones diferenciales

$$\begin{pmatrix} y_1'(x) \\ y_2'(x) \end{pmatrix} = \begin{pmatrix} 2 & 9 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} + \begin{pmatrix} 10x+1 \\ 4 \end{pmatrix}$$

con condición inicial $\mathbf{y}(0) = (6, 5)$. Las soluciones del sistema homogéneo ya las conocemos. Vamos a buscar una solución particular. Para ello planteamos como “función de prueba” la polinomial del siguiente tipo:

$$\begin{pmatrix} P_1(x) \\ P_2(x) \end{pmatrix} = \begin{pmatrix} ax^2 + bx + c \\ dx + e \end{pmatrix}$$

Obtenemos que

$$\begin{aligned} 2ax + b &= 2(ax^2 + bx + c) + 9(dx + e) + 10x + 1 \\ d &= (ax^2 + bx + c) + 2(dx + e) + 4 \end{aligned}$$

y de aquí:

$$\begin{cases} 0 = 2a \\ 2a = 2b + 9d + 10 \\ b = 2c + 9e + 1 \end{cases} \quad \begin{cases} 0 = a \\ 0 = b + 2d \\ d = c + 2e + 4 \end{cases}$$

con lo que llegamos a $a = 0$, $b = 4$, $c = -12$, $d = -2$ y $e = 3$, así

$$\mathbf{y}^P = (4x - 12, -2x + 3).$$

Ahora, las soluciones generales de este sistema son las del tipo

$$\mathbf{g}_{ab} = \left(\frac{a-3b}{2}e^{-x} + \frac{a+3b}{2}e^{5x} + 4x - 12, \frac{-a+3b}{6}e^{-x} + \frac{a+3b}{6}e^{5x} - 2x + 3 \right).$$

Al sustituir en la condición inicial llegamos a las relaciones

$$a - 12 = 6, \quad b + 3 = 5,$$

por tanto $a = 18$ y $b = 2$. Así la solución buscada es

$$\mathbf{g}_{18,2} = (6e^{-x} + 12e^{5x} + 4x - 12, -2e^{-x} + 4e^{5x} - 2x + 3).$$

5.2 Sistemas de ecuaciones diferenciales y EDO's de orden superior

Cuando hablamos de sistemas de ecuaciones diferenciales, normalmente solo consideramos sistemas de orden uno. Esto se debe a que cualquier otro sistema que considere ecuaciones diferenciales de orden superior se puede reducir a uno de los anteriores añadiendo variables adicionales.

Por ejemplo, consideremos la ecuación diferencial de orden dos

$$y'' - 3y' + 2y = 0.$$

Llamando $y_1 = y$, y añadiendo una nueva variable $y_2 = y_1' = y'$, el problema de resolver la EDO de orden dos es equivalente al de resolver el sistema de ecuaciones

$$\begin{cases} y_1' = y_2 \\ y_2' = 3y_2 - 2y_1 \end{cases}$$

La matriz de coeficientes de este sistema es

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$$

El polinomio característico de la transformación lineal asociada a A es

$$P_{f_A}(\lambda) = \lambda(\lambda - 3) + 2 = \lambda^2 - 3\lambda + 2,$$

que coincide con llamada *ecuación característica* de la ecuación diferencial de orden dos. Si resolvemos este sistema con las pautas del ejemplo 5.7 obtenemos

$$\mathbf{g}_{ab} = ((2a - b)e^x + (-a + b)e^{2x}, (2a - b)e^x + 2(-a + b)e^{2x})$$

y por tanto, $y = y_1 = Ae^x + Be^{2x}$, con $A, B \in \mathbb{R}$, que es lo que ya sabíamos.

■ **Ejemplo 5.10 — Cuerpos en caída y resistencia del aire.** A un cuerpo en caída libre en el vacío solo le afecta la gravedad, pero en general, cuando no se encuentra en el vacío, el aire ejerce una resistencia al movimiento en la caída del cuerpo. Así, el peso del cuerpo, dado por $F_g = mg$, donde m es su masa y g es la constante de gravedad, es una fuerza que actúa en la dirección de la caída, mientras que la resistencia del aire actúa en la dirección opuesta. La fuerza F_r dada por la resistencia del aire se denomina *amortiguamiento viscoso* y es proporcional a la velocidad del cuerpo, esto es, $F_r = kv(t)$, donde k es una constante que depende del cuerpo y $v(t)$ es la velocidad del cuerpo. La segunda ley del movimiento de Newton determina que la suma de las fuerzas ejercidas sobre el cuerpo es igual a su masa m por su aceleración $a(t)$. De esta manera, la distancia recorrida $s(t)$ por el cuerpo a tiempo t viene dada por la solución a la ecuación diferencial

$$m \frac{ds(t)^2}{dt^2} = -k \frac{ds(t)}{dt} + mg,$$

o de manera equivalente, como solución al sistema de ecuaciones diferenciales

$$\begin{cases} s'(t) = v(t) \\ v'(t) = -\frac{k}{m}v(t) + g, \end{cases}$$

5.3 Sistemas dinámicos discretos y continuos

Recordemos que una transformación lineal $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ se puede iterar indefinidamente (a veces también hacia atrás) y determina un *sistema dinámico discreto*; es decir, cada iteración marca un “tiempo entero”. La *órbita* $\mathcal{O}_{\vec{v}}^f$ de un vector $\vec{v} \in \mathbb{R}^n$ es la colección de puntos del espacio que recorreremos al iterar \vec{v} . Así

$$\mathcal{O}_{\vec{v}}^f = \{\dots, f^{-1}(\vec{v}), \vec{v}, f(\vec{v}), f^2(\vec{v}), \dots, f^n(\vec{v}), \dots\}.$$

Un sistema de ecuaciones diferenciales de orden uno $\mathbf{y}' = \mathbf{f}(x, \mathbf{y})$, que también podemos llamar *campo de vectores* χ , nos proporciona en cambio un *sistema dinámico continuo*; es decir, la variable independiente x , determina un “tiempo real” y por tanto continuo. Sea $\mathbf{f}(x)$ la solución del sistema para la que se cumple la condición inicial $\mathbf{y}(x_0) = \vec{v} \in \mathbb{R}^n$. Podemos hablar de la *órbita*

$$\mathcal{O}_{\vec{v}}^\chi = \{\mathbf{f}(x) : x \in \mathbb{R}\},$$

es decir, la curva de \mathbb{R}^n que define la función \mathbf{f} . Veamos un caso muy concreto.

Sea χ el campo de vectores dado por el sistema de ecuaciones diferenciales

$$\mathbf{y}' = A\mathbf{y}, \quad A = \begin{pmatrix} \ln 2 & 0 \\ 0 & \ln 2 \end{pmatrix}$$

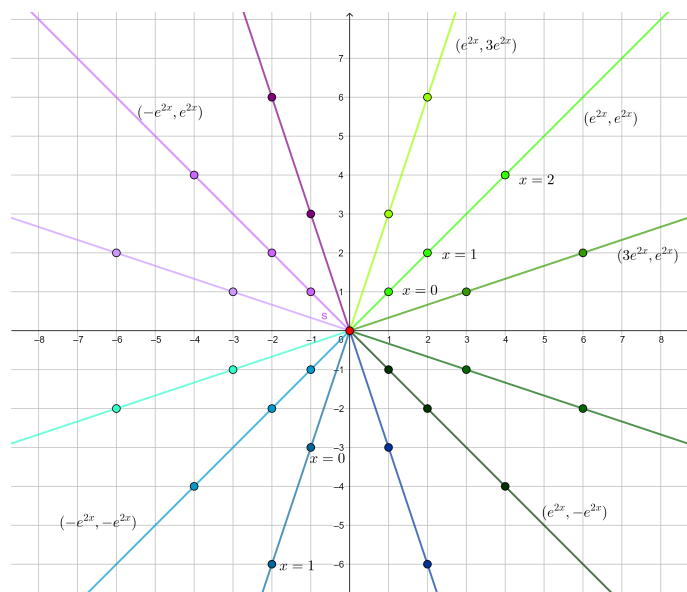
cuyas soluciones son de la forma $\mathbf{f}_{ab}(x) = (ae^{\ln(2)x}, be^{\ln(2)x}) = (a2^x, b2^x)$, para todo $(a, b) \in \mathbb{R}^2$. Por otro lado, sea f la transformación lineal dada por la exponencial

$$e^A = \begin{pmatrix} e^{\ln 2} & 0 \\ 0 & e^{\ln 2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Esto es $f(y_1, y_2) = (2y_1, 2y_2) = 2(y_1, y_2)$. Notemos que esta transformación es un múltiplo de la identidad y por tanto deja fijas todas las direcciones. Tenemos que

$$\begin{aligned} \mathcal{O}_{ab}^f &= \{\dots, (1/2a, 1/2b), (a, b), (2a, 2b), (4a, 4b), \dots, (2^n a, 2^n b), \dots\} = \\ &= \{\dots, \mathbf{f}_{ab}(-1), \mathbf{f}_{ab}(0), \mathbf{f}_{ab}(1), \mathbf{f}_{ab}(2), \dots, \mathbf{f}_{ab}(n), \dots\} \subset \mathcal{O}_{\vec{v}}^{\mathcal{X}}. \end{aligned}$$

Es decir, las soluciones a tiempo entero de la ecuación diferencial (el sistema dinámico continuo) proporcionan el mismo sistema dinámico discreto que el dado por la transformación lineal descrita por la exponencial de A . Este fenómeno es completamente general. Gráficamente lo vemos como sigue:



Notemos también que conforme x va siendo más pequeño (se acerca a $-\infty$), todas las órbitas se aproximan al $(0, 0)$. Por otro lado, el $(0, 0)$ es un punto fijo, esto es $\mathcal{O}_{00}^f = \{(0, 0)\} = \mathcal{O}_{00}^{\mathcal{X}}$. Diríamos que $(0, 0)$ es una *singularidad* del campo de vectores o un *punto fijo* del sistema dinámico.



Cálculo de varias variables

6	Funciones de varias variables	53
6.1	Funciones de varias variables. Definición	
6.2	Funciones de dos variables: Curvas de nivel	
6.3	Funciones en tres variables: Superficies de nivel	
6.4	Representación gráfica: Superficies de nivel	
6.5	Coordenadas: polares, cilíndricas, esféricas.	
6.6	Parametrización de curvas	
7	Cálculo diferencial	67
7.1	Derivadas parciales	
7.2	Derivadas direccionales	
7.3	Diferenciabilidad. Plano tangente	
7.4	Propiedades del gradiente. Puntos críticos	
7.5	Derivadas parciales de orden superior	
7.6	Matriz Hessiana. Cuádrica aproximante	
7.7	Análisis de puntos críticos	
7.8	Aplicaciones vectoriales. Matriz Jacobiana	
7.9	Cambio de variable. Regla de la cadena	
8	Cálculo integral	91
8.1	Integrales iteradas	
8.2	Parametrización de regiones	
8.3	Cambio de coordenadas	
8.4	Integrales triples. Volumen y masa	

6. Funciones de varias variables

Las funciones de varias variables surgen en contextos muy elementales:

- $F(x,y) = \sqrt{x^2 + y^2}$ da la distancia de un punto (x,y) al origen $(0,0)$.
- $F(x,y) = xy$ proporciona el área de un rectángulo de lados x e y .

Las funciones de varias variables son también esenciales en muchos problemas de la ciencia, ingeniería, economía, etc. Por poner un ejemplo, la “fórmula” que determina la presión de un gas en función del volumen V y la temperatura T , esto es, la *ley de los gases ideales*, viene dada por la función en dos variables

$$P(V,T) = \frac{cT}{V},$$

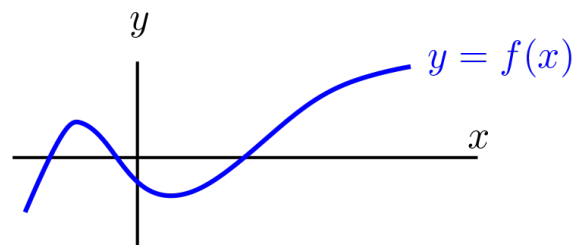
donde c es una constante.

6.1 Funciones de varias variables. Definición

Una *función de una variable*

$$\begin{aligned} f: D \subset \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto f(x) \end{aligned}$$

está definida en un subconjunto $D \subset \mathbb{R}$ que llamamos *dominio de f* .

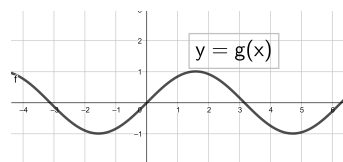
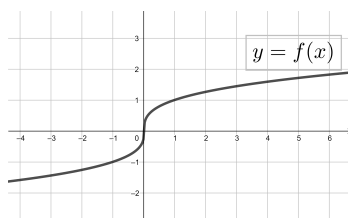


A cada valor $x \in D$, le asocia un valor $y = f(x) \in \mathbb{R}$. Acostumbramos a representar gráficamente f dentro del plano real \mathbb{R}^2 en coordenadas (x,y) . Más precisamente, representamos el conjunto

$$\text{Gr}(f) = \{(x, f(x)) : x \in D\} \subset \mathbb{R}^2$$

al cual llamamos *grafo o gráfica de f* . En general, decimos que la gráfica de una función de una variable es una *curva* en el plano \mathbb{R}^2 . El ejemplo más sencillo de curva en el plano es una recta, que (salvo si es vertical) es la gráfica de una función $f(x) = ax + b$.

■ **Ejemplo 6.1** Consideramos $f(x) = \sqrt[3]{x}$ y $g(x) = \sin(x)$. Tanto f como g tienen como dominio todo \mathbb{R} . Su representación gráfica es como sigue:

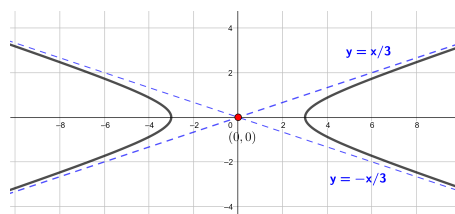
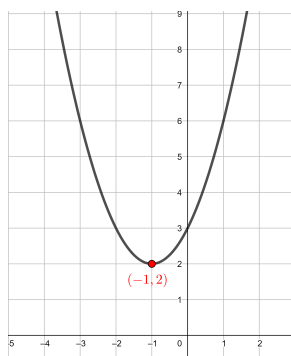


Después de las rectas, las curvas más “simples” que encontramos son las llamadas *cónicas*: elipses, parábolas e hipérbolas. Recordemos su descripción:

1. Ecuación de una **parábola**: $y = ax^2 + bx + c$ con $a, b, c \in \mathbb{R}$. Llamamos vértice de la parábola al punto $(x_0, y(x_0))$ con $x_0 = \frac{-b}{2a}$.
2. Ecuación de una **hipérbola** centrada en (a, b) :

$$\frac{(x-a)^2}{r^2} - \frac{(y-b)^2}{s^2} = \pm 1.$$

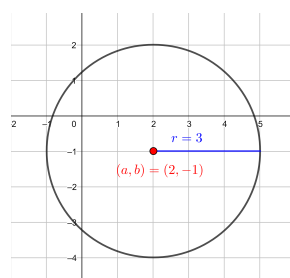
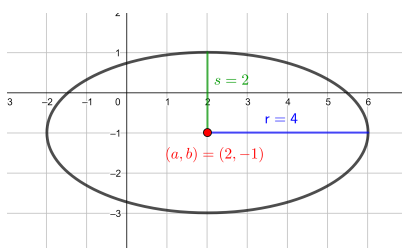
Decimos que sus asíntotas son $y = s/r(x-a) + b$ e $y = -s/r(x-a) + b$. Cuando el término a la derecha de la igualdad es $+1$, la hipérbola corta a la recta horizontal $y = y_0$ en los puntos $x = x_0 \pm r$. Cuando el término a la derecha de la igualdad es -1 , la hipérbola corta a la recta vertical $x = x_0$ en los puntos $y = y_0 \pm s$.



3. Ecuación de una **elipse** centrada en (a, b)

$$\frac{(x-a)^2}{r^2} + \frac{(y-b)^2}{s^2} = 1.$$

Tiene semiejes de longitud r y s . Cuando $r = s$ tenemos una circunferencia.



De inmediato observamos que la parábola sí es el grafo de una función, la dada por $f(x) = ax^2 + bx + c$, pero que esto no es cierto, por ejemplo, para la circunferencia. Si queremos representar una circunferencia centrada en el origen y de radio 3 tenemos que considerar la función de dos variables

$$F(x,y) = x^2 + y^2$$

y tomar uno de sus “niveles”, el nivel $z = F(x,y) = 9$.

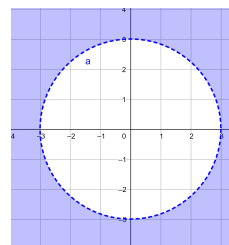
Sea $D \subset \mathbb{R}^2$ un subconjunto no vacío. Una *función de 2 variables* es una “regla” $F : D \rightarrow \mathbb{R}$ que asigna a cada punto $(x,y) \in D$ un número real $F(x,y)$. Al conjunto D lo llamamos *dominio de F* y el conjunto de todos los valores $F(x,y)$ es la *imagen de F*.

■ **Ejemplo 6.2** Determinemos el dominio y la imagen de las siguientes funciones:

- $F(x,y) = 1 + x + y$. Tiene dominio todo \mathbb{R}^2 y su imagen es todo \mathbb{R} .
- $F(x,y) = \frac{1}{\sqrt{x^2 + y^2}}$. Tiene dominio $\mathbb{R}^2 \setminus \{(0,0)\}$ y su imagen siempre es positiva.
- $F(x,y) = \frac{1}{\sqrt{x^2 + y^2 - 9}}$.

$\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 > 9\}$ es su dominio, esto es el complementario de un círculo de centro $(0,0)$ y radio 3. Su imagen siempre es positiva.

La circunferencia de centro $(0,0)$ y radio 3 no está en el dominio.



De forma general, dado un subconjunto $D \subset \mathbb{R}^n$ no vacío, una *función de n variables* es una “regla” $F : D \rightarrow \mathbb{R}$ que asigna a cada punto $(x_1, x_2, \dots, x_n) \in D$ un número real $F(x_1, x_2, \dots, x_n)$. Un poco más adelante veremos algunos ejemplos de funciones de tres variables.

6.2 Funciones de dos variables: Curvas de nivel

Como ya hemos dicho, muchas curvas planas, como la circunferencia, vienen dadas por los “niveles” de una función de dos variables. Habitualmente decimos que $x^2 + y^2 = r^2$ es una *ecuación implícita* de la circunferencia centrada en $(0,0)$ y de radio r .

En muchas ocasiones queremos representar no solamente un nivel de una función $F : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, sino todos a la vez: hablamos del *mapa o foliación de curvas de nivel*.

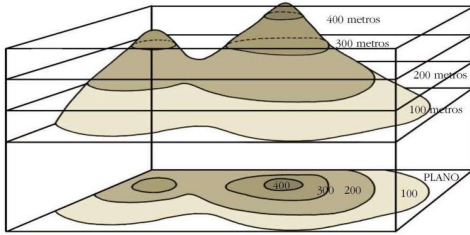
■ **Observación 6.1** Esta técnica es muy empleada en mapas:

- En topografía podemos representar terrenos montañosos dibujando curvas que unen los puntos a igual altura.
- En oceanografía se dibujan curvas, llamadas isobatas o isóbatas, que representan puntos de igual profundidad en el océano y en el mar.
- En meteorología, se dibujan curvas, llamadas isobaras o isóbaras, que representan puntos de igual presión atmosférica.

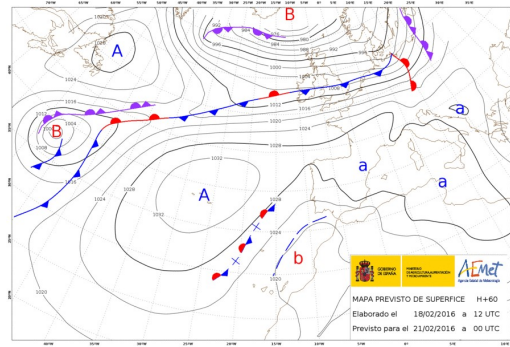
En estos ámbitos se suele hablar de *isolíneas* para referirse a las curvas de nivel (el prefijo *iso* significa *igual*).¹

¹Imágenes extraídas de internet.

PROYECCIÓN SOBRE UN PLANO DE LAS CURVAS DE NIVEL



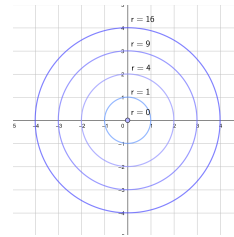
CURVAS DE NIVEL: líneas que unen puntos de igual cota o altitud.



■ **Ejemplo 6.3** Representemos el mapa de curvas de nivel de las funciones:

1) $F(x,y) = x^2 + y^2$.

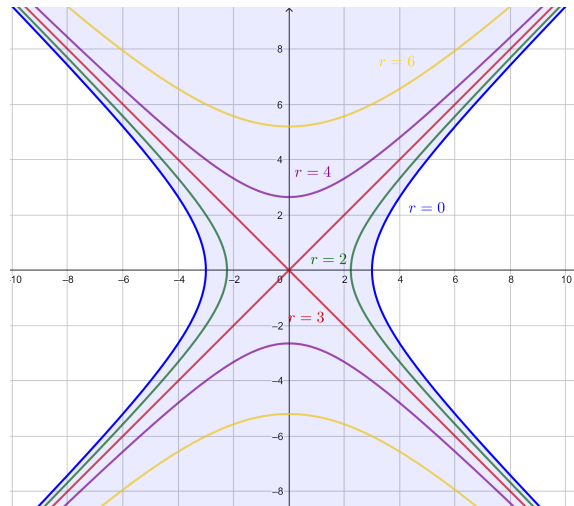
- Nivel 0: El punto $(0,0)$.
- Nivel $r > 0$: La circunferencia centrada en $(0,0)$ y de radio \sqrt{r} .
- Nivel $r < 0$: El vacío.



Observemos que el dominio de F es todo \mathbb{R}^2 . Así, cualquier punto de \mathbb{R}^2 pertenece a una y solo una curva de nivel. Podríamos decir que $D = \mathbb{R}^2$ es nuestro mapa, que cada curva de nivel es una “hoja” o “lámina” y que todo el mapa está cubierto por “hojas”, de aquí la terminología *foliación* (recubrimiento por hojas).

2) $F(x,y) = +\sqrt{9-x^2+y^2}$.

- Niveles $0 \leq r < 3$: Tenemos que $\sqrt{9-x^2+y^2} = r$ si y solamente si $x^2 - y^2 = 9 - r^2$. Esto es una hipérbola centrada en $(0,0)$ con asíntotas $y = \pm x$, que corta al eje x en los puntos $x = \pm\sqrt{9-r^2}$.
- Nivel $r = 3$: Tenemos $y^2 - x^2 = 0$. Es unión de las rectas $y = x$ e $y = -x$.
- Niveles $r > 3$: Tenemos que $\sqrt{9-x^2+y^2} = r$ si y solamente si $y^2 - x^2 = r^2 - 9$. Esto es una hipérbola centrada en $(0,0)$ con asíntotas $y = \pm x$, que corta al eje y en los puntos $x = \pm\sqrt{r^2-9}$.
- Niveles $r < 0$: El vacío.



Observemos que el dominio de F es la región $H \subset \mathbb{R}^2$ encerrada por la hipérbola $x^2 - y^2 = 9$. Más precisamente es el conjunto

$$H = \{(x,y) \in \mathbb{R}^2 : x^2 - y^2 \leq 9\}$$

que representamos de azul en la figura. Cualquier punto de H pertenece a una y sólo una curva de nivel. De nuevo observamos cómo las curvas de nivel no se cortan entre sí y recubren todo H .

6.3 Superficies

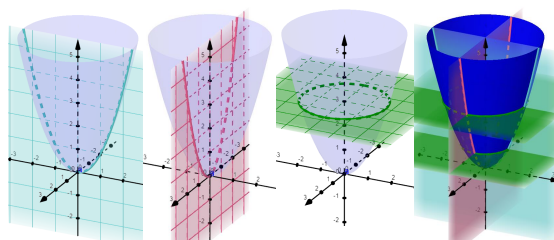
De igual modo que con las funciones de una sola variable, podemos definir la *gráfica* de una función de varias variables $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ como

$$\text{Gr}(F) = \{(\mathbf{x}, F(\mathbf{x})), \mathbf{x} = (x_1, x_2, \dots, x_n) \in D\} \subset \mathbb{R}^{n+1},$$

aunque solo seremos capaces de hacernos una imagen visual cuando tenemos dos variables. Más precisamente, cuando $n = 2$, la gráfica $\text{Gr}(F)$ es un subconjunto de \mathbb{R}^3 que define una *superficie* (dimensión dos).

Si no siempre es sencillo representar curvas en el plano, todavía lo es menos representar superficies. Para hacernos una pequeña idea de cómo son es frecuente considerar los siguientes elementos:

- Los puntos de corte de la superficie con los ejes coordenados.
 - Las *secciones planas*, esto es, los cortes con planos. Principalmente consideraremos los cortes con los ejes coordenados y los del tipo $z = \text{cte}$, es decir, los planos horizontales (recuperamos las curvas de nivel).
 - La *simetría*. Por ejemplo si $F(x, y) = F(-x, -y)$ o $F(x, y) = F(x, -y)$, estamos teniendo ciertas simetrías
 - Detectar si está generada por la *revolución* de una curva. Esto ocurre típicamente cuando detectamos que las curvas de nivel son circunferencias.
 - Si es *acotada* (se puede meter dentro de una bola) o no. Si la función puede tomar valores tan altos (o bajos) como se desee, no es acotada.
- **Ejemplo 6.4** Consideramos la función $F(x, y) = x^2 + y^2$, cuyo dominio de definición es todo \mathbb{R}^2 . Veamos cómo son los cortes de $\text{Gr}(F)$ con algunos planos:
- Con $x = 0$. Obtenemos $F(0, y) = y^2$ y por tanto nos queda una parábola.
 - Con $y = 0$. Obtenemos $F(x, 0) = x^2$ y nos queda otra parábola.
 - Con $z = s > 0$. Obtenemos $x^2 + y^2 = s$, que es una circunferencia de centro $x = y = 0$ y radio \sqrt{s} . Esto nos indica que se obtiene por revolución de cualquiera de las parábolas anteriores sobre el eje z .

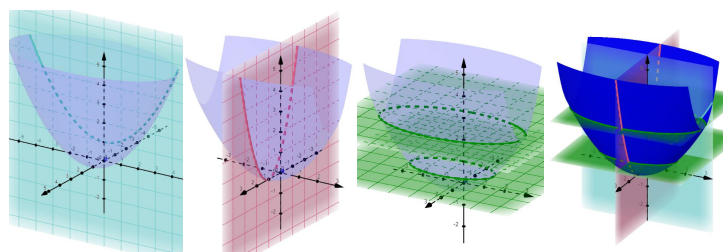


$\text{Gr}(F)$ es como un cuenco, su “nombre científico” es *paraboloide circular*.

- **Ejemplo 6.5** Consideramos la función $F(x, y) = (x - 1)^2 + \frac{y^2}{4}$ definida en todo \mathbb{R}^2 . Veamos cómo son los cortes de $\text{Gr}(F)$ con algunos planos:

- Con $x = 0$. Obtenemos $F(0,y) = \frac{y^2}{4} + 1$ y por tanto nos queda una parábola con vértice $y = 0$ y $z = 1$.
- Con $y = 0$. Obtenemos $F(x,0) = (x-1)^2$ y nos queda otra parábola con vértice $x = 1$ y $z = 0$.
- Con $z = s > 0$. Obtenemos $(x-1)^2 + \frac{y^2}{4} = s$, que es una elipse de centro $x = 1, y = 0$ y semiejes \sqrt{s} y $2\sqrt{s}$.

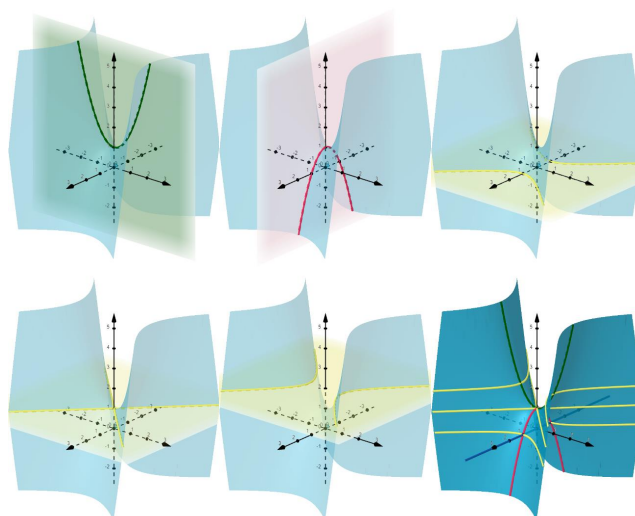
Observemos que en este caso $F(x,y) = F(x,-y)$ y $F(1+x,y) = F(1-x,y)$, es decir, la figura es simétrica respecto al plano $y = 0$ y respecto al plano $x = 1$.



De nuevo $\text{Gr}(F)$ es como un cuenco, pero con “borde” elíptico. Su “nombre científico” es *paraboloide elíptico*.

■ **Ejemplo 6.6** Consideramos $F(x,y) = y^2 - x^2 + 1$ definida en todo \mathbb{R}^2 . Veamos cómo son los cortes de $\text{Gr}(F)$ con algunos planos:

- Con $x = 0$. Obtenemos $F(0,y) = y^2 + 1$ y por tanto nos queda una parábola con vértice $y = 0$ y $z = 1$.
- Con $y = 0$. Obtenemos $F(x,0) = -x^2 + 1$ y nos queda otra parábola “hacia abajo” con vértice $x = 0$ y $z = 1$.
- Con $z = s < 1$. Obtenemos $x^2 - y^2 = 1 - s$, que es una hipérbola de centro $x = 0, y = 0$ con asíntotas $x = \pm y$ y que corta al eje x en $x = \pm\sqrt{1-s}$. Por ejemplo para $s = 0$ los cortes al eje x son en $x = \pm 1$.



- Con $z = s = 1$. Obtenemos $y^2 - x^2 = 0$, que es unión de las rectas $x = y, x = -y$.

- Con $z = s > 1$. Obtenemos $y^2 - x^2 = s - 1$, que es una hipérbola de centro $x = 0$, $y = 0$ con asíntotas $x = \pm y$ y que corta al eje y en $y = \pm\sqrt{s-1}$. Por ejemplo para $s = 2$ los cortes al eje y son en $y = \pm 1$.

Esta superficie se conoce comunmente como *silla de montar*, dado que recuerda la ensilladura de un caballo. Su nombre técnico es *paraboloide hiperbólico*.

Estos son algunos ejemplos de *superficies cuádricas*, que son las más sencillas tras los planos. Pero al igual que sucede en curvas para las cónicas, no todas las superficies cuádricas son el grafo de una función. Por ejemplo una esfera de centro $\mathcal{O} = (a, b, c)$ y radio r , cuya *ecuación implícita* está dada por el nivel r^2 de

$$F(x, y, z) = (x - a)^2 + (y - b)^2 + (z - c)^2.$$

Así, de modo similar al caso de curvas, podemos estudiar superficies dadas por niveles de funciones en tres variables.

Vamos a dar ahora por completitud el catálogo de las *superficies cuádricas*. Por comodidad en la escritura las tomamos todas ellas centradas en $\mathcal{O} = (0, 0, 0)$.

- *Paraboloide elíptico:*

$$z = \frac{x^2}{r^2} + \frac{y^2}{s^2}$$

- *Paraboloide hiperbólico:*

$$z = \frac{x^2}{r^2} - \frac{y^2}{s^2}$$

- *Elipsoide:*

$$\frac{x^2}{r^2} + \frac{y^2}{s^2} + \frac{z^2}{t^2} = 1$$

- *Cono:*

$$z^2 = \frac{x^2}{r^2} + \frac{y^2}{s^2}$$

Si $r = s = t$ recuperamos la *esfera*.

Si $r = s$ recuperamos el *cono circular*.

- *Hiperboloide de una hoja:*

$$\frac{x^2}{r^2} + \frac{y^2}{s^2} - \frac{z^2}{t^2} = 1$$

- *Hiperboloide de dos hojas:*

$$\frac{x^2}{r^2} + \frac{y^2}{s^2} - \frac{z^2}{t^2} = -1$$

- *Cilindro parabólico:*

$$x^2 = 4cy$$

- *Cilindro elíptico:*

$$\frac{x^2}{r^2} + \frac{y^2}{s^2} = -1$$

- *Cilindro hiperbólico:*

$$\frac{x^2}{r^2} - \frac{y^2}{s^2} = 1$$

Vamos a describir algunos dominios de funciones en tres variables. En general estos serán regiones del espacio \mathbb{R}^3 delimitadas por algunas superficies.

- **Ejemplo 6.7** Consideremos las siguientes funciones:

1. $F(x, y, z) = \sqrt{-x^2 - y^2 - z^2 + 6x + 4y - 4}$.

2. $G(x, y, z) = \sqrt{x^2 + y^2 - 4} + \sqrt{4 - x^2 - y^2}$.

Empecemos por F . Su dominio es el dado por el conjunto

$$D = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 - 6x - 4y + 4 \leq 0\}.$$

A priori esto no tiene una expresión conocida, pero observamos que la superficie que delimita es una cuádrica (esto es, un polinomio de grado dos en las variables x , y y z). “Completando cuadrados” tenemos que

$$x^2 + y^2 + z^2 - 6x - 4y + 4 = (x - 3)^2 + (y - 2)^2 + z^2 - 9,$$

y por tanto la superficie que delimita la región buscada es la esfera centrada en $(3, 2, 0)$ y de radio 3. Ahora debemos averiguar si D es la zona interior a la esfera, es decir la bola rellena o la parte de afuera. Como rápidamente observamos que $(0, 0, 0) \in D$, la región de definición de esta función es toda la bola (rellena) centrada en $(3, 2, 0)$ y de radio 3.

Pasamos ahora a estudiar G . Su dominio es el conjunto

$$D = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 4\}.$$

que es un cilindro recto (su eje es el “eje z ”) y base la circunferencia centrada en el origen y de radio 2. Observemos que esta función toma un único valor: el cero.

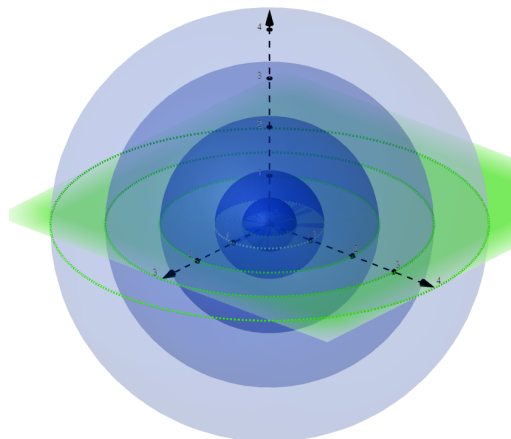
6.4 Representación gráfica: Superficies de nivel

Dada una función $F : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, podemos considerar el *mapa o foliación de superficies de nivel de F* en analogía al caso de dos variables. Esto, a parte de ser interesante *per se*, nos permite hacernos una ligera idea de la geometría del sólido 3-dimensional (que viviría dentro de \mathbb{R}^4) descrito por la gráfica de F .

■ **Ejemplo 6.8** Representemos el mapa de superficies de nivel de las funciones:

1) $F(x, y, z) = x^2 + y^2 + z^2.$

- Nivel 0: El punto $(0, 0, 0)$.
- Nivel $r > 0$: La esfera centrada en $(0, 0, 0)$ y de radio \sqrt{r} .
- Nivel $r < 0$: El vacío.



Observemos que el dominio de F es todo \mathbb{R}^3 . Así, cualquier punto de \mathbb{R}^3 pertenece a una y solo una superficie de nivel. Cada esfera es una “hoja” o “lámina” y todas ellas forman una *foliación* (recubrimiento por hojas) de \mathbb{R}^3 .

2) $F(x, y, z) = +\sqrt{9 - x^2 - y^2 + z^2}.$

- Nivel $0 \leq r < 3$: Tenemos que $\sqrt{9 - x^2 - y^2 + z^2} = r$ si y solo si

$$x^2 + y^2 - z^2 = 9 - r^2.$$

Esto es un hiperboloide “de una hoja”, obtenido al hacer girar sobre el eje z la hipérbola del plano yz centrada en $(0, 0)$ con asíntotas $z = \pm y$ y que corta al eje y en los puntos $y = \pm\sqrt{9 - r^2}$.

- Nivel $r = 3$: Tenemos

$$x^2 + y^2 = z^2.$$

Esta superficie es un cono (con “dos hojas”). Se obtiene al hacer girar sobre el eje z la unión de las rectas del plano yz dadas por $y = z$ e $y = -z$.

- Nivel $r > 3$: Tenemos que $\sqrt{9 - x^2 - y^2 + z^2} = r$ si y sólo si

$$z^2 - x^2 - y^2 = r^2 - 9.$$

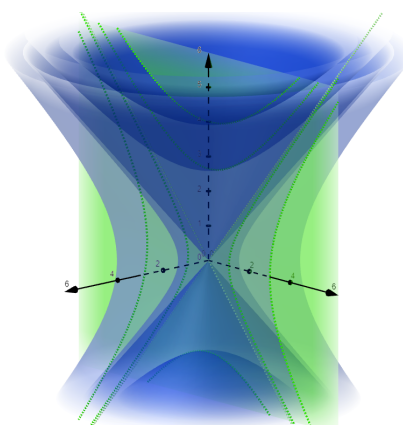
Esto es un hiperboloide “de dos hojas”, obtenido al hacer girar sobre el eje z la hipérbola del plano yz centrada en $(0,0)$ con asíntotas $z = \pm y$ y que corta al eje z en los puntos $z = \pm\sqrt{r^2 - 9}$.

- Nivel $r < 0$: El vacío.

Observemos que el dominio de F es la región $H \subset \mathbb{R}^3$ encerrada por el hiperboloide $x^2 + y^2 - z^2 = 9$. Más precisamente es el conjunto

$$H = \{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 - z^2 \leq 9\}$$

Cualquier punto de H pertenece a una y sólo una superficie de nivel (que puede ser una hoja o la unión de dos). De nuevo observamos cómo las superficies de nivel no se cortan entre sí y recubren todo H , dan una foliación de H .

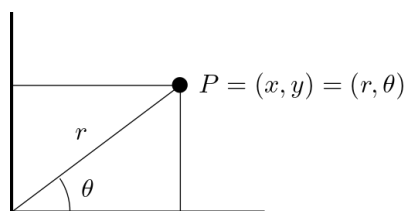


6.5 Coordenadas: polares, cilíndricas, esféricas.

Un sistema de coordenadas es una forma de definir y ubicar un punto en el plano \mathbb{R}^2 , en el espacio \mathbb{R}^3 o incluso en espacios más generales. Nosotros estamos acostumbrados a los sistemas de coordenadas llamados *cartesianos* que habitualmente designamos mediante (x,y) en el plano y (x,y,z) en el espacio. Sin embargo, existen otros sistemas de coordenadas que dependiendo de la situación pueden resultar más adecuados o convenientes.

6.5.1 Coordenadas polares

En el plano \mathbb{R}^2 podemos identificar un punto P distinto del origen mediante las coordenadas (r, θ) , donde $r > 0$ es la distancia de P al origen y θ el ángulo entre el eje x y la recta que une P con el origen, donde el ángulo lo tomamos en $[0, 2\pi)$. El origen lo representamos con $(r, \theta) = (0,0)$.



Las coordenadas polares son las naturales en la circunferencia. Por ejemplo, la circunferencia de radio 4 viene dada por tomar $r = 4$, es decir, es el conjunto de puntos en coordenadas polares de la forma $(4, \theta)$.

Pasamos de coordenadas polares a cartesianas mediante:

$$x = r \cos \theta, \quad y = r \sin \theta$$

Recíprocamente, las coordenadas cartesianas se obtienen a partir de las polares como sigue:

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arctan\left(\frac{y}{x}\right),$$

conviniendo lo siguiente:

- Si $x = 0$ e $y > 0$ tomamos $\theta = \pi/2$. Si $x = 0$ e $y < 0$ tomamos $\theta = 3\pi/2$.
- Si $y = 0$ y $x > 0$ tomamos $\theta = 0$. Si $y = 0$ y $x < 0$ tomamos $\theta = \pi$.
- Si $x \neq 0$ e $y > 0$ tomamos $\theta \in (0, \pi)$. Si $x \neq 0$ e $y < 0$ tomamos $\theta \in (\pi, 2\pi)$.

■ **Ejemplo 6.9** Consideremos los siguientes puntos en coordenadas polares:

$$P_1 = (3, \pi), \quad P_2 = (1, \pi/2), \quad P_3 = (2, \pi/4)$$

En coordenadas cartesianas se escriben como:

$$\begin{aligned} P_1 &= (3 \cos(\pi), 3 \sin(\pi)) &= (-3, 0), \\ P_2 &= (\cos(\pi/2), \sin(\pi/2)) &= (0, 1), \\ P_3 &= (2 \cos(\pi/4), 2 \sin(\pi/4)) &= (\sqrt{2}, \sqrt{2}), \end{aligned}$$

■ **Ejemplo 6.10** Consideremos los siguientes puntos en coordenadas cartesianas:

$$P_1 = (1, 1), \quad P_2 = (0, -2), \quad P_3 = (3, \sqrt{3})$$

En coordenadas polares se escriben como:

$$\begin{aligned} r = \sqrt{1+1} = \sqrt{2} & & \theta = \arctan(1) = \pi/4 & & P_1 &= (\sqrt{2}, \pi/4), \\ r = \sqrt{4} = 2 & & \theta = 3\pi/2 & & P_2 &= (2, 3\pi/2), \\ r = \sqrt{9+3} = \sqrt{12} = 2\sqrt{3} & & \theta = \arctan(\sqrt{3}/3) = \pi/6 & & P_3 &= (2\sqrt{3}, \pi/6), \end{aligned}$$

■ **Ejemplo 6.11** Describe en polares la región del plano dada por el disco

$$(x-1)^2 + y^2 \leq 1$$

Sustituyendo tenemos que

$$\begin{aligned} (r \cos \theta - 1)^2 + (r \sin \theta)^2 \leq 1 &\Leftrightarrow (r \cos \theta)^2 - 2r \cos \theta + (r \sin \theta)^2 \leq 0 \Leftrightarrow \\ &\Leftrightarrow r^2 - 2r \cos \theta \leq 0. \end{aligned}$$

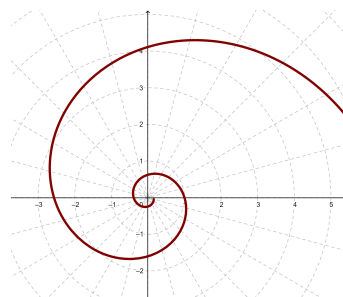
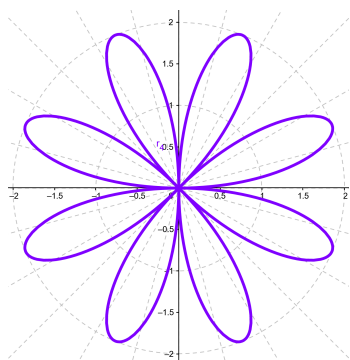
Ahora, como $r > 0$, podemos dividir a ambos lados de la desigualdad sin que esta se altere y obtenemos

$$r - 2 \cos \theta \leq 0 \Leftrightarrow r \leq 2 \cos \theta$$

Así, la descripción que buscamos es $r \leq 2 \cos \theta$. Notemos que esto en particular obliga a que $\theta \in [0, \pi/2) \cup (3\pi/2, 2\pi)$ (la coordenada x siempre es positiva o cero).

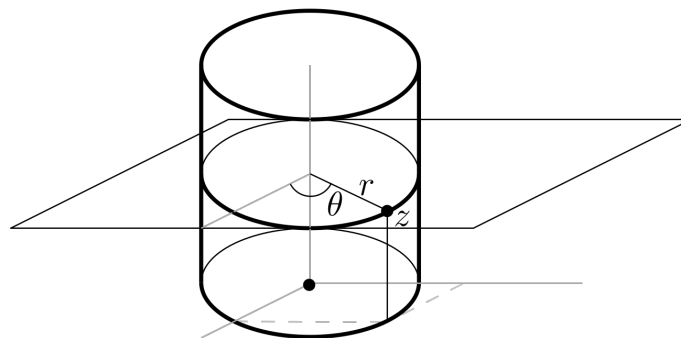
Ya hemos visto que hay curvas que tienen una expresión mucho más sencilla en coordenadas polares, como la circunferencia. Pero también las curvas cuyo radio viene descrito por:

$$r(\theta) = 2 \sin(4\theta), \quad r(\theta) = e^{-\frac{3\theta}{10}}$$



6.5.2 Coordenadas cilíndricas

En el espacio \mathbb{R}^3 al escribir (x, y, z) para dar la posición de un punto P , la coordenada z hace referencia a la “altura” del punto P . Otra manera de representar P es mediante su altura y sus coordenadas polares dentro del plano horizontal a altura z . Así obtenemos las coordenadas cilíndricas (r, θ, z) , que son las coordenadas naturales del cilindro.



■ **Ejemplo 6.12** La superficie $r = 4$ en coordenadas cartesianas viene dada por

$$r = \sqrt{x^2 + y^2} = 4 \Leftrightarrow x^2 + y^2 = 16.$$

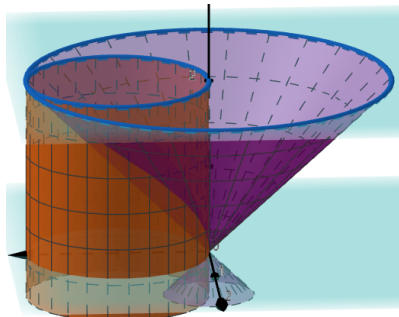
Es precisamente un cilindro. Como vemos, su descripción en cilíndricas es mucho más clara.

■ **Ejemplo 6.13** Consideramos la superficie $z^2 = r^2$. Si hacemos cortes a diferentes alturas vamos obteniendo circunferencias de radio $|z|$, es decir, cada vez más grandes conforme z aumenta (en valor absoluto). La circunferencia de radio 0 degenera a un punto. Tenemos por tanto que la superficie es un cono. En efecto, si sustituimos, vemos que

$$z^2 = r^2 = r^2(\cos^2 \theta + \sin^2 \theta) = (r \cos \theta)^2 + (r \sin \theta)^2 = x^2 + y^2$$

■ **Ejemplo 6.14** Describe en coordenadas cilíndricas el sólido contenido en el cono $z^2 = x^2 + y^2$, que es exterior al cilindro $(x-1)^2 + y^2 = 1$ y está delimitado inferiormente por el plano $z = 0$ y superiormente por el plano $z = 2$.

En la figura tenemos representada la región considerada



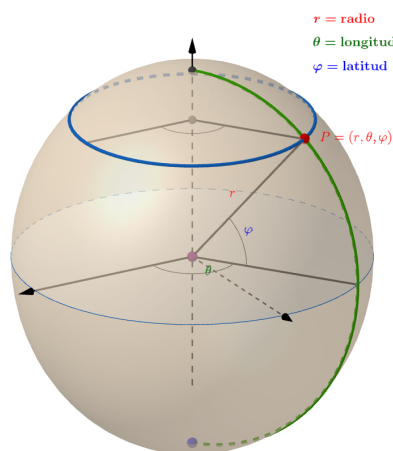
- En cilíndricas, la parte interior al cono viene descrita por $r^2 \leq z^2$; en efecto, notemos que cualquier punto sobre el eje z , es decir, con $r = 0$, cumple esta propiedad.
 - El cilindro que buscamos es el construido sobre la circunferencia centrada en $(1, 0)$ y de radio 1, que ya hemos visto que en coordenadas polares está dada por $r = 2 \cos(\theta)$. Como buscamos la región exterior al cilindro, ésta es la descrita por $r \geq 2 \cos \theta$.
- Así, la región en cilíndricas es la dada por los puntos (r, θ, z) que cumplen

$$r^2 \leq z^2, \quad r \geq 2 \cos \theta, \quad 0 \leq z \leq 2$$

6.5.3 Coordenadas esféricas

Las coordenadas esféricas son las naturales en la superficie de la esfera. De hecho, estas se corresponden con las llamadas coordenadas geográficas con las que describimos la posición de una ubicación en la superficie de la Tierra.

Las coordenadas esféricas (r, θ, φ) indican mediante $r > 0$, el radio de la esfera centrada en $(0, 0, 0)$ en la que podemos encontrar nuestro punto; después $0 \leq \theta < 2\pi$ indica la *longitud* y $0 \leq \varphi < \pi$ indica la *latitud*. El origen de \mathbb{R}^3 lo identificamos mediante $(r, \theta, \varphi) = (0, 0, 0)$. En la siguiente figura vemos una representación:



Tenemos las siguientes relaciones entre las coordenadas cartesianas y las polares

$$x = r \sin \varphi \cos \theta, \quad y = r \sin \varphi \sin \theta, \quad z = r \cos \varphi$$

Además tenemos que $r = \sqrt{x^2 + y^2 + z^2}$.

■ **Ejemplo 6.15** El punto dado en coordenadas cartesianas por $(x, y, z) = (2, 2, 0)$ se expresa en coordenadas esféricas mediante

$$r = \sqrt{4 + 4 + 0} = 2\sqrt{2}, \quad \cos \varphi = 0 \Rightarrow \varphi = \pi/2$$

$$\cos \theta = \frac{2}{2\sqrt{2}} = \frac{\sqrt{2}}{2} \text{ y } \sin \theta = \frac{2}{2\sqrt{2}} = \frac{\sqrt{2}}{2} \Rightarrow \theta = \pi/4$$

Así, en coordenadas esféricas viene dado por $(r, \varphi, \theta) = (2\sqrt{2}, \pi/2, \pi/4)$.

■ **Ejemplo 6.16** El punto dado en coordenadas esféricas por $(r, \varphi, \theta) = (4, 2\pi/3, \pi/6)$ se expresa en coordenadas cartesianas mediante

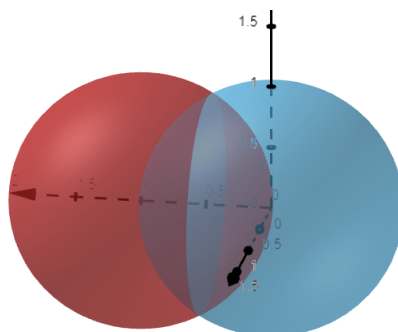
$$x = 4 \sin(\pi/6) \cos(2\pi/3) = 4 \cdot \frac{1}{2} \cdot \frac{-1}{2} = -1,$$

$$y = 4 \sin(\pi/6) \sin(2\pi/3) = 4 \cdot \frac{1}{2} \cdot \frac{\sqrt{3}}{2} = \sqrt{3},$$

$$z = 4 \cos(\pi/6) = 4 \cdot \frac{\sqrt{3}}{2} = 2\sqrt{3}$$

Así, las coordenadas cartesianas son $(x, y, z) = (-1, \sqrt{3}, 2\sqrt{3})$.

■ **Ejemplo 6.17** Describe en coordenadas esféricas el sólido dado por la intersección de las bolas $x^2 + y^2 + z^2 \leq 1$ y $(x-1)^2 + y^2 + z^2 \leq 1$



- En esféricas, la esfera $x^2 + y^2 + z^2 = 1$ se describe como $r = 1$ y por tanto su interior es $r \leq 1$.
- La esfera $(x-1)^2 + y^2 + z^2 = 1$ viene descrita por

$$(r \cos \theta \sin \varphi - 1)^2 + (r \sin \theta \sin \varphi)^2 + (r \cos \varphi)^2 = 1 \Leftrightarrow$$

$$\Leftrightarrow (r \cos \theta \sin \varphi)^2 - 2r \cos \theta \sin \varphi + 1 + (r \sin \theta \sin \varphi)^2 + (r \cos \varphi)^2 = 1 \Leftrightarrow$$

$$\Leftrightarrow r^2 \sin^2 \varphi (\cos^2 \varphi + \sin^2 \varphi) - 2r \cos \theta \sin \varphi + r^2 \cos^2 \varphi = 0 \Leftrightarrow$$

$$\Leftrightarrow r^2 = 2r \cos \theta \sin \varphi \Leftrightarrow r = 2 \cos \theta \sin \varphi$$

Así, su interior (la bola) está descrito por $r \leq 2 \cos \theta \sin \varphi$.

La región en esféricas es por tanto la dada por los puntos (r, θ, φ) que cumplen

$$r \leq 1, \quad r \leq 2 \cos \theta \sin \varphi$$

6.6 Parametrización de curvas

Una *parametrización* de una curva es una asignación de un punto $\gamma(t)$ de la curva a cada valor de un parámetro t . Esto, aparte de describir la curva, nos da una manera de recorrerla. Algunos ejemplos son los siguientes:

1. La expresión de una recta en forma paramétrica. Esto es, si consideramos la recta que pasa por el punto $P = (x_0, y_0)$ y tiene vector director (u, v) , la podemos parametrizar como

$$\gamma(t) = (x_0 + tu, y_0 + tv)$$

2. La circunferencia de radio r centrada en el origen la parametrizamos por

$$\gamma(t) = (r \cos t, r \sin t)$$

como indican las coordenadas polares.

■ **Ejemplo 6.18** Las soluciones de un sistema de ecuaciones diferenciales son ejemplos de curvas parametrizadas que ya hemos tratado. El sistema

$$y_1' = y_1, \quad y_2' = -y_2,$$

tiene como soluciones a las de la forma $\gamma(x) = (ae^x, be^{-x})$ y éstas parametrizan curvas del plano.

La curva descrita por la ecuación $x^3 - y^2 = 0$ la podemos parametrizar mediante $\gamma(t) = (t^2, t^3)$. La ventaja de las parametrizaciones es que nos permiten expresar curvas del espacio que no se pueden describir tan fácilmente mediante ecuaciones, como por ejemplo

$$\gamma(t) = (\cos t, \sin t, t)$$

que parametriza una hélice. Las curvas del espacio que se pueden dar mediante ecuaciones son aquellas que se obtienen al intersecar dos superficies. Pese a ello es interesante muchas veces tener una parametrización de éstas, y por tanto una manera de recorrerlas.

■ **Ejemplo 6.19** Proporciona una parametrización de la curva intersección del paraboloides elíptico $2z = x^2 + 4y^2$ y el cilindro elíptico $x^2 + 4y^2 = 1$.

Esta es la curva que se obtiene al tomar $z = 1/2$. Dentro de este plano vemos la elipse $x^2 + 4y^2 = 1$ que se parametriza mediante $x = \cos t$ y $2y = \sin t$, por tanto obtenemos que la curva la podemos recorrer mediante

$$\gamma(t) = (\cos t, \sin(t/2), 1/2)$$

■ **Ejemplo 6.20** Proporciona una parametrización de la curva intersección del paraboloides $z = x^2 + y^2$ y el plano $x + y = 1$.

De la expresión de plano podemos escribir $y = 1 - x$. Ahora, sustituyendo en el paraboloides tenemos

$$z = x^2 + (1 - x)^2 = 2x^2 - 2x + 1$$

Así, la curva que buscamos viene parametrizada por

$$\gamma(t) = (t, 1 - t, 2t^2 - 2t + 1)$$

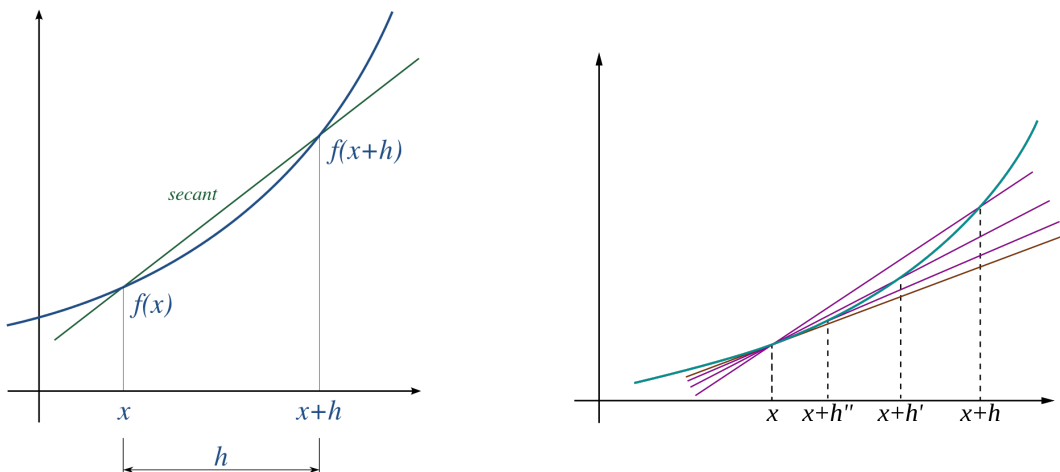
que es una parábola dentro del plano $x + y = 1$.

7. Cálculo diferencial

En este tema nos centraremos esencialmente en el estudio de funciones de dos variables, aunque algunas de las definiciones se pueden extender al caso de más variable. Para notar esto, en algunas ocasiones haremos ejemplos en tres variables.

7.1 Derivadas parciales

La derivada de una función de una variable $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ mide la razón de cambio instantánea con la que varía la función f al modificar la variable independiente x . Gráficamente la derivada de f en $x = x_0$ es la pendiente de la recta tangente al grafo de f en el punto $(x_0, f(x_0))$. Encontrar la recta tangente directamente no es claro, pues solo se conoce un punto de ésta, pero la recta tangente se puede aproximar trazando líneas secantes con puntos que están cada vez más próximos a $(x_0, f(x_0))$ como se ilustra en la figura.



Cuando se toma el límite de las pendientes de las rectas secantes de esta progresión, se consigue la pendiente de la recta tangente, y así es como se define la derivada. Más precisamente, tenemos que la derivada de $f(x)$ en $x = x_0$ viene definida mediante

$$f'(x_0) = \frac{df}{dx}(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

Observación: Puede suceder que el límite de la derecha no exista. Diríamos en ese caso que f no es derivable en $x = x_0$. Cuando f es una función derivable en todos sus puntos, podemos definir la *función derivada de f*

$$f' : I \rightarrow \mathbb{R},$$

que a cada punto x lo manda en su derivada $f'(x)$.

Del mismo modo que nos planteamos cómo varía una función dependiente de una sola variable x al cambiar x de modo infinitesimal, tiene sentido la misma pregunta cuando tenemos una función de dos variables $F : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ y queremos ver la variación de F al cambiar de modo infinitesimal una de ellas. Esto es, fijado un punto $(x_0, y_0) \in D$, definimos la *derivada parcial de F respecto a x en (x_0, y_0)* como

$$F_x(x_0, y_0) = \frac{\partial F}{\partial x}(x_0, y_0) = \lim_{h \rightarrow 0} \frac{F(x_0 + h, y_0) - F(x_0, y_0)}{h}.$$

De nuevo puede ocurrir que este límite no exista: diríamos que F no es derivable con respecto a x en (x_0, y_0) . Cuando F es derivable respecto a x en todos sus puntos, podemos definir la *función derivada parcial de F respecto a x*

$$F_x = \frac{\partial F}{\partial x} : D \rightarrow \mathbb{R}, \quad (x, y) \mapsto F_x(x, y).$$

Análogamente, definimos la *derivada parcial de F respecto a y en (x_0, y_0)* como

$$F_y(x_0, y_0) = \frac{\partial F}{\partial y}(x_0, y_0) = \lim_{h \rightarrow 0} \frac{F(x_0, y_0 + h) - F(x_0, y_0)}{h}.$$

Cuando F es derivable respecto a y en todos sus puntos, podemos definir la *función derivada parcial de F respecto a y*

$$F_y = \frac{\partial F}{\partial y} : D \rightarrow \mathbb{R}, \quad (x, y) \mapsto F_y(x, y).$$

Es muy importante observar que al definir las derivadas parciales sólo estamos mirando la variación respecto a una de las variables.

Fijemos de nuevo un punto $(x_0, y_0) \in D$. Las definiciones previas coinciden con el hecho de considerar las funciones de una variable

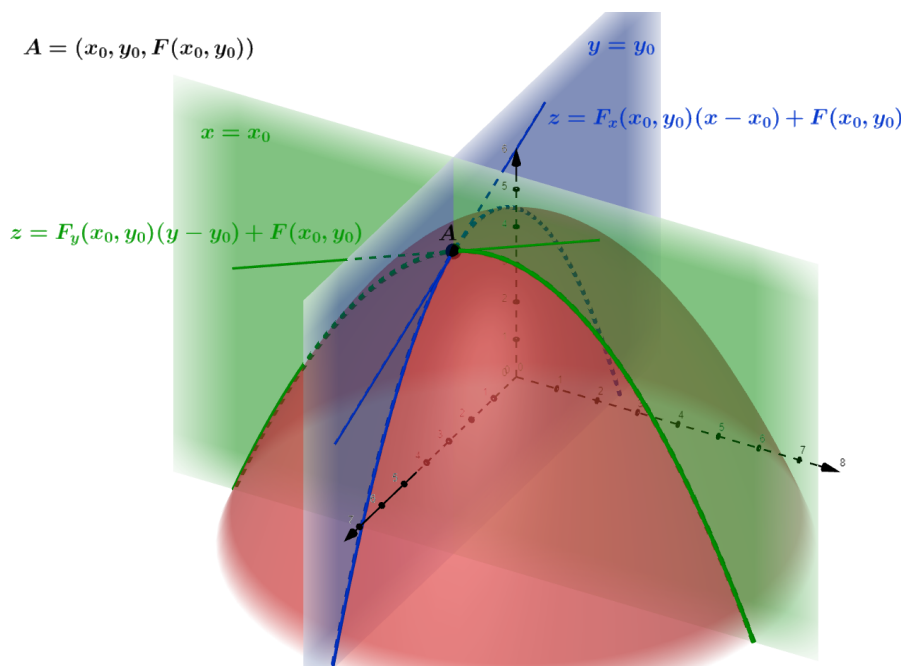
$$f_{y_0}(x) = F(x, y_0), \quad f_{x_0}(y) = F(x_0, y)$$

y tomar $F_x(x_0, y_0) = f'_{y_0}(x_0)$ y $F_y(x_0, y_0) = f'_{x_0}(y_0)$. Así, las derivadas parciales tienen una interpretación geométrica muy similar al caso de una sola variable. La vemos a

continuación. Definimos C_{y_0} como la intersección del grafo de F con el plano $y = y_0$, esto es

$$C_{y_0} = \{(x, y_0, F(x, y_0)), (x, y_0) \in D\} = \{(x, y_0, f_{y_0}(x)), (x, y_0) \in D\}$$

La derivada parcial $F_x(x_0, y_0)$ es la pendiente de la recta contenida en el plano $y = y_0$ que además es tangente a la curva C_{y_0} . Análogamente $F_y(x_0, y_0)$ es la pendiente de la recta contenida en el plano $x = x_0$ que además es tangente a la curva C_{x_0} dada por intersección del grafo de F con el plano $x = x_0$. Tenemos una representación de esta interpretación.



Aunque para funciones de más de dos variables ya no seamos capaces de visualizar su grafo, el concepto de derivada parcial se extiende por analogía y podemos pensar que tiene la misma interpretación geométrica (la pendiente de la recta tangente a la curva que se obtiene al intersectar el grafo con un plano adecuado).

■ **Ejemplo 7.1** Calcula las derivadas parciales de las siguientes funciones de dos y tres variables:

1. $F(x, y) = e^{xy} + \ln(x^2 + y^2)$, cuyo dominio es $D = \mathbb{R}^2 \setminus \{(0, 0)\}$. Tenemos:

$$\frac{\partial F}{\partial x} = ye^{xy} + \frac{2x}{x^2 + y^2}, \quad \frac{\partial F}{\partial y} = xe^{xy} + \frac{2y}{x^2 + y^2}$$

2. $F(x, y) = 3x^2y - 5x \cos(y)$. Tenemos que:

$$\frac{\partial F}{\partial x} = 6xy - 5 \cos(y), \quad \frac{\partial F}{\partial y} = 3x^2 + 5x \sin(y)$$

3. $F(x, y, z) = z^3 \cos(2x^2y)$. Tenemos que:

$$\frac{\partial F}{\partial x} = 4xyz^3 \sin(2x^2y), \quad \frac{\partial F}{\partial y} = 2x^2z^3 \sin(2x^2y), \quad \frac{\partial F}{\partial z} = 3z^2 \cos(2x^2y)$$

■ **Ejemplo 7.2** Sea $F(x, y, z) = x^3y^2z + \sin xy$. Calcula $F_x(\pi, 1, 2)$ y $F_z(1, 0, 5)$. Tenemos

$$F_x = 3x^2y^2z + y \cos xy, \quad F_z = x^3y^2$$

Así que $F_x(\pi, 1, 2) = 6\pi^2 + \cos \pi = 6\pi^2 - 1$ y $F_z(1, 0, 5) = 0$.

El *gradiente* ∇F es el vector de las derivadas parciales, esto es

$$\nabla F(x, y) = (F_x(x, y), F_y(x, y)).$$

Para que exista el gradiente en un punto, deben existir ambas derivadas parciales en dicho punto.

■ **Ejemplo 7.3** Consideremos $F(x, y) = \sqrt{x^2 + y^2}$ definida en todo el plano \mathbb{R}^2 . Su gradiente es

$$\nabla F(x, y) = \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right)$$

que tiene sentido en todo punto del dominio salvo el $(0, 0)$.

El gradiente tendrá un papel crucial en el estudio de puntos críticos y en determinar las direcciones de máximo crecimiento y decrecimiento.

7.2 Derivadas direccionales

Hemos visto que las derivadas parciales de una función $F(x, y)$ representan el cambio de la función cuando nos movemos en la dirección del eje x (variaciones de x) o en la dirección del eje y (variaciones de y). Dicho de otra manera, marcan la variación de la función al movernos en las direcciones $(1, 0)$ y $(0, 1)$, respectivamente. Nos puede interesar saber cómo varía la función al tomar una dirección distinta a estas. Por ejemplo, si estamos interesados en subir una montaña, seguramente debemos encontrar una dirección donde la pendiente no sea ni muy grande ni muy pequeña; si es muy grande quizá no seamos capaces de salvar el desnivel, si es muy pequeña rodeamos la montaña, pero apenas avanzamos hacia la cima. Esto motiva la introducción del concepto de derivada direccional.

Interpretaremos ahora un vector $\vec{v} = (a, b)$ de \mathbb{R}^2 como una velocidad. Esto es, si escribimos \vec{v} con sus coordenadas polares como

$$\vec{v} = (r \cos \theta, r \sin \theta),$$

con $0 \leq \theta < 2\pi$, tenemos que $\|\vec{v}\| = r = \sqrt{a^2 + b^2}$ marca la rapidez instantánea con la que nos movemos y el ángulo θ marca en qué dirección.

Consideremos una función $F : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$, un punto $(x_0, y_0) \in D$ y una velocidad $\vec{v} \in \mathbb{R}^2$. La *derivada de F respecto a la velocidad \vec{v} en (x_0, y_0)* es

$$F_{\vec{v}}(x_0, y_0) = \lim_{h \rightarrow 0} \frac{F((x_0, y_0) + h\vec{v}) - F(x_0, y_0)}{h}.$$

Cuando el límite en la definición no existe, decimos que F no es derivable respecto a \vec{v} en (x_0, y_0) . Cuando F es derivable para la velocidad \vec{v} en todos sus puntos, podemos definir la *función derivada de F respecto a la velocidad \vec{v} como*

$$F_{\vec{v}} : D \rightarrow \mathbb{R}, \quad (x, y) \mapsto F_{\vec{v}}(x, y).$$

Observamos que si $\vec{w} = \lambda \vec{v}$, tenemos la relación

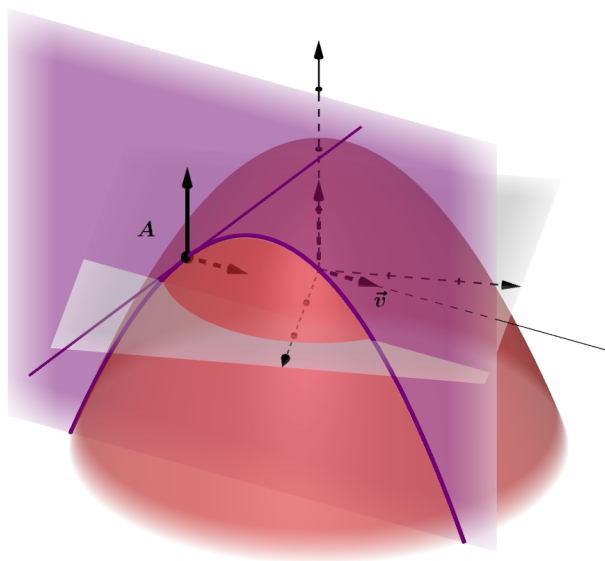
$$F_{\vec{w}}(x_0, y_0) = \lambda F_{\vec{v}}(x_0, y_0),$$

además, la derivada respecto a la velocidad \vec{w} existe si y solamente si existe la derivada respecto a la velocidad \vec{v} . Muchas veces nos interesa estudiar qué ocurre con las distintas derivadas al cambiar la dirección pero manteniendo dada una rapidez constante. Así, fijada una dirección $\theta \in [0, 2\pi)$, definimos la *derivada direccional en la dirección de θ en el punto (x_0, y_0)* como la derivada respecto a la velocidad $\vec{v} = (\cos \theta, \sin \theta)$.

Dado un vector velocidad unitario $\vec{v} = (\cos \theta, \sin \theta) \in \mathbb{R}^2$, la interpretación de la derivada direccional $F_{\vec{v}}(x_0, y_0)$ es similar a la de las derivadas parciales (pues no dejan de ser un caso particular): Consideremos el plano

$$P_\theta = \{(x, y, z) \in \mathbb{R}^3 : -\sin \theta(x - x_0) + \cos \theta(y - y_0) = 0\}$$

La derivada direccional $F_{\vec{v}}(x_0, y_0)$ es la pendiente de la recta tangente a la curva dada por $C_\theta = \text{Gr}(F) \cap P_\theta$ dentro del plano P_θ (con el signo que marque el sentido de \vec{v}).



De inmediato veremos la relación entre las derivadas respecto a un vector velocidad y las derivadas parciales mediante el uso del gradiente y cómo calcularlas, pero para eso tenemos que tener garantizada la condición de diferenciabilidad que introducimos a continuación.

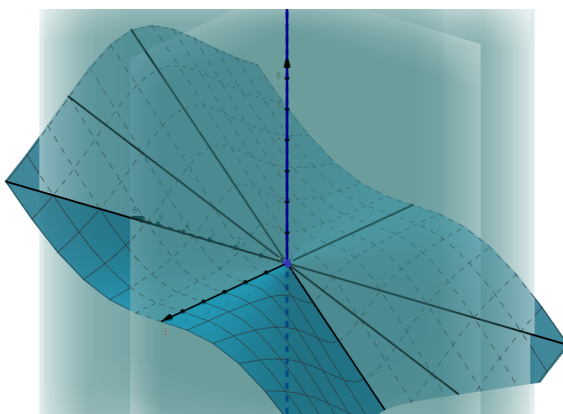
7.3 Diferenciabilidad. Plano tangente

Buscamos ahora el concepto análogo a la derivabilidad en una variable: la diferenciabilidad de funciones de varias variables. En el caso de una variable, la derivabilidad en un punto se corresponde geoméricamente con el hecho de que la curva definida por el grafo de la función tenga una recta tangente no vertical. Buscamos ahora, para una función de dos variables, una condición que garantice que su grafo tiene un *plano tangente* “no vertical” bien definido, esto es, un plano que contenga a todas las rectas tangentes de las

curvas en la superficie que pasan por el punto. La existencia de las derivadas parciales no es una condición suficiente, dado que solo garantiza que haya recta tangente en unas direcciones muy concretas (las de los ejes). Podría pensarse entonces que basta pedir que existan todas las derivadas direccionales, es decir, asegurar que hay una recta tangente en todas las direcciones posibles, pero aún así esto no resulta satisfactorio. En la figura observamos una superficie, la dada por el grafo de la función continua

$$F(x,y) = \begin{cases} x^3/(x^2+y^2) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Para F existen todas las derivadas direccionales en $(0,0)$, pues las intersecciones de la superficie con los distintos planos que contienen al eje z son rectas. En cambio, no hay plano que las contenga a todas ellas, pues para eso la propia superficie tendría que ser un plano, y no lo es. Esto nos lleva a introducir una noción más fina: la diferenciabilidad.



Una función $F : D \subset \mathbb{R}^2$ se dice *diferenciable en* $(x_0, y_0) \in D$ si se tiene que

$$\lim_{(h_1, h_2) \rightarrow (0,0)} \frac{F((x_0, y_0) + (h_1, h_2)) - F(x_0, y_0) - \langle \nabla F(x_0, y_0), (h_1, h_2) \rangle}{\|(h_1, h_2)\|} = 0.$$

Cuando F es una función diferenciable en (x_0, y_0) , podemos garantizar la existencia del *plano tangente en* $P = (x_0, y_0, F(x_0, y_0))$ *a la superficie* $Gr(F)$, que denotaremos mediante $T_{(x_0, y_0)}F$. Éste es un subespacio vectorial de \mathbb{R}^3 dado por

$$T_{(x_0, y_0)}F = \{(x, y, z) \in \mathbb{R}^3; z = F_x(x_0, y_0)x + F_y(x_0, y_0)y\}.$$

Una base de dicho plano vectorial es la dada por los vectores

$$(1, 0, F_x(x_0, y_0)), \quad (0, 1, F_y(x_0, y_0)).$$

Cualquier otro vector en este plano es de la forma

$$a(1, 0, F_x(x_0, y_0)) + b(0, 1, F_y(x_0, y_0)) = (a, b, F_{\vec{v}}(x_0, y_0)),$$

donde $F_{\vec{v}}(x_0, y_0)$ es la derivada respecto a la velocidad $\vec{v} = (a, b)$. Esto es, podemos garantizar para este caso que existen las derivadas respecto a cualquier vector velocidad $\vec{v} = (a, b) \in \mathbb{R}^2$, y que además se tiene que

$$F_{\vec{v}}(x, y) = aF_x(x, y) + bF_y(x, y) = \langle \vec{v}, \nabla F \rangle$$

■ **Ejemplo 7.4** Halla la derivada respecto a la velocidad $\vec{v} = (2, -3)$ de la función diferenciable $F(x, y) = x^2 + y^2$ en el punto $(1, 2)$. Después determina la derivada direccional en dicha dirección.

Tenemos que $\nabla F = (2x, 2y)$ y por tanto $\nabla F(1, 2) = (2, 4)$. Así que

$$F_{\vec{v}}(1, 2) = \langle (2, 4), (2, -3) \rangle = 2 \cdot 2 - 3 \cdot 4 = 4 - 12 = -8.$$

Ahora nos piden la derivada direccional, esto es, la derivada respecto a la velocidad normalizada

$$\vec{v}_{\text{un}} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{(2, -3)}{\sqrt{13}}$$

Tenemos que

$$F_{\vec{v}_{\text{un}}}(1, 2) = \frac{2}{\sqrt{13}} \cdot 2 + \frac{-3}{\sqrt{13}} \cdot 4 = \frac{-8}{\sqrt{13}} = \frac{F_{\vec{v}}(1, 2)}{\|\vec{v}\|}$$

■ **Ejemplo 7.5** Halla la derivada direccional de $F(x, y) = 4 + x^2 - 2y^2$ en el punto $P = (-1, 1)$ al desplazarse formando un ángulo de 45° con la dirección horizontal. ¿Qué ocurre si nos movemos con rapidez instantánea igual a $1/3$?

Primero nos están pidiendo la derivada direccional en la dirección de $\theta = \pi/4$, esto es la derivada respecto a la velocidad

$$\vec{v} = (\cos(45^\circ), \sin(45^\circ)) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right).$$

Tenemos que $\nabla F = (2x, -4y)$ y por tanto $\nabla F(-1, 1) = (-2, -4)$. Así que

$$F_{\vec{v}}(-1, 1) = -2 \cdot \frac{\sqrt{2}}{2} - 4 \cdot \frac{\sqrt{2}}{2} = -3\sqrt{2}.$$

Ahora, si nos movemos en la misma dirección pero con rapidez $1/3$, tenemos

$$F_{\frac{\vec{v}}{3}}(-1, 1) = \frac{F_{\vec{v}}(-1, 1)}{3} = -\sqrt{2}.$$

Usualmente nos gusta visualizar el plano tangente pegado al punto P , es decir, trasladarlo. Por eso consideramos el *plano tangente afín en $P = (x_0, y_0, F(x_0, y_0))$ a la superficie $\text{Gr}(F)$* , que denotamos mediante $T_{(x_0, y_0)}^A F$. Este viene dado por

$$T_{(x_0, y_0)}^A F = \{(x, y, z) \in \mathbb{R}^3; z = F_x(x_0, y_0)(x - x_0) + F_y(x_0, y_0)(y - y_0) + F(x_0, y_0)\}$$

■ **Ejemplo 7.6** Comprueba que el punto $P = (1, 1, 0)$ pertenece a la superficie $S = \text{Gr}(F)$, donde $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ es la función diferenciable dada por

$$F(x, y) = (x^4 - y^3)e^{xy}.$$

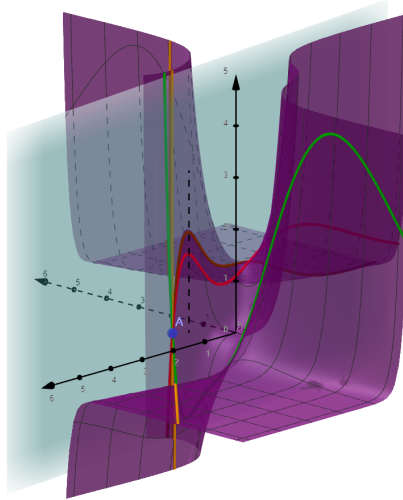
Encuentra la ecuación del plano tangente afín a S en el punto P .

Observamos que $F(1, -1) = (1^4 - 1^3)e^1 = (1 - 1)e^1 = 0$, así que efectivamente $P \in S$. Ahora necesitamos hallar el gradiente de F en $(1, 1)$. Tenemos que

$$\nabla F|_{(1,1)} = (4x^3 e^{xy} + (x^4 - y^3)ye^{xy}, -3y^2 e^{xy} + (x^4 - y^3)xe^{xy})|_{(1,1)} = (4e, -3e)$$

Ahora el plano tangente afín viene dado por

$$\begin{aligned} T_{(1,1)}^A F &= \{(x, y, z) \in \mathbb{R}^3 : z = 4e(x-1) - 3e(y-1) + 0\} = \\ &= \{(x, y, z) \in \mathbb{R}^3 : z = e(4x - 3y - 1)\}. \end{aligned}$$



El siguiente resultado nos da una condición suficiente de diferenciabilidad.

Teorema 7.3.1 Sea $F(x, y)$ una función de dos variables. Si las derivadas parciales F_x y F_y existen y son continuas en (x_0, y_0) , entonces se tiene garantizada la diferenciabilidad de F en (x_0, y_0) .

7.4 Propiedades del gradiente. Puntos críticos

A partir de este momento consideramos **funciones diferenciables**. El objetivo de esta sección es ver algunas características importantes del gradiente.

7.4.1 Direcciones de máximo crecimiento y decrecimiento

Vamos a suponer ahora que nos movemos siempre con velocidad unitaria, es decir, solo estamos considerando las derivadas direccionales (las derivadas respecto a velocidades unitarias). Vamos a determinar en qué direcciones crece más o menos rápido una función dada.

Consideremos una función diferenciable $F : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$, un punto (x_0, y_0) y una dirección $\theta \in [0, 2\pi)$. Decimos que θ es:

- Una *dirección de máximo crecimiento* en (x_0, y_0) si para todo $\phi \in [0, 2\pi)$

$$F_{(\cos \theta, \sin \theta)}(x_0, y_0) \geq F_{(\cos \phi, \sin \phi)}(x_0, y_0),$$

- Una *dirección de máximo decrecimiento* en (x_0, y_0) si para todo $\phi \in [0, 2\pi)$

$$F_{(\cos \theta, \sin \theta)}(x_0, y_0) \leq F_{(\cos \phi, \sin \phi)}(x_0, y_0),$$

- Una *dirección de crecimiento nulo* en (x_0, y_0) si

$$F_{(\cos \theta, \sin \theta)}(x_0, y_0) = 0.$$

Si hay dos direcciones linealmente independientes de crecimiento nulo en un punto, entonces necesariamente se anula el gradiente en dicho punto y por tanto todas las direcciones son de crecimiento nulo para ese punto. Diremos que se trata de un *puntos crítico* de F .

Supongamos ahora que (x_0, y_0) no es un punto crítico de F . Dada una dirección unitaria $\theta \in [0, 2\pi)$, recordemos que tenemos la igualdad

$$F_{\vec{v}}(x_0, y_0) = \langle \nabla F(x_0, y_0), \vec{v} \rangle = \|\nabla F(x_0, y_0)\| \cdot \cos \Phi,$$

donde $\vec{v} = (\cos \theta, \sin \theta)$ y Φ es el ángulo entre $\nabla F(x_0, y_0)$ y \vec{v} . Así, el gradiente $\nabla F(x_0, y_0)$ nos permite reconocer las direcciones de crecimiento y decrecimiento como sigue:

- El gradiente $\nabla F(x_0, y_0)$ determina la dirección de máximo crecimiento de la función F en el punto (x_0, y_0) , pues el ángulo en la expresión anterior es $\Phi = 0$ y por tanto $\cos(0) = 1$.
- El vector en sentido opuesto al gradiente $-\nabla F(x_0, y_0)$ determina la dirección de máximo decrecimiento de la función F en el punto (x_0, y_0) , pues el ángulo en la expresión anterior es $\Phi = \pi$ y por tanto $\cos(\pi) = -1$.
- Las direcciones ortogonales al gradiente son de crecimiento nulo, pues el ángulo que forman con $\nabla F(x_0, y_0)$ es o bien $\Phi = \pi/2$ o bien $\Phi = 3\pi/2$ y sabemos que $\cos(\pi/2) = \cos(3\pi/2) = 0$. Las direcciones de crecimiento nulo marcan la recta tangente a la curva de nivel de F que contiene a (x_0, y_0) .

- **Ejemplo 7.7** La temperatura en cada punto de una hoja de metal viene dada mediante la función

$$T(x, y) = e^x \cos(y) + e^y \cos(x)$$

- ¿En qué dirección crece la temperatura más rápidamente en el punto $(0, 0)$? ¿Cuál es la tasa de incremento?
 - ¿En qué dirección decrece la temperatura más rápido en el punto $(0, 0)$?
 - ¿Qué direcciones son tangentes a la curva de nivel de la función T en el punto $(0, 0)$?
- En primer lugar necesitamos calcular las derivadas parciales para obtener el gradiente en $(0, 0)$. Tenemos que

$$T_x(x, y) = e^x \cos(y) + e^y \sin(x), \quad T_y(x, y) = e^x \sin(y) + e^y \cos(x)$$

- La dirección de máximo crecimiento es la que marca

$$\nabla T(0, 0) = (T_x(0, 0), T_y(0, 0)) = (1, 1).$$

Esto si lo escribimos en coordenadas polares es $r = \|\nabla T(0, 0)\| = \sqrt{1+1} = \sqrt{2}$ y $\theta = \pi/4$. Así, la dirección de máximo crecimiento es $\pi/4$. La tasa de incremento viene dada por la derivada direccional en dicha dirección, que no es más que la norma del gradiente, esto es $r = \sqrt{2}$.

- La dirección de máximo decrecimiento es la opuesta al gradiente, esto es la dada por

$$-\nabla T(0, 0) = (-1, -1).$$

es decir, la marcada por $\pi/4 + \pi = 5\pi/4$.

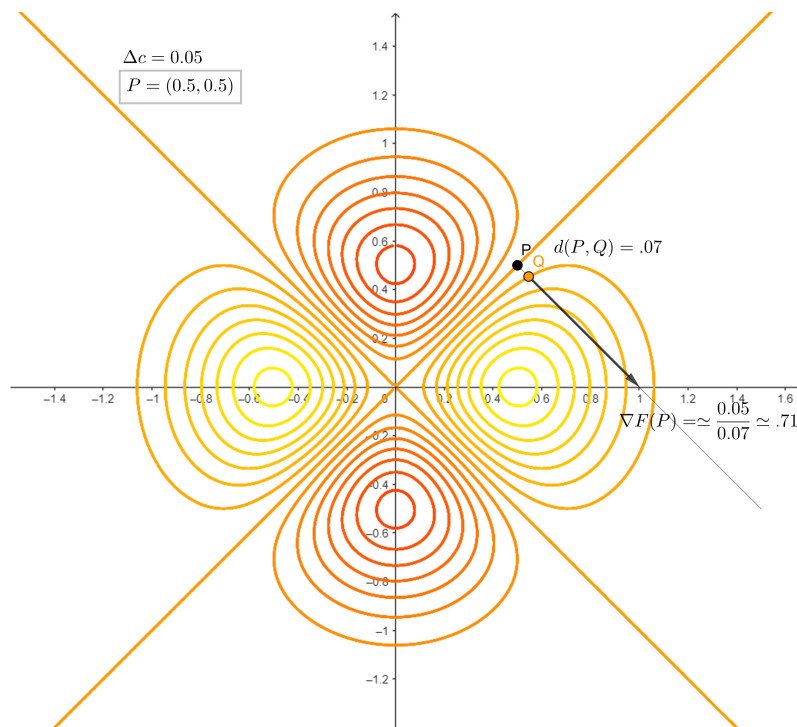
c) Las direcciones tangentes a la curva de nivel en $(0,0)$ son las de crecimiento nulo: las ortogonales al gradiente. Así son las marcadas por $(1, -1)$ y $(-1, 1)$. Es decir, las dadas por los ángulos

$$\pi/4 + \pi/2 = 3\pi/4, \quad \pi/4 + 3\pi/2 = 7\pi/4.$$

7.4.2 El gradiente mediante curvas de nivel

Vamos a ver cómo identificar la dirección del gradiente y su tamaño a partir del mapa de curvas de nivel de una función. En primer lugar lo haremos de modo visual, sin conocer de forma explícita la expresión analítica de la función considerada. Después comprobaremos que aquello que intuíamos es lo correcto.

En el mapa de curvas de nivel de la siguiente figura se han equiespaciado los niveles a $\Delta c = 0,05$, marcando las curvas de menor nivel con un tono rojo y las de mayor nivel con un tono amarillo, cambiando el tono de progresivamente al aumentar o disminuir el nivel. Buscamos determinar el gradiente en el punto $P = (0,5, 0,5)$.



1. Lo primero que sabemos es que las direcciones ortogonales al gradiente son de crecimiento nulo, y por lo tanto son tangentes a una curva de nivel, más precisamente a la curva de nivel que contiene al punto. En este caso la curva de nivel que contiene a P es una recta (con vector director $(1, 1)$) y por tanto el gradiente tiene una dirección ortogonal a dicha recta.
2. Debemos averiguar ahora el sentido del vector gradiente. Conforme las curvas van siendo más amarillas estamos en niveles más altos, por lo tanto el gradiente debe apuntar hacia las curvas amarillas. Con esto averiguamos que el gradiente tiene el mismo sentido que el $(1, -1)$.

3. Nos queda ahora estimar el tamaño del gradiente. Para ello consideramos el punto Q de intersección de la recta que pasa por P y está dirigida por el gradiente y la siguiente curva de nivel: esta es la mínima distancia entre P y la curva de nivel que contiene a Q . El tamaño del gradiente es el cociente entre el incremento de nivel y la distancia recorrida, esto es

$$\|\nabla F(0,0)\| = \frac{\Delta c}{d(P,Q)}$$

Para este caso estimamos que la distancia entre P y Q es aproximadamente 0,07. Como $\Delta c = 0,05$, concluimos que el tamaño del gradiente es

$$\|\nabla F(0,0)\| \simeq \frac{0,05}{0,07} \simeq 0,71$$

La función cuyo mapa de curvas de nivel es el anterior, es la dada por

$$F(x,y) = 4(x^2 - y^2)e^{-4(x^2+y^2)}.$$

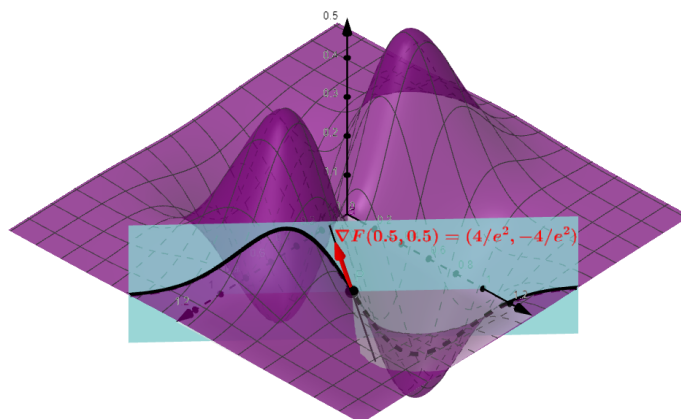
Tenemos que su gradiente es

$$\nabla F(x,y) = (-8x(4x^2 - 4y^2 - 1)e^{-4(x^2+y^2)}, -8y(4x^2 - 4y^2 + 1)e^{-4(x^2+y^2)})$$

y por tanto $\nabla F(0,5,0,5) = (\frac{4}{e^2}, -\frac{4}{e^2})$, que efectivamente apunta en la dirección y sentido considerados. Observamos también que

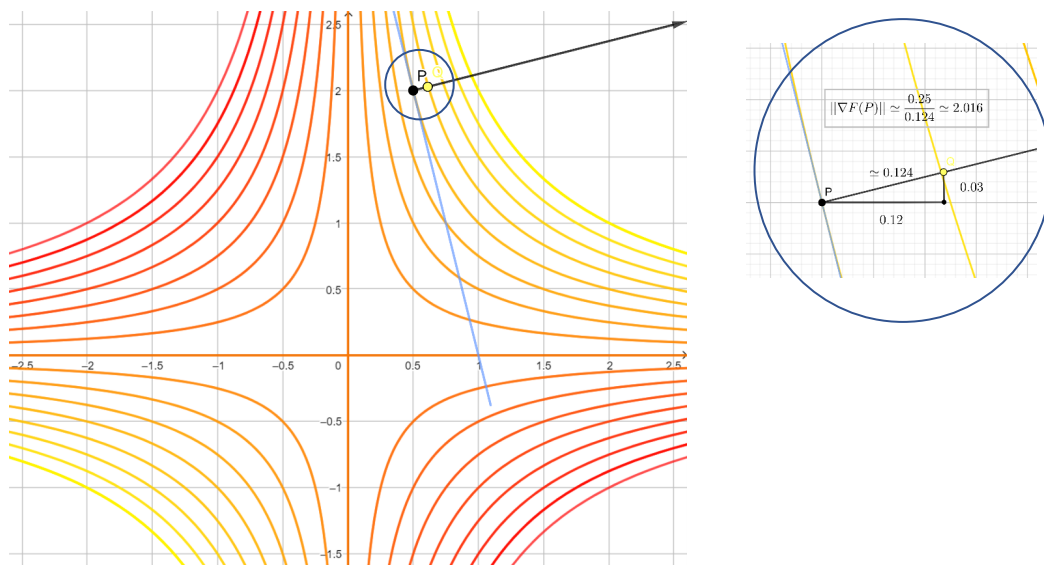
$$\|\nabla F(0,5,0,5)\| = \sqrt{\frac{16}{e^4} + \frac{16}{e^4}} = \frac{\sqrt{32}}{e^2} = \frac{\sqrt{24}}{e^2} \simeq 0,76$$

El grafo de la función F es la superficie



donde se puede apreciar que la dirección marcada por el gradiente en $(0,5,0,5)$ es la que proporciona el máximo crecimiento en dicho punto.

■ **Ejemplo 7.8** Consideremos el siguiente mapa de curvas de nivel



Con las indicaciones anteriores, hemos representado en la figura $\nabla F(1/2, 2)$, que estimamos que tiene la dirección y el sentido del vector $(4, 1)$, pero con módulo aproximado de 2,016. Es decir, estimamos que

$$\nabla F(1/2, 2) \simeq \left(\frac{2,016 \cdot 4}{\sqrt{17}}, \frac{2,016 \cdot 1}{\sqrt{17}} \right) \simeq (1,46, 0,49).$$

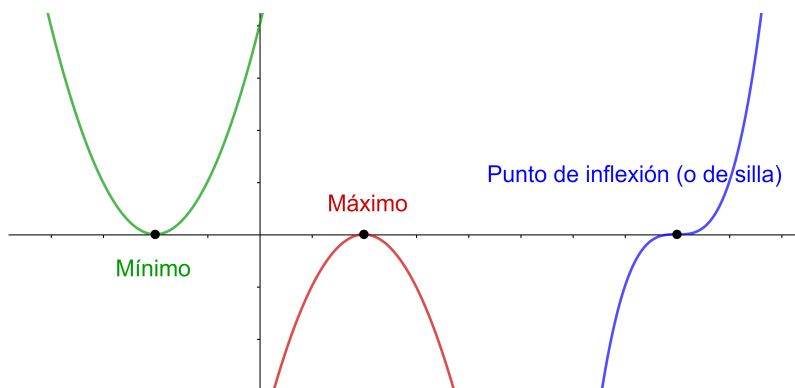
Ahora, éste es el mapa de curvas de nivel de la función

$$F(x, y) = xy$$

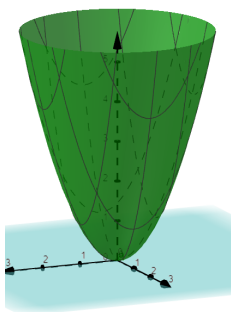
Por tanto $\nabla F(x, y) = (y, x)$ y en particular $\nabla F(1/2, 2) = (2, 1/2)$, que es similar a lo estimado.

7.4.3 Puntos críticos

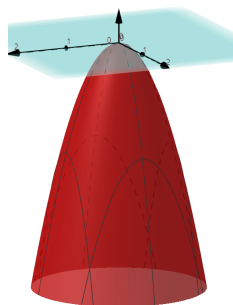
Consideremos $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ una función de una variable. Recordemos que $x = x_0$ es un *punto crítico* de f si la recta tangente a la gráfica de f es horizontal en $(x_0, f(x_0))$ o equivalentemente si $f'(x_0) = 0$. Las situaciones que se pueden dar son las siguientes:



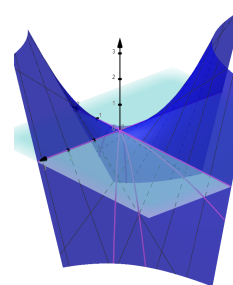
Consideremos ahora una función $F : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$. Como ya hemos dicho, un punto (x_0, y_0) es *crítico para F* si su gradiente $\nabla F(x_0, y_0)$ se anula. Esto equivale a pedir que el plano tangente a la gráfica de F en el punto $(x_0, y_0, F(x_0, y_0))$ sea $z = F(x_0, y_0)$. Estas son las situaciones que podemos encontrar:



Mínimo relativo



Máximo relativo



Punto de silla

Observación: En un punto de silla algunas de las trayectorias (curvas) contenidas en la superficie tienen un máximo relativo y otras tienen un mínimo relativo. También podemos encontrar trayectorias que tengan un punto de inflexión.

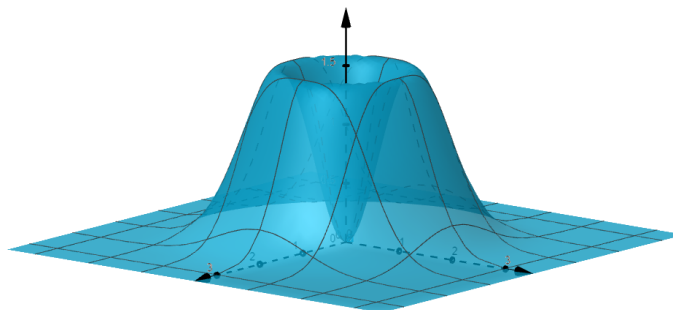
■ **Ejemplo 7.9** Vamos a hallar los puntos críticos de la función

$$F(x, y) = 4(x^2 + y^2)e^{-(x^2 + y^2)}.$$

Para ello veamos dónde se anula el gradiente. Tenemos que

$$F_x = 8x(1 - x^2 - y^2)e^{-(x^2 + y^2)}, \quad F_y = 8y(1 - x^2 - y^2)e^{-(x^2 + y^2)}$$

Los puntos críticos de F son el $(0, 0)$ y todos aquellos que satisfacen la ecuación de la circunferencia $x^2 + y^2 = 1$.



En el caso de funciones de una variable la segunda derivada nos da un criterio para discernir si los puntos críticos son máximos, mínimos o puntos de silla. Queremos tener un criterio similar para funciones de dos variables. Para ello debemos hablar de derivadas parciales de orden superior.

7.5 Derivadas parciales de orden superior

Como en las derivadas de las funciones de una variable, es posible calcular derivadas parciales segundas, terceras, o de orden mayor. Además, ahora podemos tener derivadas de orden superior mixtas, es decir, primero derivar respecto de x y después respecto de y o viceversa. Usamos la siguiente notación para describir las *derivadas parciales de segundo orden* de $F(x, y)$:

$$(F_x)_x = F_{xx} = \frac{\partial F_x}{\partial x} = \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial x} \right) = \frac{\partial^2 F}{\partial x^2}$$

$$(F_x)_y = F_{xy} = \frac{\partial F_x}{\partial y} = \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial x} \right) = \frac{\partial^2 F}{\partial y \partial x}$$

$$(F_y)_x = F_{yx} = \frac{\partial F_y}{\partial x} = \frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right) = \frac{\partial^2 F}{\partial x \partial y}$$

$$(F_y)_y = F_{yy} = \frac{\partial F_y}{\partial y} = \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial y} \right) = \frac{\partial^2 F}{\partial y^2}$$

■ **Observación 7.1** Nótese que en las distintas notaciones para las derivadas parciales mixtas $F_{xy} = \partial^2 F / \partial yx$ se invierte el orden entre la x y la y . Para recordar cuál es el orden de derivación adecuado conviene tener en mente las igualdades de arriba.

Teorema 7.5.1 — Schwarz. Si las derivadas parciales mixtas F_{xy} y F_{yx} son continuas entorno a un punto (x_0, y_0) entonces son iguales en dicho punto, esto es

$$F_{xy}(x_0, y_0) = F_{yx}(x_0, y_0).$$

■ **Ejemplo 7.10** Halla las derivadas parciales de segundo orden de la función $F(x, y) = \sin(x)y^2$. Evalúa $F_{xx}(\pi/2, 1)$. Para ello, en primer lugar buscamos las derivadas parciales primeras

$$F_x = \frac{\partial F}{\partial x} = \cos(x)y^2, \quad F_y = \frac{\partial F}{\partial y} = 2 \sin(x)y$$

Ahora derivamos a su vez cada una de estas respecto a x y respecto a y :

$$\begin{aligned} F_{xx} &= \frac{\partial F_x}{\partial x} = \frac{\partial^2 F}{\partial x^2} = -\sin(x)y^2, & F_{yx} &= \frac{\partial F_y}{\partial x} = \frac{\partial^2 F}{\partial xy} = 2 \cos(x)y. \\ F_{xy} &= \frac{\partial F_x}{\partial y} = \frac{\partial^2 F}{\partial yx} = 2 \cos(x)y, & F_{yy} &= \frac{\partial F_y}{\partial y} = \frac{\partial^2 F}{\partial y^2} = 2 \sin(x). \end{aligned}$$

Obsérvese que $F_{xy} = F_{yx}$. Por último, tenemos $F_{xx}(\pi/2, 1) = -\sin(\pi/2) \cdot 1^2 = -1$.

Para el caso de funciones de tres variables tenemos 9 derivadas parciales de segundo orden:

$$F_{xx}, F_{xy}, F_{xz}, F_{yx}, F_{yy}, F_{yz}, F_{zx}, F_{zy} \text{ y } F_{zz}$$

De nuevo, si las derivadas parciales mixtas son continuas, tenemos las igualdades

$$F_{xy} = F_{yx}, \quad F_{xz} = F_{zx}, \quad \text{y} \quad F_{yz} = F_{zy}.$$

■ **Ejemplo 7.11** Consideremos la función de tres variables $F(x, y, z) = xe^y \sin(\pi z)$. Vamos a proporcionar las derivadas parciales de segundo orden. Vemos primero las de primer orden:

$$F_x = e^y \sin(\pi z), \quad F_y = xe^y \sin(\pi z), \quad F_z = \pi xe^y \cos(\pi z)$$

Ahora tenemos que:

$$F_{xx} = 0, \quad F_{yy} = xe^y \sin(\pi z), \quad F_{zz} = -\pi^2 xe^y \sin(\pi z)$$

$$F_{xy} = e^y \sin(\pi z) = F_{yx}, \quad F_{yz} = \pi xe^y \cos(\pi z) = F_{zy}, \quad F_{zx} = \pi e^y \cos(\pi z) = F_{xz}$$

7.6 Matriz Hessiana. Cuádrica aproximante

La serie de Taylor en un punto x_0 de una función $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ suficientemente derivable viene dada por

$$s_f(x_0) = f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2!} + \cdots + f^{(n)}(x_0) \frac{x^n}{n!} + \cdots$$

Esta expansión es infinita si la función es infinitamente derivable. La aproximación de primer orden recupera la ecuación de la recta tangente. Es decir, la recta tangente es la mejor aproximación lineal de la curva $y = f(x)$. La aproximación de orden dos (parabólica) nos permite dar un criterio para determinar el tipo de puntos críticos (mínimo máximo o inflexión).

De modo análogo, se puede definir la serie de Taylor para una función de más variables. Escribimos aquí cómo es el *desarrollo de Taylor de orden 2 en (x_0, y_0)* de una función $F(x, y)$.

$$s_{F, (x_0, y_0)} = F(x_0, y_0) + \nabla F(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \frac{1}{2} (x - x_0, y - y_0) H_F(x_0, y_0) \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

donde $H_F(x_0, y_0)$ es la *matriz Hessiana* formada por las derivadas parciales segundas. Más precisamente

$$H_F(x_0, y_0) = \begin{pmatrix} F_{xx}(x_0, y_0) & F_{yx}(x_0, y_0) \\ F_{xy}(x_0, y_0) & F_{yy}(x_0, y_0) \end{pmatrix}$$

El desarrollo de Taylor de orden uno recupera la ecuación del plano tangente y por tanto es la mejor aproximación lineal de $\text{Gr}(F)$ en el punto $P = (x_0, y_0, F(x_0, y_0))$. El desarrollo de Taylor de orden dos nos proporciona la cuádrica que mejor se ajusta a la superficie $\text{Gr}(F)$ en el punto P : la dada por el grafo de $s_{F, (x_0, y_0)}$. Llamaremos a ésta *la cuádrica aproximante de F en (x_0, y_0)* .

■ **Ejemplo 7.12** Calcula la matriz Hessiana de la función $F(x, y) = \sin(x)y^2$ en el punto $(\pi/2, 1)$. Determina después la ecuación de la cuádrica que mejor aproxima al grafo de F en el punto $(\pi/2, 1)$.

En el ejercicio 7.10 hemos calculado las derivadas parciales de primer y segundo orden de F , así que

$$\nabla F(x, y) = (\cos(x)y^2, 2\sin(x)y), \quad H_F(x, y) = \begin{pmatrix} -\sin(x)y^2 & 2\cos(x)y \\ 2\cos(x)y & 2\sin(x) \end{pmatrix}.$$

Sustituyendo en el punto $(\pi/2, 1)$ llegamos a

$$\nabla F\left(\frac{\pi}{2}, 1\right) = (0, 2), \quad H_F\left(\frac{\pi}{2}, 1\right) = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$$

Así que la ecuación de la cuádrica que buscamos es

$$\begin{aligned} s_{F,(\frac{\pi}{2},1)} &= \sin\left(\frac{\pi}{2}\right)1^2 + (0,2) \begin{pmatrix} x - \frac{\pi}{2} \\ y - 1 \end{pmatrix} + \frac{1}{2}(x - \frac{\pi}{2}, y - 1) \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x - \frac{\pi}{2} \\ y - 1 \end{pmatrix} = \\ &= 1 + 2(y - 1) - \frac{1}{2}\left(x - \frac{\pi}{2}\right)^2 + (y - 1)^2 = -\frac{1}{2}x^2 + y^2 + \pi x - \frac{\pi^2}{8}. \end{aligned}$$

7.7 Análisis de puntos críticos

En general, vamos a considerar funciones F con derivadas parciales mixtas continuas. Notemos que para estas tenemos que $F_{xy} = F_{yx}$ y por tanto la matriz Hessiana es simétrica. En particular es diagonalizable y tiene una base de vectores propios ortonormales.

Fijemos un punto crítico (x_0, y_0) , esto es, un punto tal que $\nabla F(x_0, y_0) = 0$. Diagonalizando escribimos

$$H_F(x_0, y_0) = QD_FQ^t, \quad D_F = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

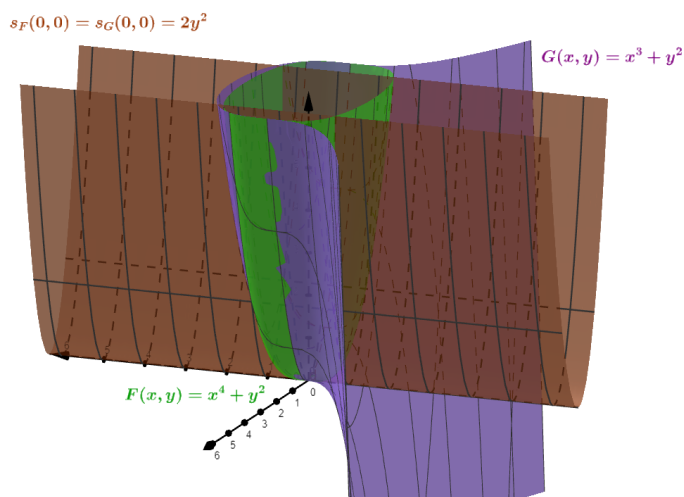
y definimos $(u, v) = (x - x_0, y - y_0)Q$. Sustituyendo estas expresiones en el desarrollo de Taylor de orden dos de F en (x_0, y_0) llegamos a que

$$s_{F,(x_0,y_0)} = F(x_0, y_0) + \frac{1}{2}(u, v)D_F(u, v)^t = F(x_0, y_0) + \frac{1}{2}(\lambda_1 u^2 + \lambda_2 v^2)$$

Notemos que $\det H_F(x_0, y_0) = |H_F(x_0, y_0)| = |QD_FQ^t| = |D_F| = \lambda_1 \lambda_2$.

Distinguiamos los siguientes casos:

- Si los dos valores propios λ_1 y λ_2 son positivos, la cuádrica que mejor aproxima a $\text{Gr}(F)$ es un paraboloide y por tanto tenemos la situación de mínimo relativo. Observamos que en este caso $|H_F(x_0, y_0)| > 0$ y se puede probar que además se tiene que $F_{xx}(x_0, y_0) > 0$.
- Si los dos valores propios λ_1 y λ_2 son negativos, la cuádrica que mejor aproxima a $\text{Gr}(F)$ es un paraboloide invertido y por tanto tenemos la situación de máximo relativo. Observamos que en este caso $|H_F(x_0, y_0)| > 0$ y se puede probar que además se tiene que $F_{xx}(x_0, y_0) < 0$.
- Si hay un valor propio positivo y uno negativo, la cuádrica que mejor aproxima a $\text{Gr}(F)$ es una silla de montar y por lo tanto tenemos la situación de punto de silla. Observamos que en este caso $|H_F(x_0, y_0)| < 0$.
- Si uno de los valores propios es nulo o equivalentemente si $|H_F(x_0, y_0)| = 0$, no podemos concluir. A continuación representamos dos superficies, definidas por los grafos de $F(x, y) = x^4 + y^2$ y $G(x, y) = x^3 + y^2$, para las cuales observamos comportamientos distintos en el $(0, 0)$. En cambio tienen la misma matriz Hessiana en $(0, 0)$ y también la misma cuádrica aproximante. En estos casos es cuando diríamos que la cuádrica aproximante es “degenerada”. En el ejemplo de la figura es un cilindro parabólico.



En resumen, podemos considerar el siguiente criterio *de la cuádrica aproximante* para distinguir de qué tipo son los puntos críticos:

- Si $\det[H_F(x_0, y_0)] = 0$ no podemos garantizar nada.
- Si $\det[H_F(x_0, y_0)] < 0$, tenemos que F tiene un punto de silla en (x_0, y_0) .
- Si $\det[H_F(x_0, y_0)] > 0$, diferenciamos dos casos:
 - Si $F_{xx}(x_0, y_0) > 0$, entonces F tiene un mínimo relativo en (x_0, y_0) .
 - Si $F_{xx}(x_0, y_0) < 0$, entonces F tiene un máximo relativo en (x_0, y_0) .

■ **Ejemplo 7.13** Determina los puntos críticos de la superficie definida por la función $F(x, y) = xy(1 - x - y)$. Después clasifícalos con el criterio de la cuádrica aproximante.

En primer lugar necesitamos las derivadas parciales de F . Tenemos que

$$F_x = y(1 - x - y) - xy = y(1 - 2x - y)$$

$$F_y = x(1 - x - y) - xy = x(1 - x - 2y)$$

Ahora los puntos críticos son aquellos donde se anulan ambas derivadas parciales a la vez, es decir, donde se anula el gradiente. Así, se debe cumplir que

$$y(1 - 2x - y) = 0, \quad x(1 - x - 2y) = 0$$

Aquí vemos que si $y = 0$, entonces x es o bien 0 o bien 1. Así $P_1 = (0, 0)$ y $P_2 = (1, 0)$ son puntos críticos. Si $x = 0$ tenemos que o bien $y = 0$ o bien $y = 1$. De aquí rescatamos el punto crítico $P_3 = (0, 1)$. Finalmente podemos tener

$$1 - 2x - y = 1 - x - 2y = 0 \Rightarrow x = y \text{ y } 1 - 3x = 0 \Rightarrow x = y = \frac{1}{3}.$$

Por lo que todavía hay un cuarto punto crítico, el punto $P_4 = (1/3, 1/3)$. Para poder determinar de qué tipo es cada uno de estos puntos, necesitamos la matriz hessiana de las derivadas segundas. Tenemos que

$$F_{xx} = -2y, \quad F_{xy} = F_{yx} = 1 - 2x - y - y = 1 - 2x - 2y, \quad F_{yy} = -2x$$

Por tanto la matriz Hessiana es

$$H_F(x, y) = \begin{pmatrix} -2y & 1 - 2x - 2y \\ 1 - 2x - 2y & -2x \end{pmatrix}$$

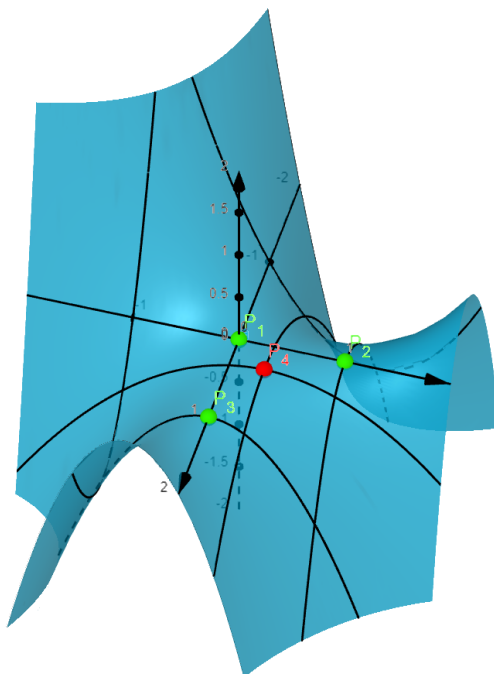
y su determinante es $\det[H_F(x,y)] = 4xy - (1 - 2x - 2y)^2$. Por tanto:

- En $P_1 = (0,0)$ tenemos $|H_F(0,0)| = -1^2 = -1 < 0$. Punto de silla.
- En $P_2 = (0,1)$ tenemos $|H_F(0,1)| = -(1-2)^2 = -1 < 0$. Punto de silla.
- En $P_3 = (1,0)$ tenemos $|H_F(1,0)| = -(1-2)^2 = -1 < 0$. Punto de silla.
- En $P_4 = (1/3, 1/3)$ tenemos

$$\left| H_F \left(\frac{1}{3}, \frac{1}{3} \right) \right| = \frac{4}{9} - \left(1 - \frac{2}{3} - \frac{2}{3} \right)^2 = \frac{1}{3} > 0.$$

Como además $F_{xx}(1/3, 1/3) = -2/3 < 0$, tenemos que P_4 es un máximo relativo.

Tenemos a continuación la representación de la superficie definida por esta función. En ella observamos los tres puntos de sillas y el máximo relativo.



■ **Ejemplo 7.14** Determina los puntos críticos de la superficie definida por la función $F(x,y) = (xy + y^2)e^{-x^2}$. Después clasifícalos con el criterio de la cuádrica aproximante.

En primer lugar necesitamos las derivadas parciales de F . Tenemos que

$$F_x = ye^{-x^2} - 2x(xy + y^2)e^{-x^2} = (y - 2x^2y - 2xy^2)e^{-x^2}, \quad F_y = (x + 2y)e^{-x^2}$$

Ahora los puntos críticos son aquellos donde se anulan ambas derivadas parciales a la vez, es decir, donde se anula el gradiente. Así, se debe cumplir que

$$x + 2y = 0, \quad y - 2x^2y - 2xy^2 = 0$$

Tenemos que $x = -2y$ y sustituyendo en la segunda ecuación llegamos a que

$$y - 2(-2y)^2y - 2(-2y)y^2 = 0,$$

esto es $y - 4y^3 = 0$. De aquí sacamos $y = 0$, $y = -1/2$ o $y = 1/2$. De esta manera, los puntos críticos son

$$P_1 = (0,0), \quad P_2 = (1, -1/2), \quad P_3 = (-1, 1/2)$$

Para poder determinar de qué tipo es cada uno de estos puntos, necesitamos la matriz hessiana de las derivadas segundas. Tenemos que

$$F_{xx} = (-4xy - 2y^2 - 2x(y - 2x^2y - 2xy^2))e^{-x^2} = (-6xy - 2y^2 + 4x^3y + 4x^2y^2)e^{-x^2}$$

$$F_{xy} = F_{yx} = (1 - 4x^2 - 4xy)e^{-x^2}, \quad F_{yy} = 2e^{-x^2}$$

Por tanto tenemos que

- En $P_1 = (0,0)$ la matriz Hessiana es

$$H_F(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

cuyo determinante es $-1 < 0$. Estamos ante un punto de silla.

- En $P_2 = (1, -1/2)$ y $P_3 = (-1, 1/2)$ la matriz Hessiana es

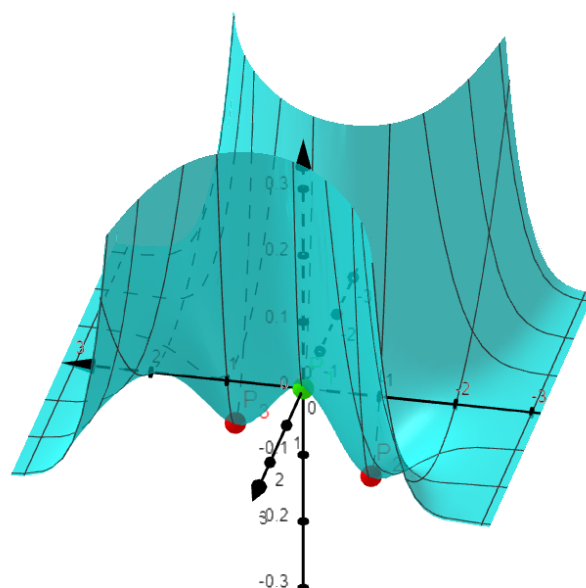
$$H_F(1, -1/2) = H_F(-1, 1/2) = \begin{pmatrix} \frac{3}{2e} & \frac{-1}{e} \\ \frac{-1}{e} & \frac{2}{e} \end{pmatrix}$$

cuyo determinante es $2/e^2 > 0$. Como además

$$F_{xx}(1, -1/2, -1) = F_{xx}(-1, 1/2) = 3/2e > 0,$$

estamos ante mínimos relativos.

Tenemos a continuación la representación de la superficie definida por esta función. En ella observamos los dos mínimos relativos y el punto de silla.



7.8 Aplicaciones vectoriales. Matriz Jacobiana

Una aplicación vectorial de una o más variables es una asignación

$$\varphi : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$$

que a cada elemento $(x_1, x_2, \dots, x_n) \in D$ le asigna un vector de \mathbb{R}^m (en lugar de un escalar). Al igual que siempre, diremos que D es el *dominio de φ* y que

$$\text{Im}(\varphi) = \{\varphi(x_1, x_2, \dots, x_n) : (x_1, x_2, \dots, x_n) \in D\} \subset \mathbb{R}^m$$

es su *imagen*. Hemos tratado ya en este curso con aplicaciones vectoriales:

- Las transformaciones lineales de $\mathbb{R}^n \rightarrow \mathbb{R}^n$, como por ejemplo la dada por $\varphi(x, y) = (2x - 3y, x - y)$. En este caso $n = m = 2$.
- Las parametrizaciones de curvas en el plano y en el espacio. Por ejemplo, la dada por $\gamma(t) = (\cos t, \sin t, t)$ que recorre una hélice. En este caso $n = 1$ y $m = 3$. Aquí se incluyen las curvas parametrizadas obtenidas como solución de sistemas de ecuaciones diferenciales.
- También el gradiente de una función $F : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ es una aplicación vectorial

$$\nabla F : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (x, y, z) \mapsto \nabla F(x, y, z).$$

En este caso $n = m = 3$.

- Los cambios de coordenadas a polares, cilíndricas y esféricas.

Las aplicaciones vectoriales, entendiéndolas como “campos de vectores”, las estudiaremos con mayor detalle en el último tema de esta asignatura. Ahora las introducimos para poder hablar con más precisión de cambios de coordenadas y de la regla de la cadena para funciones de varias variables.

Por comodidad tomamos el caso $n = m = 3$, aunque las definiciones que vamos a dar sirven en general, como veremos en los ejemplos. Consideramos una aplicación vectorial

$$\varphi = (\varphi_1, \varphi_2, \varphi_3) : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

La *matriz jacobiana de φ en un punto $P = (x_0, y_0, z_0)$* viene dada por

$$\text{Jac}_\varphi(P) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x} & \frac{\partial \varphi_1}{\partial y} & \frac{\partial \varphi_1}{\partial z} \\ \frac{\partial \varphi_2}{\partial x} & \frac{\partial \varphi_2}{\partial y} & \frac{\partial \varphi_2}{\partial z} \\ \frac{\partial \varphi_3}{\partial x} & \frac{\partial \varphi_3}{\partial y} & \frac{\partial \varphi_3}{\partial z} \end{pmatrix} \Big|_P$$

El subíndice P en la expresión anterior significa que se debe evaluar en el punto P cada una de las derivadas parciales. En general, la matriz Jacobiana se construye al colocar por **filas** las distintas derivadas parciales de cada una de las componentes de φ . Así, es una matriz de tamaño $m \times n$, donde m es la dimensión del espacio de llegada de φ y n la de su dominio, es decir, el número de variables.

■ **Ejemplo 7.15** Determina la matriz Jacobiana de la aplicación $\varphi(x) = (\sin(x), e^x)$ en el punto $x = 0$.

Podemos escribir $\varphi_1(x) = \sin(x)$ y $\varphi_2(x) = e^x$. Tenemos que

$$\varphi_1'(x) = \cos(x), \quad \varphi_2'(x) = e^x$$

Así la matriz Jacobiana en el punto $x = 0$ es

$$\text{Jac}_\varphi(0) = \begin{pmatrix} \cos(x) \\ e^x \end{pmatrix}_{|_0} = \begin{pmatrix} \cos(0) \\ e^0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

■ **Ejemplo 7.16** Determina la matriz Jacobiana de $\varphi(x, y) = (x^2 + y^2, xy)$ en el punto $(0, 1)$.

Podemos escribir $\varphi_1(x, y) = x^2 + y^2$ y $\varphi_2(x, y) = xy$. Tenemos

$$\varphi_{1,x}(x, y) = 2x, \quad \varphi_{1,y}(x, y) = 2y, \quad \varphi_{2,x}(x, y) = y, \quad \varphi_{2,y}(x, y) = x.$$

De esta manera

$$\text{Jac}_\varphi(0, 1) = \begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}_{|(0,1)} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

■ **Ejemplo 7.17** Determina la matriz Jacobiana en el punto $(1, 1, 3\pi/2)$ de

$$\varphi(x, y, z) = (e^{xy}, y + \cos(z)).$$

Podemos escribir $\varphi_1(x, y, z) = e^{xy}$ y $\varphi_2(x, y, z) = y + \cos(z)$. Tenemos que

$$\varphi_{1,x} = ye^{xy}, \quad \varphi_{1,y} = xe^{xy}, \quad \varphi_{1,z} = 0, \quad \varphi_{2,x} = 0, \quad \varphi_{2,y} = 1, \quad \varphi_{2,z} = -\sin(z).$$

De esta forma

$$\text{Jac}_\varphi(1, 1, 3\pi/2) = \begin{pmatrix} ye^{xy} & xe^{xy} & 0 \\ 0 & 1 & -\sin(z) \end{pmatrix}_{(1,1,3\pi/2)} = \begin{pmatrix} e & e & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

■ **Ejemplo 7.18** Determina la matriz Jacobiana en el punto $(1, 3\pi/2, 4)$ de

$$\varphi(x, y, z) = (x \cos(y), x \sin(y), z).$$

Podemos escribir $\varphi_1(x, y, z) = x \cos(y)$, $\varphi_2(x, y, z) = x \sin(y)$ y $\varphi_3(x, y, z) = z$. Tenemos que

$$\text{Jac}_\varphi(1, 3\pi/2, 4) = \begin{pmatrix} \cos(y) & -x \sin(y) & 0 \\ \sin(y) & x \cos(y) & 0 \\ 0 & 0 & 1 \end{pmatrix}_{(1,3\pi/2,4)} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Notación:

- Cuando φ es una función de una sola variable, esto es $\varphi : D \subset \mathbb{R} \rightarrow \mathbb{R}^m$, denotamos por comodidad mediante φ' a la matriz Jacobiana, que en este caso es de tamaño $n \times 1$.
- En el caso en el que $\varphi = F$ es una función escalar, esto es $F : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, tenemos que Jac_F es una matriz de tamaño $1 \times n$ y de hecho

$$\text{Jac}_F = \nabla F.$$

7.9 Cambio de variable. Regla de la cadena

7.9.1 Regla de la cadena (sobre una curva)

Podemos tener dos variables llamadas x e y , pero que a su vez sepamos que dependen de otra variable llamada t . Dicho de otro modo, conocemos una aplicación $\varphi : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por

$$\varphi(t) = (x(t), y(t))$$

que determina dicho cambio de variable. De hecho, lo que estamos diciendo es que φ parametriza una curva del plano. Dada una función $F : D \subset \mathbb{R}^2$, nos puede interesar el comportamiento de F a lo largo de la curva $\varphi(t)$, esto es, la función g de una sola variable dada por

$$g(t) = F(x(t), y(t)).$$

Dicho de otro modo, g se obtiene por composición como $g = F \circ \varphi$. La *regla de la cadena* nos dice que

$$g'(t) = \nabla F(\varphi(t))\varphi'(t) = (F_x(x(t), y(t)), F_y(x(t), y(t))) \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$$

■ **Ejemplo 7.19** Sea $F(x, y) = x + y^2$. Determinemos la función $g(t)$ obtenida al recorrer la curva $\varphi(t) = (\sin t, e^t)$ y su función derivada $g'(t)$.

La función que buscamos es

$$g(t) = F \circ \varphi = F(\sin t, e^t) = \sin(t) + e^{2t}$$

Tenemos que $g'(t) = \cos(t) + 2e^{2t}$. Veamos que llegamos al mismo resultado mediante es uso de la regla de la cadena. Para ello calculamos:

$$\nabla F(x, y) = (1, 2y), \quad \varphi'(t) = (\cos t, e^t)$$

Ahora tenemos que

$$g'(t) = (1, 2e^t) \begin{pmatrix} \cos t \\ e^t \end{pmatrix} = \cos t + 2e^{2t}.$$

■ **Ejemplo 7.20 — Derivadas direccionales.** Si F es una función diferenciable en (x_0, y_0) y $\vec{v} = (\cos \theta, \sin \theta) \in \mathbb{R}^2$ es una velocidad unitaria, hemos dicho que se tiene la relación

$$F_{\vec{v}}(x_0, y_0) = \cos \theta F_x(x_0, y_0) + \sin \theta F_y(x_0, y_0) = \langle \nabla F(x_0, y_0), \vec{v} \rangle$$

Tenemos que la derivada $F_{\vec{v}}(x_0, y_0)$, respecto a \vec{v} en el punto (x_0, y_0) , no es más que la derivada en $t = 0$ de la función de una variable dada por

$$f(t) = F(\varphi(t)),$$

donde $\varphi(t) = (x_0 + t \cos \theta, y_0 + t \sin \theta)$ es una parametrización de la recta

$$-\sin \theta(x - x_0) + \cos \theta(y - y_0) = 0,$$

En efecto, si aplicamos la regla de la cadena, tenemos

$$f'(0) = \nabla F(\varphi(0))\varphi'(0) = (F_x(x_0, y_0), F_y(x_0, y_0)) \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$$

7.9.2 Regla de la cadena (a lo largo de una superficie)

De manera similar, puede suceder que nuestras dos variables x e y dependan a su vez de otras dos variables r y s . Esto significa que conocemos una aplicación $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que determina dicho cambio de variable. Escribiríamos

$$\varphi(r, s) = (x(r, s), y(r, s)),$$

En este caso $\varphi(r, s)$ determina una nueva parametrización del plano \mathbb{R}^2 . Dada una función $F : D \subset \mathbb{R}^2$, nos puede interesar el comportamiento de F respecto de las variables r y s , esto es, la función G de nuevo de dos variables dada por

$$G(r, s) = F(x(r, s), y(r, s))$$

Dicho de otro modo, G se obtiene por composición como $G = F \circ \varphi$. La *regla de la cadena* nos dice que

$$\begin{aligned} \nabla G(r, s) &= \nabla F(\varphi(r, s)) \text{Jac}_{\varphi}(r, s) = \\ &= \left(F_x(x(r, s), y(r, s)), F_y(x(r, s), y(r, s)) \right) \begin{pmatrix} \partial x / \partial r & \partial x / \partial s \\ \partial y / \partial r & \partial y / \partial s \end{pmatrix} \end{aligned}$$

■ **Ejemplo 7.21** Sea $F(x, y) = 2xy$. Determinemos el gradiente de la función $G(r, s)$ obtenida al realizar los cambios de variable $x = s^2 + r^2$ e $y = s/r$.

Para esto necesitamos calcular por un lado el gradiente de la función F y por otro la matriz Jacobiana de la aplicación vectorial

$$\varphi(r, s) = (x(r, s), y(r, s)) = (s^2 + r^2, s/r)$$

Tenemos que $F_x = 2y$ y $F_y = 2x$, por tanto $\nabla F = (2y, 2x)$. Por otro lado

$$\frac{\partial x}{\partial r} = 2r, \quad \frac{\partial x}{\partial s} = 2s, \quad \frac{\partial y}{\partial r} = -s/r^2, \quad \frac{\partial y}{\partial s} = 1/r.$$

De esta manera tenemos que

$$\text{Jac}_{\varphi}(r, s) = \begin{pmatrix} 2r & 2s \\ -s/r^2 & 1/r \end{pmatrix}$$

Y así obtenemos

$$\begin{aligned} (G_r(r, s), G_s(r, s)) &= \nabla G(r, s) = (2s/r, 2(s^2 + r^2)) \begin{pmatrix} 2r & 2s \\ -s/r^2 & 1/r \end{pmatrix} = \\ &= \left(4s + \frac{-2s}{r^2}(r^2 + s^2), 4\frac{s^2}{r} + \frac{2}{r}(s^2 + r^2) \right) = \left(2s - \frac{2s^3}{r^2}, 2r + \frac{6s^2}{r} \right) \end{aligned}$$

Si estamos trabajando con tres variables x, y, z y tenemos una función $F(x, y, z)$, también podemos considerar un cambio de variable $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dado por

$$\varphi(r, s) = (x(r, s), y(r, s), z(r, s)),$$

esto es, una superficie parametrizada de \mathbb{R}^3 . Considerar como varía la función

$$G(r, s) = F(x(r, s), y(r, s), z(r, s))$$

es estudiar la función F a lo largo de la superficie determinada por $\varphi(r, s)$.

7.9.3 Derivación implícita

Finalmente, podemos usar la regla de la cadena para derivar de forma implícita. Supongamos que tenemos una función de tres variables $F(x, y, z)$ y queremos considerar la superficie dada por uno de sus niveles, esto es

$$F(x, y, z) = k$$

para determinar la variación de la variable z respecto a las variables x e y . Esto es, queremos ver z como una función $z(x, y)$. Tenemos que

$$F_x \frac{\partial x}{\partial x} + F_y \frac{\partial y}{\partial x} + F_z \frac{\partial z}{\partial x} = 0, \quad F_x \frac{\partial x}{\partial y} + F_y \frac{\partial y}{\partial y} + F_z \frac{\partial z}{\partial y} = 0$$

De aquí, y dado que sabemos que $\partial x/\partial x = \partial y/\partial y = 1$ y $\partial x/\partial y = \partial y/\partial x = 0$, obtenemos que si $F_z \neq 0$, entonces

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}, \quad \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

■ **Ejemplo 7.22** Consideremos la esfera de radio 1 dada de forma implícita por $x^2 + y^2 + z^2 = 1$. Veamos la variación de z respecto a x e y en el punto $(0, 1/\sqrt{2}, -1/\sqrt{2})$. Tenemos que

$$\partial z/\partial x = -\frac{2x}{2z}, \quad \partial z/\partial y = -\frac{2y}{2z}$$

Así, $\partial z/\partial x(0, 1/\sqrt{2}) = 0$ y $\partial z/\partial y(0, 1/\sqrt{2}) = -1$

Todo lo que hemos dicho en esta sección funciona también en general. Es decir, tenemos n variables x_1, x_2, \dots, x_n y estas a su vez dependen de m variables r_1, r_2, \dots, r_m . Esto significa que tenemos una aplicación vectorial

$$\varphi(r_1, r_2, \dots, r_m) = (x_1(r_1, r_2, \dots, r_m), x_2(r_1, r_2, \dots, r_m), \dots, x_n(r_1, r_2, \dots, r_m))$$

Dada una función F en las n variables x_1, x_2, \dots, x_n , consideramos asociada la función $G = F \circ \varphi$ que depende de las m variables r_1, r_2, \dots, r_m . Tenemos que

$$\nabla G(r_1, r_2, \dots, r_m) = \nabla F(\varphi(r_1, r_2, \dots, r_m)) \text{Jac}_\varphi(r_1, r_2, \dots, r_m).$$

8. Cálculo integral

La *integral de Riemman* de una función de una variable $f(x)$ en un intervalo $I = [a, b]$ es (si existe) un valor, que representamos habitualmente por

$$\int_a^b f(x)dx$$

Hablando sin mucho rigor, este valor se obtiene como límite de una suma de áreas de rectángulos cada vez más pequeños y por tanto describe un área.

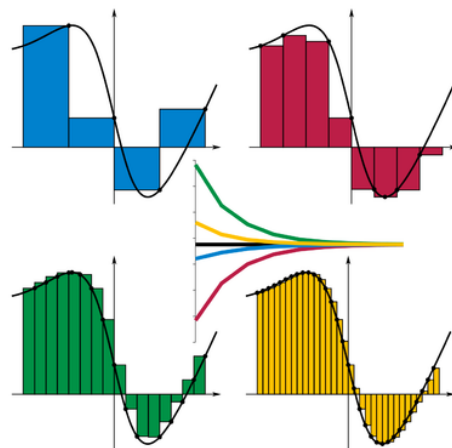
Más precisamente, si f es una función continua y no negativa en el intervalo $I = [a, b]$, entonces el área de la región R acotada por el grafo de f , el eje horizontal y las rectas $x = a$ y $x = b$ viene dada por

$$\text{Área}(R) = \int_a^b f(x)dx$$

En particular, si nuestra función es $f(x) = 1$ recuperamos la longitud de I , ya que ésta coincide con el área del rectángulo de base I y altura 1. Esto es

$$\text{Long}(I) = \int_a^b dx = b - a$$

Recordemos ahora un resultado crucial: el *teorema fundamental del cálculo*. Grosso modo este enunciado establece que derivación e integración son operaciones inversas una de la

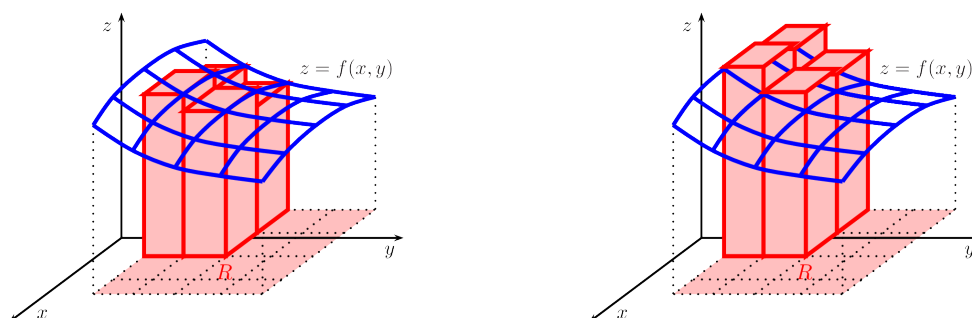


otra. Así, si f es una función continua en el intervalo $[a, b]$ y consideramos la función

$$F(x) = \int_a^x f(x)dx,$$

entonces F es derivable y $F'(x) = f(x)$.

Consideremos ahora una función $F(x, y)$ y sea $R = [a, b] \times [c, d]$ un rectángulo en el plano. La misma construcción de integral de Riemman que se hace para el caso de funciones de una variable se puede extender a dos variables. Esto es, se puede ir subdividiendo R en rectangulos cada vez más pequeños y estimar el volumen que encierra la gráfica de $F(x, y)$ mediante sumas de volúmenes de ortoedros. La siguiente imagen ilustra esta idea:



El valor obtenido por este procedimiento (si existe) es a lo que llamamos la *integral de $F(x, y)$ a lo largo de R* y lo denotamos por

$$\int_R F(x, y) dx dy.$$

De un modo más complicado aunque con ideas similares, se puede hablar también de la integral de una función a lo largo de regiones del plano más generales. Más concretamente, si Ω es una región acotada (cabe dentro de una circunferencia) del plano y F es una función no negativa en Ω , entonces

$$\text{Vol}(S) = \int_{\Omega} F(x, y) dx dy,$$

donde S es el sólido dado por la intersección del cilindro de base Ω , el plano $z = 0$ y aquello que queda por debajo de la gráfica de $F(x, y)$. Observamos también que si consideramos $F(x, y) = 1$, entonces

$$\text{Área}(\Omega) = \int_{\Omega} dx dy,$$

ya que ésta coincide con el volumen del cilindro de base Ω y altura 1.

8.1 Integrales iteradas

Llamamos *integral iterada* al anidamiento de varias integrales de una sola variable. Esto es, en el caso de dos variables a una expresión del tipo

$$I_1 = \int_a^b \left(\int_{f_1(x)}^{f_2(x)} F(x, y) dy \right) dx, \quad I_2 = \int_c^d \left(\int_{g_1(y)}^{g_2(y)} F(x, y) dx \right) dy$$

Habitualmente, y si no hay lugar a confusión, se prescindirá de los paréntesis en las expresiones anteriores. Consideremos por ejemplo I_1 . Es cierto que $F(x, y)$ es una función en dos variables, pero estamos haciendo lo siguiente:

1. En primer lugar, una integración solamente respecto la variable y , y por tanto x está actuando como si fuera constante (de modo análogo a lo que hacíamos para las derivadas parciales).
2. Después de esto nos queda una integral que depende solamente de la variable x y por tanto tras resolverla el resultado quedará un valor real.

En la expresión en I_2 se intercambian los papeles de x e y . Es decir, primero integraríamos respecto a x y después respecto a y .

■ **Ejemplo 8.1** Calcúlese la siguiente integral iterada

$$I = \int_0^1 \int_0^3 (x^2 + y^4) dx dy$$

En primer lugar integramos respecto de x con los límites de 0 a 3. Así

$$I = \int_0^1 \left[\frac{x^3}{3} + y^4 x \right]_0^3 dy = \int_0^1 (9 + 3y^4) dy$$

Ahora nos queda ya una integral que depende solamente de y . Así

$$I = \left[9y + \frac{3y^5}{5} \right]_0^1 = 9 + \frac{3}{5} = \frac{48}{5}.$$

Aprovechamos este ejemplo para comprobar que si cambiamos el orden en la integración, es decir, si consideramos

$$J = \int_0^3 \int_0^1 (x^2 + y^4) dy dx,$$

obtenemos el mismo resultado. En efecto, tenemos que

$$J = \int_0^3 \left[x^2 y + \frac{y^5}{5} \right]_0^1 dx = \int_0^3 \left(x^2 + \frac{1}{5} \right) dx = \left[\frac{x^3}{3} + \frac{x}{5} \right]_0^3 = 9 + \frac{3}{5} = \frac{48}{5}.$$

En breve hablaremos sobre este hecho, que es general.

■ **Ejemplo 8.2** Calcúlese la siguiente integral iterada

$$I = \int_{-1}^1 \int_{-x^2}^{x^2} (x^4 - 2y) dy dx$$

Antes de empezar observamos que en este caso no podemos intercambiar las variables (o no directamente), ya que los límites de integración en la variable y dependen de x . Por eso es importante realizar la integración en el orden dado.

Tenemos que

$$I = \int_{-1}^1 \left[x^4 y - y^2 \right]_{-x^2}^{x^2} dx = \int_{-1}^1 \left[(x^4 x^2 - x^4) - (x^4(-x^2) - (-x^2)^2) \right] dx =$$

$$= \int_{-1}^1 (2x^6 - x^4 + x^4) dx = \int_{-1}^1 2x^6 dx.$$

Ahora nos queda ya una integral que depende solamente de x . Así

$$I = \left. \frac{2x^7}{7} \right]_{-1}^1 = 1 + \frac{2}{7} - \frac{-2}{7} = \frac{4}{7}$$

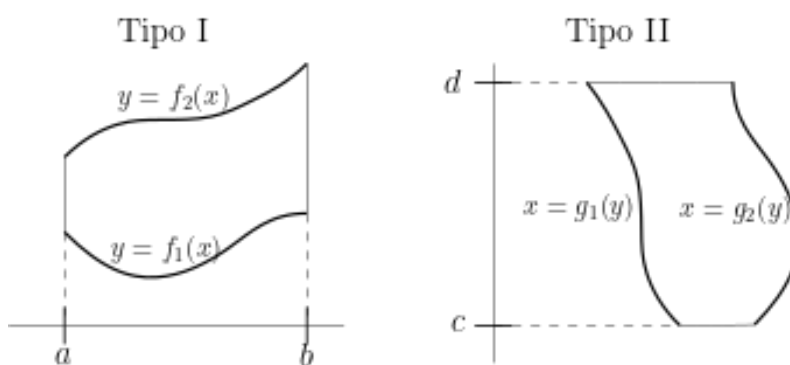
El *teorema de Fubini* nos proporciona una forma de calcular la integral de Riemann de una función de dos variables como integrales iteradas. Más precisamente, tenemos que si $R = [a, b] \times [c, d]$, entonces

$$\int_R F(x, y) dx dy = \int_a^b \left(\int_c^d F(x, y) dy \right) dx = \int_c^d \left(\int_a^b F(x, y) dx \right) dy$$

El teorema de Fubini también se puede aplicar al cálculo de la integral de una función de dos variables sobre una región más general. La dificultad en este caso reside en parametrizar la región para poderla expresar como integrales iteradas. Sobre esto vamos a hablar en la próxima sección.

8.2 Parametrización de regiones

Vamos a considerar primero el caso de regiones que tiene una estructura sencilla: *regiones básicas*. Son aquellas que tienen una estructura como la que se puede observar en la siguiente figura.



Cuando la región es de Tipo I tenemos que

$$\int_R F(x, y) dx dy = \int_a^b \int_{f_1(x)}^{f_2(x)} F(x, y) dy dx$$

Cuando la región es como de Tipo II tenemos que

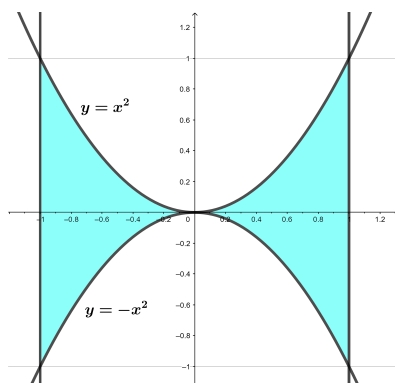
$$\int_R F(x, y) dx dy = \int_c^d \int_{g_1(y)}^{g_2(y)} F(x, y) dx dy$$

■ **Ejemplo 8.3** Considera la región Ω dada en la figura.

Plantea la integral

$$\int_{\Omega} (x^4 - 2y) dx dy$$

como integral iterada observando de qué tipo de los anteriores es. Después resuélvela.



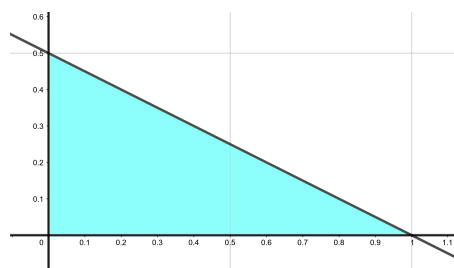
La región es de Tipo I con $a = -1$, $b = -1$, $f_1(x) = -x^2$ y $f_2(x) = x^2$. Así, tenemos que

$$\int_{\Omega} (x^4 - 2y) dx dy = \int_{-1}^1 \int_{-x^2}^{x^2} (x^4 - 2y) dy dx$$

Esta integral la hemos resuelto en el ejemplo 8.2.

■ **Ejemplo 8.4** Calcula el área del triángulo T de vértices $(0,0)$, $(1,0)$ y $(0,1/2)$.

El triángulo lo podemos ver tanto como una figura de Tipo I o de Tipo II. Vamos a hacerlo de cualquiera de ellas y ver que además coincide con el resultado ya conocido: *base por altura partido por dos*, que este caso es $1/4$.



La hipotenusa del triángulo viene marcada por la recta $x + 2y = 1$. Si planteamos los límites como de Tipo I, tenemos que $a = 0$, $b = 1$, $f_1(x) = 0$ y $f_2(x) = (1-x)/2$. Y por tanto

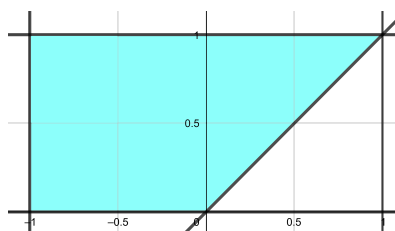
$$\int_T dx dy = \int_0^1 \int_0^{(1-x)/2} dy dx = \int_0^1 y \Big|_0^{(1-x)/2} dx = \int_0^1 \frac{1-x}{2} dx = \left[\frac{x}{2} - \frac{x^2}{4} \right]_0^1 = \frac{1}{4}$$

■ **Ejemplo 8.5** Considera la región Ω dada en la figura.

Plantea la integral

$$\int_{\Omega} (xy - y^3) dx dy$$

como integral iterada. Posteriormente resuélvela.



La región es de Tipo II con límites de integración para y dados por $c = 0$ y $d = 1$, y límites de integración para x dados por $g_1(y) = -1$ y $g_2(y) = y$. Así, tenemos que

$$\int_{\Omega} (xy - y^3) dx dy = \int_0^1 \int_{-1}^y (xy - y^3) dx dy = \int_0^1 \left[\frac{x^2 y}{2} - xy^3 \right]_{-1}^y dy =$$

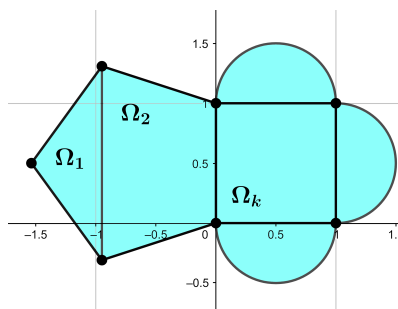
$$\begin{aligned} \int_0^1 \left[\left(\frac{y^3}{2} - y^4 \right) - \left(\frac{y}{2} + y^3 \right) \right] dy &= \int_0^1 \left(\frac{-y^3}{2} - y^4 - \frac{y}{2} \right) dy = \\ &= \left[\frac{-y^4}{8} - \frac{y^5}{5} - \frac{y^2}{4} \right]_0^1 = \frac{-1}{8} - \frac{1}{5} - \frac{1}{4} = \frac{-5 - 8 - 10}{40} = \frac{-23}{40} \end{aligned}$$

Observemos que también podríamos haber intentado parametrizar la región como si fuese de Tipo I, pero en tal caso tendríamos que dividir la región en dos partes: el cuadrado $[-1, 0] \times [0, 1]$ y el triángulo de vértices $(0, 0)$, $(0, 1)$ y $(1, 1)$. Más concretamente tendríamos que

$$\int_{\Omega} (xy - y^3) dx dy = \int_{-1}^0 \int_0^1 (xy - y^3) dy dx + \int_0^1 \int_x^1 (xy - y^3) dy dx$$

Muchas veces puede ocurrir que la región de integración Ω no es una región básica, pero que en cambio se puede subdividir en un número finito de regiones $\Omega_1, \Omega_2, \dots, \Omega_k$ cada una de ellas básicas, de Tipo I o de Tipo II. Así

$$\begin{aligned} \int_{\Omega} F(x, y) dx dy &= \int_{\Omega_1} F(x, y) dx dy + \\ &+ \int_{\Omega_2} F(x, y) dx dy + \dots + \int_{\Omega_k} F(x, y) dx dy \end{aligned}$$



■ **Ejemplo 8.6** Halla el volumen del sólido comprendido entre el paraboloides $z = x^2 + y^2$ y el cilindro $x^2 + y^2 = 1$.

Nos están pidiendo calcular la integral de la función $x^2 + y^2$ a lo largo del disco $D = \{x^2 + y^2 \leq 1\}$. Podemos ver el disco mediante una parametrización de tipo I como

$$-1 \leq x \leq 1, \quad -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2}$$

Tenemos así que

$$I = \int_D (x^2 + y^2) dx dy = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2 + y^2) dy dx$$

Podemos comprobar que a su vez

$$I = 2 \int_{-1}^1 \int_0^{\sqrt{1-x^2}} (x^2 + y^2) dy dx = 4 \int_0^1 \int_0^{\sqrt{1-x^2}} (x^2 + y^2) dy dx$$

Esto sucede por la simetrías que tiene el disco D y también la función F . En cualquier caso, tenemos que

$$I = \int_{-1}^1 \left[x^2 y + \frac{y^3}{3} \right]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dx = 2 \int_{-1}^1 \frac{2x^2 + 1}{3} \sqrt{1-x^2} dx$$

y esta última integral no es tan sencilla de calcular. De inmediato veremos una forma mucho más sencilla de resolver esta integral

8.3 Cambio de coordenadas

A veces es más sencillo parametrizar las regiones al considerar otras coordenadas distintas a las cartesianas, o incluso la función a integrar se simplifica bastante con estos cambios. Consideremos un cambio de variable

$$\varphi : (u, v) \mapsto (x(u, v), y(u, v))$$

Dada una función $F : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ en coordenadas x, y y una región de integración $\Omega \subset \mathbb{R}^2$, tenemos la siguiente relación

$$\int_{\Omega} F(x, y) dx dy = \int_{\Omega} F(x(u, v), y(u, v)) |\text{Jac}_{\varphi}(u, v)| du dv,$$

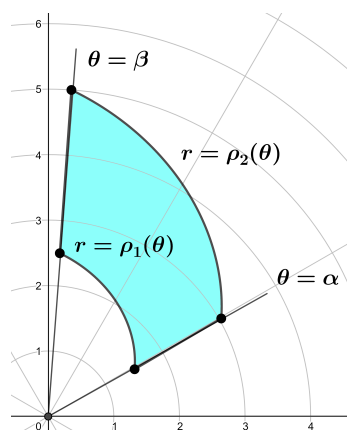
donde $\text{Jac}_{\varphi}(u, v)$ es la matriz jacobiana de φ , que viene dada por:

$$\text{Jac}_{\varphi}(u, v) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix}$$

En general suele ser difícil encontrar cambios de coordenadas adecuados en integrales dobles. Veamos el caso del cambio a polares, que es bastante útil, y otro caso más a modo de ejemplo.

8.3.1 Coordenadas polares

Es más cómodo emplear las coordenadas polares si la región considerada tiene un aspecto como el de la figura



El cambio de variable para este caso es

$$\varphi : (r, \theta) \mapsto (r \cos \theta, r \sin \theta)$$

Por tanto, tenemos que

$$|\text{Jac}_{\varphi}(r, \theta)| = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$

y de esta manera

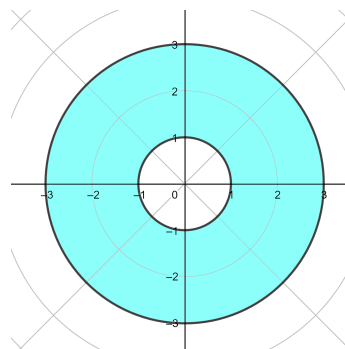
$$\int_{\Omega} F(x,y) dx dy = \int_{\Omega} rG(r,\theta) dr d\theta = \int_{\alpha}^{\beta} \int_{\rho_1(\theta)}^{\rho_2(\theta)} rG(r,\theta) dr d\theta$$

donde $G(r,\theta) = F(r\cos\theta, r\sin\theta)$.

■ **Ejemplo 8.7** Calcula el área de la corona circular $C = D_{R_2} \setminus D_{R_1}$, donde D_{R_i} es el círculo centrado en el origen y de radio R_i .

Nos piden calcular la integral

$$\begin{aligned} \int_C dx dy &= \int_C r dr d\theta = \int_0^{2\pi} \int_{R_1}^{R_2} r dr d\theta = \\ &= \int_{R_1}^{R_2} \int_0^{2\pi} r d\theta dr = \int_{R_1}^{R_2} r\theta \Big|_0^{2\pi} dr = \\ &= \int_{R_1}^{R_2} 2\pi r dr = \pi r^2 \Big|_{R_1}^{R_2} = \pi(R_2^2 - R_1^2). \end{aligned}$$



Notemos que si $R_1 = 0$ obtenemos el área del círculo de radio R_2 .

■ **Ejemplo 8.8** Vamos a reconsiderar el ejemplo 8.6 utilizando el cambio de coordenadas a polares. El paraboloides definido por el grafo $F(x,y) = x^2 + y^2$ se expresa en coordenadas polares como

$$G(r,\theta) = F(r\cos\theta, r\sin\theta) = r^2.$$

La región a lo largo de la cual queremos integrar es el círculo C de radio 1 centrado en el origen, por tanto viene dado por

$$0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi.$$

Así, tenemos que

$$\begin{aligned} \int_C (x^2 + y^2) dx dy &= \int_C r^2 \cdot r dr d\theta = \int_0^{2\pi} \int_0^1 r^3 dr d\theta = \int_0^{2\pi} \left[\frac{r^4}{4} \right]_0^1 d\theta = \\ &= \int_0^{2\pi} \frac{1}{4} d\theta = \left[\frac{\theta}{4} \right]_0^{2\pi} = \frac{\pi}{2} \end{aligned}$$

8.3.2 Otros cambios de variable

Veamos otro cambio de coordenadas un poco distinto al cambio a coordenadas polares mediante un ejemplo. Consideremos la elipse E que tiene por ecuación $\frac{x^2}{9} + y^2 = 1$. Vamos a utilizar el cambio de variable

$$x = 3r\cos\theta, \quad y = r\sin\theta$$

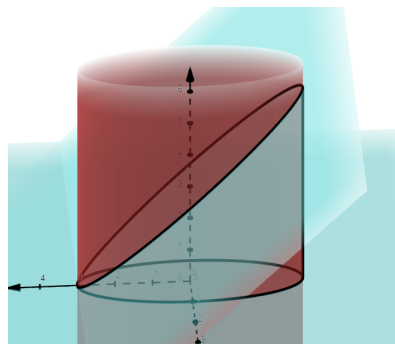
para determinar el volumen del sólido delimitado lateralmente por un cilindro de base E , inferiormente por el plano $z = 0$ y superiormente por el plano $x + z = 3$.

Nos piden calcular la integral a lo largo de E de la función $z = F(x, y) = 3 - x$, i.e.

$$\int_E (3 - x) dx dy$$

Con el cambio de variable proporcionado, tenemos que E está delimitada por

$$0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi$$



El determinante de la matriz jacobiana del cambio de variable es

$$\begin{vmatrix} 3 \cos \theta & -3r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = 3r(\sin^2 \theta + \cos^2 \theta) = 3r$$

Por otro lado $F(3r \cos \theta, r \sin \theta) = 3(1 - r \cos \theta)$. De esta manera

$$\begin{aligned} \int_E (3 - x) dx dy &= \int_0^1 \int_0^{2\pi} 3(1 - r \cos \theta) 3r d\theta dr = \int_0^1 \int_0^{2\pi} 9r(1 - r \cos \theta) d\theta dr = \\ &= \int_0^1 [9r\theta - 9r^2 \sin \theta]_0^{2\pi} dr = \int_0^1 18\pi r dr = 9\pi r^2 \Big|_0^1 = 9\pi \end{aligned}$$

8.4 Integrales triples. Volumen y masa

Del mismo modo que hemos hecho con funciones de dos variables, podríamos hablar de integrales triples, esto es, integrar a lo largo de un sólido S contenido en \mathbb{R}^3 una función de tres variables $F(x, y, z)$. Esto es

$$\int_S F(x, y, z) dx dy dz$$

El teorema de Fubini también es aplicable también en este caso, y en general, esta integral la podemos ver como una suma de integrales iteradas del tipo

$$\int_S F(x, y, z) dx dy dz = \int_a^b \int_{g_1(x)}^{g_2(x)} \int_{G_1(x,y)}^{G_2(x,y)} F(x, y, z) dz dy dx$$

(o intercambiando los papeles de las variables x , y y z). La interpretación geométrica de esta cantidad no es nada clara, hablaríamos del *hipervolumen* de un *sólido 4-dimensional*. En cambio, en física o química tenemos una interpretación más clara de este valor. Supongamos que nuestro sólido S no es densidad constante, sino que esta va cambiando de acuerdo a la función $F(x, y, z)$. Tenemos así que

$$\text{Masa}(S) = \int_S F(x, y, z) dx dy dz$$

Por supuesto, cuando $F(x, y, z) = 1$, esto es, cuando la densidad es constante e igual a 1, recuperamos el volumen del sólido S . Así

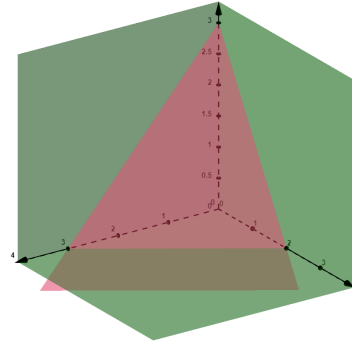
$$\text{Vol}(S) = \int_S dx dy dz$$

■ **Ejemplo 8.9** Considera el sólido S que encierran los planos coordenados y que queda por debajo del plano $2x + 3y + 2z = 6$.

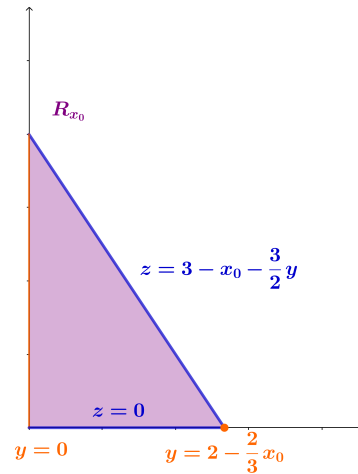
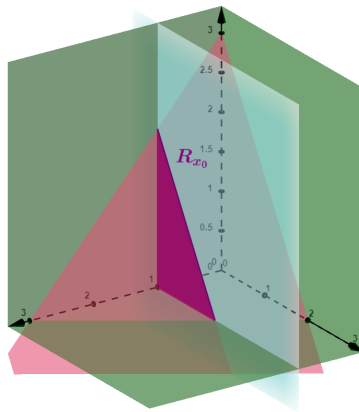
1. Calcula el volumen de S
2. Determina la masa de S si su densidad viene dada por $F(x, y, z) = x$.

En primer lugar debemos determinar los límites de integración de S como integral iterada. Realizamos una representación de nuestro sólido S para apoyarnos en la imagen.

Por la forma que tiene S (sus paredes son planas), podemos escoger como última variable de integración la que nos plazca. Escojamos por ejemplo la x y la hacemos variar: x toma como valor mínimo el 0 y su valor máximo se alcanza cuando $y = z = 0$. Así $0 \leq x \leq 3$.



Ahora, nos colocamos en $x = x_0$ y debemos parametrizar la región R_{x_0} del plano $x = x_0$ dada por $R_{x_0} = (x = x_0) \cap S$.



Tenemos que R_{x_0} es un triángulo y por tanto la podemos considerar como una región de tipo I o de tipo II. Por ejemplo, tomamos y y variando entre dos valores fijos (dependiendo de x_0) y hacemos que z varíe en función de y . El valor mínimo que toma la y es cero, y el valor máximo se alcanza cuando $z = 0$. Así, el límite superior para y es su valor en la intersección

$$(x = x_0) \cap (z = 0) \cap (2x + 3y + 2z = 6),$$

esto es $y = 2 - \frac{2}{3}x_0$. Finalmente tenemos que z varía entre 0 y $3 - x_0 - \frac{3}{2}y$. De esta manera tenemos que

$$S = \left\{ \begin{array}{l} 0 \leq x \leq 3 \\ 0 \leq y \leq 2 - \frac{2}{3}x \\ 0 \leq z \leq 3 - x - \frac{3}{2}y \end{array} \right\},$$

y por lo tanto

$$\begin{aligned}\text{Vol}(S) &= \int_S dx dy dz = \int_0^3 \int_0^{2-\frac{2}{3}x} \int_0^{3-x-\frac{3}{2}y} dz dy dx = \int_0^3 \int_0^{2-\frac{2}{3}x} z \Big|_0^{3-x-\frac{3}{2}y} dy dx = \\ &= \int_0^3 \int_0^{2-\frac{2}{3}x} \left(3-x-\frac{3}{2}y\right) dy dx = \int_0^3 \left[3y - xy - \frac{3}{4}y^2\right]_0^{2-\frac{2}{3}x} dx = \int_0^3 \left(3-2x+\frac{1}{3}x^2\right) dx = \\ &= \left[3x - x^2 + \frac{1}{9}x^3\right]_0^3 = 9 - 3^2 + \frac{3^3}{9} = 3\end{aligned}$$

Notemos que el resultado obtenido coincide con el volumen de una pirámide:

$$\text{Vol}(S) = \frac{\text{Área base} \times \text{altura}}{3}$$

El apartado 2 nos pide determinar la integral

$$\begin{aligned}\text{Masa}(S) &= \int_S x dx dy dz = \int_0^3 \int_0^{2-\frac{2}{3}x} \int_0^{3-x-\frac{3}{2}y} x dz dy dx = \int_0^3 \int_0^{2-\frac{2}{3}x} xz \Big|_0^{3-x-\frac{3}{2}y} dy dx = \\ &= \int_0^3 \int_0^{2-\frac{2}{3}x} x \left(3-x-\frac{3}{2}y\right) dy dx = \int_0^3 x \left[3y - xy - \frac{3}{4}y^2\right]_0^{2-\frac{2}{3}x} dx = \int_0^3 x \left(3-2x+\frac{1}{3}x^2\right) dx = \\ &= \left[\frac{3}{2}x^2 - \frac{2}{3}x^3 + \frac{1}{12}x^4\right]_0^3 = \frac{9}{4}\end{aligned}$$

Al igual que para las integrales dobles, en muchas ocasiones un cambio de variable puede facilitar mucho el cálculo de una integral triple. Veamos alguna situación donde esto sucede.

8.4.1 Coordenadas cilíndricas

Tenemos el cambio de variable

$$x = r \cos \theta, \quad y = r \sin \theta, \quad z = z$$

De esta manera el determinante de la matriz jacobiana es

$$\begin{vmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = r$$

y por lo tanto

$$\int_S F(x, y, z) dx dy dz = \int_S F(r \cos \theta, r \sin \theta, z) r dr d\theta dz$$

■ **Ejemplo 8.10** Calcula el volumen del tronco de cono T obtenido al considerar el cono $z^2 = x^2 + y^2$ delimitado inferiormente por el plano $z = 1$ y superiormente por $z = 3$.

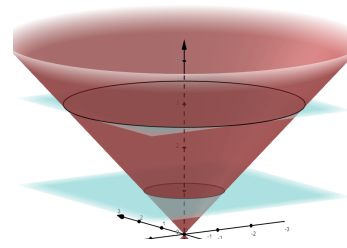
El cono en coordenadas esféricas está dado por $z^2 = r^2$. Así, como nos estamos quedando con valores de z positivos ($1 \leq z \leq 3$), estamos pidiendo que

$$0 \leq r \leq z$$

Por otro lado no tenemos ninguna restricción sobre el ángulo θ .

De esta manera

$$\begin{aligned} \int_T r dr d\theta dz &= \int_1^3 \int_0^{2\pi} \int_0^z r dr d\theta dz = \int_1^3 \int_0^{2\pi} r\theta \Big|_0^{2\pi} dr dz = \\ &= \int_1^3 \int_0^{2\pi} 2\pi r dr dz = \int_1^3 \pi r^2 \Big|_0^z dz = \int_1^3 \pi z^2 dz = \left. \frac{\pi z^3}{3} \right|_1^3 = \\ &= \frac{\pi 3^3}{3} - \frac{\pi 1^3}{3} = 9\pi - \frac{1}{3}\pi = \frac{26}{3}\pi. \end{aligned}$$



8.4.2 Coordenadas esféricas

Tenemos el cambio de variable

$$x = r \cos \theta \sin \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \varphi$$

De esta manera el determinante de la matriz jacobiana es

$$\begin{vmatrix} \cos \theta \sin \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \sin \theta \cos \varphi & r \cos \theta \sin \varphi \\ \cos \varphi & -r \sin \varphi & 0 \end{vmatrix} = r^2 \sin \varphi$$

y por lo tanto

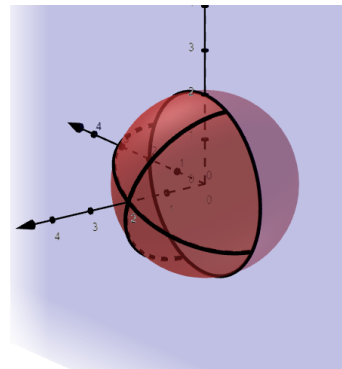
$$\int_S F(x,y,z) dx dy dz = \int_S F(r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi) r^2 \sin \varphi dr d\varphi d\theta.$$

■ **Ejemplo 8.11** Sea S es la media bola centrada en $\mathcal{O} = (0,0,0)$ y radio 4 contenida en el semiespacio $y \geq 0$. Calcula el volumen de S y la densidad de S cuando ésta varía de acuerdo a la función $F(x,y,z) = \sqrt{x^2 + y^2 + z^2}$.

Nuestro sólido S tiene un radio

$$0 \leq r \leq 4.$$

Por otro lado, pedir que la coordenada y siempre sea positiva, significa que la longitud θ siempre es un ángulo entre 0 y π . Finalmente para la latitud φ no tenemos ninguna restricción y por ello va de 0 a π .



Calculamos primero el volumen de S , esto es la integral triple

$$\begin{aligned} \int_S r^2 \sin \varphi dr d\theta d\varphi &= \int_0^4 \int_0^\pi \int_0^\pi r^3 \sin \varphi d\theta d\varphi dr = \int_0^4 \int_0^\pi r^2 \sin \varphi \theta \Big|_0^\pi d\varphi dr = \\ &= \int_0^4 \int_0^\pi r^2 \sin \varphi \pi d\varphi dr = \int_0^4 -\pi r^2 \cos \varphi \Big|_0^\pi dr = \int_0^4 2\pi r^2 dr = \\ &= \frac{2\pi r^3}{3} \Big|_0^4 = \frac{2\pi 4^3}{3}. \end{aligned}$$

Observemos que hemos llegado un resultado ya conocido, que es el volumen de la bola (el doble de lo que hemos obtenido), dado por

$$\frac{4\pi r^3}{3}$$

Vamos a calcular ahora la densidad. Notemos que la densidad en coordenadas esféricas es $F(r \cos \theta \sin \varphi, r \sin \theta \sin \varphi, r \cos \varphi) = r$ y por tanto, estamos considerando

$$\begin{aligned} \int_S r r^2 \sin \varphi dr d\theta d\varphi &= \int_0^4 \int_0^\pi \int_0^\pi r^3 \sin \varphi d\theta d\varphi dr = \int_0^4 \int_0^\pi r^3 \sin \varphi \theta \Big|_0^\pi d\varphi dr = \\ &= \int_0^4 \int_0^\pi r^3 \sin \varphi \pi d\varphi dr = \int_0^4 -\pi r^3 \cos \varphi \Big|_0^\pi dr = \int_0^4 2\pi r^3 dr = \frac{\pi r^4}{2} \Big|_0^4 = 2\pi 4^3. \end{aligned}$$



Cálculo vectorial

9	Aplicaciones vectoriales	107
9.1	Rotacional y divergencia. Potencial	
9.2	Integrales de línea. Trabajo	

9. Aplicaciones vectoriales

Una *aplicación vectorial* es una asignación $\mathbf{F} : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ que a cada vector $\vec{x} = (x_1, x_2, \dots, x_n) \in D$, le hace corresponder un vector

$$\mathbf{F}(\vec{x}) = (F_1(\vec{x}), F_2(\vec{x}), \dots, F_m(\vec{x})).$$

Las $F_i : D \rightarrow \mathbb{R}$ son funciones escalares y las llamamos *componentes o proyecciones* de la función vectorial \mathbf{F} . Para las aplicaciones vectoriales tienen sentido los conceptos de *continuidad, diferenciabilidad e integrabilidad*. Así, una aplicación vectorial es continua (diferenciable, integrable) si y solo si lo es cada una de sus componentes. Cuando $m = 1$ la aplicación vectorial es, de hecho, una función escalar, como las que hemos estudiado en los últimos temas del curso.

Hemos tratado ya en este curso con aplicaciones vectoriales:

- Las transformaciones lineales de $\mathbb{R}^n \rightarrow \mathbb{R}^n$, como por ejemplo la dada por $\mathbf{F}(x, y) = (2x - 3y, x - y)$. En este caso $n = m = 2$.
- Las parametrizaciones de curvas en el plano o en el espacio ($n = 1$ y $m = 2, 3$). Por ejemplo, la dada por $\gamma(t) = (\cos t, \sin t, t)$ que recorre una hélice. Aquí se incluyen las curvas parametrizadas obtenidas como solución de sistemas de ecuaciones diferenciales de tamaños 2×2 o 3×3 .
- Las parametrizaciones de superficies en el espacio ($n = 2$ y $m = 3$). Por ejemplo, la dada por $\mathbf{S}(u, v) = (u, v, u^2 + v^2)$, que describe un paraboloides circular.
- El gradiente $\nabla F : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ de una función escalar $F : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ es una aplicación vectorial (con $n = m = 3$) dada por $(x, y, z) \mapsto \nabla F(x, y, z)$.
- Los cambios de coordenadas, en particular los cambios a polares, cilíndricas o esféricas.

A partir de ahora nos centramos, principalmente, en el caso $n = m = 3$.

9.1 Rotacional y divergencia. Potencial

Un *campo de vectores* o *campo vectorial* en \mathbb{R}^3 es una asignación de un “vector tangente” a cada punto del espacio \mathbb{R}^3 . Para nosotros una aplicación vectorial representa a un campo de vectores y frecuentemente usamos ambos nombres indistintamente. Los campos vectoriales se utilizan en física, para representar la velocidad y la dirección de un fluido en el espacio, o la intensidad y la dirección de fuerzas como la gravitatoria o la fuerza electromagnética.

Los libros de física e ingeniería están repletos de operadores diferenciales. El más importante es quizá el *operador diferencial vectorial* ∇ definido como

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

que “actúa” u “opera” sobre campos vectoriales. El operador diferencial ∇ puede actuar de dos formas sobre un campo vectorial $\mathbf{F} = (F_1, F_2, F_3)$:

- La *divergencia* $\text{div}(\mathbf{F}) = \nabla \cdot \mathbf{F}$ se define como

$$\text{div}(\mathbf{F}) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

- El *rotacional* $\text{rot}(\mathbf{F}) = \nabla \times \mathbf{F}$ se define como

$$\begin{aligned} \text{rot}(\mathbf{F}) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \\ &= \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right). \end{aligned}$$

Usaremos a veces por comodidad la notación $\frac{\partial}{\partial x} = \partial_x$, $\frac{\partial}{\partial y} = \partial_y$ y $\frac{\partial}{\partial z} = \partial_z$

Observación: La divergencia es una función escalar (devuelve un escalar), mientras que el rotacional es de nuevo un campo vectorial (devuelve un vector).

- **Ejemplo 9.1** Calcula el rotacional y la divergencia de la aplicación vectorial

$$\mathbf{F}(x, y, z) = (2y + z^2, y^2 + x, xyz)$$

Necesitamos para ello las derivadas parciales de cada una de las componentes de $\mathbf{F}(x, y, z)$. Esto es, si escribimos $F_1 = 2y + z^2$, $F_2 = y^2 + x$ y $F_3 = xyz$, la *matriz Jacobiana* es

$$J_{\mathbf{F}} = \begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \\ \frac{\partial F_3}{\partial x} & \frac{\partial F_3}{\partial y} & \frac{\partial F_3}{\partial z} \end{pmatrix} = \begin{pmatrix} 0 & 2 & 2z \\ 1 & 2y & 0 \\ yz & xz & xy \end{pmatrix}.$$

De esta manera tenemos que

$$\text{Div}(\mathbf{F})(x, y, z) = y(2 + x)$$

$$\text{y Rot}(\mathbf{F})(x, y, z) = (xz, 2z - yz, -1).$$

Una propiedad muy importante del rotacional es que es capaz de detectar los gradientes, por lo menos cerca de los puntos en los que estamos mirando. Más precisamente, consideremos un campo vectorial \mathbf{F} . Tenemos que $\text{rot}(\mathbf{F}) = \vec{0}$ si y solamente si existe una función escalar $F(x, y, z)$ tal que $\mathbf{F} = \nabla F$. Esta función escalar se suele llamar *potencial*. Los campos que cumplen esta propiedad se llaman *campos conservativos*. De hecho, en física, las fuerzas llamadas conservativas son precisamente aquellas dadas por el gradiente de una energía potencial (conservan la energía potencial).

■ **Ejemplo 9.2** Comprueba que el rotacional un gradiente es nulo.

Recordemos que el gradiente de una función $F(x, y, z)$ está definido como $\nabla F = (F_x, F_y, F_z)$. Ahora para calcular el rotacional tenemos que

$$\begin{aligned} \text{Rot}(\nabla F) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} = \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z}, \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x}, \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right) = \\ &= (F_{yz} - F_{zy}, F_{xz} - F_{zx}, F_{xy} - F_{yx}) = (0, 0, 0) \end{aligned}$$

■ **Ejemplo 9.3 — Campo conservativo.** Comprueba que el campo vectorial

$$\mathbf{F} = (3x^2 + yz, 2yz + xz, y^2 + xy + 1)$$

es un campo conservativo.

Debemos ver que $\text{Rot}(\mathbf{F}) = (0, 0, 0)$. Tenemos que

$$\begin{aligned} \text{Rot}(\mathbf{F}) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 3x^2 + yz & 2yz + xz & y^2 + xy + 1 \end{vmatrix} = \\ &= ((2y + x) - (2y + x), y - y, z - z) = (0, 0, 0). \end{aligned}$$

Como ya hemos dicho, cuando el rotacional de un campo o aplicación vectorial $\mathbf{F} = (F_1, F_2, F_3)$ es cero, es porque el campo es un gradiente. Es interesante saber reconstruir un potencial para un campo conservativo. Esto es, buscamos una función escalar F tal que

$$F_1 = \frac{\partial F}{\partial x}, F_2 = \frac{\partial F}{\partial y}, F_3 = \frac{\partial F}{\partial z}$$

Para ello, hacemos lo siguiente:

1. Integramos F_1 respecto a la variable x , asumiendo que y y z actúan como constantes, esto es, escribiremos

$$F(x, y, z) = \int F_1(x, y, z) dx = H_1(x, y, z) + G_1(y, z).$$

Esto es, la constante de integración depende de las variables y y z .

2. Consideramos la igualdad

$$F_2(x, y, z) = \frac{\partial F}{\partial y}(x, y, z) = \frac{\partial H_1}{\partial y}(x, y, z) + \frac{\partial G_1}{\partial y}(y, z).$$

De aquí nos queda una condición $\frac{\partial G_1}{\partial y}(y, z) = L_2(y, z)$, que integrando respecto a y nos proporciona una igualdad $G_1(y, z) = G_2(y, z) + H_2(z)$.

3. Por último igualamos

$$F_3(x, y, z) = \frac{\partial F}{\partial z}(x, y, z) = \frac{\partial H_1}{\partial z}(x, y, z) + \frac{\partial G_2}{\partial z}(y, z) + H_2'(z),$$

lo cual nos permite determinar (salvo constante) la función $H_2(z)$.

Veamos este proceso mediante un ejemplo:

■ **Ejemplo 9.4 — Reconstrucción de un potencial.** Consideremos la fuerza conservativa del ejemplo 9.3 y calculemos un potencial para ella.

1. Comenzamos igualando

$$F(x, y, z) = \int (3x^2 + yz) dx = x^3 + xyz + G_1(y, z), \quad (H_1(x, y, z) = x^3 + xyz)$$

2. Ahora tenemos que

$$F_2(x, y, z) = 2yz + xz = xz + G_{1y}, \quad (L_2(y, z) = 2yz)$$

y de esta manera $G_1(y, z) = \int 2yz dy = y^2 z + H_2(z)$ (con $G_2(y, z) = y^2 z$).

3. Finalmente igualamos

$$F_3(x, y, z) = y^2 + xy + 1 = xy + y^2 + H_2'(z),$$

y de aquí obtenemos que $H_2'(z) = 1$ y por tanto $H_2(z) = z + c$, con $c \in \mathbb{R}$. Así, tenemos que

$$F(x, y, z) = x^3 + xyz + y^2 z + z + c, \quad c \in \mathbb{R}$$

es un potencial para \mathbf{F} . En efecto, podemos comprobar que $\mathbf{F} = \nabla F$.

Lo dicho en dimensión 3 se puede aplicar para un campo en dos dimensiones $\mathbf{F} = (F_1, F_2)$ actuando como que $F_3 = 0$. Por ejemplo, tenemos que

$$\text{rot}(\mathbf{F}) = \left(0, 0, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right).$$

La condición para que un campo sea conservativo se reduce a que esta diferencia de derivadas parciales sea nula.

El *laplaciano* ΔF de un potencial F es la divergencia de su gradiente, esto es, la función escalar dada por

$$\Delta F = \text{div}(\nabla F) = \frac{\partial F}{\partial x^2} + \frac{\partial F}{\partial y^2} + \frac{\partial F}{\partial z^2}.$$

El laplaciano aparece en muchos problemas de la física matemática. En particular cuando se resuelve el átomo de hidrógeno a través de la ecuación de Schrödinger y se calculan los orbitales atómicos.

9.2 Integrales de línea. Trabajo

En física es natural introducir una manera de integrar que indique, en cierto modo, qué cantidad del campo pasa a través de una curva o una superficie: las integrales de línea y de superficie. Si $\mathbf{F} = (F_1, F_2, F_3)$ es un campo vectorial en el espacio, entonces una partícula

(por ejemplo una unidad de carga en un campo eléctrico o una unidad de masa en un campo gravitacional) experimentará la fuerza \mathbf{F} . Un concepto fundamental es el de *trabajo realizado por \mathbf{F} a lo largo de la trayectoria γ* .

Supongamos que tenemos una curva parametrizada $\gamma: [a, b] \rightarrow \mathbb{R}^3$. Esto es, γ asigna a cada valor $a \leq t \leq b$ un punto $\gamma(t) = (x(t), y(t), z(t))$, y describe una manera de recorrer cierta curva en el espacio. Se define la *integral de línea de una aplicación vectorial \mathbf{F} a lo largo de la trayectoria γ* como

$$\int_{\gamma} \mathbf{F} d\gamma = \int_a^b \left(F_1(\gamma(t))x'(t) + F_2(\gamma(t))y'(t) + F_3(\gamma(t))z'(t) \right) dt.$$

Lo primero que debemos observar es que, en general, el valor de la integral de línea depende de la curva del espacio que estamos recorriendo, y no solo de los extremos de la curva. En cambio, de lo que sí no depende el valor de la integral de línea es de la manera en cómo se recorre una curva, siempre que se fije una orientación, o lo que es lo mismo que se especifique desde qué punto hacia qué punto vamos.

■ **Ejemplo 9.5** Calculemos el trabajo ejercido por la fuerza $\mathbf{F} = (x, y, yz)$ a lo largo de los siguientes caminos que empiezan en el punto $P_0 = (0, 0, 0)$ y terminan en el punto $P_1 = (1, 1, 1)$.

- El segmento S de recta que los une. La parametrización más natural del segmento S con la orientación fijada es $\gamma(t) = (t, t, t)$, con la t variando de 0 a 1. Así, tenemos que

$$\int_{\gamma} \mathbf{F} d\gamma = \int_0^1 \langle (t, t, t^2), (1, 1, 1) \rangle dt = \int_0^1 (2t + t^2) dt = \left[t^2 + \frac{t^3}{3} \right]_0^1 = 1 + \frac{1}{3} = \frac{4}{3}.$$

También la parametrización $\gamma(t) = (t^3, t^3, t^3)$, con $0 < t < 1$ recorre el segmento S . Tenemos que

$$\int_0^1 \langle (t^3, t^3, t^6), (3t^2, 3t^2, 3t^2) \rangle dt = \int_0^1 (6t^5 + 3t^8) dt = \left[t^6 + \frac{t^9}{3} \right]_0^1 = 1 + \frac{1}{3} = \frac{4}{3}.$$

como ya sabíamos que debía quedar.

- La parábola P parametrizada por $\gamma(t) = (t^2, t^2, t)$, con $0 < t < 1$. Tenemos

$$\int_{\gamma} \mathbf{F} d\gamma = \int_0^1 \langle (t^2, t^2, t^3), (2t, 2t, 1) \rangle dt = \int_0^1 (2t^3 + 2t^3 + t^3) dt = \left[\frac{5t^4}{4} \right]_0^1 = \frac{5}{4}.$$

También la parametrización $\gamma(t) = (\sin^2 t, \sin^2 t, \sin t)$, con $0 < t < \pi/2$ recorre la parábola P . Tenemos que

$$\begin{aligned} & \int_0^1 \langle (\sin^2 t, \sin^2 t, \sin^3 t), (2 \cos t \sin t, 2 \cos t \sin t, \cos t) \rangle dt = \\ & = \int_0^1 (2 \sin^3 t \cos t + 2 \sin^3 t \cos t + \sin^3 t \cos t) dt = \left[\frac{5 \sin^4 t}{4} \right]_0^1 = \frac{5}{4}. \end{aligned}$$

Al cambiar de orientación, cambia el signo del valor de la integral como observamos en el siguiente ejemplo.

■ **Ejemplo 9.6** Si considero el trabajo que realiza la fuerza $\mathbf{F} = (xy, x, yz)$ a lo largo de la trayectoria $\gamma(t) = (t^3, t, 1)$, cuando t varía entre 0 y 1, obtengo que

$$\int_{\gamma} \mathbf{F} d\gamma = \int_0^1 \langle (t^4, t^3, t), (3t^2, 1, 0) \rangle dt = \int_0^1 (3t^6 + t^3) dt = \left[\frac{3t^7}{7} + \frac{t^4}{4} \right]_0^1 = \frac{3}{7} + \frac{1}{4} = \frac{19}{28}$$

Notemos que este camino comienza en el punto $(0, 0, 1)$ y termina en $(1, 1, 1)$. Si ahora tomamos el camino $(-t^3, -t, 1)$, con t variando entre -1 y 0 , observamos que la curva que recorre es la misma que anteriormente, pero ahora el punto de partida es $(1, 1, 1)$ y el de llegada $(0, 0, 1)$. Así, nos queda que

$$\begin{aligned} \int_{\gamma} \mathbf{F} d\gamma &= \int_{-1}^0 \langle (t^4, -t^3, -t), (-3t^2, -1, 0) \rangle dt = \int_{-1}^0 (-3t^6 + t^3) dt = \left[\frac{-3t^7}{7} + \frac{t^4}{4} \right]_{-1}^0 = \\ &= - \left(\frac{3}{7} + \frac{1}{4} \right) = -\frac{19}{28} \end{aligned}$$

9.2.1 El caso de un campo conservativo

Consideremos un campo $\mathbf{F} = \nabla F$. El *teorema de gradiente*, también conocido como *teorema fundamental de cálculo para integrales de línea*, nos dice que la integral de línea de \mathbf{F} a lo largo de una trayectoria $\gamma: [a, b] \rightarrow \mathbb{R}^3$ es

$$\int_{\gamma} \mathbf{F} d\gamma = F(\gamma(b)) - F(\gamma(a)).$$

Es decir, este valor no depende de la curva a lo largo de la cual integramos, sino solamente de sus extremos. En particular, si la trayectoria es cerrada (ambos extremos coinciden), el valor de la integral de línea es cero.

■ **Ejemplo 9.7** Calcula la integral del campo $\mathbf{F} = (2x, 2y, 2z)$ a lo largo de la circunferencia C de radio 1 y centrada en el origen contenida en el plano $y = z$.

Observamos que el campo es conservativo, de hecho tenemos que $\mathbf{F} = \nabla F$, con $F(x, y, z) = x^2 + y^2 + z^2$. De esta manera, para cualquier parametrización de C , tenemos que

$$\int_{\gamma} \mathbf{F} d\gamma = 0,$$

pues C es una curva cerrada. Si tomamos una parametrización, por ejemplo la dada por $\gamma(t) = (\cos t, \sin t, \sin t)$, con $0 < t < 2\pi$, y por tanto

$$\gamma(0) = \gamma(2\pi) = (1, 0, 0), \quad (\text{trayectoria cerrada}),$$

vemos que

$$\int_0^{2\pi} \langle (2 \cos t, 2 \sin t, 2 \sin t), (-\sin t, \cos t, \cos t) \rangle dt = \int_0^{2\pi} 2 \cos t \sin t dt = \sin^2 t \Big|_0^{2\pi} = 0.$$